



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁷ : C07H 21/04, C12N 1/21, 15/00, 15/63, C07K 14/705, G01N 33/53, C12Q 1/68	A1	(11) International Publication Number: WO 00/26227 (43) International Publication Date: 11 May 2000 (11.05.00)
(21) International Application Number: PCT/US99/25178 (22) International Filing Date: 27 October 1999 (27.10.99) (30) Priority Data: 09/183,175 30 October 1998 (30.10.98) US (63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Application US 09/183,175 (CON) Filed on 30 October 1998 (30.10.98) (71) Applicant (for all designated States except US): MILLENNIUM PHARMACEUTICALS, INC. [US/US]; 75 Sidney Street, Cambridge, MA 02139 (US). (72) Inventor; and (75) Inventor/Applicant (for US only): MCCARTHY, Sean, A. [GB/US]; 62 Commonwealth Avenue, #4, Boston, MA 02116 (US). (74) Agents: MEIKLEJOHN, Anita, L. et al.; Fish & Richardson P.C., 225 Franklin Street, Boston, MA 02110-2804 (US).	(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, 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, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, 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), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG). Published <i>With international search report.</i>	
(54) Title: LDL RELATED PROTEIN AND USES THEREOF		
(57) Abstract <p>The invention provides isolated nucleic acid molecules, designated TANGO 136 nucleic acid molecules, which encode transmembrane proteins with homology to members of the LDL receptor family of proteins. The invention also provides antisense nucleic acid molecules, recombinant expression vectors containing the nucleic acid molecules of the invention, host cells into which the expression vectors have been introduced, and non-human transgenic animals in which a nucleic acid molecule of the invention has been introduced or disrupted. The invention still further provides isolated polypeptides, fusion polypeptides, antigenic peptides and antibodies. Diagnostic, screening, and therapeutic methods utilizing compositions of the invention are also provided.</p>		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

- 1 -

LDL RELATED PROTEIN AND USES THEREOFBACKGROUND OF THE INVENTION

5 Members of the low density lipoprotein (LDL) receptor gene family include: the LDL receptor, the very low density lipoprotein (VLDL) receptor, LDL receptor-related protein 2 (LRP-2/megalin/gp330), LDL receptor-related protein 3 (LRP-3/LRp-105), LDL receptor-related
10 protein 5 (LRP-5), LDL receptor-related protein 6 (LRP-6), and LR8B

Common to many members of the LDL receptor family is the endocytic uptake of ligands. The ligands bound by the members of the LDL receptor family are diverse. For
15 example, LDL receptor is thought to bind lipoproteins containing apoB or apoE; VLDL receptor is thought to bind lipoproteins containing apoE; and LRP-2 is thought to bind apoE-containing lipoproteins, plasminogen, lactoferrin, and lipoprotein lipase. Significantly, LRP-
20 2 is also thought to mediate the uptake of complexes of clusterin/ApoJ and the amyloid beta protein.

Summary of the Invention

The present invention is based, at least in part, on the discovery of cDNA molecules encoding TANGO 136, a
25 type I membrane protein. Described below are cDNA molecules encoding both human and murine TANGO 136. These proteins, fragments, derivatives, and variants thereof are collectively referred to as "polypeptides of the invention" or "proteins of the invention." Nucleic
30 acid molecules encoding polypeptides of the invention are collectively referred to as "nucleic acids of the invention."

The nucleic acids and polypeptides of the present invention are useful as modulating agents in regulating a

- 2 -

variety of cellular processes. Accordingly, in one aspect, this invention provides isolated nucleic acid molecules encoding a polypeptide of the invention or a biologically active portion thereof. The present
5 invention also provides nucleic acid molecules which are suitable as primers or hybridization probes for the detection of nucleic acids encoding a polypeptide of the invention.

The invention features nucleic acid molecules
10 which are at least 45% (or 55%, 65%, 75%, 85%, 95%, or 98%) identical to the nucleotide sequence of SEQ ID NO:1, 3, 4 or 6, or the nucleotide sequence of the cDNA insert of a clone deposited with ATCC as Accession Number 98880 (the "cDNA of ATCC 98880"), or a complement thereof.

15 The invention features nucleic acid molecules which include a fragment of at least 300 (325, 350, 375, 400, 425, 450, 500, 550, 600, 650, 700, 800, 900, 1000, or 1200) nucleotides of the nucleotide sequence of SEQ ID NO:1, 3, 4, or 6, or the nucleotide sequence of the cDNA
20 of ATCC Accession Number 98880, or a complement thereof.

The invention also features nucleic acid molecules which include a nucleotide sequence encoding a protein having an amino acid sequence that is at least 45% (or 55%, 65%, 75%, 85%, 95%, or 98%) identical to the amino
25 acid sequence of SEQ ID NO:2 or 5, or the amino acid sequence encoded by the cDNA of ATCC Accession Number 98880, or a complement thereof.

In various embodiments, the nucleic acid molecules have the nucleotide sequence of SEQ ID NO:1, 3, 4, or 6,
30 or the nucleotide sequence of the cDNA of ATCC Accession Number 98880.

Also within the invention are nucleic acid molecules which encode a fragment of a polypeptide having the amino acid sequence of SEQ ID NO:2 or 5, the fragment
35 including at least 15 (25, 30, 50, 100, 150, 300, or 400)

- 3 -

contiguous amino acids of SEQ ID NO:2 or 5 or the amino acid sequence encoded by the cDNA of ATCC Accession Number 98880.

The invention includes nucleic acid molecules
5 which encode a naturally occurring allelic variant of a polypeptide comprising the amino acid sequence of SEQ ID NO:2 or 5 or the amino acid sequence encoded by the cDNA of ATCC Accession Number 98880, wherein the nucleic acid molecule hybridizes to a nucleic acid molecule comprising
10 the nucleic acid sequence encoding SEQ ID NO:2 or 5, or a complement thereof under stringent conditions.

Also within the invention are isolated polypeptides or proteins having an amino acid sequence that is at least about 65%, preferably 75%, 85%, 95%, or
15 98% identical to the amino acid sequence of SEQ ID NO:2 or 5.

Also within the invention are isolated polypeptides or proteins which are encoded by a nucleic acid molecule having a nucleotide sequence that is at
20 least about 65%, preferably 75%, 85%, or 95% identical the nucleic acid sequence encoding SEQ ID NO:2 or 5 and isolated polypeptides or proteins which are encoded by a nucleic acid molecule having a nucleotide sequence which hybridizes under stringent hybridization conditions to a
25 nucleic acid molecule having the nucleotide sequence of SEQ ID NO:1, 3, 4, or 6, or a complement thereof or the non-coding strand of the cDNA of ATCC Accession Number 98880.

Also within the invention are polypeptides which
30 are naturally occurring allelic variants of a polypeptide that includes the amino acid sequence of SEQ ID NO:2 or 5 or the amino acid sequence encoded by the cDNA of ATCC Accession Number 98880, wherein the polypeptide is encoded by a nucleic acid molecule which hybridizes to a

- 4 -

nucleic acid molecule comprising SEQ ID NO:1, 3, 4, or 6, or a complement thereof under stringent conditions.

The invention also features nucleic acid molecules that hybridize under stringent conditions to a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO: 1, 3, 4, or 6, or the cDNA of ATCC Accession Number 98880, or a complement thereof. In other embodiments, the nucleic acid molecules are at least 300 (325, 350, 375, 400, 425, 450, 500, 550, 600, 650, 700, 10 800, 900, 1000, or 1290) nucleotides in length and hybridize under stringent conditions to a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1, 3, 4, or 6, or the cDNA of ATCC Accession Number 98880, or a complement thereof.

15 In other embodiments, the isolated nucleic acid molecules encode one or more domains, e.g., a cytoplasmic domain (SEQ ID NO:11 or 18), a transmembrane domain (SEQ ID NO:10 or 17), or an extracellular domain (SEQ ID NO:9 or 16), of a polypeptide of the invention or a complement thereof. In another embodiment, the invention provides an isolated nucleic acid molecule which is antisense to the coding strand of a nucleic acid of the invention.

Another aspect of the invention provides vectors, e.g., recombinant expression vectors, comprising a 25 nucleic acid molecule of the invention. In another embodiment, the invention provides host cells containing such a vector. The invention also provides methods for producing a polypeptide of the invention by culturing, in a suitable medium, a host cell of the invention 30 containing a recombinant expression vector such that a polypeptide is produced.

Another aspect of the invention features isolated or recombinant proteins and polypeptides of the invention. Preferred proteins and polypeptides possess 35 at least one biological activity possessed by the

- 5 -

corresponding naturally-occurring human polypeptide. An activity, a biological activity, and a functional activity of a polypeptide of the invention refers to an activity exerted by a protein, polypeptide or nucleic acid molecule of the invention on a responsive cell as determined *in vivo*, or *in vitro*, according to standard techniques. Such activities can be a direct activity, such as an association with or an enzymatic activity on a second protein or an indirect activity, such as a cellular signaling activity mediated by interaction of the protein with a second protein. Thus, such activities include, e.g., (1) the ability to form protein-protein interactions with proteins in the signaling pathway of the naturally-occurring polypeptide; (2) the ability to bind a ligand of the naturally-occurring polypeptide; (3) the ability to bind to an intracellular target of the naturally-occurring polypeptide. Other activities include: (1) the ability to modulate cellular proliferation; (2) the ability to modulate cellular differentiation; and (3) the ability to modulate cell death.

In one embodiment, a polypeptide of the invention has an amino acid sequence sufficiently identical to an identified domain of a polypeptide of the invention. As used herein, the term "sufficiently identical" refers to a first amino acid or nucleotide sequence which contains a sufficient or minimum number of identical or equivalent (e.g., with a similar side chain) amino acid residues or nucleotides to a second amino acid or nucleotide sequence such that the first and second amino acid or nucleotide sequences have a common structural domain and/or common functional activity. For example, amino acid or nucleotide sequences which contain a common structural domain having about 65% identity, preferably 75%

- 6 -

identity, more preferably 85%, 95%, or 98% identity are defined herein as sufficiently identical.

In one embodiment, the isolated polypeptides lack both a transmembrane and a cytoplasmic domain. In
5 another embodiment the polypeptide lacks both a transmembrane domain and a cytoplasmic domain and is soluble under physiological conditions.

The polypeptides of the present invention, or biologically active portions thereof, can be operably
10 linked to a heterologous amino acid sequence to form fusion proteins. The invention further features antibodies that specifically bind a polypeptide of the invention such as monoclonal or polyclonal antibodies. In addition, the polypeptides of the invention or
15 biologically active portions thereof can be incorporated into pharmaceutical compositions, which optionally include pharmaceutically acceptable carriers.

In another aspect, the present invention provides methods for detecting the presence of the activity or
20 expression of a polypeptide of the invention in a biological sample by contacting the biological sample with an agent capable of detecting an indicator of activity such that the presence of activity is detected in the biological sample.

In another aspect, the invention provides methods
25 for modulating activity of a polypeptide of the invention comprising contacting a cell with an agent that modulates (inhibits or stimulates) the activity or expression of a polypeptide of the invention such that activity or
30 expression in the cell is modulated. In one embodiment, the agent is an antibody that specifically binds to a polypeptide of the invention.

In another embodiment, the agent modulates
expression of a polypeptide of the invention by
35 modulating transcription, splicing, or translation of an

- 7 -

mRNA encoding a polypeptide of the invention. In yet another embodiment, the agent is a nucleic acid molecule having a nucleotide sequence that is antisense to the coding strand of an mRNA encoding a polypeptide of the invention.

5 The present invention also provides methods to treat a subject having a disorder characterized by aberrant activity of a polypeptide of the invention or aberrant expression of a nucleic acid of the invention
10 by administering an agent which is a modulator of the activity of a polypeptide of the invention or a modulator of the expression of a nucleic acid of the invention to the subject. In one embodiment, the modulator is a
15 protein of the invention. In another embodiment, the modulator is a nucleic acid of the invention. In other embodiments, the modulator is a peptide, peptidomimetic, or other small molecule.

The present invention also provides diagnostic assays for identifying the presence or absence of a
20 genetic lesion or mutation characterized by at least one of: (i) aberrant modification or mutation of a gene encoding a polypeptide of the invention, (ii) mis-regulation of a gene encoding a polypeptide of the invention, and (iii) aberrant post-translational
25 modification of the invention wherein a wild-type form of the gene encodes a protein having the activity of the polypeptide of the invention.

In another aspect, the invention provides a method for identifying a compound that binds to or modulates the
30 activity of a polypeptide of the invention. In general, such methods entail measuring a biological activity of the polypeptide in the presence and absence of a test compound and identifying those compounds which alter the activity of the polypeptide.

- 8 -

The invention also features methods for identifying a compound which modulates the expression of a polypeptide or nucleic acid of the invention by measuring the expression of the polypeptide or nucleic acid in the presence and absence of the compound.

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

Brief Description of the Drawings

10 Figure 1 depicts a partial cDNA sequence (SEQ ID NO:1) and predicted partial amino acid sequence (SEQ ID NO:2) of murine TANGO 136. The open reading frame of SEQ ID NO:1 extends from nucleotide 89 to nucleotide 1813 (SEQ ID NO:3).

15 Figure 2 depicts a hydropathy plot of a portion of murine TANGO 136. Relatively hydrophobic residues are above the horizontal line, and relatively hydrophilic residues are below the horizontal line. The cysteine residues (cys) and potential N-glycosylation sites (Ngly) 20 are indicated by short vertical lines just below the hydropathy trace.

Figure 3 depicts the cDNA sequence (SEQ ID NO:4) and predicted amino acid sequence (SEQ ID NO:5) of human TANGO 136. The open reading frame of SEQ ID NO:4 extends 25 from nucleotide 541 to 2679 (SEQ ID NO:6).

Figure 4 depicts a hydropathy plot of human TANGO 136. Relatively hydrophobic residues are above the horizontal line, and relatively hydrophilic residues are below the horizontal line. The cysteine residues (cys) 30 and potential N-glycosylation sites (Ngly) are indicated by short vertical lines just below the hydropathy trace.

Figure 5 depicts an alignment of the amino acid sequences of murine TANGO 136 (partial sequence; SEQ ID

- 9 -

NO:2), human TANGO 136 (SEQ ID NO:5), human LRp105 (SEQ ID NO:11) and rat LRp105 (SEQ ID NO:13).

Figure 6 depicts an alignment of the nucleic acid sequences of murine TANGO 136 (partial sequence; SEQ ID NO:1) and human TANGO 136 (SEQ ID NO:4).

Figure 7 depicts an alignment of the amino acid sequences of murine TANGO 136 (partial sequence; SEQ ID NO:2) and human TANGO 136 (SEQ ID NO:5).

Figure 8 depicts alignments of the CUB-like domains of murine TANGO 136 with a consensus CUB domain. In these alignments an uppercase letter between the two sequences indicates an exact match, and a "+" indicates a similarity.

Figure 9 depicts alignments of the CUB-like domains of human TANGO 136 with a consensus CUB domain. In these alignments an uppercase letter between the two sequences indicates an exact match, and a "+" indicates a similarity.

Figure 10 depicts alignments of the LDL class A domains of human TANGO 136 with a consensus LDL class A domain. In these alignments an uppercase letter between the two sequences indicates an exact match, and a "x" indicates a similarity.

Detailed Description of the Invention

The present invention is based on the discovery of cDNA molecules encoding murine and human TANGO 136, a transmembrane protein. Various features of murine and human TANGO 136 are summarized below.

Murine TANGO 136

A cDNA encoding a portion of murine TANGO 136 was identified using a screening process which selects for nucleotide sequences which encode secreted proteins. A detailed description of this method, called "signal

- 10 -

trapping" is provided in PCT Publication No. WO 98/22491, published May 28, 1998. In brief, a randomly primed cDNA library was prepared using cDNA prepared from mRNA extracted from lipopolysaccharide-stimulated mouse
5 macrophages. To prepare this library, the cDNA was inserted into the mammalian expression vector pMEAP adjacent to a cDNA encoding placental alkaline phosphatase which lacks a secretory signal. Next, the cDNA library was amplified in bacteria. The amplified
10 cDNA was then isolated and transfected into human 293T cells. After 28 hours, cell supernatants were collected and assayed for alkaline phosphatase activity. Clones giving rise to detectable alkaline phosphatase activity in the supernatant of transfected cells were isolated and
15 analyzed further by sequencing and the novel clones subjected to further sequencing

Once such clone, murine TANGO 136, was identified. This clone includes a 1813 nucleotide cDNA (Figure 1; SEQ ID NO:1). The open reading frame of this cDNA,
20 nucleotides 89 to 1813 (SEQ ID NO:3), encodes a 575 amino acid putative type I membrane protein (Figure 1; SEQ ID NO:2). Because no translation stop codon occurs at the end of the open reading frame, this cDNA is likely to be a partial cDNA which does not encode the most carboxy
25 terminal portion of murine TANGO 136.

The signal peptide prediction program SIGNALP (Nielsen et al. (1997) *Protein Engineering* 10:1-6) predicted that murine TANGO 136 includes a 17 amino acid signal peptide (amino acid 1 to about amino acid 17 of
30 SEQ ID NO:2; SEQ ID NO:7) preceding the 558 amino acid (partial) mature protein (about amino acid 18 to amino acid 575; SEQ ID NO:8). Mature murine TANGO 136 has an extracellular domain (amino acids 18 to 441 of SEQ ID NO:2; SEQ ID NO:9); a transmembrane domain (amino acids
35 442 to 462 of SEQ ID NO:2; SEQ ID NO:10); and a

- 11 -

cytoplasmic domain (amino acids 463 to 575 of SEQ ID NO:2; SEQ ID NO:11).

The extracellular region of murine TANGO 136 includes two CUB-like domains (amino acids 32 to 86 and amino acids 193 to 306 of SEQ ID NO:2). CUB domains are extracellular domains found in a number of functionally diverse, developmentally regulated proteins including the dorsal-ventral patterning protein tolloid, bone morphogenetic protein 1, a family of spermadhesins, complement subcomponents Cls/Clr and the neuronal recognition molecule A5. The majority of CUB domains contain four conserved cysteines which are thought to form two disulfide bridges (C1-C2 and C3-C4) (Bork et al. (1993) *J. Mol. Biol.* 231:539-545). The first CUB-like domain of murine TANGO 136 (amino acids 32 to 86) includes two cysteines, and the second CUB-like domain of murine TANGO 136 (amino acids 193 to 306) includes two cysteines. Alignments of the CUB-like domains of murine TANGO 136 with a CUB domain consensus sequence are depicted in Figure 8.

Figure 2 depicts a hydropathy plot of a portion of murine TANGO 136. Relatively hydrophobic residues are above the horizontal line, and relatively hydrophilic residues are below the horizontal line. The cysteine residues (cys) and potential N-glycosylation sites (Ngly) are indicated by short vertical lines just below the hydropathy trace.

Human TANGO 136

The murine TANGO 136 cDNA described above was used to screen a human placental cDNA library to identify human clones encoding TANGO 136. One clone identified by this screening was sequenced fully. This human TANGO 136 cDNA (Figure 4; SEQ ID NO:4) includes an open reading frame (nucleotides 541 to 2679 of SEQ ID NO:4; SEQ ID

- 12 -

NO:6) encoding a 713 amino acid putative type I transmembrane protein (Figure 4; SEQ ID NO:5).

The signal peptide prediction program SIGNALP (Nielsen et al. (1997) *Protein Engineering* 10:1-6) predicted that human TANGO 136 includes a 16 amino acid signal peptide (amino acid 1 to about amino acid 16 of SEQ ID NO:5; SEQ ID NO:14) preceding the 697 amino acid mature protein (about amino acid 17 to amino acid 713 of SEQ ID NO:5; SEQ ID NO:15). Human TANGO 136 has an extracellular domain (amino acids 17 to 440 of SEQ ID NO:5; SEQ ID NO:16); a transmembrane domain (amino acids 441 to 461 of SEQ ID NO:5; SEQ ID NO:17); and a cytoplasmic domain (amino acids 462 to 713 of SEQ ID NO:5; SEQ ID NO:18).

A clone, pT136, which encodes human TANGO 136 was deposited with the American Type Culture Collection (10801 University Boulevard, Manassas, VA 20110-2209) on September 11, 1998 and assigned Accession Number 98880. This deposit will be maintained under the terms of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure. This deposit was made merely as a convenience for those of skill in the art and is not an admission that a deposit is required under 35 U.S.C. §112.

The extracellular region of human TANGO 136 includes two CUB-like domains (amino acids 31 to 136 and amino acids 192 to 305 of SEQ ID NO:5). Both of the CUB-like domains of human TANGO 136 include two cysteines. Alignments of the CUB-like domains of human TANGO 136 with a CUB domain consensus sequence are depicted in Figure 9.

The extracellular region of human TANGO 136 also includes four LDL receptor class A domains (amino acids 138 to 176, amino acids 328 to 355; amino acids 380 to 398; and amino acids 399 to 435 of SEQ ID NO:5). The

- 13 -

LDL receptor class A domain is an approximately 40 amino acid cysteine-rich domain having a found in LDL receptor and other members of the LDL receptor family. Repeats of this domain are thought to involved in ligand binding
5 (Yamamoto et al. (1984) *Cell* 39:27-38; and Fass et al. (1997) *Nature* 388:691-693). The LDL receptor class A domain extending from amino acid 380 to 398 of human TANGO 136 has relatively weak homology to the consensus LDL receptor type A domain compared to the other three
10 LDL receptor class A domains. Alignments of the LDL receptor class A domains of human TANGO 136 with a LDL receptor class A domain consensus sequence are depicted in Figure 10.

Figure 4 depicts a hydropathy plot of human TANGO
15 136. Relatively hydrophobic residues are above the horizontal line, and relatively hydrophilic residues are below the horizontal line. The cysteine residues (cys) and potential N-glycosylation sites (Ngly) are indicated by short vertical lines just below the hydropathy trace.

20 Mature human TANGO 136 has a predicted MW of 76.7 kDa (78.4 kDa for immature human TANGO 136), not including post-translational modifications.

Human TANGO 136 maps to chromosome 14 near D14S283.

25 The amino acid sequence of human TANGO 136 (SEQ ID NO:5) was used to search public databases (using BLASTP; Altschul et al. (1990) *J. Mol. Biol.* 215:403-410) in order to identify proteins having homology to human TANGO 136. This analysis revealed that both murine and human
30 TANGO 136 has considerable homology to human LDL receptor related protein LRp105/LRP-3 (Ishii et al. (1998) *Genomics* 51:132-135). Figure 5 depicts an alignment of the amino acids sequences of murine TANGO 136, human TANGO 136, human LRp105/LRP-3, and rat Lrp105/LRP-3.

- 14 -

When compared using the algorithm of Myers and Miller ((1988) *CABIOS* 4:11-17; PAM120 scoring matrix, -12 gap opening penalty, -4 gap extension penalty) murine TANGO 136 is 34.4% identical to human LRp105/LRP-3 and
5 34% identical to rat LRp105/LRP-3; human TANGO 136 is 38% identical to human LRp105/LRP-3 and 37.6% identical to rat Lrp105/LRP-3; and human TANGO 136 is 72.6 identical to murine TANGO 136.

The full length human TANGO 136 nucleotide
10 sequence is 86.1% identical (FASTA version 2.0u53; Pearson and Lipman (1988) *Proc. Natl Acad. Sci. USA* 85:2444-2448) to the partial murine TANGO 136 nucleotide sequence (Figure 6). The full length human TANGO 136 amino acid sequence is 90.8% identical (FASTA version
15 2.0u53; Pearson and Lipman (1988) *Proc. Natl Acad. Sci.* 85:2444-2448) to the partial murine TANGO 136 amino acid sequence (Figure 7). As shown in Figure 7, the protein domain structure (described above) is highly conserved between the human and murine proteins.

20 Human multiple tissue northern (MTN) blots (Clontech, Palo Alto, CA), containing 2 μ g of poly A+ RNA per lane were probed with a murine TANGO 136 cDNA probe. This analysis revealed that TANGO 136 mRNA is relatively highly expressed in spleen, prostate, uterus,
25 peripheral blood leukocytes, heart, placenta, kidney and pancreas. This analysis also revealed that TANGO 136 mRNA is expressed at a somewhat lower level in thymus, testis, colon, lung, liver and skeletal muscle. TANGO 136 nucleic acids, polypeptides, agonists, and
30 antagonists can be used to modulate the activities of the tissues in which it is expressed and thus treat disorders of these tissues. For example, TANGO 136 is expressed in prostate and testis and may be involved in spermatogenesis.

- 15 -

Use of TANGO 136 Nucleic Acids, Polypeptides, and TANGO 136 Agonists or Antagonists

Due to the homology between TANGO 136 and Lrp105/LRP-3, TANGO 136 is predicted to be a member of the low density lipoprotein receptor family, which includes LDLR, LRP-2 (megalin/gp330), LRP-3 (Lrp105), LRP-5, LRP-6, and LR8B. Members of this family are endocytic receptors that bind and internalize ligands from the circulation and extracellular space. Since TANGO 136 is predicted to be a member of the low density lipoprotein receptor family, it may function similarly to other members of the low density lipoprotein receptor family.

LDLR binds plasma lipoproteins that contain apolipoprotein B-100 (apoB-100) or apoE on their surface. LDLR is critical for the uptake of these lipoproteins, and mutations in LDLR are the cause of familial hypercholesterolemia, a disorder characterized by high levels of cholesterol-rich LDL in the plasma. The elevation of plasma cholesterol levels in patients afflicted with familial hypercholesterolemia leads to atherosclerosis and increased risk for myocardial infarction. TANGO 136 potentially plays a role in disorders of lipoprotein metabolism and transport, e.g., cardiovascular diseases such as atherosclerosis. Accordingly, TANGO 136 nucleic acids, polypeptides and TANGO 136 antagonists and agonists are useful for treatment of disorders of lipoprotein metabolism and transport, e.g., cardiovascular diseases such as atherosclerosis.

In vitro studies have shown that LRP-2 is capable of binding and mediating the cellular uptake of a large number of different ligands including apoE-enriched very low density lipoproteins (Willnow et al. (1992) *J. Biol. Chem.* 267:26172-26180), complexes of urokinase

- 16 -

plasminogen activator and plasminogen activator inhibitor-1 (tPA:PAI-1) (Willnow et al., *supra*), lipoprotein lipase (Willnow et al., *supra*), and lactoferrin. A receptor associated protein known as RAP
5 (Orlando et al. (1992) *Proc. Natl Acad. Sci.* 89:6698-6702) inhibits the binding of these ligands to LRP-2. Some or all of these ligands may bind TANGO 136. Accordingly, TANGO 136 nucleic acids, polypeptides, antagonists and agonists are useful for treatment of
10 clotting disorders, e.g., inhibiting clot formation or dissolving clots.

A few specific and physiologically relevant ligands for LRP-2 have been identified, including apolipoprotein J(apoJ)/clusterin (Kounnas et al. (1995)
15 *J. Biol. Chem.* 22:13070-13075) and thyroglobulin (Zheng et al. (1998) *Endocrinology* 139:1462-1465). ApoJ has been reported to bind several proteins, including the β A4 peptide of the Alzheimer's precursor protein, a subclass of high density lipoprotein, and the complement membrane
20 attack complex C5-C9 (Kounnas et al., *supra*). The clearance of apoJ complexed with these and other molecules is expected to occur via LRP-2. Thus, LRP-2 may play an important functional role in the clearance of these complexes. For example, LRP-2 may function to
25 target lipoproteins for clearance or may inhibit the cytolytic activity of the complement membrane C5b-C9 by clearing the apoJ/C5b-C9 complex. The fact that LRP-2 can bind the apoJ/amyloid- β complex suggests that LRP-2 may be involved in regulating the pathogenesis of
30 Alzheimer's disease. A role for LRP-2 in Alzheimer's disease is further supported by another study that showed that LRP-2 may be involved in transporting the apoJ/amyloid β complex across the blood-brain-barrier (Zlokovic et al. (1996) *Proc. Natl Acad. Sci.* 93:4229-
35 4234). Thus, TANGO 136 nucleic acids, proteins,

- 17 -

agonists, and antagonists are useful for the treatment of Alzheimer's disease and other neurodegenerative disorders, e.g., Huntington's disease and Parkinson's disease.

5 LRP-2 is involved in participating in the endocytosis of thyroglobulin, which results in the release of thyroid hormones (Zheng et al. (1998) *Endocrinolgy* 139:1462-65). TANGO 136 may also be involved in the regulating the release of thyroid
10 hormones. Thus, TANGO 136 nucleic acids, proteins, agonists, and antagonists are useful for the treatment of thyroid disorders, e.g., thyroid hormone release disorders..

LRP-2 is also predicted to play a role as a drug
15 receptor and is thought to be involved in the uptake of polybasic drugs, e.g., aprotinin, aminoglycosides and polymyxin B. The uptake of polybasic drugs can be toxic, e.g., the administration of aminoglycosides is often associated with nephro- and ototoxicity. TANGO 136 may
20 also mediate uptake of polybasic drugs, and TANGO 136 nucleic acids, proteins, agonists, and antagonists are useful for the modulating the uptake of such drugs. TANGO 136 can also be used to design less toxic versions of such drugs.

25 In addition, LRP-2 is involved in the pathogenesis of Heymann Nephritis nephropathy (HN), an autoimmune glomerular disease, which is similar to human membranous nephropathy. It is thought that LRP-2 is the major pathogenic antigen and forms an antigen-antibody complex
30 between the glomular basement membrane and the foot processes of glomerular epithelial cells. The presence of the antigen-antibody complex leads to extensive damage of the basement membrane and proteinuria (Farquhar et al. (1994) *Ann. N.Y. Acad. Sci.* 97-106). Similar to LRP-2,
35 TANGO 136 may play a pathogenic role in autoimmune

- 18 -

glomerular disease. Thus, TANGO 136 nucleic acids, proteins, agonists, and antagonists are useful for the treatment of autoimmune glomerular disease.

LRP-5 and LRP-6 are thought to function in
5 endocytosis. Based on genetic evidence, LRP-5 and possibly LRP-6 are thought to play a role in the molecular pathogenesis of type I diabetes (Brown et al. (1998) *Biochem. Biophys. Res. Comm.* 248:879-888). TANGO 136 is also likely plays a role in type I diabetes.
10 Thus, TANGO 136 nucleic acids, proteins, agonists, and antagonists are useful for the treatment of type I diabetes.

LR8B is expressed in brain and might be involved in brain-specific lipid transport. Brain-specific lipid
15 transport may involve apoE4, which is associated with Alzheimer's disease. TANGO 136 may also be involved in brain-specific lipid transport, and TANGO 136 nucleic acids, proteins, agonists, and antagonists are useful for the treatment of Alzheimer's disease.

20 In general, TANGO 136 nucleic acids, proteins, agonists, and antagonists may be useful for the treatment of neurological disorders, e.g., neurodegenerative disorders and neuropsychiatric disorders. Examples of neurodegenerative disorders
25 include Alzheimer's disease, Parkinson's disease, and Huntington's disease. Examples of neuropsychiatric disorders include schizophrenia, attention deficit disorder, unipolar affective (mood) disorder, bipolar affective (mood) disorders (e.g., severe bipolar
30 affective disorder (BP-I) and bipolar affective disorder with hypomania and major depression (BP-II)), and schizoaffective disorders.

- 19 -

TABLE 1: Summary of Murine and Human TANGO 136 Sequence Information

Gene	cDNA	ORF	Protein	Figure	Accession No.
5 MURINE TANGO 136	SEQ ID NO:1	SEQ ID NO:3	SEQ ID NO:2	Fig. 1	
HUMAN TANGO 136	SEQ ID NO:4	SEQ ID NO:5	SEQ ID NO:5	Fig. 3	ATCC 98880

TABLE 2: Summary of Domains of Murine and Human TANGO 136 Protein

Protein	Signal Sequence	Mature Protein	Extracellular Domain	Transmembrane Domain	Cytoplasmic Domain
10 MURINE TANGO 136	aa 1-17	aa 18-575	aa 18-441	aa 442-462	aa 463-575
SEQ ID NO:2	SEQ ID NO:7	SEQ ID NO:8	SEQ ID NO:9	SEQ ID NO:10	SEQ ID NO:11
15 HUMAN TANGO 136	aa 1-16	aa 17-713	aa 17-440	aa 441-461	aa 462-713
SEQ ID NO:5	SEQ ID NO:14	SEQ ID NO:15	SEQ ID NO:16	SEQ ID NO:17	SEQ ID NO:18

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

I. Isolated Nucleic Acid Molecules

One aspect of the invention pertains to isolated nucleic acid molecules that encode a polypeptide of the invention or a biologically active portion thereof, as

- 20 -

well as nucleic acid molecules sufficient for use as hybridization probes to identify nucleic acid molecules encoding a polypeptide of the invention and fragments of such nucleic acid molecules suitable for use as PCR
5 primers for the amplification or mutation of nucleic acid molecules. As used herein, the term "nucleic acid molecule" is intended to include DNA molecules (e.g., cDNA or genomic DNA) and RNA molecules (e.g., mRNA) and analogs of the DNA or RNA generated using nucleotide
10 analogs. The nucleic acid molecule can be single-stranded or double-stranded, but preferably is double-stranded DNA.

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

A nucleic acid molecule of the present invention, e.g., a nucleic acid molecule having the nucleotide sequence of SEQ ID NO:1, 3, 4, or 6 or the cDNA of ATCC
35 98880, or a complement thereof, can be isolated using

- 21 -

standard molecular biology techniques and the sequence information provided herein. Using all or a portion of the nucleic acid sequences of SEQ ID NO:1, 3, 4, or 6 or the cDNA of ATCC 98880 as a hybridization probe, nucleic acid molecules of the invention can be isolated using standard hybridization and cloning techniques (e.g., as described in Sambrook et al., eds., *Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989*).

A nucleic acid molecule of the invention can be amplified using cDNA, mRNA or genomic DNA as a template and appropriate oligonucleotide primers according to standard PCR amplification techniques. The nucleic acid so amplified can be cloned into an appropriate vector and characterized by DNA sequence analysis. Furthermore, oligonucleotides corresponding to all or a portion of a nucleic acid molecule of the invention can be prepared by standard synthetic techniques, e.g., using an automated DNA synthesizer.

In another embodiment, an isolated nucleic acid molecule of the invention comprises a nucleic acid molecule which is a complement of the nucleotide sequence of SEQ ID NO:1, 3, 4, or 6 or the cDNA of ATCC 98880, or a portion thereof. A nucleic acid molecule which is complementary to a given nucleotide sequence is one which is sufficiently complementary to the given nucleotide sequence that it can hybridize to the given nucleotide sequence thereby forming a stable duplex.

Moreover, a nucleic acid molecule of the invention can comprise only a portion of a nucleic acid sequence encoding a full length polypeptide of the invention for example, a fragment which can be used as a probe or primer or a fragment encoding a biologically active portion of a polypeptide of the invention. The nucleotide

- 22 -

sequence determined from the cloning one gene allows for the generation of probes and primers designed for use in identifying and/or cloning homologues in other cell types, e.g., from other tissues, as well as homologues from other mammals. The probe/primer typically comprises substantially purified oligonucleotide. The oligonucleotide typically comprises a region of nucleotide sequence that hybridizes under stringent conditions to at least about 12, preferably about 25, more preferably about 50, 75, 100, 125, 150, 175, 200, 250, 300, 350 or 400 consecutive nucleotides of the sense or anti-sense sequence of SEQ ID NO:1, 3, 4, or 6, or the cDNA of ATCC 98880 or of a naturally occurring mutant of SEQ ID NO:1, 3, 4, or 6 or the cDNA of ATCC 98880.

Probes based on the sequence of a nucleic acid molecule of the invention can be used to detect transcripts or genomic sequences encoding the same protein molecule encoded by a selected nucleic acid molecule. The probe comprises a label group attached thereto, e.g., a radioisotope, a fluorescent compound, an enzyme, or an enzyme co-factor. Such probes can be used as part of a diagnostic test kit for identifying cells or tissues which mis-express the protein, such as by measuring levels of a nucleic acid molecule encoding the protein in a sample of cells from a subject, e.g., detecting mRNA levels or determining whether a gene encoding the protein has been mutated or deleted.

A nucleic acid fragment encoding a "biologically active portion" of a polypeptide of the invention can be prepared by isolating a portion of any of SEQ ID NOS:3 or 6 or the nucleotide sequence of the cDNA of ATCC 98880 which encodes a polypeptide having a biological activity, expressing the encoded portion of the polypeptide protein (e.g., by recombinant expression *in vitro*) and assessing the activity of the encoded portion of the polypeptide.

- 23 -

The invention further encompasses nucleic acid molecules that differ from the nucleotide sequence of SEQ ID NO:1, 3, 4, or 6 or the cDNA of ATCC 98880 due to degeneracy of the genetic code and thus encode the same
5 protein as that encoded by the nucleotide sequence of SEQ ID NO:3 or 6 or the cDNA of ATCC 98880.

In addition to the nucleotide sequences of SEQ ID NOs:3 and 6 and present in cDNA of ATCC 98880, it will be appreciated by those skilled in the art that DNA sequence
10 polymorphisms that lead to changes in the amino acid sequence may exist within a population (e.g., the human population). Such genetic polymorphisms may exist among individuals within a population due to natural allelic variation. An allele is one of a group of genes which
15 occur alternatively at a given genetic locus. As used herein, the phrase "allelic variant" refers to a nucleotide sequence which occurs at a given locus or to a polypeptide encoded by the nucleotide sequence. As used herein, the terms "gene" and "recombinant gene" refer to
20 nucleic acid molecules comprising an open reading frame encoding a polypeptide of the invention. Such natural allelic variations can typically result in 1-5% variance in the nucleotide sequence of a given gene. Alternative alleles can be identified by sequencing the gene of
25 interest in a number of different individuals. This can be readily carried out by using hybridization probes to identify the same genetic locus in a variety of individuals. Any and all such nucleotide variations and resulting amino acid polymorphisms or variations that are
30 the result of natural allelic variation and that do not alter the functional activity are intended to be within the scope of the invention.

Moreover, nucleic acid molecules encoding proteins of the invention from other species (homologues), which
35 have a nucleotide sequence which differs from that of the

- 24 -

human protein described herein are intended to be within the scope of the invention. Nucleic acid molecules corresponding to natural allelic variants and homologues of a cDNA of the invention can be isolated based on their
5 identity to the human nucleic acid molecule disclosed herein using the human cDNAs, or a portion thereof, as a hybridization probe according to standard hybridization techniques under stringent hybridization conditions. For example, a cDNA encoding a soluble form of a membrane-
10 bound protein of the invention isolated based on its hybridization to a nucleic acid molecule encoding all or part of the membrane-bound form. Likewise, a cDNA encoding a membrane-bound form can be isolated based on its hybridization to a nucleic acid molecule encoding all
15 or part of the soluble form.

Accordingly, in another embodiment, an isolated nucleic acid molecule of the invention is at least 300 (325, 350, 375, 400, 425, 450, 500, 550, 600, 650, 700, 800, 900, 1000, or 1290) nucleotides in length and
20 hybridizes under stringent conditions to the nucleic acid molecule comprising the nucleotide sequence, preferably the coding sequence, of SEQ ID NO:1, 3, 4, or 6, or the cDNA of ATCC 98880, or a complement thereof.

As used herein, the term "hybridizes under
25 stringent conditions" is intended to describe conditions for hybridization and washing under which nucleotide sequences at least 60% (65%, 70%, preferably 75%) identical to each other typically remain hybridized to each other. Such stringent conditions are known to those
30 skilled in the art and can be found in *Current Protocols in Molecular Biology*, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. A non-limiting example of stringent hybridization conditions are hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45°C, followed by
35 one or more washes in 0.2 X SSC, 0.1% SDS at 50-65°C.

- 25 -

Preferably, an isolated nucleic acid molecule of the invention that hybridizes under stringent conditions to the sequence of SEQ ID NO:1, 3, 4, or 6, or the cDNA of ATCC 98880, or the complement thereof, corresponds to a naturally-occurring nucleic acid molecule. As used herein, a "naturally-occurring" nucleic acid molecule refers to an RNA or DNA molecule having a nucleotide sequence that occurs in nature (e.g., encodes a natural protein).

10 In addition to naturally-occurring allelic variants of a nucleic acid molecule of the invention sequence that may exist in the population, the skilled artisan will further appreciate that changes can be introduced by mutation thereby leading to changes in the amino acid sequence of the encoded protein, without altering the biological activity of the protein. For example, one can make nucleotide substitutions leading to amino acid substitutions at "non-essential" amino acid residues. A "non-essential" amino acid residue is a residue that can be altered from the wild-type sequence without altering the biological activity, whereas an "essential" amino acid residue is required for biological activity. For example, amino acid residues that are not conserved or only semi-conserved among homologues of various species may be non-essential for activity and thus would be likely targets for alteration. Alternatively, amino acid residues that are conserved among the homologues of various species (e.g., murine and human) may be essential for activity and thus would not be likely targets for alteration.

30 Accordingly, another aspect of the invention pertains to nucleic acid molecules encoding a polypeptide of the invention that contain changes in amino acid residues that are not essential for activity. Such polypeptides differ in amino acid sequence from SEQ ID

- 26 -

NO:2, 5, 8, or 15 yet retain biological activity. In one embodiment, the isolated nucleic acid molecule includes a nucleotide sequence encoding a protein that includes an amino acid sequence that is at least about 45% identical, 5 65%, 75%, 85%, 95%, or 98% identical to the amino acid sequence of SEQ ID NO:2, 5, 8, or 15.

An isolated nucleic acid molecule encoding a variant protein can be created by introducing one or more nucleotide substitutions, additions or deletions into the 10 nucleotide sequence of SEQ ID NO:1, 3, 4, or 6 or the cDNA of ATCC 98880 such that one or more amino acid substitutions, additions or deletions are introduced into the encoded protein. Mutations can be introduced by standard techniques, such as site-directed mutagenesis 15 and PCR-mediated mutagenesis. Preferably, conservative amino acid substitutions are made at one or more predicted non-essential amino acid residues. A "conservative amino acid substitution" is one in which the amino acid residue is replaced with an amino acid 20 residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art. These families include amino acids with basic side chains (e.g., lysine, arginine, histidine), acidic side chains (e.g., aspartic acid, 25 glutamic acid), uncharged polar side chains (e.g., glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (e.g., alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (e.g., 30 threonine, valine, isoleucine) and aromatic side chains (e.g., tyrosine, phenylalanine, tryptophan, histidine). Alternatively, mutations can be introduced randomly along all or part of the coding sequence, such as by saturation mutagenesis, and the resultant mutants can be 35 screened for biological activity to identify mutants that

- 27 -

retain activity. Following mutagenesis, the encoded protein can be expressed recombinantly and the activity of the protein can be determined.

In one embodiment, a mutant polypeptide that is a
5 variant of a polypeptide of the invention can be assayed for: (1) the ability to form protein:protein interactions with proteins in a signaling pathway of the polypeptide of the invention; (2) the ability to bind a ligand of the polypeptide of the invention; or (3) the ability to bind
10 to an intracellular target protein of the polypeptide of the invention. In yet another embodiment, the mutant polypeptide can be assayed for the ability to modulate cellular proliferation or cellular differentiation.

The present invention encompasses antisense
15 nucleic acid molecules, i.e., molecules which are complementary to a sense nucleic acid encoding a polypeptide of the invention, e.g., complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence. Accordingly, an
20 antisense nucleic acid can hydrogen bond to a sense nucleic acid. The antisense nucleic acid can be complementary to an entire coding strand, or to only a portion thereof, e.g., all or part of the protein coding region (or open reading frame). An antisense nucleic
25 acid molecule can be antisense to all or part of a non-coding region of the coding strand of a nucleotide sequence encoding a polypeptide of the invention. The non-coding regions ("5' and 3' untranslated regions") are the 5' and 3' sequences which flank the coding region and
30 are not translated into amino acids.

An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An antisense nucleic acid of the invention can be constructed using chemical synthesis and enzymatic
35 ligation reactions using procedures known in the art.

- 28 -

For example, an antisense nucleic acid (e.g., an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, e.g., phosphorothioate derivatives and acridine substituted nucleotides can be used. Examples of modified nucleotides which can be used to generate the antisense nucleic acid include 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxymethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine. Alternatively, the antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (i.e., RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

- 29 -

The antisense nucleic acid molecules of the invention are typically administered to a subject or generated *in situ* such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding a selected polypeptide of the invention to thereby inhibit expression, e.g., by inhibiting transcription and/or translation. The hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule which binds to DNA duplexes, through specific interactions in the major groove of the double helix. An example of a route of administration of antisense nucleic acid molecules of the invention includes direct injection at a tissue site. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then administered systemically. For example, for systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface, e.g., by linking the antisense nucleic acid molecules to peptides or antibodies which bind to cell surface receptors or antigens. The antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. To achieve sufficient intracellular concentrations of the antisense molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a strong pol II or pol III promoter are preferred.

An antisense nucleic acid molecule of the invention can be an α -anomeric nucleic acid molecule. An α -anomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual β -units, the strands run parallel to each other (Gaultier et al. (1987) *Nucleic Acids Res.* 15:6625-6641). The antisense nucleic acid molecule can

- 30 -

also comprise a 2'-o-methylribonucleotide (Inoue et al. (1987) *Nucleic Acids Res.* 15:6131-6148) or a chimeric RNA-DNA analogue (Inoue et al. (1987) *FEBS Lett.* 215:327-330).

5 The invention also encompasses ribozymes. Ribozymes are catalytic RNA molecules with ribonuclease activity which are capable of cleaving a single-stranded nucleic acid, such as an mRNA, to which they have a complementary region. Thus, ribozymes (e.g., hammerhead
10 ribozymes (described in Haselhoff and Gerlach (1988) *Nature* 334:585-591)) can be used to catalytically cleave mRNA transcripts to thereby inhibit translation of the protein encoded by the mRNA. A ribozyme having
15 specificity for a nucleic acid molecule encoding a polypeptide of the invention can be designed based upon the nucleotide sequence of a cDNA disclosed herein. For example, a derivative of a *Tetrahymena* L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence
20 to be cleaved in a Cech et al. U.S. Patent No. 4,987,071; and Cech et al. U.S. Patent No. 5,116,742. Alternatively, an mRNA encoding a polypeptide of the invention can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA
25 molecules. See, e.g., Bartel and Szostak (1993) *Science* 261:1411-1418.

 The invention also encompasses nucleic acid molecules which form triple helical structures. For example, expression of a polypeptide of the invention can
30 be inhibited by targeting nucleotide sequences complementary to the regulatory region of the gene encoding the polypeptide (e.g., the promoter and/or enhancer) to form triple helical structures that prevent transcription of the gene in target cells. See generally
35 Helene (1991) *Anticancer Drug Des.* 6(6):569-84; Helene

- 31 -

(1992) *Ann. N.Y. Acad. Sci.* 660:27-36; and Maher (1992) *Bioassays* 14(12):807-15.

In various embodiments, the nucleic acid molecules of the invention can be modified at the base moiety, 5 sugar moiety or phosphate backbone to improve, e.g., the stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate backbone of the nucleic acids can be modified to generate peptide nucleic acids (see Hyrup et al. (1996) *Bioorganic & Medicinal* 10 *Chemistry* 4(1): 5-23). As used herein, the terms "peptide nucleic acids" or "PNAs" refer to nucleic acid mimics, e.g., DNA mimics, in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are 15 retained. The neutral backbone of PNAs has been shown to allow for specific hybridization to DNA and RNA under conditions of low ionic strength. The synthesis of PNA oligomers can be performed using standard solid phase peptide synthesis protocols as described in Hyrup et al. 20 (1996), *supra*; Perry-O'Keefe et al. (1996) *Proc. Natl. Acad. Sci. USA* 93: 14670-675.

PNAs can be used in therapeutic and diagnostic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of 25 gene expression by, e.g., inducing transcription or translation arrest or inhibiting replication. PNAs can also be used, e.g., in the analysis of single base pair mutations in a gene by, e.g., PNA directed PCR clamping; as artificial restriction enzymes when used in 30 combination with other enzymes, e.g., S1 nucleases (Hyrup (1996), *supra*; or as probes or primers for DNA sequence and hybridization (Hyrup (1996), *supra*; Perry-O'Keefe et al. (1996) *Proc. Natl. Acad. Sci. USA* 93:14670-675).

In another embodiment, PNAs can be modified, e.g., 35 to enhance their stability or cellular uptake, by

- 32 -

attaching lipophilic or other helper groups to PNA, by the formation of PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. For example, PNA-DNA chimeras can be generated
5 which may combine the advantageous properties of PNA and DNA. Such chimeras allow DNA recognition enzymes, e.g., RNase H and DNA polymerases, to interact with the DNA portion while the PNA portion would provide high binding affinity and specificity. PNA-DNA chimeras can be linked
10 using linkers of appropriate lengths selected in terms of base stacking, number of bonds between the nucleobases, and orientation (Hyrup (1996), *supra*). The synthesis of PNA-DNA chimeras can be performed as described in Hyrup (1996), *supra*, and Finn et al. (1996) *Nucleic Acids Res.*
15 24:3357-63. For example, a DNA chain can be synthesized on a solid support using standard phosphoramidite coupling chemistry and modified nucleoside analogs. Compounds such as 5'-(4-methoxytrityl)amino-5'-deoxy-thymidine phosphoramidite can be used as a link between
20 the PNA and the 5' end of DNA (Mag et al. (1989) *Nucleic Acids Res.* 17:5973-88). PNA monomers are then coupled in a stepwise manner to produce a chimeric molecule with a 5' PNA segment and a 3' DNA segment (Finn et al. (1996) *Nucleic Acids Res.* 24(17):3357-63). Alternatively,
25 chimeric molecules can be synthesized with a 5' DNA segment and a 3' PNA segment (Peterser et al. (1975) *Bioorganic Med. Chem. Lett.* 5:1119-11124).

In other embodiments, the oligonucleotide may include other appended groups such as peptides (e.g., for
30 targeting host cell receptors *in vivo*), or agents facilitating transport across the cell membrane (see, e.g., Letsinger et al. (1989) *Proc. Natl. Acad. Sci. USA* 86:6553-6556; Lemaitre et al. (1987) *Proc. Natl. Acad. Sci. USA* 84:648-652; PCT Publication No. W0 88/09810) or
35 the blood-brain barrier (see, e.g., PCT Publication No.

- 33 -

WO 89/10134). In addition, oligonucleotides can be modified with hybridization-triggered cleavage agents (see, e.g., Krol et al. (1988) *Bio/Techniques* 6:958-976) or intercalating agents (see, e.g., Zon (1988) *Pharm. Res.* 5:539-549). To this end, the oligonucleotide may be conjugated to another molecule, e.g., a peptide, hybridization triggered cross-linking agent, transport agent, hybridization-triggered cleavage agent, etc.

II. Isolated Proteins and Antibodies

10 One aspect of the invention pertains to isolated proteins, and biologically active portions thereof, as well as polypeptide fragments suitable for use as immunogens to raise antibodies directed against a polypeptide of the invention. In one embodiment, the
15 native polypeptide can be isolated from cells or tissue sources by an appropriate purification scheme using standard protein purification techniques. In another embodiment, polypeptides of the invention are produced by recombinant DNA techniques. Alternative to recombinant
20 expression, a polypeptide of the invention can be synthesized chemically using standard peptide synthesis techniques.

An "isolated" or "purified" protein or biologically active portion thereof is substantially free
25 of cellular material or other contaminating proteins from the cell or tissue source from which the protein is derived, or substantially free of chemical precursors or other chemicals when chemically synthesized. The language "substantially free of cellular material"
30 includes preparations of protein in which the protein is separated from cellular components of the cells from which it is isolated or recombinantly produced. Thus, protein that is substantially free of cellular material includes preparations of protein having less than about

- 34 -

30%, 20%, 10%, or 5% (by dry weight) of heterologous protein (also referred to herein as a "contaminating protein"). When the protein or biologically active portion thereof is recombinantly produced, it is also preferably substantially free of culture medium, i.e., 5 culture medium represents less than about 20%, 10%, or 5% of the volume of the protein preparation. When the protein is produced by chemical synthesis, it is preferably substantially free of chemical precursors or 10 other chemicals, i.e., it is separated from chemical precursors or other chemicals which are involved in the synthesis of the protein. Accordingly such preparations of the protein have less than about 30%, 20%, 10%, 5% (by dry weight) of chemical precursors or compounds other 15 than the polypeptide of interest.

Biologically active portions of a polypeptide of the invention include polypeptides comprising amino acid sequences sufficiently identical to or derived from the amino acid sequence of the protein (e.g., the amino acid 20 sequence shown in any of SEQ ID Nos:2, 5, 8, and 15), which include fewer amino acids than the full length protein, and exhibit at least one activity of the corresponding full-length protein. Typically, biologically active portions comprise a domain or motif 25 with at least one activity of the corresponding protein. A biologically active portion of a protein of the invention can be a polypeptide which is, for example, 10, 25, 50, 100 or more amino acids in length. Moreover, other biologically active portions, in which other 30 regions of the protein are deleted, can be prepared by recombinant techniques and evaluated for one or more of the functional activities of the native form of a polypeptide of the invention.

Polypeptides can have the amino acid sequence of 35 SEQ ID NO:2, 5, 7-11, or 14-18. Other useful proteins

- 35 -

are substantially identical (e.g., at least about 45%, preferably 55%, 65%, 75%, 85%, 95%, or 99%) to any of SEQ ID NO:2, 5, 7-11, or 14-18 and retain the functional activity of the protein of the corresponding naturally-occurring protein yet differ in amino acid sequence due to natural allelic variation or mutagenesis.

To determine the percent identity of two amino acid sequences or of two nucleic acids, the sequences are aligned for optimal comparison purposes (e.g., gaps can be introduced in the sequence of a first amino acid or nucleic acid sequence for optimal alignment with a second amino or nucleic acid sequence). The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are identical at that position. The percent identity between the two sequences is a function of the number of identical positions shared by the sequences (i.e., % identity = # of identical positions/total # of positions (e.g., overlapping positions) x 100). In one embodiment the two sequences are the same length.

The determination of percent identity between two sequences can be accomplished using a mathematical algorithm. A preferred, non-limiting example of a mathematical algorithm utilized for the comparison of two sequences is the algorithm of Karlin and Altschul (1990) *Proc. Natl. Acad. Sci. USA* 87:2264-2268, modified as in Karlin and Altschul (1993) *Proc. Natl. Acad. Sci. USA* 90:5873-5877. Such an algorithm is incorporated into the NBLAST and XBLAST programs of Altschul, et al. (1990) *J. Mol. Biol.* 215:403-410. BLAST nucleotide searches can be performed with the NBLAST program, score = 100, wordlength = 12 to obtain nucleotide sequences homologous

- 36 -

to a nucleic acid molecules of the invention. BLAST protein searches can be performed with the XBLAST program, score = 50, wordlength = 3 to obtain amino acid sequences homologous to a protein molecules of the invention. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul et al. (1997) *Nucleic Acids Res.* 25:3389-3402. Alternatively, PSI-Blast can be used to perform an iterated search which detects distant relationships between molecules. *Id.* When utilizing BLAST, Gapped BLAST, and PSI-Blast programs, the default parameters of the respective programs (e.g., XBLAST and NBLAST) can be used. See <http://www.ncbi.nlm.nih.gov>. Another preferred, non-limiting example of a mathematical algorithm utilized for the comparison of sequences is the algorithm of Myers and Miller, (1988) *CABIOS* 4:11-17. Such an algorithm is incorporated into the ALIGN program (version 2.0) which is part of the GCG sequence alignment software package. When utilizing the ALIGN program for comparing amino acid sequences, a PAM120 weight residue table, a gap length penalty of 12, and a gap penalty of 4 can be used.

The percent identity between two sequences can be determined using techniques similar to those described above, with or without allowing gaps. In calculating percent identity, only exact matches are counted.

The invention also provides chimeric or fusion proteins. As used herein, a "chimeric protein" or "fusion protein" comprises all or part (preferably biologically active) of a polypeptide of the invention operably linked to a heterologous polypeptide (i.e., a polypeptide other than the same polypeptide of the invention). Within the fusion protein, the term "operably linked" is intended to indicate that the polypeptide of the invention and the heterologous

- 37 -

polypeptide are fused in-frame to each other. The heterologous polypeptide can be fused to the N-terminus or C-terminus of the polypeptide of the invention. One useful fusion protein is a GST fusion protein in which the polypeptide of the invention is fused to the C-terminus of GST sequences. Such fusion proteins can facilitate the purification of a recombinant polypeptide of the invention.

In another embodiment, the fusion protein contains a heterologous signal sequence at its N-terminus. For example, the native signal sequence of a polypeptide of the invention can be removed and replaced with a signal sequence from another protein. For example, the gp67 secretory sequence of the baculovirus envelope protein can be used as a heterologous signal sequence (*Current Protocols in Molecular Biology*, Ausubel et al., eds., John Wiley & Sons, 1992). Other examples of eukaryotic heterologous signal sequences include the secretory sequences of melittin and human placental alkaline phosphatase (Stratagene; La Jolla, California). In yet another example, useful prokaryotic heterologous signal sequences include the phoA secretory signal (Sambrook et al., *supra*) and the protein A secretory signal (Pharmacia Biotech; Piscataway, New Jersey).

In yet another embodiment, the fusion protein is an immunoglobulin fusion protein in which all or part of a polypeptide of the invention is fused to sequences derived from a member of the immunoglobulin protein family. The immunoglobulin fusion proteins of the invention can be incorporated into pharmaceutical compositions and administered to a subject to inhibit an interaction between a ligand (soluble or membrane-bound) and a protein on the surface of a cell (receptor), to thereby suppress signal transduction *in vivo*. The immunoglobulin fusion protein can be used to affect the

- 38 -

bioavailability of a cognate ligand of a polypeptide of the invention. Inhibition of ligand/receptor interaction may be useful therapeutically, both for treating proliferative and differentiative disorders and for
5 modulating (e.g. promoting or inhibiting) cell survival. Moreover, the immunoglobulin fusion proteins of the invention can be used as immunogens to produce antibodies directed against a polypeptide of the invention in a subject, to purify ligands and in screening assays to
10 identify molecules which inhibit the interaction of receptors with ligands.

Chimeric and fusion proteins of the invention can be produced by standard recombinant DNA techniques. In another embodiment, the fusion gene can be synthesized by
15 conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers which give rise to complementary overhangs between two consecutive gene fragments which can subsequently be
20 annealed and reamplified to generate a chimeric gene sequence (see, e.g., Ausubel et al., *supra*). Moreover, many expression vectors are commercially available that already encode a fusion moiety (e.g., a GST polypeptide). A nucleic acid encoding a polypeptide of the invention
25 can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the polypeptide of the invention.

A signal sequence of a polypeptide of the invention (SEQ ID NO:7 or 14) can be used to facilitate
30 secretion and isolation of the secreted protein or other proteins of interest. Signal sequences are typically characterized by a core of hydrophobic amino acids which are generally cleaved from the mature protein during secretion in one or more cleavage events. Such signal
35 peptides contain processing sites that allow cleavage of

- 39 -

the signal sequence from the mature proteins as they pass through the secretory pathway. Thus, the invention pertains to the described polypeptides having a signal sequence, as well as to the signal sequence itself and to the polypeptide in the absence of the signal sequence (i.e., the cleavage products). In one embodiment, a nucleic acid sequence encoding a signal sequence of the invention can be operably linked in an expression vector to a protein of interest, such as a protein which is ordinarily not secreted or is otherwise difficult to isolate. The signal sequence directs secretion of the protein, such as from a eukaryotic host into which the expression vector is transformed, and the signal sequence is subsequently or concurrently cleaved. The protein can then be readily purified from the extracellular medium by art recognized methods. Alternatively, the signal sequence can be linked to the protein of interest using a sequence which facilitates purification, such as with a GST domain.

In another embodiment, the signal sequences of the present invention can be used to identify regulatory sequences, e.g., promoters, enhancers, repressors. Since signal sequences are the most amino-terminal sequences of a peptide, it is expected that the nucleic acids which flank the signal sequence on its amino-terminal side will be regulatory sequences which affect transcription. Thus, a nucleotide sequence which encodes all or a portion of a signal sequence can be used as a probe to identify and isolate signal sequences and their flanking regions, and these flanking regions can be studied to identify regulatory elements therein.

The present invention also pertains to variants of the polypeptides of the invention. Such variants have an altered amino acid sequence which can function as either agonists (mimetics) or as antagonists. Variants can be

- 40 -

generated by mutagenesis, e.g., discrete point mutation or truncation. An agonist can retain substantially the same, or a subset, of the biological activities of the naturally occurring form of the protein. An antagonist
5 of a protein can inhibit one or more of the activities of the naturally occurring form of the protein by, for example, competitively binding to a downstream or upstream member of a cellular signaling cascade which includes the protein of interest. Thus, specific
10 biological effects can be elicited by treatment with a variant of limited function. Treatment of a subject with a variant having a subset of the biological activities of the naturally occurring form of the protein can have fewer side effects in a subject relative to treatment
15 with the naturally occurring form of the protein.

Variants of a protein of the invention which function as either agonists (mimetics) or as antagonists can be identified by screening combinatorial libraries of mutants, e.g., truncation mutants, of the protein of the
20 invention for agonist or antagonist activity. In one embodiment, a variegated library of variants is generated by combinatorial mutagenesis at the nucleic acid level and is encoded by a variegated gene library. A variegated library of variants can be produced by, for
25 example, enzymatically ligating a mixture of synthetic oligonucleotides into gene sequences such that a degenerate set of potential protein sequences is expressible as individual polypeptides, or alternatively, as a set of larger fusion proteins (e.g., for phage
30 display). There are a variety of methods which can be used to produce libraries of potential variants of the polypeptides of the invention from a degenerate oligonucleotide sequence. Methods for synthesizing degenerate oligonucleotides are known in the art (see,
35 e.g., Narang (1983) *Tetrahedron* 39:3; Itakura et al.

- 41 -

(1984) *Annu. Rev. Biochem.* 53:323; Itakura et al. (1984) *Science* 198:1056; Ike et al. (1983) *Nucleic Acid Res.* 11:477).

In addition, libraries of fragments of the coding
5 sequence of a polypeptide of the invention can be used to
generate a variegated population of polypeptides for
screening and subsequent selection of variants. For
example, a library of coding sequence fragments can be
generated by treating a double stranded PCR fragment of
10 the coding sequence of interest with a nuclease under
conditions wherein nicking occurs only about once per
molecule, denaturing the double stranded DNA, renaturing
the DNA to form double stranded DNA which can include
sense/antisense pairs from different nicked products,
15 removing single stranded portions from reformed duplexes
by treatment with S1 nuclease, and ligating the resulting
fragment library into an expression vector. By this
method, an expression library can be derived which
encodes N-terminal and internal fragments of various
20 sizes of the protein of interest.

Several techniques are known in the art for
screening gene products of combinatorial libraries made
by point mutations or truncation, and for screening cDNA
libraries for gene products having a selected property.
25 The most widely used techniques, which are amenable to
high through-put analysis, for screening large gene
libraries typically include cloning the gene library into
replicable expression vectors, transforming appropriate
cells with the resulting library of vectors, and
30 expressing the combinatorial genes under conditions in
which detection of a desired activity facilitates
isolation of the vector encoding the gene whose product
was detected. Recursive ensemble mutagenesis (REM), a
technique which enhances the frequency of functional
35 mutants in the libraries, can be used in combination with

- 42 -

the screening assays to identify variants of a protein of the invention (Arkin and Yourvan (1992) *Proc. Natl. Acad. Sci. USA* 89:7811-7815; Delgrave et al. (1993) *Protein Engineering* 6:327-331).

5 An isolated polypeptide of the invention, or a fragment thereof, can be used as an immunogen to generate antibodies using standard techniques for polyclonal and monoclonal antibody preparation. The full-length polypeptide or protein can be used or, alternatively, the
10 invention provides antigenic peptide fragments for use as immunogens. The antigenic peptide of a protein of the invention comprises at least 8 (preferably 10, 15, 20, or 30) amino acid residues of the amino acid sequence of SEQ ID NO:2 or 5 and encompasses an epitope of the protein
15 such that an antibody raised against the peptide forms a specific immune complex with the protein.

Preferred epitopes encompassed by the antigenic peptide are regions that are located on the surface of the protein, e.g., hydrophilic regions. Figures 2 and 4
20 are hydrophobicity plots of the proteins of the invention. These plots or similar analyses can be used to identify hydrophilic regions.

An immunogen typically is used to prepare antibodies by immunizing a suitable subject, (e.g.,
25 rabbit, goat, mouse or other mammal). An appropriate immunogenic preparation can contain, for example, recombinantly expressed or chemically synthesized polypeptide. The preparation can further include an adjuvant, such as Freund's complete or incomplete
30 adjuvant, or similar immunostimulatory agent.

Accordingly, another aspect of the invention pertains to antibodies directed against a polypeptide of the invention. The term "antibody" as used herein refers to immunoglobulin molecules and immunologically active
35 portions of immunoglobulin molecules, i.e., molecules

- 43 -

that contain an antigen binding site which specifically binds an antigen, such as a polypeptide of the invention. A molecule which specifically binds to a given polypeptide of the invention is a molecule which binds
5 the polypeptide, but does not substantially bind other molecules in a sample, e.g., a biological sample, which naturally contains the polypeptide. Examples of immunologically active portions of immunoglobulin molecules include F(ab) and F(ab')₂ fragments which can be
10 generated by treating the antibody with an enzyme such as pepsin. The invention provides polyclonal and monoclonal antibodies. The term "monoclonal antibody" or "monoclonal antibody composition", as used herein, refers to a population of antibody molecules that contain only
15 one species of an antigen binding site capable of immunoreacting with a particular epitope.

Polyclonal antibodies can be prepared as described above by immunizing a suitable subject with a polypeptide of the invention as an immunogen. The antibody titer in
20 the immunized subject can be monitored over time by standard techniques, such as with an enzyme linked immunosorbent assay (ELISA) using immobilized polypeptide. If desired, the antibody molecules can be isolated from the mammal (e.g., from the blood) and
25 further purified by well-known techniques, such as protein A chromatography to obtain the IgG fraction. At an appropriate time after immunization, e.g., when the specific antibody titers are highest, antibody-producing cells can be obtained from the subject and used to
30 prepare monoclonal antibodies by standard techniques, such as the hybridoma technique originally described by Kohler and Milstein (1975) *Nature* 256:495-497, the human B cell hybridoma technique (Kozbor et al. (1983) *Immunol. Today* 4:72), the EBV-hybridoma technique (Cole et al.
35 (1985), *Monoclonal Antibodies and Cancer Therapy*, Alan R.

- 44 -

Liss, Inc., pp. 77-96) or trioma techniques. The technology for producing hybridomas is well known (see generally *Current Protocols in Immunology* (1994) Coligan et al. (eds.) John Wiley & Sons, Inc., New York, NY).

5 Hybridoma cells producing a monoclonal antibody of the invention are detected by screening the hybridoma culture supernatants for antibodies that bind the polypeptide of interest, e.g., using a standard ELISA assay.

Alternative to preparing monoclonal antibody-
10 secreting hybridomas, a monoclonal antibody directed against a polypeptide of the invention can be identified and isolated by screening a recombinant combinatorial immunoglobulin library (e.g., an antibody phage display library) with the polypeptide of interest. Kits for
15 generating and screening phage display libraries are commercially available (e.g., the Pharmacia *Recombinant Phage Antibody System*, Catalog No. 27-9400-01; and the Stratagene *SurfZAP™ Phage Display Kit*, Catalog No. 240612). Additionally, examples of methods and reagents
20 particularly amenable for use in generating and screening antibody display library can be found in, for example, U.S. Patent No. 5,223,409; PCT Publication No. WO 92/18619; PCT Publication No. WO 91/17271; PCT Publication No. WO 92/20791; PCT Publication No. WO
25 92/15679; PCT Publication No. WO 93/01288; PCT Publication No. WO 92/01047; PCT Publication No. WO 92/09690; PCT Publication No. WO 90/02809; Fuchs et al. (1991) *Bio/Technology* 9:1370-1372; Hay et al. (1992) *Hum. Antibod. Hybridomas* 3:81-85; Huse et al. (1989) *Science*
30 246:1275-1281; Griffiths et al. (1993) *EMBO J.* 12:725-734.

Additionally, recombinant antibodies, such as chimeric and humanized monoclonal antibodies, comprising both human and non-human portions, which can be made
35 using standard recombinant DNA techniques, are within the

- 45 -

scope of the invention. Such chimeric and humanized monoclonal antibodies can be produced by recombinant DNA techniques known in the art, for example using methods described in PCT Publication No. WO 87/02671; European
5 Patent Application 184,187; European Patent Application 171,496; European Patent Application 173,494; PCT Publication No. WO 86/01533; U.S. Patent No. 4,816,567; European Patent Application 125,023; Better et al. (1988) *Science* 240:1041-1043; Liu et al. (1987) *Proc. Natl.*
10 *Acad. Sci. USA* 84:3439-3443; Liu et al. (1987) *J. Immunol.* 139:3521-3526; Sun et al. (1987) *Proc. Natl. Acad. Sci. USA* 84:214-218; Nishimura et al. (1987) *Canc. Res.* 47:999-1005; Wood et al. (1985) *Nature* 314:446-449; and Shaw et al. (1988) *J. Natl. Cancer Inst.* 80:1553-
15 1559; Morrison (1985) *Science* 229:1202-1207; Oi et al. (1986) *Bio/Techniques* 4:214; U.S. Patent 5,225,539; Jones et al. (1986) *Nature* 321:552-525; Verhoeyan et al. (1988) *Science* 239:1534; and Beidler et al. (1988) *J. Immunol.* 141:4053-4060.

20 Completely human antibodies are particularly desirable for therapeutic treatment of human patients. Such antibodies can be produced using transgenic mice which are incapable of expressing endogenous immunoglobulin heavy and light chains genes, but which
25 can express human heavy and light chain genes. The transgenic mice are immunized in the normal fashion with a selected antigen, e.g., all or a portion of a polypeptide of the invention. Monoclonal antibodies directed against the antigen can be obtained using
30 conventional hybridoma technology. The human immunoglobulin transgenes harbored by the transgenic mice rearrange during B cell differentiation, and subsequently undergo class switching and somatic mutation. Thus, using such a technique, it is possible to produce
35 therapeutically useful IgG, IgA and IgE antibodies. For

- 46 -

an overview of this technology for producing human antibodies, see Lonberg and Huszar (1995, *Int. Rev. Immunol.* 13:65-93). For a detailed discussion of this technology for producing human antibodies and human
5 monoclonal antibodies and protocols for producing such antibodies, see, e.g., U.S. Patent 5,625,126; U.S. Patent 5,633,425; U.S. Patent 5,569,825; U.S. Patent 5,661,016; and U.S. Patent 5,545,806. In addition, companies such as Abgenix, Inc. (Freemont, CA), can be engaged to
10 provide human antibodies directed against a selected antigen using technology similar to that described above.

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

An antibody directed against a polypeptide of the
20 invention (e.g., monoclonal antibody) can be used to isolate the polypeptide by standard techniques, such as affinity chromatography or immunoprecipitation. Moreover, such an antibody can be used to detect the protein (e.g., in a cellular lysate or cell supernatant)
25 in order to evaluate the abundance and pattern of expression of the polypeptide. The antibodies can also be used diagnostically to monitor protein levels in tissue as part of a clinical testing procedure, e.g., to, for example, determine the efficacy of a given treatment
30 regimen. Detection can be facilitated by coupling the antibody to a detectable substance. Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, and radioactive materials.

- 47 -

Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, β -galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and
5 avidin/biotin; examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a luminescent material includes luminol; examples of
10 bioluminescent materials include luciferase, luciferin, and aequorin, and examples of suitable radioactive material include ^{125}I , ^{131}I , ^{35}S or ^3H .

III. Recombinant Expression Vectors and Host Cells

Another aspect of the invention pertains to
15 vectors, preferably expression vectors, containing a nucleic acid encoding a polypeptide of the invention (or a portion thereof). As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One
20 type of vector is a "plasmid", which refers to a circular double stranded DNA loop into which additional DNA segments can be ligated. Another type of vector is a viral vector, wherein additional DNA segments can be ligated into the viral genome. Certain vectors are
25 capable of autonomous replication in a host cell into which they are introduced (e.g., bacterial vectors having a bacterial origin of replication and episomal mammalian vectors). Other vectors (e.g., non-episomal mammalian vectors) are integrated into the genome of a host cell
30 upon introduction into the host cell, and thereby are replicated along with the host genome. Moreover, certain vectors, expression vectors, are capable of directing the expression of genes to which they are operably linked. In general, expression vectors of utility in recombinant

- 48 -

DNA techniques are often in the form of plasmids (vectors). However, the invention is intended to include such other forms of expression vectors, such as viral vectors (e.g., replication defective retroviruses, adenoviruses and adeno-associated viruses), which serve equivalent functions.

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

- 49 -

vectors of the invention can be introduced into host cells to thereby produce proteins or peptides, including fusion proteins or peptides, encoded by nucleic acids as described herein.

5 The recombinant expression vectors of the invention can be designed for expression of a polypeptide of the invention in prokaryotic or eukaryotic cells, e.g., bacterial cells such as *E. coli*, insect cells (using baculovirus expression vectors), yeast cells or
10 mammalian cells. Suitable host cells are discussed further in Goeddel, *supra*. Alternatively, the recombinant expression vector can be transcribed and translated *in vitro*, for example using T7 promoter regulatory sequences and T7 polymerase.

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

- 50 -

(Pharmacia, Piscataway, NJ) which fuse glutathione S-transferase (GST), maltose E binding protein, or protein A, respectively, to the target recombinant protein.

Examples of suitable inducible non-fusion *E. coli* expression vectors include pTrc (Amann et al., (1988) 5 *Gene* 69:301-315) and pET 11d (Studier et al., *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, California (1990) 60-89). Target gene expression from the pTrc vector relies on 10 host RNA polymerase transcription from a hybrid trp-lac fusion promoter. Target gene expression from the pET 11d vector relies on transcription from a T7 gn10-lac fusion promoter mediated by a coexpressed viral RNA polymerase (T7 gn1). This viral polymerase is supplied by host 15 strains BL21(DE3) or HMS174(DE3) from a resident λ prophage harboring a T7 gn1 gene under the transcriptional control of the lacUV 5 promoter.

One strategy to maximize recombinant protein expression in *E. coli* is to express the protein in a host 20 bacteria with an impaired capacity to proteolytically cleave the recombinant protein (Gottesman, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, California (1990) 119-128). Another strategy is to alter the nucleic acid sequence of 25 the nucleic acid to be inserted into an expression vector so that the individual codons for each amino acid are those preferentially utilized in *E. coli* (Wada et al. (1992) *Nucleic Acids Res.* 20:2111-2118). Such alteration of nucleic acid sequences of the invention can be carried 30 out by standard DNA synthesis techniques.

In another embodiment, the expression vector is a yeast expression vector. Examples of vectors for expression in yeast *S. cerevisiae* include pYepSec1 (Baldari et al. (1987) *EMBO J.* 6:229-234), pMFa (Kurjan 35 and Herskowitz, (1982) *Cell* 30:933-943), pJRY88 (Schultz

- 51 -

et al. (1987) *Gene* 54:113-123), pYES2 (Invitrogen Corporation, San Diego, CA), and pPicZ (Invitrogen Corp, San Diego, CA).

Alternatively, the expression vector is a
5 baculovirus expression vector. Baculovirus vectors available for expression of proteins in cultured insect cells (e.g., Sf 9 cells) include the pAc series (Smith et al. (1983) *Mol. Cell Biol.* 3:2156-2165) and the pVL series (Lucklow and Summers (1989) *Virology* 170:31-39).

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

In another embodiment, the recombinant mammalian expression vector is capable of directing expression of the nucleic acid preferentially in a particular cell type
25 (e.g., tissue-specific regulatory elements are used to express the nucleic acid). Tissue-specific regulatory elements are known in the art. Non-limiting examples of suitable tissue-specific promoters include the albumin promoter (liver-specific; Pinkert et al. (1987) *Genes*
30 *Dev.* 1:268-277), lymphoid-specific promoters (Calame and Eaton (1988) *Adv. Immunol.* 43:235-275), in particular promoters of T cell receptors (Winoto and Baltimore (1989) *EMBO J.* 8:729-733) and immunoglobulins (Banerji et al. (1983) *Cell* 33:729-740; Queen and Baltimore (1983)
35 *Cell* 33:741-748), neuron-specific promoters (e.g., the

- 52 -

neurofilament promoter; Byrne and Ruddle (1989) *Proc. Natl. Acad. Sci. USA* 86:5473-5477), pancreas-specific promoters (Edlund et al. (1985) *Science* 230:912-916), and mammary gland-specific promoters (e.g., milk whey
5 promoter; U.S. Patent No. 4,873,316 and European Application Publication No. 264,166). Developmentally-regulated promoters are also encompassed, for example the murine hox promoters (Kessel and Gruss (1990) *Science* 249:374-379) and the α -fetoprotein promoter (Campes and
10 Tilghman (1989) *Genes Dev.* 3:537-546).

The invention further provides a recombinant expression vector comprising a DNA molecule of the invention cloned into the expression vector in an antisense orientation. That is, the DNA molecule is
15 operably linked to a regulatory sequence in a manner which allows for expression (by transcription of the DNA molecule) of an RNA molecule which is antisense to the mRNA encoding a polypeptide of the invention. Regulatory sequences operably linked to a nucleic acid
20 cloned in the antisense orientation can be chosen which direct the continuous expression of the antisense RNA molecule in a variety of cell types, for instance viral promoters and/or enhancers, or regulatory sequences can be chosen which direct constitutive, tissue specific or
25 cell type specific expression of antisense RNA. The antisense expression vector can be in the form of a recombinant plasmid, phagemid or attenuated virus in which antisense nucleic acids are produced under the control of a high efficiency regulatory region, the
30 activity of which can be determined by the cell type into which the vector is introduced. For a discussion of the regulation of gene expression using antisense genes see Weintraub et al. (*Reviews - Trends in Genetics*, Vol. 1(1) 1986).

- 53 -

Another aspect of the invention pertains to host cells into which a recombinant expression vector of the invention has been introduced. The terms "host cell" and "recombinant host cell" are used interchangeably herein. 5 It is understood that such terms refer not only to the particular subject cell but to the progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may 10 not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein.

A host cell can be any prokaryotic cell (e.g., *E. coli*) or any eukaryotic cell (e.g., an insect cell, a 15 yeast cell, or a mammalian cell).

Vector DNA can be introduced into prokaryotic or eukaryotic cells via conventional transformation or transfection techniques. As used herein, the terms "transformation" and "transfection" are intended to refer 20 to a variety of art-recognized techniques for introducing foreign nucleic acid into a host cell, including calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection, lipofection, or electroporation. Suitable methods for transforming or 25 transfecting host cells can be found in Sambrook, et al. (*supra*), and other laboratory manuals.

For stable transfection of mammalian cells, it is known that, depending upon the expression vector and transfection technique used, only a small fraction of 30 cells may integrate the foreign DNA into their genome. In order to identify and select these integrants, a gene that encodes a selectable marker (e.g., for resistance to antibiotics) is generally introduced into the host cells along with the gene of interest. Preferred selectable 35 markers include those which confer resistance to drugs,

- 54 -

such as G418, hygromycin and methotrexate. Cells stably transfected with the introduced nucleic acid can be identified by drug selection (e.g., cells that have incorporated the selectable marker gene will survive, while the other cells die).

A host cell of the invention, such as a prokaryotic or eukaryotic host cell in culture, can be used to produce a polypeptide of the invention. Accordingly, the invention further provides methods for producing a polypeptide of the invention using the host cells of the invention. In one embodiment, the method comprises culturing the host cell of invention (into which a recombinant expression vector encoding a polypeptide of the invention has been introduced) in a suitable medium such that the polypeptide is produced. In another embodiment, the method further comprises isolating the polypeptide from the medium or the host cell.

The host cells of the invention can also be used to produce nonhuman transgenic animals. For example, in one embodiment, a host cell of the invention is a fertilized oocyte or an embryonic stem cell into which a sequences encoding a polypeptide of the invention have been introduced. Such host cells can then be used to create non-human transgenic animals in which exogenous sequences encoding a polypeptide of the invention have been introduced into their genome or homologous recombinant animals in which endogenous encoding a polypeptide of the invention sequences have been altered. Such animals are useful for studying the function and/or activity of the polypeptide and for identifying and/or evaluating modulators of polypeptide activity. As used herein, a "transgenic animal" is a non-human animal, preferably a mammal, more preferably a rodent such as a rat or mouse, in which one or more of the cells of the

- 55 -

animal includes a transgene. Other examples of transgenic animals include non-human primates, sheep, dogs, cows, goats, chickens, amphibians, etc. A transgene is exogenous DNA which is integrated into the genome of a cell from which a transgenic animal develops and which remains in the genome of the mature animal, thereby directing the expression of an encoded gene product in one or more cell types or tissues of the transgenic animal. As used herein, an "homologous recombinant animal" is a non-human animal, preferably a mammal, more preferably a mouse, in which an endogenous gene has been altered by homologous recombination between the endogenous gene and an exogenous DNA molecule introduced into a cell of the animal, e.g., an embryonic cell of the animal, prior to development of the animal.

A transgenic animal of the invention can be created by introducing nucleic acid encoding a polypeptide of the invention (or a homologue thereof) into the male pronuclei of a fertilized oocyte, e.g., by microinjection, retroviral infection, and allowing the oocyte to develop in a pseudopregnant female foster animal. Intronic sequences and polyadenylation signals can also be included in the transgene to increase the efficiency of expression of the transgene. A tissue-specific regulatory sequence(s) can be operably linked to the transgene to direct expression of the polypeptide of the invention to particular cells. Methods for generating transgenic animals via embryo manipulation and microinjection, particularly animals such as mice, have become conventional in the art and are described, for example, in U.S. Patent NOS. 4,736,866 and 4,870,009, U.S. Patent No. 4,873,191 and in Hogan, *Manipulating the Mouse Embryo*, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1986). Similar methods are used for production of other transgenic animals. A transgenic

- 56 -

founder animal can be identified based upon the presence of the transgene in its genome and/or expression of mRNA encoding the transgene in tissues or cells of the animals. A transgenic founder animal can then be used to
5 breed additional animals carrying the transgene. Moreover, transgenic animals carrying the transgene can further be bred to other transgenic animals carrying other transgenes.

To create an homologous recombinant animal, a
10 vector is prepared which contains at least a portion of a gene encoding a polypeptide of the invention into which a deletion, addition or substitution has been introduced to thereby alter, e.g., functionally disrupt, the gene. In a preferred embodiment, the vector is designed such that,
15 upon homologous recombination, the endogenous gene is functionally disrupted (i.e., no longer encodes a functional protein; also referred to as a "knock out" vector). Alternatively, the vector can be designed such that, upon homologous recombination, the endogenous gene
20 is mutated or otherwise altered but still encodes functional protein (e.g., the upstream regulatory region can be altered to thereby alter the expression of the endogenous protein). In the homologous recombination vector, the altered portion of the gene is flanked at its
25 5' and 3' ends by additional nucleic acid of the gene to allow for homologous recombination to occur between the exogenous gene carried by the vector and an endogenous gene in an embryonic stem cell. The additional flanking nucleic acid sequences are of sufficient length for
30 successful homologous recombination with the endogenous gene. Typically, several kilobases of flanking DNA (both at the 5' and 3' ends) are included in the vector (see, e.g., Thomas and Capecchi (1987) *Cell* 51:503 for a description of homologous recombination vectors). The
35 vector is introduced into an embryonic stem cell line

- 57 -

(e.g., by electroporation) and cells in which the introduced gene has homologously recombined with the endogenous gene are selected (see, e.g., Li et al. (1992) *Cell* 69:915). The selected cells are then injected into
5 a blastocyst of an animal (e.g., a mouse) to form aggregation chimeras (see, e.g., Bradley in *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach*, Robertson, ed. (IRL, Oxford, 1987) pp. 113-152). A chimeric embryo can then be implanted into a
10 suitable pseudopregnant female foster animal and the embryo brought to term. Progeny harboring the homologously recombined DNA in their germ cells can be used to breed animals in which all cells of the animal contain the homologously recombined DNA by germline
15 transmission of the transgene. Methods for constructing homologous recombination vectors and homologous recombinant animals are described further in Bradley (1991) *Current Opinion in Bio/Technology* 2:823-829 and in PCT Publication NOS. WO 90/11354, WO 91/01140, WO
20 92/0968, and WO 93/04169.

In another embodiment, transgenic non-human animals can be produced which contain selected systems which allow for regulated expression of the transgene. One example of such a system is the *cre/loxP* recombinase
25 system of bacteriophage P1. For a description of the *cre/loxP* recombinase system, see, e.g., Lakso et al. (1992) *Proc. Natl. Acad. Sci. USA* 89:6232-6236. Another example of a recombinase system is the FLP recombinase system of *Saccharomyces cerevisiae* (O'Gorman et al.
30 (1991) *Science* 251:1351-1355. If a *cre/loxP* recombinase system is used to regulate expression of the transgene, animals containing transgenes encoding both the *Cre* recombinase and a selected protein are required. Such animals can be provided through the construction of
35 "double" transgenic animals, e.g., by mating two

- 58 -

transgenic animals, one containing a transgene encoding a selected protein and the other containing a transgene encoding a recombinase.

Clones of the non-human transgenic animals described herein can also be produced according to the methods described in Wilmut et al. (1997) *Nature* 385:810-813 and PCT Publication NOS. WO 97/07668 and WO 97/07669.

IV. Pharmaceutical Compositions

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

The invention includes methods for preparing pharmaceutical compositions for modulating the expression or activity of a polypeptide or nucleic acid of the invention. Such methods comprise formulating a pharmaceutically acceptable carrier with an agent which modulates expression or activity of a polypeptide or nucleic acid of the invention. Such compositions can

- 59 -

further include additional active agents. Thus, the invention further includes methods for preparing a pharmaceutical composition by formulating a pharmaceutically acceptable carrier with an agent which modulates expression or activity of a polypeptide or nucleic acid of the invention and one or more additional active compounds.

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

Pharmaceutical compositions suitable for injectable use include sterile aqueous solutions (where water soluble) or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersions. For intravenous administration, suitable carriers include physiological

- 60 -

saline, bacteriostatic water, Cremophor EL™ (BASF; Parsippany, NJ) or phosphate buffered saline (PBS). In all cases, the composition must be sterile and should be fluid to the extent that easy syringability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, and liquid polyethylene glycol, and the like), and suitable mixtures thereof. The proper fluidity can be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. Prevention of the action of microorganisms can be achieved by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, ascorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars, polyalcohols such as mannitol, sorbitol, sodium chloride in the composition. Prolonged absorption of the injectable compositions can be brought about by including in the composition an agent which delays absorption, for example, aluminum monostearate and gelatin.

Sterile injectable solutions can be prepared by incorporating the active compound (e.g., a polypeptide or antibody) in the required amount in an appropriate solvent with one or a combination of ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the active compound into a sterile vehicle which contains a basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile

- 61 -

injectable solutions, the preferred methods of preparation are vacuum drying and freeze-drying which yields a powder of the active ingredient plus any additional desired ingredient from a previously sterile-
5 filtered solution thereof.

Oral compositions generally include an inert diluent or an edible carrier. They can be enclosed in gelatin capsules or compressed into tablets. For the purpose of oral therapeutic administration, the active
10 compound can be incorporated with excipients and used in the form of tablets, troches, or capsules. Oral compositions can also be prepared using a fluid carrier for use as a mouthwash, wherein the compound in the fluid carrier is applied orally and swished and expectorated or
15 swallowed.

Pharmaceutically compatible binding agents, and/or adjuvant materials can be included as part of the composition. The tablets, pills, capsules, troches and the like can contain any of the following ingredients, or
20 compounds of a similar nature: a binder such as microcrystalline cellulose, gum tragacanth or gelatin; an excipient such as starch or lactose, a disintegrating agent such as alginic acid, Primogel, or corn starch; a lubricant such as magnesium stearate or Sterotes; a
25 glidant such as colloidal silicon dioxide; a sweetening agent such as sucrose or saccharin; or a flavoring agent such as peppermint, methyl salicylate, or orange flavoring.

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

Systemic administration can also be by
35 transmucosal or transdermal means. For transmucosal or

- 62 -

transdermal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration, 5 detergents, bile salts, and fusidic acid derivatives. Transmucosal administration can be accomplished through the use of nasal sprays or suppositories. For transdermal administration, the active compounds are formulated into ointments, salves, gels, or creams as 10 generally known in the art.

The compounds can also be prepared in the form of suppositories (e.g., with conventional suppository bases such as cocoa butter and other glycerides) or retention enemas for rectal delivery.

15 In one embodiment, the active compounds are prepared with carriers that will protect the compound against rapid elimination from the body, such as a controlled release formulation, including implants and microencapsulated delivery systems. Biodegradable, 20 biocompatible polymers can be used, such as ethylene vinyl acetate, polyanhydrides, polyglycolic acid, collagen, polyorthoesters, and polylactic acid. Methods for preparation of such formulations will be apparent to those skilled in the art. The materials can also be 25 obtained commercially from Alza Corporation and Nova Pharmaceuticals, Inc. Liposomal suspensions (including liposomes targeted to infected cells with monoclonal antibodies to viral antigens) can also be used as pharmaceutically acceptable carriers. These can be 30 prepared according to methods known to those skilled in the art, for example, as described in U.S. Patent No. 4,522,811.

It is especially advantageous to formulate oral or parenteral compositions in dosage unit form for ease of 35 administration and uniformity of dosage. Dosage unit

- 63 -

form as used herein refers to physically discrete units suited as unitary dosages for the subject to be treated; each unit containing a predetermined quantity of active compound calculated to produce the desired therapeutic effect in association with the required pharmaceutical carrier. The specification for the dosage unit forms of the invention are dictated by and directly dependent on the unique characteristics of the active compound and the particular therapeutic effect to be achieved, and the limitations inherent in the art of compounding such an active compound for the treatment of individuals.

For antibodies, the preferred dosage is 0.1 mg/kg to 100 mg/kg of body weight (generally 10 mg/kg to 20 mg/kg). If the antibody is to act in the brain, a dosage of 50 mg/kg to 100 mg/kg is usually appropriate.

Generally, partially human antibodies and fully human antibodies have a longer half-life within the human body than other antibodies. Accordingly, lower dosages and less frequent administration is often possible. Modifications such as lipidation can be used to stabilize antibodies and to enhance uptake and tissue penetration (e.g., into the brain). A method for lipidation of antibodies is described by Cruikshank et al. ((1997) *J. Acquired Immune Deficiency Syndromes and Human Retrovirology* 14:193).

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

- 64 -

Alternatively, where the complete gene delivery vector can be produced intact from recombinant cells, e.g. retroviral vectors, the pharmaceutical preparation can include one or more cells which produce the gene delivery system.

The pharmaceutical compositions can be included in a container, pack, or dispenser together with instructions for administration.

V. Uses and Methods of the Invention

10 The nucleic acid molecules, proteins, protein homologues, and antibodies described herein can be used in one or more of the following methods: a) screening assays; b) detection assays (e.g., chromosomal mapping, tissue typing, forensic biology); c) predictive medicine
15 (e.g., diagnostic assays, prognostic assays, monitoring clinical trials, and pharmacogenomics); and d) methods of treatment (e.g., therapeutic and prophylactic). For example, polypeptides of the invention can be used to (i) modulate cellular proliferation; (ii) modulate cellular
20 differentiation; and (iii) modulate cell survival. The isolated nucleic acid molecules of the invention can be used to express proteins (e.g., via a recombinant expression vector in a host cell in gene therapy applications), to detect mRNA (e.g., in a biological
25 sample) or a genetic lesion, and to modulate activity of a polypeptide of the invention. In addition, the polypeptides of the invention can be used to screen drugs or compounds which modulate activity or expression of a polypeptide of the invention as well as to treat
30 disorders characterized by insufficient or excessive production of a protein of the invention or production of a form of a protein of the invention which has decreased or aberrant activity compared to the wild type protein. In addition, the antibodies of the invention

- 65 -

can be used to detect and isolate a protein of the and modulate activity of a protein of the invention.

This invention further pertains to novel agents identified by the above-described screening assays and
5 uses thereof for treatments as described herein.

A. Screening Assays

The invention provides a method (also referred to herein as a "screening assay") for identifying
10 modulators, i.e., candidate or test compounds or agents (e.g., peptides, peptidomimetics, small molecules or other drugs) which bind to polypeptide of the invention or have a stimulatory or inhibitory effect on, for example, expression or activity of a polypeptide of the invention.

15 In one embodiment, the invention provides assays for screening candidate or test compounds which bind to or modulate the activity of the membrane-bound form of a polypeptide of the invention or biologically active portion thereof. The test compounds of the present
20 invention can be obtained using any of the numerous approaches in combinatorial library methods known in the art, including: biological libraries; spatially addressable parallel solid phase or solution phase libraries; synthetic library methods requiring
25 deconvolution; the "one-bead one-compound" library method; and synthetic library methods using affinity chromatography selection. The biological library approach is limited to peptide libraries, while the other four approaches are applicable to peptide, non-peptide
30 oligomer or small molecule libraries of compounds (Lam (1997) *Anticancer Drug Des.* 12:145).

Examples of methods for the synthesis of molecular libraries can be found in the art, for example in: DeWitt et al. (1993) *Proc. Natl. Acad. Sci. USA* 90:6909;

- 66 -

Erb et al. (1994) *Proc. Natl. Acad. Sci. USA* 91:11422;
Zuckermann et al. (1994). *J. Med. Chem.* 37:2678; Cho et
al. (1993) *Science* 261:1303; Carrell et al. (1994) *Angew.
Chem. Int. Ed. Engl.* 33:2059; Carell et al. (1994) *Angew.
5 Chem. Int. Ed. Engl.* 33:2061; and Gallop et al. (1994) *J.
Med. Chem.* 37:1233.

Libraries of compounds may be presented in
solution (e.g., Houghten (1992) *Bio/Techniques* 13:412-
421), or on beads (Lam (1991) *Nature* 354:82-84), chips
10 (Fodor (1993) *Nature* 364:555-556), bacteria (U.S. Patent
No. 5,223,409), spores (Patent NOS. 5,571,698; 5,403,484;
and 5,223,409), plasmids (Cull et al. (1992) *Proc. Natl.
Acad. Sci. USA* 89:1865-1869) or phage (Scott and Smith
(1990) *Science* 249:386-390; Devlin (1990) *Science*
15 249:404-406; Cwirla et al. (1990) *Proc. Natl. Acad. Sci.
USA* 87:6378-6382; and Felici (1991) *J. Mol. Biol.*
222:301-310).

In one embodiment, an assay is a cell-based assay
in which a cell which expresses a membrane-bound form of
20 a polypeptide of the invention, or a biologically active
portion thereof, on the cell surface is contacted with a
test compound and the ability of the test compound to
bind to the polypeptide determined. The cell, for
example, can be a yeast cell or a cell of mammalian
25 origin. Determining the ability of the test compound to
bind to the polypeptide can be accomplished, for example,
by coupling the test compound with a radioisotope or
enzymatic label such that binding of the test compound to
the polypeptide or biologically active portion thereof
30 can be determined by detecting the labeled compound in a
complex. For example, test compounds can be labeled with
 ^{125}I , ^{35}S , ^{14}C , or ^3H , either directly or indirectly, and
the radioisotope detected by direct counting of
radioemmission or by scintillation counting.
35 Alternatively, test compounds can be enzymatically

- 67 -

labeled with, for example, horseradish peroxidase, alkaline phosphatase, or luciferase, and the enzymatic label detected by determination of conversion of an appropriate substrate to product. In a preferred
5 embodiment, the assay comprises contacting a cell which expresses a membrane-bound form of a polypeptide of the invention, or a biologically active portion thereof, on the cell surface with a known compound which binds the polypeptide to form an assay mixture, contacting the
10 assay mixture with a test compound, and determining the ability of the test compound to interact with the polypeptide, wherein determining the ability of the test compound to interact with the polypeptide comprises determining the ability of the test compound to
15 preferentially bind to the polypeptide or a biologically active portion thereof as compared to the known compound.

In another embodiment, an assay is a cell-based assay comprising contacting a cell expressing a membrane-bound form of a polypeptide of the invention, or a
20 biologically active portion thereof, on the cell surface with a test compound and determining the ability of the test compound to modulate (e.g., stimulate or inhibit) the activity of the polypeptide or biologically active portion thereof. Determining the ability of the test
25 compound to modulate the activity of the polypeptide or a biologically active portion thereof can be accomplished, for example, by determining the ability of the polypeptide protein to bind to or interact with a target molecule.

30 Determining the ability of a polypeptide of the invention to bind to or interact with a target molecule can be accomplished by one of the methods described above for determining direct binding. As used herein, a "target molecule" is a molecule with which a selected

- 68 -

polypeptide (e.g., a polypeptide of the invention binds or interacts with in nature, for example, a molecule on the surface of a cell which expresses the selected protein, a molecule on the surface of a second cell, a molecule in the extracellular milieu, a molecule associated with the internal surface of a cell membrane or a cytoplasmic molecule. A target molecule can be a polypeptide of the invention or some other polypeptide or protein. For example, a target molecule can be a component of a signal transduction pathway which facilitates transduction of an extracellular signal (e.g., a signal generated by binding of a compound to a polypeptide of the invention) through the cell membrane and into the cell or a second intercellular protein which has catalytic activity or a protein which facilitates the association of downstream signaling molecules with a polypeptide of the invention. Determining the ability of a polypeptide of the invention to bind to or interact with a target molecule can be accomplished by determining the activity of the target molecule. For example, the activity of the target molecule can be determined by detecting induction of a cellular second messenger of the target (e.g., intracellular Ca^{2+} , diacylglycerol, IP₃, etc.), detecting catalytic/enzymatic activity of the target on an appropriate substrate, detecting the induction of a reporter gene (e.g., a regulatory element that is responsive to a polypeptide of the invention operably linked to a nucleic acid encoding a detectable marker, e.g. luciferase), or detecting a cellular response, for example, cellular differentiation, or cell proliferation.

In yet another embodiment, an assay of the present invention is a cell-free assay comprising contacting a polypeptide of the invention or biologically active portion thereof with a test compound and determining the

- 69 -

ability of the test compound to bind to the polypeptide or biologically active portion thereof. Binding of the test compound to the polypeptide can be determined either directly or indirectly as described above. In a preferred embodiment, the assay includes contacting the polypeptide of the invention or biologically active portion thereof with a known compound which binds the polypeptide to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with the polypeptide, wherein determining the ability of the test compound to interact with the polypeptide comprises determining the ability of the test compound to preferentially bind to the polypeptide or biologically active portion thereof as compared to the known compound.

In another embodiment, an assay is a cell-free assay comprising contacting a polypeptide of the invention or biologically active portion thereof with a test compound and determining the ability of the test compound to modulate (e.g., stimulate or inhibit) the activity of the polypeptide or biologically active portion thereof. Determining the ability of the test compound to modulate the activity of the polypeptide can be accomplished, for example, by determining the ability of the polypeptide to bind to a target molecule by one of the methods described above for determining direct binding. In an alternative embodiment, determining the ability of the test compound to modulate the activity of the polypeptide can be accomplished by determining the ability of the polypeptide of the invention to further modulate the target molecule. For example, the catalytic/enzymatic activity of the target molecule on an appropriate substrate can be determined as previously described.

- 70 -

In yet another embodiment, the cell-free assay comprises contacting a polypeptide of the invention or biologically active portion thereof with a known compound which binds the polypeptide to form an assay mixture, 5 contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with the polypeptide, wherein determining the ability of the test compound to interact with the polypeptide comprises determining the ability of the polypeptide to 10 preferentially bind to or modulate the activity of a target molecule.

The cell-free assays of the present invention are amenable to use of both a soluble form or the membrane-bound form of a polypeptide of the invention. In the 15 case of cell-free assays comprising the membrane-bound form of the polypeptide, it may be desirable to utilize a solubilizing agent such that the membrane-bound form of the polypeptide is maintained in solution. Examples of such solubilizing agents include non-ionic detergents 20 such as n-octylglucoside, n-dodecylglucoside, n-octylmaltoside, octanoyl-N-methylglucamide, decanoyl-N-methylglucamide, Triton X-100, Triton X-114, Thesit, Isotridecylpoly(ethylene glycol ether)_n, 3-[(3-cholamidopropyl)dimethylamminio]-1-propane sulfonate 25 (CHAPS), 3-[(3-cholamidopropyl)dimethylamminio]-2-hydroxy-1-propane sulfonate (CHAPSO), or N-dodecyl=N,N-dimethyl-3-ammonio-1-propane sulfonate.

In more than one embodiment of the above assay methods of the present invention, it may be desirable to 30 immobilize either the polypeptide of the invention or its target molecule to facilitate separation of complexed from uncomplexed forms of one or both of the proteins, as well as to accommodate automation of the assay. Binding of a test compound to the polypeptide, or interaction of 35 the polypeptide with a target molecule in the presence

- 71 -

and absence of a candidate compound, can be accomplished in any vessel suitable for containing the reactants. Examples of such vessels include microtitre plates, test tubes, and micro-centrifuge tubes. In one embodiment, a fusion protein can be provided which adds a domain that allows one or both of the proteins to be bound to a matrix. For example, glutathione-S-transferase fusion proteins or glutathione-S-transferase fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical; St. Louis, MO) or glutathione derivatized microtitre plates, which are then combined with the test compound or the test compound and either the non-adsorbed target protein or A polypeptide of the invention, and the mixture incubated under conditions conducive to complex formation (e.g., at physiological conditions for salt and pH). Following incubation, the beads or microtitre plate wells are washed to remove any unbound components and complex formation is measured either directly or indirectly, for example, as described above. Alternatively, the complexes can be dissociated from the matrix, and the level of binding or activity of the polypeptide of the invention can be determined using standard techniques.

Other techniques for immobilizing proteins on matrices can also be used in the screening assays of the invention. For example, either the polypeptide of the invention or its target molecule can be immobilized utilizing conjugation of biotin and streptavidin. Biotinylated polypeptide of the invention or target molecules can be prepared from biotin-NHS (N-hydroxy-succinimide) using techniques well known in the art (e.g., biotinylation kit, Pierce Chemicals; Rockford, IL), and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical). Alternatively, antibodies reactive with the polypeptide of the invention

- 72 -

or target molecules but which do not interfere with binding of the polypeptide of the invention to its target molecule can be derivatized to the wells of the plate, and unbound target or polypeptide of the invention
5 trapped in the wells by antibody conjugation. Methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies reactive with the polypeptide of the invention or target
10 molecule, as well as enzyme-linked assays which rely on detecting an enzymatic activity associated with the polypeptide of the invention or target molecule.

In another embodiment, modulators of expression of a polypeptide of the invention are identified in a method
15 in which a cell is contacted with a candidate compound and the expression of the selected mRNA or protein (i.e., the mRNA or protein corresponding to a polypeptide or nucleic acid of the invention) in the cell is determined. The level of expression of the selected mRNA or protein
20 in the presence of the candidate compound is compared to the level of expression of the selected mRNA or protein in the absence of the candidate compound. The candidate compound can then be identified as a modulator of expression of the polypeptide of the invention based on
25 this comparison. For example, when expression of the selected mRNA or protein is greater (statistically significantly greater) in the presence of the candidate compound than in its absence, the candidate compound is identified as a stimulator of the selected mRNA or
30 protein expression. Alternatively, when expression of the selected mRNA or protein is less (statistically significantly less) in the presence of the candidate compound than in its absence, the candidate compound is identified as an inhibitor of the selected mRNA or
35 protein expression. The level of the selected mRNA or

- 73 -

protein expression in the cells can be determined by methods described herein.

In yet another aspect of the invention, a polypeptide of the inventions can be used as "bait proteins" in a two-hybrid assay or three hybrid assay (see, e.g., U.S. Patent No. 5,283,317; Zervos et al. (1993) *Cell* 72:223-232; Madura et al. (1993) *J. Biol. Chem.* 268:12046-12054; Bartel et al. (1993) *Bio/Techniques* 14:920-924; Iwabuchi et al. (1993) *Oncogene* 8:1693-1696; and PCT Publication No. WO 94/10300), to identify other proteins, which bind to or interact with the polypeptide of the invention and modulate activity of the polypeptide of the invention. Such binding proteins are also likely to be involved in the propagation of signals by the polypeptide of the inventions as, for example, upstream or downstream elements of a signaling pathway involving the polypeptide of the invention.

This invention further pertains to novel agents identified by the above-described screening assays and uses thereof for treatments as described herein.

B. Detection Assays

Portions or fragments of the cDNA sequences identified herein (and the corresponding complete gene sequences) can be used in numerous ways as polynucleotide reagents. For example, these sequences can be used to: (i) map their respective genes on a chromosome and, thus, locate gene regions associated with genetic disease; (ii) identify an individual from a minute biological sample (tissue typing); and (iii) aid in forensic identification of a biological sample. These applications are described in the subsections below.

1. Chromosome Mapping

- 74 -

Once the sequence (or a portion of the sequence) of a gene has been isolated, this sequence can be used to map the location of the gene on a chromosome.

Accordingly, nucleic acid molecules described herein or
5 fragments thereof, can be used to map the location of the corresponding genes on a chromosome. The mapping of the sequences to chromosomes is an important first step in correlating these sequences with genes associated with disease.

10 Briefly, genes can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp in length) from the sequence of a gene of the invention. Computer analysis of the sequence of a gene of the invention can be used to rapidly select primers that do not span more
15 than one exon in the genomic DNA, thus complicating the amplification process. These primers can then be used for PCR screening of somatic cell hybrids containing individual human chromosomes. Only those hybrids containing the human gene corresponding to the gene
20 sequences will yield an amplified fragment. For a review of this technique, see D'Eustachio et al. ((1983) *Science* 220:919-924).

PCR mapping of somatic cell hybrids is a rapid procedure for assigning a particular sequence to a
25 particular chromosome. Three or more sequences can be assigned per day using a single thermal cycler. Using the nucleic acid sequences of the invention to design oligonucleotide primers, sublocalization can be achieved with panels of fragments from specific chromosomes.
30 Other mapping strategies which can similarly be used to map a gene to its chromosome include *in situ* hybridization (described in Fan et al. (1990) *Proc. Natl. Acad. Sci. USA* 87:6223-27), pre-screening with labeled flow-sorted chromosomes (CITE), and pre-selection by
35 hybridization to chromosome specific cDNA libraries.

- 75 -

Fluorescence *in situ* hybridization (FISH) of a DNA sequence to a metaphase chromosomal spread can further be used to provide a precise chromosomal location in one step. For a review of this technique, see Verma et al.,
5 (Human Chromosomes: A Manual of Basic Techniques (Pergamon Press, New York, 1988)).

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

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

Moreover, differences in the DNA sequences between individuals affected and unaffected with a disease associated with a gene of the invention can be
30 determined. If a mutation is observed in some or all of the affected individuals but not in any unaffected individuals, then the mutation is likely to be the causative agent of the particular disease. Comparison of affected and unaffected individuals generally involves
35 first looking for structural alterations in the

- 76 -

chromosomes such as deletions or translocations that are visible from chromosome spreads or detectable using PCR based on that DNA sequence. Ultimately, complete sequencing of genes from several individuals can be performed to confirm the presence of a mutation and to distinguish mutations from polymorphisms.

2. Tissue Typing

The nucleic acid sequences of the present invention can also be used to identify individuals from minute biological samples. The United States military, for example, is considering the use of restriction fragment length polymorphism (RFLP) for identification of its personnel. In this technique, an individual's genomic DNA is digested with one or more restriction enzymes, and probed on a Southern blot to yield unique bands for identification. This method does not suffer from the current limitations of "Dog Tags" which can be lost, switched, or stolen, making positive identification difficult. The sequences of the present invention are useful as additional DNA markers for RFLP (described in U.S. Patent 5,272,057).

Furthermore, the sequences of the present invention can be used to provide an alternative technique which determines the actual base-by-base DNA sequence of selected portions of an individual's genome. Thus, the nucleic acid sequences described herein can be used to prepare two PCR primers from the 5' and 3' ends of the sequences. These primers can then be used to amplify an individual's DNA and subsequently sequence it.

Panels of corresponding DNA sequences from individuals, prepared in this manner, can provide unique individual identifications, as each individual will have a unique set of such DNA sequences due to allelic differences. The sequences of the present invention can

- 77 -

be used to obtain such identification sequences from individuals and from tissue. The nucleic acid sequences of the invention uniquely represent portions of the human genome. Allelic variation occurs to some degree in the coding regions of these sequences, and to a greater degree in the noncoding regions. It is estimated that allelic variation between individual humans occurs with a frequency of about once per each 500 bases. Each of the sequences described herein can, to some degree, be used as a standard against which DNA from an individual can be compared for identification purposes. Because greater numbers of polymorphisms occur in the noncoding regions, fewer sequences are necessary to differentiate individuals. The noncoding sequences of SEQ ID NO:1 or 4 can comfortably provide positive individual identification with a panel of perhaps 10 to 1,000 primers which each yield a noncoding amplified sequence of 100 bases. If predicted coding sequences, such as those in SEQ ID NO:3 or 6 are used, a more appropriate number of primers for positive individual identification would be 500-2,000.

If a panel of reagents from the nucleic acid sequences described herein is used to generate a unique identification database for an individual, those same reagents can later be used to identify tissue from that individual. Using the unique identification database, positive identification of the individual, living or dead, can be made from extremely small tissue samples.

3. Use of Partial Gene Sequences in Forensic Biology

DNA-based identification techniques can also be used in forensic biology. Forensic biology is a scientific field employing genetic typing of biological evidence found at a crime scene as a means for positively identifying, for example, a perpetrator of a crime. To

- 78 -

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

The sequences of the present invention can be used to provide polynucleotide reagents, e.g., PCR primers, targeted to specific loci in the human genome, which can enhance the reliability of DNA-based forensic identifications by, for example, providing another "identification marker" (i.e. another DNA sequence that is unique to a particular individual). As mentioned above, actual base sequence information can be used for identification as an accurate alternative to patterns formed by restriction enzyme generated fragments. Sequences targeted to noncoding regions are particularly appropriate for this use as greater numbers of polymorphisms occur in the noncoding regions, making it easier to differentiate individuals using this technique. Examples of polynucleotide reagents include the nucleic acid sequences of the invention or portions thereof, e.g., fragments derived from noncoding regions having a length of at least 20 or 30 bases.

The nucleic acid sequences described herein can further be used to provide polynucleotide reagents, e.g., labeled or labelable probes which can be used in, for example, an *in situ* hybridization technique, to identify a specific tissue, e.g., brain tissue. This can be very useful in cases where a forensic pathologist is presented with a tissue of unknown origin. Panels of such probes can be used to identify tissue by species and/or by organ type.

- 79 -

C. Predictive Medicine

The present invention also pertains to the field of predictive medicine in which diagnostic assays, prognostic assays, pharmacogenomics, and monitoring
5 clinical trails are used for prognostic (predictive) purposes to thereby treat an individual prophylactically. Accordingly, one aspect of the present invention relates to diagnostic assays for determining expression of a polypeptide or nucleic acid of the invention and/or
10 activity of a polypeptide of the invention, in the context of a biological sample (e.g., blood, serum, cells, tissue) to thereby determine whether an individual is afflicted with a disease or disorder, or is at risk of developing a disorder, associated with aberrant
15 expression or activity of a polypeptide of the invention. The invention also provides for prognostic (or predictive) assays for determining whether an individual is at risk of developing a disorder associated with aberrant expression or activity of a polypeptide of the
20 invention. For example, mutations in a gene of the invention can be assayed in a biological sample. Such assays can be used for prognostic or predictive purpose to thereby prophylactically treat an individual prior to the onset of a disorder characterized by or associated
25 with aberrant expression or activity of a polypeptide of the invention.

Another aspect of the invention provides methods for expression of a nucleic acid or polypeptide of the invention or activity of a polypeptide of the invention
30 in an individual to thereby select appropriate therapeutic or prophylactic agents for that individual (referred to herein as "pharmacogenomics"). Pharmacogenomics allows for the selection of agents (e.g., drugs) for therapeutic or prophylactic treatment
35 of an individual based on the genotype of the individual

- 80 -

(e.g., the genotype of the individual examined to determine the ability of the individual to respond to a particular agent).

Yet another aspect of the invention pertains to monitoring the influence of agents (e.g., drugs or other compounds) on the expression or activity of a polypeptide of the invention in clinical trials. These and other agents are described in further detail in the following sections.

10 1. Diagnostic Assays

An exemplary method for detecting the presence or absence of a polypeptide or nucleic acid of the invention in a biological sample involves obtaining a biological sample from a test subject and contacting the biological sample with a compound or an agent capable of detecting a polypeptide or nucleic acid (e.g., mRNA, genomic DNA) of the invention such that the presence of a polypeptide or nucleic acid of the invention is detected in the biological sample. A preferred agent for detecting mRNA or genomic DNA encoding a polypeptide of the invention is a labeled nucleic acid probe capable of hybridizing to mRNA or genomic DNA encoding a polypeptide of the invention. The nucleic acid probe can be, for example, a full-length cDNA, such as the nucleic acid of SEQ ID NO: 1, 3, 4, or 6, or a portion thereof, such as an oligonucleotide of at least 15, 30, 50, 100, 250 or 500 nucleotides in length and sufficient to specifically hybridize under stringent conditions to a mRNA or genomic DNA encoding a polypeptide of the invention. Other suitable probes for use in the diagnostic assays of the invention are described herein.

A preferred agent for detecting a polypeptide of the invention is an antibody capable of binding to a polypeptide of the invention, preferably an antibody with

- 81 -

a detectable label. Antibodies can be polyclonal, or more preferably, monoclonal. An intact antibody, or a fragment thereof (e.g., Fab or F(ab')₂) can be used. The term "labeled", with regard to the probe or antibody, is intended to encompass direct labeling of the probe or antibody by coupling (i.e., physically linking) a detectable substance to the probe or antibody, as well as indirect labeling of the probe or antibody by reactivity with another reagent that is directly labeled. Examples of indirect labeling include detection of a primary antibody using a fluorescently labeled secondary antibody and end-labeling of a DNA probe with biotin such that it can be detected with fluorescently labeled streptavidin. The term "biological sample" is intended to include tissues, cells and biological fluids isolated from a subject, as well as tissues, cells and fluids present within a subject. That is, the detection method of the invention can be used to detect mRNA, protein, or genomic DNA in a biological sample *in vitro* as well as *in vivo*. For example, *in vitro* techniques for detection of mRNA include Northern hybridizations and *in situ* hybridizations. *In vitro* techniques for detection of a polypeptide of the invention include enzyme linked immunosorbent assays (ELISAs), Western blots, immunoprecipitations and immunofluorescence. *In vitro* techniques for detection of genomic DNA include Southern hybridizations. Furthermore, *in vivo* techniques for detection of a polypeptide of the invention include introducing into a subject a labeled antibody directed against the polypeptide. For example, the antibody can be labeled with a radioactive marker whose presence and location in a subject can be detected by standard imaging techniques.

In one embodiment, the biological sample contains protein molecules from the test subject. Alternatively,

- 82 -

the biological sample can contain mRNA molecules from the test subject or genomic DNA molecules from the test subject. A preferred biological sample is a peripheral blood leukocyte sample isolated by conventional means
5 from a subject.

In another embodiment, the methods further involve obtaining a control biological sample from a control subject, contacting the control sample with a compound or agent capable of detecting a polypeptide of the invention
10 or mRNA or genomic DNA encoding a polypeptide of the invention, such that the presence of the polypeptide or mRNA or genomic DNA encoding the polypeptide is detected in the biological sample, and comparing the presence of the polypeptide or mRNA or genomic DNA encoding the
15 polypeptide in the control sample with the presence of the polypeptide or mRNA or genomic DNA encoding the polypeptide in the test sample.

The invention also encompasses kits for detecting the presence of a polypeptide or nucleic acid of the
20 invention in a biological sample (a test sample). Such kits can be used to determine if a subject is suffering from or is at increased risk of developing a disorder associated with aberrant expression of a polypeptide of the invention (e.g., a disorder of lipid metabolism or
25 transport). For example, the kit can comprise a labeled compound or agent capable of detecting the polypeptide or mRNA encoding the polypeptide in a biological sample and means for determining the amount of the polypeptide or mRNA in the sample (e.g., an antibody which binds the
30 polypeptide or an oligonucleotide probe which binds to DNA or mRNA encoding the polypeptide). Kits may also include instructions for observing that the tested subject is suffering from or is at risk of developing a disorder associated with aberrant expression of the
35 polypeptide if the amount of the polypeptide or mRNA

- 83 -

encoding the polypeptide is above or below a normal level.

For antibody-based kits, the kit may comprise, for example: (1) a first antibody (e.g., attached to a solid support) which binds to a polypeptide of the invention; and, optionally, (2) a second, different antibody which binds to either the polypeptide or the first antibody and is conjugated to a detectable agent.

For oligonucleotide-based kits, the kit may comprise, for example: (1) an oligonucleotide, e.g., a detectably labeled oligonucleotide, which hybridizes to a nucleic acid sequence encoding a polypeptide of the invention or (2) a pair of primers useful for amplifying a nucleic acid molecule encoding a polypeptide of the invention.

The kit may also comprise, e.g., a buffering agent, a preservative, or a protein stabilizing agent. The kit may also comprise components necessary for detecting the detectable agent (e.g., an enzyme or a substrate). The kit may also contain a control sample or a series of control samples which can be assayed and compared to the test sample contained. Each component of the kit is usually enclosed within an individual container and all of the various containers are within a single package along with instructions for observing whether the tested subject is suffering from or is at risk of developing a disorder associated with aberrant expression of the polypeptide.

2. Prognostic Assays

The methods described herein can furthermore be utilized as diagnostic or prognostic assays to identify subjects having or at risk of developing a disease or disorder associated with aberrant expression or activity of a polypeptide of the invention. For example, the

- 84 -

assays described herein, such as the preceding diagnostic assays or the following assays, can be utilized to identify a subject having or at risk of developing a disorder associated with aberrant expression or activity
5 of a polypeptide of the invention. Alternatively, the prognostic assays can be utilized to identify a subject having or at risk for developing such a disease or disorder. Thus, the present invention provides a method in which a test sample is obtained from a subject and a
10 polypeptide or nucleic acid (e.g., mRNA, genomic DNA) of the invention is detected, wherein the presence of the polypeptide or nucleic acid is diagnostic for a subject having or at risk of developing a disease or disorder associated with aberrant expression or activity of the
15 polypeptide. As used herein, a "test sample" refers to a biological sample obtained from a subject of interest. For example, a test sample can be a biological fluid (e.g., serum), cell sample, or tissue.

Furthermore, the prognostic assays described
20 herein can be used to determine whether a subject can be administered an agent (e.g., an agonist, antagonist, peptidomimetic, protein, peptide, nucleic acid, small molecule, or other drug candidate) to treat a disease or disorder associated with aberrant expression or activity
25 of a polypeptide of the invention. For example, such methods can be used to determine whether a subject can be effectively treated with a specific agent or class of agents (e.g., agents of a type which decrease activity of the polypeptide). Thus, the present invention provides
30 methods for determining whether a subject can be effectively treated with an agent for a disorder associated with aberrant expression or activity of a polypeptide of the invention in which a test sample is obtained and the polypeptide or nucleic acid encoding the
35 polypeptide is detected (e.g., wherein the presence of

- 85 -

the polypeptide or nucleic acid is diagnostic for a subject that can be administered the agent to treat a disorder associated with aberrant expression or activity of the polypeptide).

5 The methods of the invention can also be used to detect genetic lesions or mutations in a gene of the invention, thereby determining if a subject with the lesioned gene is at risk for a disorder characterized
10 aberrant expression or activity of a polypeptide of the invention. In preferred embodiments, the methods include detecting, in a sample of cells from the subject, the presence or absence of a genetic lesion or mutation characterized by at least one of an alteration affecting the integrity of a gene encoding the polypeptide of the
15 invention, or the mis-expression of the gene encoding the polypeptide of the invention. For example, such genetic lesions or mutations can be detected by ascertaining the existence of at least one of: 1) a deletion of one or more nucleotides from the gene; 2) an addition of one or
20 more nucleotides to the gene; 3) a substitution of one or more nucleotides of the gene; 4) a chromosomal rearrangement of the gene; 5) an alteration in the level of a messenger RNA transcript of the gene; 6) an aberrant modification of the gene, such as of the methylation
25 pattern of the genomic DNA; 7) the presence of a non-wild type splicing pattern of a messenger RNA transcript of the gene; 8) a non-wild type level of a the protein encoded by the gene; 9) an allelic loss of the gene; and 10) an inappropriate post-translational modification of
30 the protein encoded by the gene. As described herein, there are a large number of assay techniques known in the art which can be used for detecting lesions in a gene.

In certain embodiments, detection of the lesion involves the use of a probe/primer in a polymerase chain
35 reaction (PCR) (see, e.g., U.S. Patent Nos. 4,683,195 and

- 86 -

4,683,202), such as anchor PCR or RACE PCR, or, alternatively, in a ligation chain reaction (LCR) (see, e.g., Landegran et al. (1988) *Science* 241:1077-1080; and Nakazawa et al. (1994) *Proc. Natl. Acad. Sci. USA* 91:360-5 364), the latter of which can be particularly useful for detecting point mutations in a gene (see, e.g., Abravaya et al. (1995) *Nucleic Acids Res.* 23:675-682). This method can include the steps of collecting a sample of cells from a patient, isolating nucleic acid (e.g., 10 genomic, mRNA or both) from the cells of the sample, contacting the nucleic acid sample with one or more primers which specifically hybridize to the selected gene under conditions such that hybridization and amplification of the gene (if present) occurs, and 15 detecting the presence or absence of an amplification product, or detecting the size of the amplification product and comparing the length to a control sample. It is anticipated that PCR and/or LCR may be desirable to use as a preliminary amplification step in conjunction 20 with any of the techniques used for detecting mutations described herein.

Alternative amplification methods include: self sustained sequence replication (Guatelli et al. (1990) *Proc. Natl. Acad. Sci. USA* 87:1874-1878), transcriptional 25 amplification system (Kwoh, et al. (1989) *Proc. Natl. Acad. Sci. USA* 86:1173-1177), Q-Beta Replicase (Lizardi et al. (1988) *Bio/Technology* 6:1197), or any other nucleic acid amplification method, followed by the detection of the amplified molecules using techniques 30 well known to those of skill in the art. These detection schemes are especially useful for the detection of nucleic acid molecules if such molecules are present in very low numbers.

In an alternative embodiment, mutations in a 35 selected gene from a sample cell can be identified by

- 87 -

alterations in restriction enzyme cleavage patterns. For example, sample and control DNA is isolated, amplified (optionally), digested with one or more restriction endonucleases, and fragment length sizes are determined
5 by gel electrophoresis and compared. Differences in fragment length sizes between sample and control DNA indicates mutations in the sample DNA. Moreover, the use of sequence specific ribozymes (see, e.g., U.S. Patent No. 5,498,531) can be used to score for the presence of
10 specific mutations by development or loss of a ribozyme cleavage site.

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

In yet another embodiment, any of a variety of sequencing reactions known in the art can be used to
35 directly sequence the selected gene and detect mutations

- 88 -

by comparing the sequence of the sample nucleic acids with the corresponding wild-type (control) sequence. Examples of sequencing reactions include those based on techniques developed by Maxim and Gilbert ((1977) *Proc. Natl. Acad. Sci. USA* 74:560) or Sanger ((1977) *Proc. Natl. Acad. Sci. USA* 74:5463). It is also contemplated that any of a variety of automated sequencing procedures can be utilized when performing the diagnostic assays ((1995) *Bio/Techniques* 19:448), including sequencing by mass spectrometry (see, e.g., PCT Publication No. WO 94/16101; Cohen et al. (1996) *Adv. Chromatogr.* 36:127-162; and Griffin et al. (1993) *Appl. Biochem. Biotechnol.* 38:147-159).

Other methods for detecting mutations in a selected gene include methods in which protection from cleavage agents is used to detect mismatched bases in RNA/RNA or RNA/DNA heteroduplexes (Myers et al. (1985) *Science* 230:1242). In general, the technique of "mismatch cleavage" entails providing heteroduplexes formed by hybridizing (labeled) RNA or DNA containing the wild-type sequence with potentially mutant RNA or DNA obtained from a tissue sample. The double-stranded duplexes are treated with an agent which cleaves single-stranded regions of the duplex such as which will exist due to basepair mismatches between the control and sample strands. RNA/DNA duplexes can be treated with RNase to digest mismatched regions, and DNA/DNA hybrids can be treated with S1 nuclease to digest mismatched regions. In other embodiments, either DNA/DNA or RNA/DNA duplexes can be treated with hydroxylamine or osmium tetroxide and with piperidine in order to digest mismatched regions. After digestion of the mismatched regions, the resulting material is then separated by size on denaturing polyacrylamide gels to determine the site of mutation. See, e.g., Cotton et al. (1988) *Proc. Natl. Acad. Sci.*

- 89 -

USA 85:4397; Saleeba et al. (1992) *Methods Enzymol.*
217:286-295. In a preferred embodiment, the control DNA
or RNA can be labeled for detection.

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

20 In other embodiments, alterations in
electrophoretic mobility will be used to identify
mutations in genes. For example, single strand
conformation polymorphism (SSCP) may be used to detect
differences in electrophoretic mobility between mutant
25 and wild type nucleic acids (Orita et al. (1989) *Proc.*
Natl. Acad. Sci. USA 86:2766; see also Cotton (1993)
Mutat. Res. 285:125-144; Hayashi (1992) *Genet. Anal.*
Tech. Appl. 9:73-79). Single-stranded DNA fragments of
sample and control nucleic acids will be denatured and
30 allowed to renature. The secondary structure of single-
stranded nucleic acids varies according to sequence, and
the resulting alteration in electrophoretic mobility
enables the detection of even a single base change. The
DNA fragments may be labeled or detected with labeled
35 probes. The sensitivity of the assay may be enhanced by

- 90 -

using RNA (rather than DNA), in which the secondary structure is more sensitive to a change in sequence. In a preferred embodiment, the subject method utilizes heteroduplex analysis to separate double stranded
5 heteroduplex molecules on the basis of changes in electrophoretic mobility (Keen et al. (1991) *Trends Genet.* 7:5).

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

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

Alternatively, allele specific amplification
35 technology which depends on selective PCR amplification

- 91 -

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

The methods described herein may be performed, for example, by utilizing pre-packaged diagnostic kits comprising at least one probe nucleic acid or antibody reagent described herein, which may be conveniently used, e.g., in clinical settings to diagnose patients exhibiting symptoms or family history of a disease or illness involving a gene encoding a polypeptide of the invention. Furthermore, any cell type or tissue, preferably peripheral blood leukocytes, in which the polypeptide of the invention is expressed may be utilized in the prognostic assays described herein.

3. Pharmacogenomics

Agents, or modulators which have a stimulatory or inhibitory effect on activity or expression of a

- 92 -

polypeptide of the invention as identified by a screening assay described herein can be administered to individuals to treat (prophylactically or therapeutically) disorders associated with aberrant activity of the polypeptide. In conjunction with such treatment, the pharmacogenomics (i.e., the study of the relationship between an individual's genotype and that individual's response to a foreign compound or drug) of the individual may be considered. Differences in metabolism of therapeutics can lead to severe toxicity or therapeutic failure by altering the relation between dose and blood concentration of the pharmacologically active drug. Thus, the pharmacogenomics of the individual permits the selection of effective agents (e.g., drugs) for prophylactic or therapeutic treatments based on a consideration of the individual's genotype. Such pharmacogenomics can further be used to determine appropriate dosages and therapeutic regimens. Accordingly, the activity of a polypeptide of the invention, expression of a nucleic acid of the invention, or mutation content of a gene of the invention in an individual can be determined to thereby select appropriate agent(s) for therapeutic or prophylactic treatment of the individual.

Pharmacogenomics deals with clinically significant hereditary variations in the response to drugs due to altered drug disposition and abnormal action in affected persons. See, e.g., Linder (1997) *Clin. Chem.* 43(2):254-266. In general, two types of pharmacogenetic conditions can be differentiated. Genetic conditions transmitted as a single factor altering the way drugs act on the body are referred to as "altered drug action." Genetic conditions transmitted as single factors altering the way the body acts on drugs are referred to as "altered drug metabolism". These pharmacogenetic conditions can occur

- 93 -

either as rare defects or as polymorphisms. For example, glucose-6-phosphate dehydrogenase deficiency (G6PD) is a common inherited enzymopathy in which the main clinical complication is haemolysis after ingestion of oxidant
5 drugs (anti-malarials, sulfonamides, analgesics, nitrofurans) and consumption of fava beans.

As an illustrative embodiment, the activity of drug metabolizing enzymes is a major determinant of both the intensity and duration of drug action. The discovery
10 of genetic polymorphisms of drug metabolizing enzymes (e.g., N-acetyltransferase 2 (NAT 2) and cytochrome P450 enzymes CYP2D6 and CYP2C19) has provided an explanation as to why some patients do not obtain the expected drug effects or show exaggerated drug response and serious
15 toxicity after taking the standard and safe dose of a drug. These polymorphisms are expressed in two phenotypes in the population, the extensive metabolizer (EM) and poor metabolizer (PM). The prevalence of PM is different among different populations. For example, the
20 gene coding for CYP2D6 is highly polymorphic and several mutations have been identified in PM, which all lead to the absence of functional CYP2D6. Poor metabolizers of CYP2D6 and CYP2C19 quite frequently experience exaggerated drug response and side effects when they
25 receive standard doses. If a metabolite is the active therapeutic moiety, a PM will show no therapeutic response, as demonstrated for the analgesic effect of codeine mediated by its CYP2D6-formed metabolite morphine. The other extreme are the so called ultra-
30 rapid metabolizers who do not respond to standard doses. Recently, the molecular basis of ultra-rapid metabolism has been identified to be due to CYP2D6 gene amplification.

Thus, the activity of a polypeptide of the
35 invention, expression of a nucleic acid encoding the

- 94 -

polypeptide, or mutation content of a gene encoding the polypeptide in an individual can be determined to thereby select appropriate agent(s) for therapeutic or prophylactic treatment of the individual. In addition, 5 pharmacogenetic studies can be used to apply genotyping of polymorphic alleles encoding drug-metabolizing enzymes to the identification of an individual's drug responsiveness phenotype. This knowledge, when applied to dosing or drug selection, can avoid adverse reactions 10 or therapeutic failure and thus enhance therapeutic or prophylactic efficiency when treating a subject with a modulator of activity or expression of the polypeptide, such as a modulator identified by one of the exemplary screening assays described herein.

15 4. Monitoring of Effects During Clinical Trials

Monitoring the influence of agents (e.g., drugs, compounds) on the expression or activity of a polypeptide of the invention (e.g., the ability to modulate aberrant cell proliferation and/or differentiation) can be applied 20 not only in basic drug screening, but also in clinical trials. For example, the effectiveness of an agent, as determined by a screening assay as described herein, to increase gene expression, protein levels or protein activity, can be monitored in clinical trials of subjects 25 exhibiting decreased gene expression, protein levels, or protein activity. Alternatively, the effectiveness of an agent, as determined by a screening assay, to decrease gene expression, protein levels or protein activity, can be monitored in clinical trials of subjects exhibiting 30 increased gene expression, protein levels, or protein activity.

For example, and not by way of limitation, genes, including those of the invention, that are modulated in cells by treatment with an agent (e.g., compound, drug or

- 95 -

small molecule) which modulates activity or expression of a polypeptide of the invention (e.g., as identified in a screening assay described herein) can be identified. Thus, to study the effect of agents on cellular
5 proliferation disorders, for example, in a clinical trial, cells can be isolated and RNA prepared and analyzed for the levels of expression of a gene of the invention and other genes implicated in the disorder. The levels of gene expression (i.e., a gene expression
10 pattern) can be quantified by Northern blot analysis or RT-PCR, as described herein, or alternatively by measuring the amount of protein produced, by one of the methods as described herein, or by measuring the levels of activity of a gene of the invention or other genes.
15 In this way, the gene expression pattern can serve as a marker, indicative of the physiological response of the cells to the agent. Accordingly, this response state may be determined before, and at various points during, treatment of the individual with the agent.

20 In a preferred embodiment, the present invention provides a method for monitoring the effectiveness of treatment of a subject with an agent (e.g., an agonist, antagonist, peptidomimetic, protein, peptide, nucleic acid, small molecule, or other drug candidate identified
25 by the screening assays described herein) comprising the steps of (i) obtaining a pre-administration sample from a subject prior to administration of the agent; (ii) detecting the level of the polypeptide or nucleic acid of the invention in the preadministration sample; (iii)
30 obtaining one or more post-administration samples from the subject; (iv) detecting the level the of the polypeptide or nucleic acid of the invention in the post-administration samples; (v) comparing the level of the polypeptide or nucleic acid of the invention in the pre-
35 administration sample with the level of the polypeptide

- 96 -

or nucleic acid of the invention in the post-administration sample or samples; and (vi) altering the administration of the agent to the subject accordingly. For example, increased administration of the agent may be desirable to increase the expression or activity of the polypeptide to higher levels than detected, i.e., to increase the effectiveness of the agent. Alternatively, decreased administration of the agent may be desirable to decrease expression or activity of the polypeptide to lower levels than detected, i.e., to decrease the effectiveness of the agent.

C. Methods of Treatment

The present invention provides for both prophylactic and therapeutic methods of treating a subject at risk of (or susceptible to) a disorder or having a disorder associated with aberrant expression or activity of a polypeptide of the invention. Such disorders can include, e.g., disorders of lipoprotein metabolism, disorders of lipoprotein transport, neurodegenerative disorders, neuropsychiatric disorders, and clotting disorders.

1. Prophylactic Methods

In one aspect, the invention provides a method for preventing in a subject, a disease or condition associated with an aberrant expression or activity of a polypeptide of the invention, by administering to the subject an agent which modulates expression or at least one activity of the polypeptide. Subjects at risk for a disease which is caused or contributed to by aberrant expression or activity of a polypeptide of the invention can be identified by, for example, any or a combination of diagnostic or prognostic assays as described herein. Administration of a prophylactic agent can occur prior to

- 97 -

the manifestation of symptoms characteristic of the aberrancy, such that a disease or disorder is prevented or, alternatively, delayed in its progression. Depending on the type of aberrancy, for example, an agonist or
5 antagonist agent can be used for treating the subject. The appropriate agent can be determined based on screening assays described herein.

2. Therapeutic Methods

Another aspect of the invention pertains to
10 methods of modulating expression or activity of a polypeptide of the invention for therapeutic purposes. The modulatory method of the invention involves contacting a cell with an agent that modulates one or more of the activities of the polypeptide. An agent that
15 modulates activity can be an agent as described herein, such as a nucleic acid or a protein, a naturally-occurring cognate ligand of the polypeptide, a peptide, a peptidomimetic, or other small molecule. In one embodiment, the agent stimulates one or more of the
20 biological activities of the polypeptide. Examples of such stimulatory agents include the active polypeptide of the invention and a nucleic acid molecule encoding the polypeptide of the invention that has been introduced into the cell. In another embodiment, the agent inhibits
25 one or more of the biological activities of the polypeptide of the invention. Examples of such inhibitory agents include antisense nucleic acid molecules and antibodies. These modulatory methods can be performed *in vitro* (e.g., by culturing the cell with
30 the agent) or, alternatively, *in vivo* (e.g., by administering the agent to a subject). As such, the present invention provides methods of treating an individual afflicted with a disease or disorder characterized by aberrant expression or activity of a

- 98 -

polypeptide of the invention. In one embodiment, the method involves administering an agent (e.g., an agent identified by a screening assay described herein), or combination of agents that modulates (e.g., upregulates
5 or downregulates) expression or activity. In another embodiment, the method involves administering a polypeptide of the invention or a nucleic acid molecule of the invention as therapy to compensate for reduced or aberrant expression or activity of the polypeptide.

10 Stimulation of activity is desirable in situations in which activity or expression is abnormally low or downregulated and/or in which increased activity is likely to have a beneficial effect. Conversely, inhibition of activity is desirable in situations in
15 which activity or expression is abnormally high or upregulated and/or in which decreased activity is likely to have a beneficial effect.

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

Equivalents

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

30 What is claimed is:

- 99 -

1. An isolated nucleic acid molecule selected from the group consisting of:

- a) a nucleic acid molecule comprising a nucleotide sequence which is at least 55% identical to
5 the nucleotide sequence of SEQ ID NO:1, 3, 4, or 6, the cDNA insert of the plasmid deposited with ATCC as Accession Number 98880, or a complement thereof;
- b) a nucleic acid molecule comprising a fragment of at least 300 nucleotides of the nucleotide sequence of
10 SEQ ID NO:1, 3, 4, or 6, the cDNA insert of the plasmid deposited with ATCC as Accession Number 98880, or a complement thereof;
- c) a nucleic acid molecule which encodes a polypeptide comprising the amino acid sequence of SEQ ID
15 NO:2 or 5, amino acid sequence encoded by the cDNA insert of the plasmid deposited with ATCC as Accession Number 98880;
- d) a nucleic acid molecule which encodes a fragment of a polypeptide comprising the amino acid
20 sequence of SEQ ID NO:2 or 5, or the polypeptide encoded by the cDNA insert of the plasmid deposited with ATCC as Accession Number 98880, wherein the fragment comprises at least 15 contiguous amino acids of SEQ ID NO:2 or 5, or the polypeptide encoded by the cDNA insert of the plasmid
25 deposited with ATCC as Accession Number 98880; and
- e) a nucleic acid molecule which encodes a naturally occurring allelic variant of a polypeptide comprising the amino acid sequence of SEQ ID NO:2 or 5, or the amino acid sequence encoded by the cDNA insert of
30 the plasmid deposited with ATCC as Accession Number 98880, wherein the nucleic acid molecule hybridizes to a nucleic acid molecule comprising SEQ ID NO:3 or 6 or a complement thereof under stringent conditions.

- 100 -

2. The isolated nucleic acid molecule of claim 1, which is selected from the group consisting of:

a) a nucleic acid comprising the nucleotide sequence of SEQ ID NO:1, 3, 4, or 5, or the cDNA insert
5 of the plasmid deposited with ATCC as Accession Number 98880, or a complement thereof; and

b) a nucleic acid molecule which encodes a polypeptide comprising the amino acid sequence of SEQ ID NO:2 or 5, the amino acid sequence encoded by the cDNA
10 insert of the plasmid deposited with ATCC as Accession Number 98880.

3. The nucleic acid molecule of claim 1 further comprising vector nucleic acid sequences.

4. The nucleic acid molecule of claim 1 further
15 comprising nucleic acid sequences encoding a heterologous polypeptide.

5. A host cell which contains the nucleic acid molecule of claim 1.

6. The host cell of claim 5 which is a mammalian
20 host cell.

7. A non-human mammalian host cell containing the nucleic acid molecule of claim 1.

8. An isolated polypeptide selected from the group consisting of:

a) a fragment of a polypeptide comprising the
25 amino acid sequence of SEQ ID NO:2 or 5, wherein the fragment comprises at least 15 contiguous amino acids of SEQ ID NO:2 or 5;

- 101 -

b) a naturally occurring allelic variant of a polypeptide comprising the amino acid sequence of SEQ ID NO:2 or 5, or the amino acid sequence encoded by the cDNA insert of the plasmid deposited with ATCC as Accession Number 98880, wherein the polypeptide is encoded by a nucleic acid molecule which hybridizes to a nucleic acid molecule comprising SEQ ID NO:1 or 4 or a complement thereof under stringent conditions; and

c) a polypeptide which is encoded by a nucleic acid molecule comprising a nucleotide sequence which is at least 55% identical to a nucleic acid comprising the nucleotide sequence of SEQ ID NO:1 or 4 or a complement thereof.

9. The isolated polypeptide of claim 8 comprising the amino acid sequence of SEQ ID NO:2 or 5, or the amino acid sequence encoded by the cDNA insert of the plasmid deposited with ATCC as Accession Number 98880.

10. The polypeptide of claim 8 further comprising heterologous amino acid sequences.

11. An antibody which selectively binds to a polypeptide of claim 8.

12. A method for producing a polypeptide selected from the group consisting of:

a) a polypeptide comprising the amino acid sequence of SEQ ID NO:2 or 5, or the amino acid sequence encoded by the cDNA insert of the plasmid deposited with ATCC as Accession Number 98880;

b) a polypeptide comprising a fragment of the amino acid sequence of SEQ ID NO:2 or 5, or the amino acid sequence encoded by the cDNA insert of the plasmid

- 102 -

deposited with ATCC as Accession Number 98880, wherein the fragment comprises at least 15 contiguous amino acids of SEQ ID NO:2 or 5, or the amino acid sequence encoded by the cDNA insert of the plasmid deposited with ATCC as
5 Accession Number 98880 or the amino acid sequence encoded by the cDNA insert of the plasmid deposited with ATCC as Accession Number 98880; and

c) a naturally occurring allelic variant of a polypeptide comprising the amino acid sequence of SEQ ID
10 NO:2 or 5, or the amino acid sequence encoded by the cDNA insert of the plasmid deposited with ATCC as Accession Number 98880, wherein the polypeptide is encoded by a nucleic acid molecule which hybridizes to a nucleic acid molecule comprising SEQ ID NO:1 or 4 or a complement
15 thereof under stringent conditions;

comprising culturing the host cell of claim 5 under conditions in which the nucleic acid molecule is expressed.

13. A method for detecting the presence of a
20 polypeptide of claim 8 in a sample, comprising:

a) contacting the sample with a compound which selectively binds to a polypeptide of claim 8; and

b) determining whether the compound binds to the polypeptide in the sample.

25 14. The method of claim 13, wherein the compound which binds to the polypeptide is an antibody.

15. A kit comprising a compound which selectively binds to a polypeptide of claim 8 and instructions for use.

- 103 -

16. A method for detecting the presence of a nucleic acid molecule of claim 1 in a sample, comprising the steps of:

- a) contacting the sample with a nucleic acid probe or primer which selectively hybridizes to the nucleic acid molecule; and
- b) determining whether the nucleic acid probe or primer binds to a nucleic acid molecule in the sample.

17. The method of claim 16, wherein the sample comprises mRNA molecules and is contacted with a nucleic acid probe.

18. A kit comprising a compound which selectively hybridizes to a nucleic acid molecule of claim 1 and instructions for use.

19. A method for identifying a compound which binds to a polypeptide of claim 8 comprising the steps of:

- a) contacting a polypeptide, or a cell expressing a polypeptide of claim 8 with a test compound; and
- b) determining whether the polypeptide binds to the test compound.

20. The method of claim 19, wherein the binding of the test compound to the polypeptide is detected by a method selected from the group consisting of:

- a) detection of binding by direct detecting of test compound/polypeptide binding;
- b) detection of binding using a competition binding assay;
- c) detection of binding using an assay for TANGO 136-mediated signal transduction.

- 104 -

21. A method for modulating the activity of a polypeptide of claim 8 comprising contacting a polypeptide or a cell expressing a polypeptide of claim 8 with a compound which binds to the polypeptide in a sufficient concentration to modulate the activity of the polypeptide.

22. A method for identifying a compound which modulates the activity of a polypeptide of claim 8, comprising:

- 10 a) contacting a polypeptide of claim 8 with a test compound; and
- b) determining the effect of the test compound on the activity of the polypeptide to thereby identify a compound which modulates the activity of the polypeptide.

GAATTCGGCAGAGACCCGGCGAGGAGCTTCCTGCCAGGTAGACAAGAGGGCGCCCGGGACCCTGCCGGCACCCTGT 79

M L S A L P L L F L L L G G A L A 17
 CAACCCAGG ATG CTA TCG GCC CTC CCT CTT CTT TTC CTC CTC CTC GGA GGT GCT CTG GCC 139

R P D R I T F P R S A C E A P P A V L S 37
 CGC CCA GAC CGG ATC ACT TTC CCA CGT TCT GCT TGT GAG GCG CCC CCA GCA GTG CTC TCA 199

E V Q G T L Q R P L G R D S R S S P A N 57
 GAA GTA CAA GGC ACT TTA CAG AGG CCT CTA GGC CGG GAC AGC CGA AGC TCC CCT GCC AAC 259

C T W V I L G S K D Q T V T V R F Q K L 77
 TGC ACC TGG GTT ATC CTG GGC AGC AAG GAT CAG ACA GTA ACT GTC AGG TTC CAG AAG CTG 319

H L A C G S E H L I L H S P L Q P P I S 97
 CAT CTG GCC TGT GGC TCA GAG CAT TTA ATC CTG CAC TCC CCT CTA CAG CCA CCA ATC TCC 379

L C E A P S G P L Q L P G G N V T I T Y 117
 CTG TGT GAG GCC CCT TCT GGC CCT TTG CAG CTA CCA GGG GGC AAT GTC ACC ATT ACA TAC 439

S Y A G A R A P M G Q G F L L T Y S Q D 137
 AGC TAT GCT GGG GCC AGA GCA CCC ATG GGC CAG GGC TTC TTG CTG ACT TAC AGT CAA GAT 499

W L L C L Q E E F Q C L N H R C I P A A 157
 TGG CTA CTG TGC CTG CAA GAA GAA TTC CAG TGC CTG AAC CAC CGC TGC ATT CCA GCT GCT 559

Q R C D G I D A C G D G S D E A G C S S 177
 CAG CGC TGT GAT GGG ATT GAT GCC TGC GGG GAT GGC TCA GAC GAG GCA GGC TGC AGC TCA 619

D P F P N L N P A P A P T L A C N L T L 197
 GAT CCA TTC CCT AAC CTG AAC CCA GCC CCC GCA CCA ACT CTG GCC TGC AAT CTC ACC TTG 679

E D F Y G V F S S P G Y S H L A S V S H 217
 GAG GAC TTT TAT GGG GTC TTT TCT TCC CCT GGA TAT TCA CAC CTG GCC TCA GTC TCC CAC 739

P Q S C L W L L D P H D G R R L A V R F 237
 CCC CAG TCC TGC CTG TGG CTG CTG GAC CCC CAT GAT GGC CGG AGG CTG GCA GTG CGC TTC 799

T A L D L S Y G D A V H V Y D G A G P P 257
 ACA GCC CTG GAC TTG AGT TAC GGA GAT GCA GTG CAT GTG TAT GAT GGT GCT GGA CCC CCC 859

E T P R L L R S L T H F S N G K A V T V 277
 GAG ACC CCT CGA CTG CTA CGT AGC CTC ACC CAC TTC AGC AAT GGC AAG GCT GTC ACT GTG 919

E T L S G Q A V V S Y H T V A W S S G R 297
 GAG ACC CTG TCT GGT CAG GCT GTT GTG TCC TAC CAC ACA GTT GCT TGG AGC AGT GGC CGG 979

G F N A T Y H V R G Y C L P W D R P C G 317
 GGC TTT AAT GCT ACC TAC CAT GTC CGG GGC TAC TGT TTA CCT TGG GAC AGA CCC TGT GGC 1039

L J S G L G A S E N L G E R C Y S E A Q 337
 TTG GGC TCT GGC CTG GGG GCT AGT GAG AAC CTA GGT GAG CGT TGC TAT AGC GAG GCA CAG 1099

R C D G S W D C A D G T D E E G C P G C 357
 CGC TGT GAT GGC TCA TGG GAC TGT GCC GAT GGC ACA GAT GAG GAG GGT TCC CCT GGC TGC 1159

P P G H F P C G A A J T P G A T A C Y L 377
 CCA CCA GGG CAC TTC CCC TGT GGA GCT GCA GGC ACC CCT GGT GCC ACA GCC TCC TAC CTG 1219

FIGURE 1 (SHEET 1 OF 2)

P	A	D	R	C	N	Y	Q	T	F	C	A	D	G	A	D	E	R	R	C	397
CCT	GCT	GAC	CGC	TGC	AAC	TAC	CAG	ACG	TTC	TGC	GCC	GAT	GGA	GCG	GAT	GAG	AGG	CGC	TGC	1279
R	H	C	Q	P	G	N	F	R	C	R	D	E	K	C	V	Y	E	T	W	417
CGG	CAT	TGC	CAG	CCC	GGC	AAC	TTC	CSG	TGC	CGG	GAT	GAG	AAG	TGT	GTG	TAT	GAG	ACA	TGG	1339
V	C	D	G	Q	P	D	C	T	D	G	S	D	E	W	D	C	S	Y	A	437
JTG	TGT	GAT	GGG	CAG	CCA	GAC	TGT	ACT	GAC	GGC	AGT	GAT	GAG	TGG	GAC	TGC	TCC	TAC	GCC	1399
L	P	R	K	V	I	T	A	A	V	I	G	S	L	V	C	G	L	L	L	457
CTG	CCC	CGA	AAA	GTC	ATC	ACA	GCA	GCA	GTC	ATT	GGC	AGC	CTG	GTG	TGT	GGC	CTG	TTG	CTG	1459
V	I	A	L	G	C	T	C	K	L	Y	A	I	R	T	Q	E	Y	S	M	477
GTC	ATC	GCT	CTC	GGC	TGC	ACC	TGC	AAA	CTC	TAT	GCC	ATC	CGC	ACC	CAG	GAA	TAC	AGC	ATG	1519
F	A	P	L	C	R	M	E	A	E	I	V	Q	Q	Q	A	P	P	S	Y	497
TTT	GCC	CCG	CTC	TGC	CGG	ATG	GAG	GCT	GAG	ATT	GTG	CAG	CAA	CAG	GCA	CCC	CCT	TCC	TAT	1579
G	Q	L	I	A	Q	G	A	I	P	P	V	E	D	F	P	T	E	N	P	517
GGG	CAG	CTC	ATT	GCC	CAG	GGT	GCC	ATC	CCG	CCT	GTG	GAA	GAC	TTC	CCC	ACA	GAG	AAC	CCT	1639
N	D	N	S	V	L	G	N	L	R	S	L	L	Q	I	L	R	Q	D	M	537
AAC	GAT	AAC	TCT	GTG	CTG	GGA	AAC	CTA	CGT	TCT	CTG	CTT	CAG	ATC	TTA	CGC	CAG	GAT	ATG	1699
T	P	G	G	T	S	G	G	R	R	R	Q	R	G	R	S	V	R	R	L	557
ACT	CCA	GGT	GGC	ACT	TCA	GGG	GGC	CGC	CGT	CGC	CAG	CGT	GGA	CGC	TCC	GTC	CGC	CGT	CTG	1759
V	R	R	L	R	R	W	G	L	L	P	R	T	N	T	L	L	E			575
GTT	CGC	CGT	CTC	CGT	CGT	TGG	GGC	CTG	CTT	CCT	CGA	ACT	AAT	ACC	CTG	CTC	GAG			1813

FIGURE 1 (SHEET 2 OF 2)

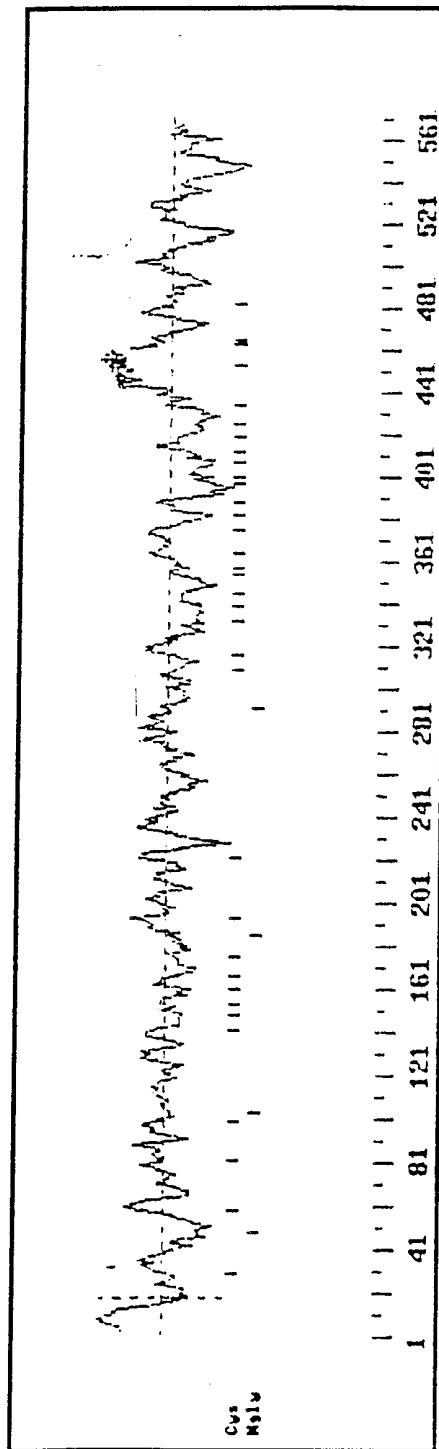


FIGURE 2

GTCCGACCCACGCGTCCGCGCCTGGGCAAGGGCCGGGGCCGGGGCCGAGCCACCTCTTCCCCTCCCCCGCTTCCCTGTC	79
GCGCTCCGCTGGCTGGACGCGCTGGAGGAGTGGAGCAGCACCCGGCCGGCCCTGGGGGCTGACAGTCGGCAAAGTTTGG	158
CCCGAAGAGGAAGTGGTCTCAAACCCCGGCAGGTGGCGACCAGGCCAGACCAGGGGGCGCTCGCTGCCTGCGGGCGGGCT	237
GTAGGCGAGGGCGCGCCCCAGTGCCGAGACCCGGGGCTTCAGGAGCCCGGCCCGGGAGAGAAGAGTGC GGCGGGCGGACG	316
GAGAAAACAACCTCCAAAGTTGGCGAAAGGCACCCGCCCTACTCCCGGGCTGCCGCGCCCTCCCCGCCCCAGCCCTGGC	395
ATCCAGAGTACGGGTTCGAGCCCGGGCCATGGAGCCCCCTGGGGAGGCGGCACCAGGGAGCCTGGGGCGCCCGGGGCTCC	474
GCCGCGACCCCATCGGGTAGACCACAGAAGCTCCGGGACCCTTCCGGCACCTCTGGACAGCCCAGG	549
A T L L L L L L G G A L A H P D R I I F	23
GCC ACC CTC CTC CTC CTC CTT GGA GGC GCT CTG GCC CAT CCA GAC CGG ATT ATT TTT	609
P N H A C E D P P A V L L E V Q G T L Q	43
CCA AAT CAT GCT TGT GAG GAC CCC CCA GCA GTG CTC TTA GAA GTG CAG GGC ACC TTA CAG	669
R P L V R D S R T S P A N C T W L I L G	63
AGG CCC CTG GTC CGG GAC AGC CGC ACC TCC CCT GCC AAC TGC ACC TGG CTC ATC CTG GGC	729
S K E Q T V T I R F Q K L H L A C G S E	83
AGC AAG GAA CAG ACT GTC ACC ATC AGG TTC CAG AAG CTA CAC CTG GCC TGT GGC TCA GAG	789
R L T L R S P L Q P L I S L C E A P P S	103
CGC TTA ACC CTA CGC TCC CCT CTC CAG CCA CTG ATC TCC CTG TGT GAG GCA CCT CCC AGC	849
P L Q L P G G N V T I T Y S Y A G A R A	123
CCT CTG CAG CTG CCC GGG GGC AAC GTC ACC ATC ACT TAC AGC TAT GCT GGG GCC AGA GCA	909
P M G Q G F L L S Y S Q D W L M C L Q E	143
CCC ATG GGC CAG GGC TTC CTG CTC TCC TAC AGC CAA GAT TGG CTG ATG TGC CTG CAG GAA	969
E F Q C L N H R C V S A V Q R C D G V D	163
GAG TTT CAG TGC CTG AAC CAC CGC TGT GTA TCT GCT GTC CAG CGC TGT GAT GGG GTT GAT	1029
A C G D G S D E A G C S S D P F P G L T	183
GCC TGT GGC GAT GGC TCT GAT GAA GCA GGT TGC AGC TCA GAC CCC TTC CCT GGC CTG ACC	1089
P R P V P S L P C N V T L E D F Y G V F	203
CCA AGA CCC GTC CCC TCC CTG CCT TGC AAT GTC ACC TTG GAG GAC TTC TAT GGG GTC TTC	1149
S S P G Y T H L A S V S H P Q S C H W L	223
TCC TCT CCT GGA TAT ACA CAC CTA GCC TCA GTC TCC CAC CCC CAG TCC TGC CAT TGG CTG	1209
L D P H D G R R L A V R F T A L D L G F	243
CTG GAC CCC CAT GAT GGC CGG CGG CTG GCC GTG CGC TTC ACA GCC CTG GAC TTG GGC TTT	1269
G D A V H V Y D G P G P P E S S R L L R	263
GGA GAT GCA GTG CAT GTG TAT GAC GGC CCT GGG CCC CCT GAG AGC TCC CGA CTA CTG CGT	1329
S L T H F S N G K A V T V E T L S G Q A	283
AGT CTC ACC CAC TTC AGC AAT GGC AAG GCT GTC ACT GTG GAG ACA CTG TCT GGC CAG GCT	1389
V V S Y H T V A W S N J R J F N A T Y H	303

FIGURE 3 (SHEET 1 OF 3)

GTT GTG TCC TAC CAC ACA GTT GCT TGG AGC/AAT GGT CGT GGC TTC AAT GCC ACC TAC CAT 1449
V R G Y C L P W D R P C G L G S G L G A 323
GTG CCG GGC TAT TGC TTG CCT TGG GAC AGA CCC TGT GGC TTA GGC TCT GGC CTG GGA GCT 1509
G E G L G E R C Y S E A Q R C D G S W D 343
GGC GAA GGC CTA GGT GAG CGC TGC TAC AGT GAG GCA CAG CGC TGT GAC GGC TCA TGG GAC 1569
C A D G T D E E D C P G C P P G H F P C 363
TGT GCT GAC GGC ACA GAT GAG GAG GAC TGC CCA GGC TGC CCA CCT GGA CAC TTC CCC TGT 1629
G A A G T S G A T A C Y L P A D R C N Y 383
GGG GCT GCT GGC ACC TCT GGT GCC ACA GCC TGC TAC CTG CCT GCT GAC CGC TGC AAC TAC 1689
Q T F C A D G A D E R R C R H C Q P G N 403
CAG ACT TTC TGT GCT GAT GGA GCA GAT GAG AGA CGC TGT CGG CAT TGC CAG CCT GGC AAT 1749
F R C R D E K C V Y E T W V C D G Q P D 423
TTC CGA TGC CCG GAC GAG AAG TGC GTG TAT GAG ACG TGG GTG TGC GAT GGG CAG CCA GAC 1809
C A D G S D E W D C S Y V L P R K V I T 443
TGT GCG GAC GGC AGT GAT GAG TGG GAC TGC TCC TAT GTT CTG CCC CGC AAG GTC ATT ACA 1869
A A V I G S L V C G L L L V I A L G C T 463
GCT GCA GTC ATT GGC AGC CTA GTG TGC GGC CTG CTC CTG GTC ATC GCC CTG GGC TGC ACC 1929
C K L Y A I R T Q E Y S I F A P L S R M 483
TGC AAG CTC TAT GCC ATT CGC ACC CAG GAG TAC AGC ATC TTT GCC CCC CTC TCC CGG ATG 1989
E A E I V Q Q Q A P P S Y G Q L I A Q G 503
GAG GCT GAG ATT GTG CAG CAG CAG GCA CCC CCT TCC TAC GGG CAG CTC ATT GCC CAG GGT 2049
A I P P V E D F P T E N P N D N S V L G 523
GCC ATC CCA CCT GTA GAA GAC TTT CCT ACA GAG AAT CCT AAT GAT AAC TCA GTG CTG GGC 2109
N L R S L L Q I L R Q D M T P G G G P G 543
AAC CTG CGT TCT CTG CTA CAG ATC TTA CGC CAG GAT ATG ACT CCA GGA GGT GGC CCA GGT 2169
A R R R Q R G R L M R R L V R R L R R W 563
JCC CGC CGT CGT CAG CGG GGC CGC TTG ATG CGA CGC CTG GTA CCC CGT CTC CGC CGC TGG 2229
G L L P R T N T P A R A S E A R S Q V T 583
GGC TTG CTC CCT CGA ACC AAC ACC CCG GCT CCG GCC TCT GAG GCC AGA TCC CAG GTC ACA 2289
P S A A P L E A L D G G T G P A R E G G 603
CCT TCT GCT GCT CCC CTT GAG GCC CTA GAT GGT GGC ACA GGT CCA GCC CGT GAG GGC GGG 2349
A V G G Q D G E Q A P P L P I K A P L P 623
GCA GTG GGT GGG CAA GAT GGG GAG CAG GCA CCC CCA CTG CCC ATC AAG GCT CCC CTC CCA 2409
S A S T S P A P T T V P E A P G P L P S 643
TCT GCT AGC ACG TCT CCA GCC CCC ACT ACT GTC CCT GAA GCC CCA GGG CCA CTG CCC TCA 2469
L P L E P S L L S G V V Q A L R G R L L 663
CTG CCC CTA GAG CCA TCA CTA TTG TCT GGA GTG GTG CAG GCC CTG CCA GGC CGC CTG TTG 2529
P S L G P P G P T R S P P G P H T A V L 683
JCC AGC CTG GGG CCC CCA GGA CCA ACC CCG AGC CCC CCT GGA CCC CAC ACA GCA GTC CTG 2589
A L E D E D D V L L V P L A E P J V W V 703

FIGURE 3 (SHEET 2 OF 3)

6/17

3CC CTG GAA GAT GAG GAC GAT GTG CTA CTG' GTG CCA CTG GCT GAG CCG GGG GTG TGG GTA	2649
A E A E D E P L L . T *	714
3CT GAG GCA GAG GAT GAG CCA CTG CTT ACC TGA	2682
GGGGACCTGGGGCTCTACTGAGGCCTCTCCCTGGGGCTCTACTCATAGTGGCACACCTTTTACAGGTGGGTCAGC	2761
TTCCCTCCACCACCTTCCTTCCTCTCCCTGGATTCAGGGACTTGGTGGGCCTCCCGTTGACCCTATGTAGCTGCTAT	2840
AAAGTTAAGTGTCCCTCAGGCAGGGAGAGGGCTCACAGAGTCTCCTCTTACGTGGCCATGGCCAGACCCCCAGTCCC	2919
TTCCACCACCACCTGCTCCCCACGCCACCACCATTGGGGTGGCTGTTTTTAAAAAAGTAAAGTTCTTAGAGGATAAAAAAA	2998
AAAAAAAAAGGGCGGCCGC	3017

FIGURE 3 (SHEET 3 of 3)

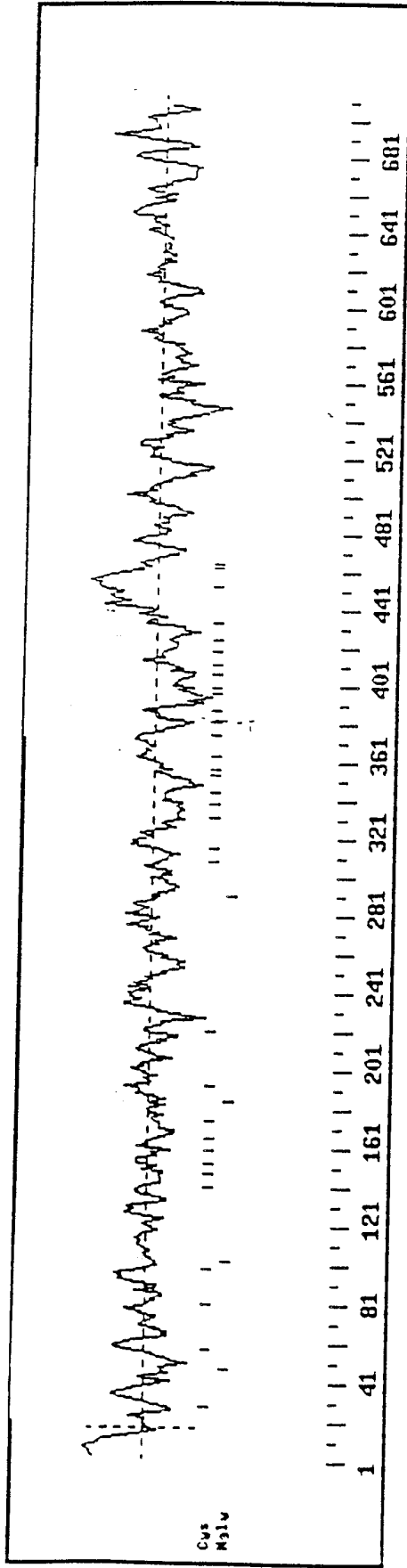


FIGURE 4

```

          10      20      30
      GAATTCGGCACGAGAC-CCGGCGAGGAGCTTCCTGCC
          : : : : : : : : : : : : : : : : : : :
CCTGGGGAGGCGGCACCAGGGAGCCTGGGCGCCCGGGGCTCCGCGCG--AC--CCCATC
      440      450      460      470      480

      40      50      60      70      80      90
AGGTAGACAAGAGGGCGCCCGGGACCCTGCCGGCACCCCTGTTCAACCCAGGATGCTATC
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
GGGTAGACCACAGAAGCTCCGGGACCCTCCGGCACCTCTGGACAGCCAGGATGCTGTT
      490      500      510      520      530      540

      100     110     120     130     140     150
GGCCCTCCCTCTTTCTTCTCCTCCTCGGAGGTGCTCTGGCCCGCCAGACCCGGATCAC
: : V : ^ : : : : : : : : : : : : : : : : : : : : : : : : : : :
GGCC-ACCCTCCTCCT--CCTCCTCCTTGGAGGCGCTCTGGCCCATCCAGACCGGATTAT
      550     560     570     580     590     600

      160     170     180     190     200     210
TTTTCCACGTTCTGCTTGTGAGGGCGCCCCAGCAGTGCTCTCAGAAGTACAAGGCACTTT
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
TTTTCCAAATCATGCTTGTGAGGACCCCGCAGCAGTGCTCTTAGAAGTGCAGGGCACCTT
      610     620     630     640     650     660

      220     230     240     250     260     270
ACAGAGGCCTCTAGGCCGGGACAGCCGAAGCTCCCCTGCCAACTGCACCTGGGTTATCCT
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
ACAGAGGCCCCCTGGTCCGGGACAGCCGCACCTCCCCTGCCAACTGCACCTGGCTCATCCT
      670     680     690     700     710     720

      280     290     300     310     320     330
GGGCAGCAAGGATCAGACAGTAACTGTCAGGTTCCAGAAGCTGCATCTGGCCTGTGGCTC
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
GGGCAGCAAGGAACAGACTGTCACCATCAGGTTCCAGAAGCTACACCTGGCCTGTGGCTC
      730     740     750     760     770     780

      340     350     360     370     380     390
AGAGCATTTAATCCTGCACTCCCTCTACAGCCACCAATCTCCCTGTGTGAGGCCCTTC
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
AGAGCCCTTAACCTACGCTCCCTCTCCAGCCACTGATCTCCCTGTGTGAGGCACCTCC
      790     800     810     820     830     840

      400     410     420     430     440     450

```

FIGURE 6 (1 of 4)

```

TGGCCCTTTGCAGCTACCAGGGGGCAATGTCAACCATACATACAGCTATGCTGGGGCCAG
.....
CAGCCCTCTGCAGCTGCCCGGGGCAACGTCAACCATCACTTACAGCTATGCTGGGGCCAG
  850      860      870      880      890      900

  460      470      480      490      500      510
AGCACCCATGGGCCAGGGCTTCTTGCTGACTTACAGTCAAGATTGGCTACTGTGCCTGCA
.....
AGCACCCATGGGCCAGGGCTTCTTGCTGACTTACAGTCAAGATTGGCTACTGTGCCTGCA
  910      920      930      940      950      960

  520      530      540      550      560      570
AGAAGAATTCAGTGCCTGAACCACCGCTGCATTCCAGCTGCTCAGCGCTGTGATGGGAT
.....
GGAAGAGTTTTCAGTGCCTGAACCACCGCTGTGTATCTGCTGTCAGCGCTGTGATGGGGT
  970      980      990      1000     1010     1020

  580      590      600      610      620      630
TGATGCCTGCGGGGATGGCTCAGACGAGGCAGGCTGCAGCTCAGATCCATTCCCTAACCT
.....
TGATGCCTGCGGGGATGGCTCAGACGAGGCAGGCTGCAGCTCAGATCCATTCCCTAACCT
  1030     1040     1050     1060     1070     1080

  640      650      660      670      680      690
GAACCCAGCCCCCGCACCACTCTGGCCTGCAATCTCACCTTGGAGGACTTTTATGGGGT
.....
GACCCCAAGACCCGTCCTCCCTGCCTTGCATGTCACTTGGAGGACTTCTATGGGGT
  1090     1100     1110     1120     1130     1140

  700      710      720      730      740      750
CTTTTCTTCCCTGGATATCACACCTGGCCTCAGTCTCCCACCCCGAGTCTGCCTGTG
.....
CTTCTCCTCTCCTGGATATACACACCTAGCCTCAGTCTCCCACCCCGAGTCTGCCTGTG
  1150     1160     1170     1180     1190     1200

  760      770      780      790      800      810
GCTGCTGGACCCCATGATGGCCGGAGGCTGGCAGTGGCCTTCCAGCCCTGGACTTGAG
.....
GCTGCTGGACCCCATGATGGCCGGAGGCTGGCAGTGGCCTTCCAGCCCTGGACTTGAG
  1210     1220     1230     1240     1250     1260

  820      830      840      850      860      870
TTACGGAGATGCAGTGCATGTGTATGATGGTGGACCCCGGAGACCCCTCGACTGCT
.....
CTTTGGAGATGCAGTGCATGTGTATGACGGCCCTGGGCCCCCTGAGAGCTCCCAGTACT
  1270     1280     1290     1300     1310     1320

  880      890      900      910      920      930
ACGTAGCCTCACCCACTTCAAGCAATGGCAAGGCTGCACTGTGGAGACCCTGTCTGGTCA
.....
GCGTAGTCTCACCCACTTCAAGCAATGGCAAGGCTGCACTGTGGAGACCCTGTCTGGTCA
  1330     1340     1350     1360     1370     1380

  940      950      960      970      980      990
GGCTGTTGTGTCTTACCACACAGTTGCTTGGAGCAGTGGCCGGGGCTTTAATGCTACCTA
.....
GGCTGTTGTGTCTTACCACACAGTTGCTTGGAGCAATGGTCTGGCTTCAATGCCACCTA
  1390     1400     1410     1420     1430     1440

1000     1010     1020     1030     1040     1050

```

FIGURE 6 (2 OF 4)


```

      10      20      30      40      50      60
MLSALPLLFLLLGGALARPDRIITFPRSACEAPPAVLSEVQGTLQRPLGRDSRSSPANCTW
X:.....:.....:.....:.....:.....:.....:
MLLATLLLLLLGGALAHPDRIIFPNHACEDPPAVLLEVQGTLQRPLVRDSRTSPANCTW
10      20      30      40      50

      70      80      90      100     110     120
VILGSKDQTVTVRFQKLHLACGSEHLILHSPLQPPISLCEAPSGPLQLPGGNVTTITYSYA
.....:.....:.....:.....:.....:.....:
LILGSKEQTVTIRFQKLHLACGSERLTLRSPLOPLISLCEAPPSPLQLPGGNVTTITYSYA
60      70      80      90      100     110

      130     140     150     160     170     180
GARAPMGQGFLLTYSQDWLLCLOEEFQCLNHRCPAAQRCDGIDACGDGSDGDEAGCSSDPF
.....:.....:.....:.....:.....:.....:
GARAPMGQGFLLSYSQDWLMCLQEEFQCLNHRCSVSAVQRCDGVDACGDGSDGDEAGCSSDPF
120     130     140     150     160     170

      190     200     210     220     230     240
PNLNPAPAPTLACNLTLEDYFVGFSSPGYSHLASVSHPOSCLWLLDPHDGRRRLAVRFTAL
.....:.....:.....:.....:.....:.....:
PGLTPRPVPSLPCNVLTLEDYFVGFSSPGYTHLASVSHPOSCHWLLDPHDGRRRLAVRFTAL
180     190     200     210     220     230

      250     260     270     280     290     300
DLSYGDVAVHVDGAGPPETPRLLRSLTHFSNGKAVTVETLSGQAVVSYHTVAWSSGRGFN
.....:.....:.....:.....:.....:.....:
DLGFGDAVHVDGPGPPPESSRLLRSLTHFSNGKAVTVETLSGQAVVSYHTVAWSNGRGFN
240     250     260     270     280     290

      310     320     330     340     350     360
ATYHVRYGYCLPWDRPCGLGSLGASENLGERCYSEAQRCDGSDCADGTDEEGCPGCPPG
.....:.....:.....:.....:.....:.....:
ATYHVRYGYCLPWDRPCGLGSLGAGEGLGERCYSEAQRCDGSDCADGTDEEDCPGCPPG
300     310     320     330     340     350

      370     380     390     400     410     420
HFPCGAAGTPGATACYLPADRCNYQTFCADGADERRCRHCQPGNFRRCRDEKCVYETWVCD
.....:.....:.....:.....:.....:.....:
HFPCGAAGTSGATACYLPADRCNYQTFCADGADERRCRHCQPGNFRRCRDEKCVYETWVCD
360     370     380     390     400     410

      430     440     450     460     470     480
QQPDCDEGSDWDCSYALPRKVITA AVIGSLVCGLLLVIALGCTCKLYAIRTQEYSMFAP
.....:.....:.....:.....:.....:.....:
QQPDCADGSDWDCSYVLPKRVITA AVIGSLVCGLLLVIALGCTCKLYAIRTQEYSIFAP
420     430     440     450     460     470

      490     500     510     520     530     540
LCRMEAEIVQQQAPPSYGQLIAQGAIPPVEDFPTENPNDNSVLGNLRSLLQLRQDMTPG
.....:.....:.....:.....:.....:.....:
LSRMEAEIVQQQAPPSYGQLIAQGAIPPVEDFPTENPNDNSVLGNLRSLLQLRQDMTPG
480     490     500     510     520     530

```

FIGURE 7 (cont)

1 30
MEKRAAAGLEGAPGARAQLAVVCLWNLFLTGRLS---SAVPALAACSGKLEQHTERRGVIYSPAWPLNYPPGTNCSWYIQ
MEKRAAAGPEGAPGARAPLAVVCLWNLFLTGRLS---SAVPALAACSGKLEQHTERRGVIYSPAWPLNYPPGTNCSWYIQ
-----MLSALPLLFLLLGGALARPDRITFPRSACEAPPAVLSEVQGTLQRPLGRDSRSSPANCTWVIL
-----MLLAT-LLLLLLGGALAHPDRIIFPNHACEDPPAVLLEVQGTLQRPLVRDSRTSPANCTWLIL

31 160
GDRGDMITISFRNFDVEESHQCSLDWLLGPAAPPRQEAFLRCGSAIPPAFISARDHVWIFFHSDASSSGQAQGFRLSYI
GDRGDMITISFRNFDVEESHQCSLDWLLGPAAPPRQEAFLRCGSAIPPAFISARDHVWIFFHSDASSSGQAQGFRLSYI
GSKDQTVTVRFQKLHLA----CGSEHLILH--SP-LQPPISLCEAPSGPLQLPG-GNVTITYSYAGARAPMGQGFLLTYS
GSKEQTVTIRFQKLHLA----CGSERLTLR--SP-LQPLISLCEAPSPPLQLPG-GNVTITYSYAGARAPMGQGFLLSYS

161 240
RGKLGQASCQADEFRCDNGKCLPGPWQCNTVDECGDGSDEGNCSAPASEPPGSLCPGGTFPCSGARSTRCLPVERRCDGL
RGKLGQASCQDDEFRCNDGKCLPGPWQCNMVDCEGDSDEGNCSAPASEPPGSLCPGGTFPCSGARSTRCLPVERRCDGT
QDWL--L-CLQEEFQCLNHRCIPAAQRCDGIDACGDSDEAGCS---SDP-----FP-----
QDWL--M-CLQEEFQCLNHRCVSAVQRCDGVDACGDSDEAGCS---SDP-----FP-----

241 320
QDCGEGSDEAGCPDLACGRRLGSFYGSFASPDLFGAARGPSDLHCTWLVDTQDSRRVLLQL-ELRLGYDDYVQVYEGGLGE
QDCGEGSDEAGCPDLACGRRLGSFYGSFASPDLFGAARGPSDLHCTWLVDTQDPRRVLLQL-ELRLGYDDYVQVYEGGLGE
-----NLNPAPAPTLACNLTLEDFYGVFSSPGYSHLASVSHQSCWLWLLDPHDGRRLAVRFTALDLSYGDVAVHYVDGAGP
-----GLTPRPVPSLPCNVTLEDFYGVFSSPGYTHLASVSHQSCHWLLDPHDGRRLAVRFTALDLGFGDAVHYVDGPGP

321 400
-RGDRLQLTSLYRSNHRPVSLEAAQGRITVAYHARARSAGHGFNATYQVKGYCLPWEQPCGSSS-DSDGGSGLDQGCFFSE
-RGDRLQLTSLYRSNHRPVSLEAAQGRITVAYHARARSAGHGFNATYQVKGYCLPWEQPCGSSS-EGDDGSTGEQGCFFSE
PETRLLRSLTHFSNGKAVTVETLSGQAVVSYHTVAWSSGRGFNATYHVRGYCLPWDRPCGLGSGLGASENLGER-CYSE
PESSRLLRSLTHFSNGKAVTVETLSGQAVVSYHTVAWSSGRGFNATYHVRGYCLPWDRPCGLGSGLGAGEGLGER-CYSE

401 480
PORCDGWWHCASGRDEQGCPCPPDQYPC--EG--GSGLCYTPADRCNNQKSCPDGADEKNCFCQPGTFHCGTNLCIFE
PORCDGWWHCASGRDEQGCPCPPDQYPC--EG--GSGLCYAPADRCNNQKSCPDGADEKNCFCQPGTFHCGTNLCIFE
AQRCDGSDWDCADGTDEEGCPCPPGHFPCGAAGTSGATACYLPAADRCNYQTFCADGADERRRRCQPGNFRRCRDEKCVYE
AQRCDGSDWDCADGTDEEDCPCPPGHFPCGAAGTSGATACYLPAADRCNYQTFCADGADERRRRCQPGNFRRCRDEKCVYE

481 560
TWRCDGQEDCQDGSDEHGCLAAVPRKVITAALIGSLVCGLLLVIAGLCAFKLYSLRTQEQYRAFETQMRLEAEFVRREAP
TWRCDGQEDCQDGSDEHGCLAAVPRKVITAALIGSLVCGLLLVIAGLCAFKLYSLRTQEQYRAFETQMRLEAEFVRREAP
TWVCDGQPDCTDGSDEWDCSYALPRKVITAAVIGSLVCGLLLVIAGLCTCKLYAIRTQEQYSMFAP-LCRMEAEIVQQQAP
TWVCDGQPDCAWDCSYVLPVKVITAAVIGSLVCGLLLVIAGLCTCKLYAIRTQEQYSIFAP-LSRMEAEIVQQQAP

561 640
PSYGQLIAQGLIPPVEDFPVYSASQASVQLNLRMTAMR--RQMRRHASRRGSPRRRLGRLWNRLFHRPRAPRGQIPLLTAA
PSYGQLIAQGLIPPVEDFPVYSASQASVQLNLRMTAMR--RQMRRHASRRGSPRRRLGRLWNRLFHRPRAPRGQIPLLTAA
PSYGQLIAQGAIPPVEDFPTEPNPDHNSVLGNLRSLLQILRQDMTPGGTSGGRRRQRGRSVRRLVRRLLR-WGLLPRNTNL
PSYGQLIAQGAIPPVEDFPTEPNPDHNSVLGNLRSLLQILRQDMTPGGGPGARRRQRGRLMRRLVRRLLR-WGLLPRNTNP

641 720
RPSQTVLGDGFLQAPGAAPDP-PAPLMDTGSTR-AAGDRPPSAPGRAPEVGPSPGFLPSGLRD-PECRPVDKDRKVCRE
RPSQTVLGDGLLQAAPGPPVDP-PVPNTDTGSPR-EAGDGPSPSGSHAPEVGPSPVPPPLNLRD-PEYRPEDKERRKACVD
LE-----
ARASEARSCVTPSAAPLEALEGGTTPAREGGAVGGQDGEQAPPLPIKAPLPSASTSPAPTTPPEAPGPLEPLPLEPSLLS

721 785
PLADGFPAPADAPREPCACDPHPQVSTASSTLQPHSPEPLGVCRNPPPPCSFMLENSDDEALLVC
PLEDSPAPVDTFPEPCACDPHPQVSTASSTLQPHSPEPLGVCRNPPPPCSFMLENSDDEALLVC

VVVQALRGRLLPFLSPPGSTRDPDPHTAVLAGEDEDDVLLVPLAEPG--VVVAEADEDEPLT--

FIGURE 5

CGGnFfTeMnGtIqSPNYPNpYPNNMeCVWrIrVPPGYRIeLqFQdFDI ErHDeCWYDYV
 ** E GT Q P ** + **C+W+I + *** + FQ + + C+ + +
 32 PPAVL-SEVCGTLQRPLGRDRSSPANCTWVILGSKDQTVTVRFQKLHL A----CGSEHL

86

*CGGnFfTeMnGtIqSPNYPNpYPNNMe..CVWrIrVPPGYRIeLqFQdF DIErHDeCWYDYVEIRDGmWddSPliGRFCG.....deiPEdiIStsNR
 * ** + *G+++SP+Y + CW + *G+R + F+ + D+ + D V **DG ** P R *** ** + S +
 193 NWLTL-EDFYGVFSSPGYSHLASVSHPOSQLWLLDPHDGRRLAVRFTAL DLS-----YGDAVHVYDG--AGPPETPRLLRSLTHFSNGKAVTVETLSGQ

MWIEFrSDnSWqkRCfkAyymai*
 + *** + **RGF+A+Y+
 284 AVVSYHTVAWSSGRGFNATYHVR

306

FIGURE 8

*CGGnFfTeMnGtIqSPNYPNpYPNNMeCVWrIrVPPGYRIeLqFQdFDI
 ++ E GT Q P ++ + ++C+W I + +++ + FQ + +
 31 PPAVL-LEVQGTLQRPLVRDSRTSPANCTWLILGSKEQVTIRFQKLHL 78

ErHDeCWYDYVEIRDGmWddSPliGRFCGdeiPEdIiStsNRMWIEFrSD
 C+ + + R+ + + ++ +P +++ ++ +++I++
 79 A----CGSERLTLRSP---L--QPLISLCEAPPSPLQLPGGNVTITYSYA 119

nSWq..kRGFKAYYmai*
 + + GF ++Y +
 120 GARAPMGQGFLLSYSQD 136

*CGGnFfTeMnGtIqSPNYPNpYPNNMe..CVWrIrVPPGYRIeLqFQdF
 C ++ + +G+++SP+Y + C W + +G+R + F+ +
 192 CNVTL-EDFYGVFSSPGYTHLASVSHPQSCHWLLDPHDGRRLAVRFTAL 239

DIeRDeCWYDYVEIRDGm..WddSPliGRFCG..deiPEdIiStsNRMW
 D+ D V ++DG+ +++S+L++ + +++ ++ + S +
 240 DLG-----FGDAVHVYDGPPESSRLLRSLTHFSNGKAVTVETLSGQAV 284

IeFrSDnSWqkRGFKAYYmai*
 + +++ + ++RGF+A+Y+
 285 VSYHTVAWSNGRCFNATYHVR 305

FIGURE 9

tTCrPDEFQcnsGeRCIPmswvCDGDpDCeDWSDEWpeNCmp
 C EFQC + + C++ +CDG + C+D+SDE ++C +
 138 LMCLQEEFQCLNHR-CVSAVQRCDGVDACGDGSDE--AGCSS 176

GeRCIPmswvCDGDpDCeDWSDEWpeNCmp
 GERC + +CDG++DC D+ DE E+C++
 328 GERCYSEAQRCDGSWDCADGTDE--EDCPG 355

vCDGDpDCeDWSDEWpeNCmp
 +C+ + C D+ DE ++C+
 380 RCNYQTFCADGADE--RRCRH 398

CrPDEFQcnsGeRCIPmswvCDGDpDCeDWSDEWpeNCmp*
 C+P F+C++ + C+ wvCDG+pDC D+SDE +C
 399 CQPGNFRRCRDEK-CVYETWVCDGQPDCADGSDE--wDCSY 435

FIGURE 10

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/25178

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) : Please See Extra Sheet.
US CL : 536/23.5; 435/6, 7.1, 69.1, 69.7, 320.1, 252.3; 530/350, 388.22
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
U.S. : 536/23.5; 435/6, 7.1, 69.1, 69.7, 320.1, 252.3; 530/350, 388.22

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
APS, MEDLINE
search terms: ldl receptor

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X -- Y	ISHII et al. cDNA cloning of a new low-density lipoprotein receptor-related protein and mapping of its gene (LRP3) to chromosome bands 19q12-q13.2. Genomics. July 1998, Vol. 51, No. 1, pages 132-135, see paragraph bridging pages 132 and 134; Figure 1; page 134, paragraph bridging columns 1-2, column 2, full paragraph 1; Figure 2.	1, 8, 11, 13, 14, 16, 17, 19, 20 -- 3-7, 10, 12, 15, 18, 21, 22
A	YAMAMOTO et al. Close encounters with apolipoprotein E receptors. Current Opinion in Lipidology. 1996, Vol. 7, pages 298-302.	1-22

Further documents are listed in the continuation of Box C. See patent family annex.

* Special categories of cited documents:	*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
A document defining the general state of the art which is not considered to be of particular relevance	*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
E earlier document published on or after the international filing date	*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
L document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	*G* document member of the same patent family
O document referring to an oral disclosure, use, exhibition or other means	
P document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search
30 JANUARY 2000

Date of mailing of the international search report
09 FEB 2000

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231
Facsimile No. (703) 305-3230

Authorized officer
DAVID S. ROMEO
Telephone No. (703) 308-0196

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/25178

A. CLASSIFICATION OF SUBJECT MATTER:

IPC (7):

C07H 21/04; C12N 1/21, 15/00, 15/63; C07K 14/705; G01N 33/53; C12Q 1/68