

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
28 March 2002 (28.03.2002)

PCT

(10) International Publication Number
WO 02/24911 A2

(51) International Patent Classification⁷: C12N 15/12,
C07K 14/47, C12Q 1/68, C12N 5/10, 15/62, C07K 16/18,
G01N 33/68

(21) International Application Number: PCT/US01/29373

(22) International Filing Date:
19 September 2001 (19.09.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/233,790 19 September 2000 (19.09.2000) US

(71) Applicant (for all designated States except US): MIL-
LENNIUM PHARMACEUTICALS, INC. [US/US]; 75
Sydney Street, Cambridge, MA 02139 (US).

(72) Inventor; and

(75) Inventor/Applicant (for US only): CURTIS, Rory, A., J.
[GB/US]; 31 Constitution Drive, Southborough, MA 01772
(US).

(74) Agents: MANDRAGOURAS, Amy, E.; Lahive & Cock-
field, LLP, 28 State Street, Boston, MA 02109 et al. (US).

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

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

Published:

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

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



WO 02/24911 A2

(54) Title: 63760, A NOVEL HUMAN TRANSPORTER AND USES THEREOF

(57) Abstract: The invention provides isolated nucleic acids molecules, designated TP-2 nucleic acid molecules, which encode novel transporter family molecules. The invention also provides antisense nucleic acid molecules, recombinant expression vectors containing TP-2 nucleic acid molecules, host cells into which the expression vectors have been introduced, and nonhuman transgenic animals in which a TP-2 gene has been introduced or disrupted. The invention still further provides isolated TP-2 polypeptides, fusion polypeptides, antigenic peptides and anti-TP-2 antibodies. Diagnostic methods utilizing compositions of the invention are also provided.

63760, A NOVEL HUMAN TRANSPORTER AND USES THEREOF**Related Applications**

This application claims priority to U.S. Provisional Patent Application No. 5 60/233,790 entitled "63760, A NOVEL HUMAN TRANSPORTER AND USES THEREOF," filed on September 19, 2000. The contents of the above-referenced patent application are incorporated herein in their entirety by reference.

Background of the Invention

10 Cellular membranes serve to differentiate the contents of a cell from the surrounding environment, and may also serve as effective barriers against the unregulated influx of hazardous or unwanted compounds, and the unregulated efflux of desirable compounds. Membranes are by nature impervious to the unfacilitated diffusion of hydrophilic compounds such as proteins, water molecules, and ions due to 15 their structure: a bilayer of lipid molecules in which the polar head groups face outwards (towards the exterior and interior of the cell) and the nonpolar tails face inwards (at the center of bilayer, forming a hydrophobic core). Membranes enable a cell to maintain a relatively higher intra-cellular concentration of desired compounds and a relatively lower intra-cellular concentration of undesired compounds than are contained within the 20 surrounding environment.

Membranes also present a structural difficulty for cells, in that most desired compounds cannot readily enter the cell, nor can most waste products readily exit the cell through this lipid bilayer. The import and export of such compounds is facilitated by proteins which are embedded (singly or in complexes) in the cellular membrane. 25 There are several general classes of membrane transport proteins: channels/pores, permeases, and transporters. The former are integral membrane proteins which form a regulated passage through a membrane. This regulation, or "gating" is generally specific to the molecules to be transported by the pore or channel, rendering these transmembrane constructs selectively permeable to a specific class of substrates. For 30 example, a calcium channel is constructed such that only ions having a like charge and size to that of calcium may pass through. Channel and pore proteins tend to have discrete hydrophobic and hydrophilic domains, such that the hydrophobic face of the protein may associate with the interior of the membrane while the hydrophilic face lines the interior of the channel, thus providing a sheltered hydrophilic environment through 35 which the selected hydrophilic molecule may pass. This pore/channel-mediated system of facilitated diffusion is limited to ions and other very small molecules, due to the fact that pores or channels sufficiently large to permit the passage of whole proteins by

facilitated diffusion would be unable to prevent the simultaneous passage of smaller hydrophilic molecules.

Transport of larger molecules takes place by the action of “permeases” and “transporters”, two other classes of membrane-localized proteins which serve to move charged molecules from one side of a cellular membrane to the other. Unlike channel molecules, which permit diffusion-limited solute movement of a particular solute, these proteins require an energetic input, either in the form of a diffusion gradient (permeases) or through coupling to hydrolysis of an energy providing molecule (*e.g.*, ATP or GTP) (transporters). The permeases (integral membrane proteins often having between 6-14 membrane-spanning α -helices) enable the facilitated diffusion of molecules such as glucose or other sugars into the cell when the concentration of these molecules on one side of the membrane is greater than that on the other. Permeases do not form open channels through the membrane, but rather bind to the target molecule at the surface of the membrane and then undergo a conformational shift such that the target molecule is released on the opposite side of the membrane.

Transporters, in contrast, permit the movement of target molecules across membranes against the existing concentration gradient (active transport), a situation in which facilitated diffusion cannot occur. There are two general mechanisms used by cells for this type of membrane transport: symport/antiport, and energy-coupled transport, such as that mediated by the ABC transporters. Symport and antiport systems couple the movement of two different molecules across the membrane (via molecules having two separate binding sites for the two different molecules); in symport, both molecules are transported in the same direction, while in antiport, one molecule is imported while the other is exported. This is possible energetically because one of the two molecules moves in accordance with a concentration gradient, and this energetically favorable event is permitted only upon concomitant movement of a desired compound against the prevailing concentration gradient.

Single molecules may also be transported across the membrane against the concentration gradient in an energy-driven process, such as that utilized by the ABC transporters. In this ABC transporter system, the transport protein located in the membrane has an ATP-binding cassette; upon binding of the target molecule, the ATP is converted to ADP and inorganic phosphate (P_i), and the resulting release of energy is used to drive the movement of the target molecule to the opposite face of the membrane, facilitated by the transporter.

Transport molecules are specific for a particular target solute or class of solutes, and are also present in one or more specific membranes. Transport molecules localized to the plasma membrane permit an exchange of solutes with the surrounding

environment, while transport molecules localized to intra-cellular membranes (*e.g.*, membranes of the mitochondrion, peroxisome, lysosome, endoplasmic reticulum, nucleus, or vacuole) permit import and export of molecules from organelle to organelle or to the cytoplasm. For example, in the case of the mitochondrion, transporters in the
5 inner and outer mitochondrial membranes permit the import of sugar molecules, calcium ions, and water (among other molecules) into the organelle and the export of newly synthesized ATP to the cytosol.

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

Summary of the Invention

20 The present invention is based, at least in part, on the discovery of novel human transporter family members, referred to herein as “transporter-2” or “TP-2” nucleic acid and polypeptide molecules. The TP-2 nucleic acid and polypeptide molecules of the present invention are useful as modulating agents in regulating a variety of cellular processes, *e.g.*, cellular growth, migration, or proliferation. Accordingly, in one aspect,
25 this invention provides isolated nucleic acid molecules encoding TP-2 polypeptides or biologically active portions thereof, as well as nucleic acid fragments suitable as primers or hybridization probes for the detection of TP-2-encoding nucleic acids.

In one embodiment, the invention features an isolated nucleic acid molecule that includes the nucleotide sequence set forth in SEQ ID NO:1 or SEQ ID NO:3. In another
30 embodiment, the invention features an isolated nucleic acid molecule that encodes a polypeptide including the amino acid sequence set forth in SEQ ID NO:2. In another embodiment, the invention features an isolated nucleic acid molecule that includes the nucleotide sequence contained in the plasmid deposited with ATCC® as Accession Number ____.

35 In still other embodiments, the invention features isolated nucleic acid molecules including nucleotide sequences that are substantially identical (*e.g.*, 60% identical) to the nucleotide sequence set forth as SEQ ID NO:1 or SEQ ID NO:3. The invention further features isolated nucleic acid molecules including at least 50 contiguous

nucleotides of the nucleotide sequence set forth as SEQ ID NO:1 or SEQ ID NO:3. In another embodiment, the invention features isolated nucleic acid molecules which encode a polypeptide including an amino acid sequence that is substantially identical (*e.g.*, 60% identical) to the amino acid sequence set forth as SEQ ID NO:2. The present invention also features nucleic acid molecules which encode allelic variants of the polypeptide having the amino acid sequence set forth as SEQ ID NO:2. In addition to isolated nucleic acid molecules encoding full-length polypeptides, the present invention also features nucleic acid molecules which encode fragments, for example, biologically active or antigenic fragments, of the full-length polypeptides of the present invention (*e.g.*, fragments including at least 10 contiguous amino acid residues of the amino acid sequence of SEQ ID NO:2). In still other embodiments, the invention features nucleic acid molecules that are complementary to, antisense to, or hybridize under stringent conditions to the isolated nucleic acid molecules described herein.

In another aspect, the invention provides vectors including the isolated nucleic acid molecules described herein (*e.g.*, TP-2-encoding nucleic acid molecules). Such vectors can optionally include nucleotide sequences encoding heterologous polypeptides. Also featured are host cells including such vectors (*e.g.*, host cells including vectors suitable for producing TP-2 nucleic acid molecules and polypeptides).

In another aspect, the invention features isolated TP-2 polypeptides and/or biologically active or antigenic fragments thereof. Exemplary embodiments feature a polypeptide including the amino acid sequence set forth as SEQ ID NO:2, a polypeptide including an amino acid sequence at least 60% identical to the amino acid sequence set forth as SEQ ID NO:2, a polypeptide encoded by a nucleic acid molecule including a nucleotide sequence at least 60% identical to the nucleotide sequence set forth as SEQ ID NO:1 or SEQ ID NO:3. Also featured are fragments of the full-length polypeptides described herein (*e.g.*, fragments including at least 10 contiguous amino acid residues of the sequence set forth as SEQ ID NO:2) as well as allelic variants of the polypeptide having the amino acid sequence set forth as SEQ ID NO:2.

The TP-2 polypeptides and/or biologically active or antigenic fragments thereof, are useful, for example, as reagents or targets in assays applicable to treatment and/or diagnosis of TP-2 mediated or related disorders. In one embodiment, a TP-2 polypeptide or fragment thereof, has a TP-2 activity. In another embodiment, a TP-2 polypeptide or fragment thereof, includes at least one of the following domains: a transmembrane domain, a sugar transporter domain, a LacY proton/sugar symporter domain, and optionally, has a TP-2 activity. In a related aspect, the invention features antibodies (*e.g.*, antibodies which specifically bind to any one of the polypeptides described herein) as well as fusion polypeptides including all or a fragment of a polypeptide described herein.

The present invention further features methods for detecting TP-2 polypeptides and/or TP-2 nucleic acid molecules, such methods featuring, for example, a probe, primer or antibody described herein. Also featured are kits *e.g.*, kits for the detection of TP-2 polypeptides and/or TP-2 nucleic acid molecules. In a related aspect, the invention
5 features methods for identifying compounds which bind to and/or modulate the activity of a TP-2 polypeptide or TP-2 nucleic acid molecule described herein. Further featured are methods for modulating a TP-2 activity.

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

Brief Description of the Drawings

Figures 1A-B depicts the cDNA sequence and predicted amino acid sequence of human TP-2. The nucleotide sequence corresponds to nucleic acids 1 to 1963 of SEQ
15 ID NO:1. The amino acid sequence corresponds to amino acids 1 to 474 of SEQ ID NO: 2. The coding region without the 5' and 3' untranslated regions of the human TP-2 gene is shown in SEQ ID NO: 3.

Figure 2 depicts a structural, hydrophobicity, and antigenicity analysis of the human TP-2 polypeptide.

Figures 3A-C depicts the results of a search which was performed against the
20 HMM database in PFAM and which resulted in the identification of one "sugar transporter domain" and one "LacY proton/sugar symporter domain" in the human TP-2 polypeptide (SEQ ID NO:2).

Figure 4 depicts the results of a search which was performed against the
25 MEMSAT database and which resulted in the identification of twelve "transmembrane domains" in the human TP-2 polypeptide (SEQ ID NO:2).

Figure 5 depicts an alignment of the human TP-2 amino acid sequence (SEQ ID NO:2) with the amino acid sequences of the *Salmonella typhi* tetracycline-6-hydroxylase/oxygenase homolog gene (SEQ ID NO:4) using the CLUSTAL W™ (1.74)
30 alignment program.

Detailed Description of the Invention

The present invention is based, at least in part, on the discovery of novel molecules, referred to herein as "transporter-2" or "TP-2" nucleic acid and polypeptide
35 molecules, which are novel members of the transporter family. These novel molecules are capable of, for example, transporting ions, proteins, sugars, and small molecules across biological membranes both within a cell and between the cell and the environment and, thus, play a role in or function in a variety of cellular processes, *e.g.*,

proliferation, growth, differentiation, migration, immune responses, hormonal responses, and inter- or intra-cellular communication.

As used herein, the term "transporter" includes a molecule which is involved in the movement of a biochemical molecule from one side of a lipid bilayer to the other, for example, against a pre-existing concentration gradient. Transporters are usually involved in the movement of biochemical compounds which would normally not be able to cross a membrane (*e.g.*, a protein; an ion; a sugar; or other small molecule, such as ATP; signaling molecules; vitamins; and cofactors). Transporter molecules are involved in the growth, development, and differentiation of cells, in the regulation of cellular homeostasis, in the metabolism and catabolism of biochemical molecules necessary for energy production or storage, in intra- or inter-cellular signaling, in metabolism or catabolism of metabolically important biomolecules, and in the removal of potentially harmful compounds from the interior of the cell. Examples of transporters include GSH transporters, ATP transporters, sugar transporters, and fatty acid transporters. As transporters, the TP-2 molecules of the present invention provide novel diagnostic targets and therapeutic agents to control transporter-associated disorders.

As used herein, a "transporter-associated disorder" includes a disorder, disease or condition which is caused or characterized by a misregulation (*e.g.*, downregulation or upregulation) of a transporter-mediated activity. Transporter-associated disorders can detrimentally affect cellular functions such as cellular proliferation, growth, differentiation, or migration, cellular regulation of homeostasis, inter- or intra-cellular communication; tissue function, such as cardiac function or musculoskeletal function; systemic responses in an organism, such as nervous system responses, hormonal responses (*e.g.*, insulin response), or immune responses; and protection of cells from toxic compounds (*e.g.*, carcinogens, toxins, mutagens, and toxic byproducts of metabolic activity (*e.g.*, reactive oxygen species)). Examples of transporter-associated disorders include CNS disorders such as cognitive and neurodegenerative disorders, examples of which include, but are not limited to, Alzheimer's disease, dementias related to Alzheimer's disease (such as Pick's disease), Parkinson's and other Lewy diffuse body diseases, senile dementia, Huntington's disease, Gilles de la Tourette's syndrome, multiple sclerosis, amyotrophic lateral sclerosis, progressive supranuclear palsy, epilepsy, and Jakob-Creutzfeldt disease; autonomic function disorders such as hypertension and sleep disorders, and neuropsychiatric disorders, such as depression, schizophrenia, schizoaffective disorder, korsakoff's psychosis, mania, anxiety disorders, or phobic disorders; learning or memory disorders, *e.g.*, amnesia or age-related memory loss, attention deficit disorder, dysthymic disorder, major depressive disorder, mania, obsessive-compulsive disorder, psychoactive substance use disorders, anxiety, phobias, panic disorder, as well as bipolar affective disorder, *e.g.*, severe bipolar affective (mood)

disorder (BP-1), and bipolar affective neurological disorders, *e.g.*, migraine and obesity. Further CNS-related disorders include, for example, those listed in the American Psychiatric Association's Diagnostic and Statistical manual of Mental Disorders (DSM), the most current version of which is incorporated herein by reference in its entirety.

5 Further examples of transporter-associated disorders include cardiac-related disorders. Cardiovascular system disorders in which the TP-2 molecules of the invention may be directly or indirectly involved include arteriosclerosis, ischemia reperfusion injury, restenosis, arterial inflammation, vascular wall remodeling, ventricular remodeling, rapid ventricular pacing, coronary microembolism, tachycardia, 10 bradycardia, pressure overload, aortic bending, coronary artery ligation, vascular heart disease, atrial fibrillation, Jervell syndrome, Lange-Nielsen syndrome, long-QT syndrome, congestive heart failure, sinus node dysfunction, angina, heart failure, hypertension, atrial fibrillation, atrial flutter, dilated cardiomyopathy, idiopathic cardiomyopathy, myocardial infarction, coronary artery disease, coronary artery spasm, 15 and arrhythmia. TP-2-mediated or related disorders also include disorders of the musculoskeletal system such as paralysis and muscle weakness, *e.g.*, ataxia, myotonia, and myokymia.

Transporter-associated disorders also include cellular proliferation, growth, differentiation, or migration disorders. Cellular proliferation, growth, differentiation, or 20 migration disorders include those disorders that affect cell proliferation, growth, differentiation, or migration processes. As used herein, a "cellular proliferation, growth, differentiation, or migration process" is a process by which a cell increases in number, size or content, by which a cell develops a specialized set of characteristics which differ from that of other cells, or by which a cell moves closer to or further from a particular 25 location or stimulus. The TP-2 molecules of the present invention are involved in signal transduction mechanisms, which are known to be involved in cellular growth, differentiation, and migration processes. Thus, the TP-2 molecules may modulate cellular growth, differentiation, or migration, and may play a role in disorders characterized by aberrantly regulated growth, differentiation, or migration. Such 30 disorders include cancer, *e.g.*, carcinoma, sarcoma, or leukemia; tumor angiogenesis and metastasis; skeletal dysplasia; hepatic disorders; and hematopoietic and/or myeloproliferative disorders.

TP-2-associated disorders also include hormonal disorders, such as conditions or diseases in which the production and/or regulation of hormones in an organism is 35 aberrant. Examples of such disorders and diseases include type I and type II diabetes mellitus, pituitary disorders (*e.g.*, growth disorders), thyroid disorders (*e.g.*, hypothyroidism or hyperthyroidism), and reproductive or fertility disorders (*e.g.*, disorders which affect the organs of the reproductive system, *e.g.*, the prostate gland, the

uterus, or the vagina; disorders which involve an imbalance in the levels of a reproductive hormone in a subject; disorders affecting the ability of a subject to reproduce; and disorders affecting secondary sex characteristic development, *e.g.*, adrenal hyperplasia).

5 TP-2-associated disorders also include immune disorders, such as autoimmune disorders or immune deficiency disorders, *e.g.*, congenital X-linked infantile hypogammaglobulinemia, transient hypogammaglobulinemia, common variable immunodeficiency, selective IgA deficiency, chronic mucocutaneous candidiasis, or severe combined immunodeficiency.

10 TP-2-associated disorders also include disorders associated with sugar homeostasis, such as obesity, anorexia, hypoglycemia, glycogen storage disease (Von Gierke disease), type I glycogenosis, seasonal affective disorder, and cluster B personality disorders.

15 TP-2-associated disorders also include disorders affecting tissues in which TP-2 protein is expressed.

As used herein, a "transporter-mediated activity" includes an activity which involves the facilitated movement of one or more molecules from one side of a biological membrane to the other. Transporter-mediated activities include the import or export across internal or external cellular membranes of biochemical molecules
20 necessary for energy production or storage, intra- or inter-cellular signaling, metabolism or catabolism of metabolically important biomolecules, and removal of potentially harmful compounds from the cell.

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

For example, the family of TP-2 polypeptides comprise at least one "transmembrane domain" and preferably twelve transmembrane domains. As used herein, the term "transmembrane domain" includes an amino acid sequence of about 15-
35 45 amino acid residues in length which spans the plasma membrane. More preferably, a transmembrane domain includes about at least 15, 20, 25, 30, 35, 40, or 45 amino acid residues and spans the plasma membrane. Transmembrane domains are rich in hydrophobic residues, and typically have an alpha-helical structure. In a preferred

embodiment, at least 50%, 60%, 70%, 80%, 90%, 95% or more of the amino acids of a transmembrane domain are hydrophobic, *e.g.*, leucines, isoleucines, alanines, valines, phenylalanines, prolines or methionines. Transmembrane domains are described in, for example, Zagotta W.N. *et al.*, (1996) *Annual Rev. Neurosci.* 19: 235-263, the contents of
5 which are incorporated herein by reference. A MEMSAT analysis and a structural, hydrophobicity, and antigenicity analysis resulted in the identification of twelve transmembrane domains in the amino acid sequence of human TP-2 (SEQ ID NO:2) at about residues 45-69, 80-102, 112-126, 133-156, 167-190, 197-218, 288-310, 323-343, 352-368, 375-391, 409-433, and 442-458 as set forth in Figures 2 and 4.

10 Accordingly, TP-2 polypeptides having at least 50-60% homology, preferably about 60-70%, more preferably about 70-80%, or about 80-90% homology with a transmembrane domain of human TP-2 are within the scope of the invention.

For example, in one embodiment, members of the TP-2 family of proteins include at least one "sugar transporter domain" in the protein or corresponding nucleic
15 acid molecule. As used herein, the term "sugar transporter domain" includes a protein domain having at least about 350-500 amino acid residues and a sugar transporter mediated activity. Preferably, a sugar transporter domain includes a polypeptide having an amino acid sequence of about 350-450, 400-450, or more preferably about 417 amino acid residues, and a sugar transporter mediated activity. To identify the presence of a
20 sugar transporter domain in a TP-2 protein, and make the determination that a protein of interest has a particular profile, the amino acid sequence of the protein may be searched against a database of known protein domains (*e.g.*, the PFAM HMM database). A PFAM sugar transporter domain has been assigned the PFAM Accession PF00083. A search was performed against the PFAM HMM database resulting in the identification
25 of a sugar transporter domain in the amino acid sequence of human TP-2 (SEQ ID NO:2) at about residues 37-454 of SEQ ID NO:2. The results of the search are set forth in Figures 3A-C.

As used herein, a "sugar transporter mediated activity" includes the ability to bind a monosaccharide, such as D-glucose, D-fructose, and/or D-galactose; the ability to
30 transport a monosaccharide such as D-glucose, D-fructose, and/or D-galactose, across a cell membrane (*e.g.*, a liver cell membrane, fat cell membrane, muscle cell membrane, and/or blood cell membrane, such as an erythrocyte membrane); and the ability to modulate sugar homeostasis in a cell. Accordingly, identifying the presence of a "sugar transporter domain" can include isolating a fragment of a TP-2 molecule (*e.g.*, a TP-2
35 polypeptide) and assaying for the ability of the fragment to exhibit one of the aforementioned sugar transporter mediated activities.

In another embodiment, a TP-2 molecule of the present invention is identified based on the presence of at least one "LacY proton/sugar symporter domain." As used herein, the term "LacY proton/sugar symporter domain" includes a protein domain having at least about 350-500 amino acid residues and a LacY proton/sugar symporter mediated activity. Preferably, a LacY proton/sugar symporter domain includes a protein domain having an amino acid sequence of about 300-400, 300-350, or more preferably, about 344 amino acid residues and a LacY proton/sugar symporter mediated activity. To identify the presence of a LacY proton/sugar symporter domain in a TP-2 protein, and make the determination that a protein of interest has a particular profile, the amino acid sequence of the protein may be searched against a database of known protein domains (*e.g.*, the PFAM HMM database). A PFAM LacY proton/sugar symporter domain has been assigned the PFAM Accession PF01306. A search was performed against the PFAM HMM database resulting in the identification of a LacY proton/sugar symporter domain in the amino acid sequence of human TP-2 (SEQ ID NO:2) at about residues 39-383 of SEQ ID NO:2. The results of the search are set forth in Figures 3A-C.

As used herein, a "LacY proton/sugar symporter mediated activity" includes the ability to mediate the transport of a variety of sugars (*e.g.*, D-glucose, D-fructose, and/or D-galactose) with the concomitant transport of hydrogen ions across a biological membrane. Accordingly, identifying the presence of a "LacY proton/sugar symporter domain" can include isolating a fragment of a TP-2 molecule (*e.g.*, a TP-2 polypeptide) and assaying for the ability of the fragment to exhibit one of the aforementioned LacY proton/sugar symporter mediated activities.

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

In a preferred embodiment, the TP-2 molecules of the invention include at least one, preferably two, even more preferably twelve transmembrane domain(s), and/or at least one sugar transporter domain, and/or at least one LacY proton/sugar symporter domain.

Isolated polypeptides of the present invention, preferably TP-2 polypeptides, have an amino acid sequence sufficiently identical to the amino acid sequence of SEQ ID NO:2 or are encoded by a nucleotide sequence sufficiently identical to SEQ ID NO:1 or 3. As used herein, the term "sufficiently identical" refers to a first amino acid or nucleotide sequence which contains a sufficient or minimum number of identical or

equivalent (*e.g.*, an amino acid residue which has a similar side chain) amino acid residues or nucleotides to a second amino acid or nucleotide sequence such that the first and second amino acid or nucleotide sequences share common structural domains or motifs and/or a common functional activity. For example, amino acid or nucleotide sequences which share common structural domains having at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more homology or identity across the amino acid sequences of the domains and contain at least one and preferably two structural domains or motifs, are defined herein as sufficiently identical. Furthermore, amino acid or nucleotide sequences which share at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more homology or identity and share a common functional activity are defined herein as sufficiently identical.

In a preferred embodiment, a TP-2 polypeptide includes at least one or more of the following domains: a transmembrane domain, and/or a sugar transporter domain, and/or a LacY proton/sugar symporter domain, and has an amino acid sequence at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more homologous or identical to the amino acid sequence of SEQ ID NO:2, or the amino acid sequence encoded by the DNA insert of the plasmid deposited with ATCC as Accession Number _____. In yet another preferred embodiment, a TP-2 polypeptide includes at least one or more of the following domains: a transmembrane domain, and/or a sugar transporter domain, and/or a LacY proton/sugar symporter domain, and is encoded by a nucleic acid molecule having a nucleotide sequence which hybridizes under stringent hybridization conditions to a complement of a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1 or SEQ ID NO:3. In another preferred embodiment, a TP-2 polypeptide includes at least one or more of the following domains: a transmembrane domain, and/or a sugar transporter domain, and/or a LacY proton/sugar symporter domain, and has a TP-2 activity.

As used interchangeably herein, a "TP-2 activity", "biological activity of TP-2" or "functional activity of TP-2", refers to an activity exerted by a TP-2 protein, polypeptide or nucleic acid molecule on a TP-2 responsive cell or tissue, or on a TP-2 protein substrate, as determined *in vivo*, or *in vitro*, according to standard techniques. In one embodiment, a TP-2 activity is a direct activity, such as an association with a TP-2-target molecule. As used herein, a "target molecule" or "binding partner" is a molecule with which a TP-2 protein binds or interacts in nature, such that TP-2-mediated function is achieved. A TP-2 target molecule can be a non-TP-2 molecule or a TP-2 protein or polypeptide of the present invention (*e.g.*, a molecule to be transported, *e.g.*, a monosaccharide). In an exemplary embodiment, a TP-2 target molecule is a TP-2 ligand (*e.g.*, an energy molecule, a metabolite, a monosaccharide or an ion). Alternatively, a

TP-2 activity is an indirect activity, such as a cellular signaling activity mediated by interaction of the TP-2 protein with a TP-2 ligand. The biological activities of TP-2 are described herein. For example, the TP-2 proteins of the present invention can have one or more of the following activities: 1) modulate the import and export of molecules, *e.g.*,
5 hormones, ions, cytokines, neurotransmitters, monosaccharides, and metabolites, from cells, 2) modulate intra- or inter-cellular signaling, 3) modulate removal of potentially harmful compounds from the cell, or facilitate the compartmentalization of these molecules into a sequestered intra-cellular space (*e.g.*, the peroxisome), and 4) modulate transport of biological molecules across membranes, *e.g.*, the plasma membrane, or the
10 membrane of the mitochondrion, the peroxisome, the lysosome, the endoplasmic reticulum, the nucleus, or the vacuole.

The nucleotide sequence of the isolated human TP-2 cDNA and the predicted amino acid sequence of the human TP-2 polypeptide are shown in Figures 1A-B and in SEQ ID NOs:1 and 2, respectively. A plasmid containing the nucleotide sequence
15 encoding human TP-2 was deposited with the American Type Culture Collection (ATCC), 10801 University Boulevard, Manassas, VA 20110-2209, on _____ and assigned Accession Number _____. 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
20 for those of skill in the art and is not an admission that a deposit is required under 35 U.S.C. §112.

The human TP-2 gene, which is approximately 1963 nucleotides in length, encodes a polypeptide which is approximately 474 amino acid residues in length.

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

I. Isolated Nucleic Acid Molecules

One aspect of the invention pertains to isolated nucleic acid molecules that encode TP-2 polypeptides or biologically active portions thereof, as well as nucleic acid
30 fragments sufficient for use as hybridization probes to identify TP-2-encoding nucleic acid molecules (*e.g.*, TP-2 mRNA) and fragments for use as PCR primers for the amplification or mutation of TP-2 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
35 using nucleotide analogs. The nucleic acid molecule can be single-stranded or double-stranded, but preferably is double-stranded DNA.

The term "isolated nucleic acid molecule" includes nucleic acid molecules which are separated from other nucleic acid molecules which are present in the natural source of the nucleic acid. For example, with regards to genomic DNA, the term "isolated" includes nucleic acid molecules which are separated from the chromosome with which the genomic DNA is naturally associated. Preferably, an "isolated" nucleic acid is free of sequences which naturally flank the nucleic acid (*i.e.*, sequences located at the 5' and 3' ends of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. For example, in various embodiments, the isolated TP-2 nucleic acid molecule can contain less than about 5 kb, 4kb, 3kb, 2kb, 1 kb, 0.5 kb or 0.1 kb of nucleotide sequences which naturally flank the nucleic acid molecule in genomic DNA of the cell from which the nucleic acid is derived. Moreover, an "isolated" nucleic acid molecule, such as a cDNA molecule, can be substantially free of other cellular material, or culture medium when produced by recombinant techniques, or substantially free of chemical precursors or other chemicals when chemically synthesized.

A nucleic acid molecule of the present invention, *e.g.*, a nucleic acid molecule having the nucleotide sequence of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, or a portion thereof, can be isolated using standard molecular biology techniques and the sequence information provided herein. Using all or a portion of the nucleic acid sequence of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, as a hybridization probe, TP-2 nucleic acid molecules can be isolated using standard hybridization and cloning techniques (*e.g.*, as described in Sambrook, J., Fritsh, E. F., and Maniatis, T. *Molecular Cloning: A Laboratory Manual. 2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989*).

Moreover, a nucleic acid molecule encompassing all or a portion of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____ can be isolated by the polymerase chain reaction (PCR) using synthetic oligonucleotide primers designed based upon the sequence of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____.

A nucleic acid of the invention can be amplified using cDNA, mRNA or alternatively, genomic DNA, as a template and appropriate oligonucleotide primers according to standard PCR amplification techniques. The nucleic acid so amplified can be cloned into an appropriate vector and characterized by DNA sequence analysis. Furthermore, oligonucleotides corresponding to TP-2 nucleotide sequences can be prepared by standard synthetic techniques, *e.g.*, using an automated DNA synthesizer.

In one embodiment, an isolated nucleic acid molecule of the invention comprises the nucleotide sequence shown in SEQ ID NO:1. The sequence of SEQ ID NO:1 corresponds to the human TP-2 cDNA. This cDNA comprises sequences encoding the human TP-2 polypeptide (*i.e.*, "the coding region", from nucleotides 67-1491) as well as 5' untranslated sequences (nucleotides 1-66) and 3' untranslated sequences (nucleotides 1492-1963). Alternatively, the nucleic acid molecule can comprise only the coding region of SEQ ID NO:1 (*e.g.*, nucleotides 67-1491, corresponding to SEQ ID NO:3). Accordingly, in another embodiment, the isolated nucleic acid molecule comprises SEQ ID NO:3 and nucleotides 1-66 and 1492-1963 of SEQ ID NO:1. In yet another embodiment, the nucleic acid molecule consists of the nucleotide sequence set forth as SEQ ID NO:1 or SEQ ID NO:3.

In still another embodiment, an isolated nucleic acid molecule of the invention comprises a nucleic acid molecule which is a complement of the nucleotide sequence shown in SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, or a portion of any of these nucleotide sequences. A nucleic acid molecule which is complementary to the nucleotide sequence shown in SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, is one which is sufficiently complementary to the nucleotide sequence shown in SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, such that it can hybridize to the nucleotide sequence shown in SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, thereby forming a stable duplex.

In still another preferred embodiment, an isolated nucleic acid molecule of the present invention comprises a nucleotide sequence which is at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more identical to the nucleotide sequence shown in SEQ ID NO:1 or 3 (*e.g.*, to the entire length of the nucleotide sequence), or to the nucleotide sequence (*e.g.*, the entire length of the nucleotide sequence) of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, or a portion of any of these nucleotide sequences. In one embodiment, a nucleic acid molecule of the present invention comprises a nucleotide sequence which is at least (or no greater than) 50, 100, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 1950 or more nucleotides in length and hybridizes under stringent hybridization conditions to a complement of a nucleic acid molecule of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____.

Moreover, the nucleic acid molecule of the invention can comprise only a portion of the nucleic acid sequence of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, for example, a fragment which can be used as a probe or primer or a fragment encoding a portion of a TP-2 polypeptide, *e.g.*, a biologically active portion of a TP-2 polypeptide. The nucleotide sequence determined from the cloning of the TP-2 gene allows for the generation of probes and primers designed for use in identifying and/or cloning other TP-2 family members, as well as TP-2 homologues from other species. The probe/primer typically comprises substantially purified oligonucleotide. The probe/primer (*e.g.*, oligonucleotide) typically comprises a region of nucleotide sequence that hybridizes under stringent conditions to at least about 12 or 15, preferably about 20 or 25, more preferably about 30, 35, 40, 45, 50, 55, 60, 65, 75, 80, 85, 90, 95, or 100 or more consecutive nucleotides of a sense sequence of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, of an anti-sense sequence of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, or of a naturally occurring allelic variant or mutant of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____.

Exemplary probes or primers are at least 12, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75 or more nucleotides in length and/or comprise consecutive nucleotides of an isolated nucleic acid molecule described herein. Probes based on the TP-2 nucleotide sequences can be used to detect (*e.g.*, specifically detect) transcripts or genomic sequences encoding the same or homologous polypeptides. In preferred embodiments, the probe further comprises a label group attached thereto, *e.g.*, the label group can be a radioisotope, a fluorescent compound, an enzyme, or an enzyme co-factor. In another embodiment a set of primers is provided, *e.g.*, primers suitable for use in a PCR, which can be used to amplify a selected region of a TP-2 sequence, *e.g.*, a domain, region, site or other sequence described herein. The primers should be at least 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100 or more nucleotides in length. Such probes can be used as a part of a diagnostic test kit for identifying cells or tissue which misexpress a TP-2 polypeptide, such as by measuring a level of a TP-2-encoding nucleic acid in a sample of cells from a subject *e.g.*, detecting TP-2 mRNA levels or determining whether a genomic TP-2 gene has been mutated or deleted.

A nucleic acid fragment encoding a "biologically active portion of a TP-2 polypeptide" can be prepared by isolating a portion of the nucleotide sequence of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, which encodes a polypeptide having a TP-2

biological activity (the biological activities of the TP-2 polypeptides are described herein), expressing the encoded portion of the TP-2 polypeptide (*e.g.*, by recombinant expression *in vitro*) and assessing the activity of the encoded portion of the TP-2 polypeptide. In an exemplary embodiment, the nucleic acid molecule is at least 50, 100, 5
200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 1950 or more nucleotides in length and encodes a polypeptide having a TP-2 activity (as described herein).

The invention further encompasses nucleic acid molecules that differ from the nucleotide sequence shown in SEQ ID NO:1 or 3, or the nucleotide sequence of the
10 DNA insert of the plasmid deposited with ATCC as Accession Number _____. Such differences can be due to due to degeneracy of the genetic code, thus resulting in a nucleic acid which encodes the same TP-2 polypeptides as those encoded by the nucleotide sequence shown in SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____. In
15 another embodiment, an isolated nucleic acid molecule of the invention has a nucleotide sequence encoding a polypeptide having an amino acid sequence which differs by at least 1, but no greater than 5, 10, 20, 50 or 100 amino acid residues from the amino acid sequence shown in SEQ ID NO:2, or the amino acid sequence encoded by the DNA insert of the plasmid deposited with the ATCC as Accession Number _____. In yet
20 another embodiment, the nucleic acid molecule encodes the amino acid sequence of human TP-2. If an alignment is needed for this comparison, the sequences should be aligned for maximum homology.

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

Allelic variants result, for example, from DNA sequence polymorphisms within a population (*e.g.*, the human population) that lead to changes in the amino acid sequences of the TP-2 polypeptides. Such genetic polymorphism in the TP-2 genes may exist among individuals within a population due to natural allelic variation. As used
35 herein, the terms "gene" and "recombinant gene" refer to nucleic acid molecules which include an open reading frame encoding a TP-2 polypeptide, preferably a mammalian TP-2 polypeptide, and can further include non-coding regulatory sequences, and introns.

Accordingly, in one embodiment, the invention features isolated nucleic acid molecules which encode a naturally occurring allelic variant of a polypeptide comprising the amino acid sequence of SEQ ID NO:2, or an amino acid sequence encoded by the DNA insert of the plasmid deposited with ATCC as Accession Number _____, wherein the nucleic acid molecule hybridizes to a complement of a nucleic acid molecule comprising SEQ ID NO:1 or SEQ ID NO:3, for example, under stringent hybridization conditions.

Allelic variants of human TP-2 include both functional and non-functional TP-2 polypeptides. Functional allelic variants are naturally occurring amino acid sequence variants of the human TP-2 polypeptide that have a TP-2 activity, *e.g.*, maintain the ability to bind a TP-2 ligand or substrate and/or modulate the import and export of molecules from cells or across membranes, *e.g.*, monosaccharides. Functional allelic variants will typically contain only conservative substitution of one or more amino acids of SEQ ID NO:2, or substitution, deletion or insertion of non-critical residues in non-critical regions of the polypeptide.

Non-functional allelic variants are naturally occurring amino acid sequence variants of the human TP-2 polypeptide that do not have a TP-2 activity, *e.g.*, they do not have the ability to transport molecules into and out of cells or across membranes. Non-functional allelic variants will typically contain a non-conservative substitution, a deletion, or insertion or premature truncation of the amino acid sequence of SEQ ID NO:2, or a substitution, insertion or deletion in critical residues or critical regions.

The present invention further provides non-human orthologues of the human TP-2 polypeptide. Orthologues of human TP-2 polypeptides are polypeptides that are isolated from non-human organisms and possess the same TP-2 activity, *e.g.*, ligand binding and/or modulation of import and export of molecules from cells or across membranes, *e.g.*, monosaccharides, as the human TP-2 polypeptide. Orthologues of the human TP-2 polypeptide can readily be identified as comprising an amino acid sequence that is substantially identical to SEQ ID NO:2.

Moreover, nucleic acid molecules encoding other TP-2 family members and, thus, which have a nucleotide sequence which differs from the TP-2 sequences of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____ are intended to be within the scope of the invention. For example, another TP-2 cDNA can be identified based on the nucleotide sequence of human TP-2. Moreover, nucleic acid molecules encoding TP-2 polypeptides from different species, and which, thus, have a nucleotide sequence which differs from the TP-2 sequences of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____ are

intended to be within the scope of the invention. For example, a mouse TP-2 cDNA can be identified based on the nucleotide sequence of a human TP-2.

Nucleic acid molecules corresponding to natural allelic variants and homologues of the TP-2 cDNAs of the invention can be isolated based on their homology to the TP-2
5 nucleic acids disclosed herein using the cDNAs disclosed herein, or a portion thereof, as a hybridization probe according to standard hybridization techniques under stringent hybridization conditions. Nucleic acid molecules corresponding to natural allelic variants and homologues of the TP-2 cDNAs of the invention can further be isolated by mapping to the same chromosome or locus as the TP-2 gene.

10 Orthologues, homologues and allelic variants can be identified using methods known in the art (*e.g.*, by hybridization to an isolated nucleic acid molecule of the present invention, for example, under stringent hybridization conditions). In one embodiment, an isolated nucleic acid molecule of the invention is at least 15, 20, 25, 30 or more nucleotides in length and hybridizes under stringent conditions to the nucleic
15 acid molecule comprising the nucleotide sequence of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____. In other embodiment, the nucleic acid is at least 50, 100, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 1950 or more nucleotides in
20 length.

As used herein, the term "hybridizes under stringent conditions" is intended to describe conditions for hybridization and washing under which nucleotide sequences that are significantly identical or homologous to each other remain hybridized to each other. Preferably, the conditions are such that sequences at least about 70%, more
25 preferably at least about 80%, even more preferably at least about 85% or 90% identical to each other remain hybridized to each other. Such stringent conditions are known to those skilled in the art and can be found in *Current Protocols in Molecular Biology*, Ausubel *et al.*, eds., John Wiley & Sons, Inc. (1995), sections 2, 4 and 6. Additional stringent conditions can be found in *Molecular Cloning: A Laboratory Manual*,
30 Sambrook *et al.*, Cold Spring Harbor Press, Cold Spring Harbor, NY (1989), chapters 7, 9 and 11. A preferred, non-limiting example of stringent hybridization conditions includes hybridization in 4X sodium chloride/sodium citrate (SSC), at about 65-70°C (or hybridization in 4X SSC plus 50% formamide at about 42-50°C) followed by one or more washes in 1X SSC, at about 65-70°C. A preferred, non-limiting example of highly
35 stringent hybridization conditions includes hybridization in 1X SSC, at about 65-70°C (or hybridization in 1X SSC plus 50% formamide at about 42-50°C) followed by one or more washes in 0.3X SSC, at about 65-70°C. A preferred, non-limiting example of reduced stringency hybridization conditions includes hybridization in 4X SSC, at about

50-60°C (or alternatively hybridization in 6X SSC plus 50% formamide at about 40-45°C) followed by one or more washes in 2X SSC, at about 50-60°C. Ranges intermediate to the above-recited values, *e.g.*, at 65-70°C or at 42-50°C are also intended to be encompassed by the present invention. SSPE (1xSSPE is 0.15M NaCl, 10mM NaH₂PO₄, and 1.25mM EDTA, pH 7.4) can be substituted for SSC (1xSSC is 0.15M NaCl and 15mM sodium citrate) in the hybridization and wash buffers; washes are performed for 15 minutes each after hybridization is complete. The hybridization temperature for hybrids anticipated to be less than 50 base pairs in length should be 5-10°C less than the melting temperature (T_m) of the hybrid, where T_m is determined according to the following equations. For hybrids less than 18 base pairs in length, T_m(°C) = 2(# of A + T bases) + 4(# of G + C bases). For hybrids between 18 and 49 base pairs in length, T_m(°C) = 81.5 + 16.6(log₁₀[Na⁺]) + 0.41(%G+C) - (600/N), where N is the number of bases in the hybrid, and [Na⁺] is the concentration of sodium ions in the hybridization buffer ([Na⁺] for 1xSSC = 0.165 M). It will also be recognized by the skilled practitioner that additional reagents may be added to hybridization and/or wash buffers to decrease non-specific hybridization of nucleic acid molecules to membranes, for example, nitrocellulose or nylon membranes, including but not limited to blocking agents (*e.g.*, BSA or salmon or herring sperm carrier DNA), detergents (*e.g.*, SDS), chelating agents (*e.g.*, EDTA), Ficoll, PVP and the like. When using nylon membranes, in particular, an additional preferred, non-limiting example of stringent hybridization conditions is hybridization in 0.25-0.5M NaH₂PO₄, 7% SDS at about 65°C, followed by one or more washes at 0.02M NaH₂PO₄, 1% SDS at 65°C, see *e.g.*, Church and Gilbert (1984) *Proc. Natl. Acad. Sci. USA* 81:1991-1995, (or alternatively 0.2X SSC, 1% SDS).

Preferably, an isolated nucleic acid molecule of the invention that hybridizes under stringent conditions to the sequence of SEQ ID NO:1 or 3 and 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 polypeptide).

In addition to naturally-occurring allelic variants of the TP-2 sequences that may exist in the population, the skilled artisan will further appreciate that changes can be introduced by mutation into the nucleotide sequences of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, thereby leading to changes in the amino acid sequence of the encoded TP-2 polypeptides, without altering the functional ability of the TP-2 polypeptides. For example, nucleotide substitutions leading to amino acid substitutions at "non-essential" amino acid residues can be made in the sequence of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____. A "non-essential" amino acid residue is a residue that can be

altered from the wild-type sequence of TP-2 (*e.g.*, the sequence of SEQ ID NO:2) without altering the biological activity, whereas an "essential" amino acid residue is required for biological activity. For example, amino acid residues that are conserved among the TP-2 polypeptides of the present invention, *e.g.*, those present in a
5 transmembrane domain, and/or a sugar transporter domain, and/or a LacY proton/sugar symporter are predicted to be particularly unamenable to alteration. Furthermore, additional amino acid residues that are conserved between the TP-2 polypeptides of the present invention and other members of the TP-2 family are not likely to be amenable to alteration.

10 Accordingly, another aspect of the invention pertains to nucleic acid molecules encoding TP-2 polypeptides that contain changes in amino acid residues that are not essential for activity. Such TP-2 polypeptides differ in amino acid sequence from SEQ ID NO:2, yet retain biological activity. In one embodiment, the isolated nucleic acid molecule comprises a nucleotide sequence encoding a polypeptide, wherein the
15 polypeptide comprises an amino acid sequence at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more identical to SEQ ID NO:2 (*e.g.*, to the entire length of SEQ ID NO:2).

An isolated nucleic acid molecule encoding a TP-2 polypeptide identical to the polypeptide of SEQ ID NO:2, can be created by introducing one or more nucleotide
20 substitutions, additions or deletions into the nucleotide sequence of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, such that one or more amino acid substitutions, additions or deletions are introduced into the encoded polypeptide. Mutations can be introduced into SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid
25 deposited with ATCC as Accession Number _____ by standard techniques, such as site-directed mutagenesis and PCR-mediated mutagenesis. Preferably, conservative amino acid substitutions are made at one or more predicted non-essential amino acid residues. A "conservative amino acid substitution" is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid
30 residues having similar side chains have been defined in the art. These families include amino acids with basic side chains (*e.g.*, lysine, arginine, histidine), acidic side chains (*e.g.*, aspartic acid, glutamic acid), uncharged polar side chains (*e.g.*, glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine, tryptophan), nonpolar side chains (*e.g.*, alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine), beta-branched
35 side chains (*e.g.*, threonine, valine, isoleucine) and aromatic side chains (*e.g.*, tyrosine, phenylalanine, tryptophan, histidine). Thus, a predicted nonessential amino acid residue in a TP-2 polypeptide is preferably replaced with another amino acid residue from the same side chain family. Alternatively, in another embodiment, mutations can be

introduced randomly along all or part of a TP-2 coding sequence, such as by saturation mutagenesis, and the resultant mutants can be screened for TP-2 biological activity to identify mutants that retain activity. Following mutagenesis of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as

5 Accession Number _____, the encoded polypeptide can be expressed recombinantly and the activity of the polypeptide can be determined.

In a preferred embodiment, a mutant TP-2 polypeptide can be assayed for the ability to 1) modulate the import and export of molecules, *e.g.*, hormones, ions, cytokines, neurotransmitters, monosaccharides, and metabolites, from cells, 2) modulate
10 intra- or inter-cellular signaling, 3) modulate removal of potentially harmful compounds from the cell, or facilitate the compartmentalization of these molecules into a sequestered intra-cellular space (*e.g.*, the peroxisome), and 4) modulate transport of biological molecules across membranes, *e.g.*, the plasma membrane, or the membrane of the mitochondrion, the peroxisome, the lysosome, the endoplasmic reticulum, the
15 nucleus, or the vacuole.

In addition to the nucleic acid molecules encoding TP-2 polypeptides described above, another aspect of the invention pertains to isolated nucleic acid molecules which are antisense thereto. In an exemplary embodiment, the invention provides an isolated nucleic acid molecule which is antisense to a TP-2 nucleic acid molecule (*e.g.*, is
20 antisense to the coding strand of a TP-2 nucleic acid molecule). An "antisense" nucleic acid comprises a nucleotide sequence which is complementary to a "sense" nucleic acid encoding a polypeptide, *e.g.*, complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence. Accordingly, an antisense nucleic acid can hydrogen bond to a sense nucleic acid. The antisense nucleic acid can
25 be complementary to an entire TP-2 coding strand, or to only a portion thereof. In one embodiment, an antisense nucleic acid molecule is antisense to a "coding region" of the coding strand of a nucleotide sequence encoding TP-2. The term "coding region" refers to the region of the nucleotide sequence comprising codons which are translated into amino acid residues (*e.g.*, the coding region of human TP-2 corresponds to SEQ ID
30 NO:3). In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence encoding TP-2. The term "noncoding region" refers to 5' and 3' sequences which flank the coding region that are not translated into amino acids (*i.e.*, also referred to as 5' and 3' untranslated regions).

35 Given the coding strand sequences encoding TP-2 disclosed herein (*e.g.*, SEQ ID NO:3), antisense nucleic acids of the invention can be designed according to the rules of Watson and Crick base pairing. The antisense nucleic acid molecule can be complementary to the entire coding region of TP-2 mRNA, but more preferably is an

oligonucleotide which is antisense to only a portion of the coding or noncoding region of TP-2 mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of TP-2 mRNA (*e.g.*, between the -10 and +10 regions of the start site of a gene nucleotide sequence). An antisense oligonucleotide
5 can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An antisense nucleic acid of the invention can be constructed using chemical synthesis and enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (*e.g.*, an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to
10 increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, *e.g.*, phosphorothioate derivatives and acridine substituted nucleotides can be used. Examples of modified nucleotides which can be used to generate the antisense nucleic acid include 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xantine, 4-acetylcytosine, 5-(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'-
20 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).

30 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 TP-2 polypeptide to thereby inhibit expression of the polypeptide, *e.g.*, by inhibiting transcription and/or translation. The hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for example,
35 in the case of an antisense nucleic acid molecule which binds to DNA duplexes, through specific interactions in the major groove of the double helix. An example of a route of administration of antisense nucleic acid molecules of the invention include direct injection at a tissue site. Alternatively, antisense nucleic acid molecules can be modified

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

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

In still another embodiment, an antisense nucleic acid of the invention is a ribozyme. Ribozymes are catalytic RNA molecules with ribonuclease activity which are capable of cleaving a single-stranded nucleic acid, such as an mRNA, to which they have a complementary region. Thus, ribozymes (*e.g.*, hammerhead ribozymes (described in Haselhoff and Gerlach (1988) *Nature* 334:585-591)) can be used to catalytically cleave TP-2 mRNA transcripts to thereby inhibit translation of TP-2 mRNA. A ribozyme having specificity for a TP-2-encoding nucleic acid can be designed based upon the nucleotide sequence of a TP-2 cDNA disclosed herein (*i.e.*, SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____). For example, a derivative of a *Tetrahymena* L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in a TP-2-encoding mRNA. See, *e.g.*, Cech *et al.* U.S. Patent No. 4,987,071; and Cech *et al.* U.S. Patent No. 5,116,742. Alternatively, TP-2 mRNA can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. See, *e.g.*, Bartel, D. and Szostak, J.W. (1993) *Science* 261:1411-1418.

Alternatively, TP-2 gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region of the TP-2 (*e.g.*, the TP-2 promoter and/or enhancers) to form triple helical structures that prevent transcription of the TP-2 gene in target cells. See generally, Helene, C. (1991) *Anticancer Drug Des.* 6(6):569-84; Helene, C. *et al.* (1992) *Ann. N.Y. Acad. Sci.* 660:27-36; and Maher, L.J. (1992) *Bioassays* 14(12):807-15.

In yet another embodiment, the TP-2 nucleic acid molecules of the present invention can be modified at the base moiety, sugar moiety or phosphate backbone to improve, *e.g.*, the stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate backbone of the nucleic acid molecules can be modified to generate peptide nucleic acids (see Hyrup B. *et al.* (1996) *Bioorganic & Medicinal Chemistry* 4 (1): 5-23). As used herein, the terms "peptide nucleic acids" or "PNAs" refer to nucleic acid mimics, *e.g.*, DNA mimics, in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of PNAs has been shown to allow for specific hybridization to DNA and RNA under conditions of low ionic strength. The synthesis of PNA oligomers can be performed using standard solid phase peptide synthesis protocols as described in Hyrup B. *et al.* (1996) *supra*; Perry-O'Keefe *et al. Proc. Natl. Acad. Sci.* 93: 14670-675.

PNAs of TP-2 nucleic acid molecules can be used in therapeutic and diagnostic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, for example, inducing transcription or translation arrest or inhibiting replication. PNAs of TP-2 nucleic acid molecules can also be used in the analysis of single base pair mutations in a gene, (*e.g.*, by PNA-directed PCR clamping); as 'artificial restriction enzymes' when used in combination with other enzymes, (*e.g.*, S1 nucleases (Hyrup B. (1996) *supra*)); or as probes or primers for DNA sequencing or hybridization (Hyrup B. *et al.* (1996) *supra*; Perry-O'Keefe *supra*).

In another embodiment, PNAs of TP-2 can be modified, (*e.g.*, to enhance their stability or cellular uptake), by attaching lipophilic or other helper groups to PNA, by the formation of PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. For example, PNA-DNA chimeras of TP-2 nucleic acid molecules can be generated which may combine the advantageous properties of PNA and DNA. Such chimeras allow DNA recognition enzymes, (*e.g.*, RNase H and DNA polymerases), to interact with the DNA portion while the PNA portion would provide high binding affinity and specificity. PNA-DNA chimeras can be linked using linkers of appropriate lengths selected in terms of base stacking, number of bonds between the nucleobases, and orientation (Hyrup B. (1996) *supra*). The synthesis of PNA-DNA chimeras can be performed as described in Hyrup B. (1996) *supra* and Finn P.J. *et al.* (1996) *Nucleic Acids Res.* 24 (17): 3357-63. For example, a DNA chain can be synthesized on a solid support using standard phosphoramidite coupling chemistry and modified nucleoside analogs, *e.g.*, 5'-(4-methoxytrityl)amino-5'-deoxy-thymidine phosphoramidite, can be used as a between the PNA and the 5' end of DNA (Mag, M. *et al.* (1989) *Nucleic Acid Res.* 17: 5973-88). PNA monomers are then coupled in a

stepwise manner to produce a chimeric molecule with a 5' PNA segment and a 3' DNA segment (Finn P.J. *et al.* (1996) *supra*). Alternatively, chimeric molecules can be synthesized with a 5' DNA segment and a 3' PNA segment (Peterser, K.H. *et al.* (1975) *Bioorganic Med. Chem. Lett.* 5: 1119-11124).

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

15 Alternatively, the expression characteristics of an endogenous TP-2 gene within a cell line or microorganism may be modified by inserting a heterologous DNA regulatory element into the genome of a stable cell line or cloned microorganism such that the inserted regulatory element is operatively linked with the endogenous TP-2 gene. For example, an endogenous TP-2 gene which is normally "transcriptionally
20 silent", *i.e.*, a TP-2 gene which is normally not expressed, or is expressed only at very low levels in a cell line or microorganism, may be activated by inserting a regulatory element which is capable of promoting the expression of a normally expressed gene product in that cell line or microorganism. Alternatively, a transcriptionally silent, endogenous TP-2 gene may be activated by insertion of a promiscuous regulatory
25 element that works across cell types.

A heterologous regulatory element may be inserted into a stable cell line or cloned microorganism, such that it is operatively linked with an endogenous TP-2 gene, using techniques, such as targeted homologous recombination, which are well known to those of skill in the art, and described, *e.g.*, in Chappel, U.S. Patent No. 5,272,071; PCT
30 publication No. WO 91/06667, published May 16, 1991.

II. Isolated TP-2 Polypeptides and Anti-TP-2 Antibodies

One aspect of the invention pertains to isolated TP-2 or recombinant polypeptides and polypeptides, and biologically active portions thereof, as well as
35 polypeptide fragments suitable for use as immunogens to raise anti-TP-2 antibodies. In one embodiment, native TP-2 polypeptides can be isolated from cells or tissue sources by an appropriate purification scheme using standard protein purification techniques. In another embodiment, TP-2 polypeptides are produced by recombinant DNA techniques.

Alternative to recombinant expression, a TP-2 polypeptide or polypeptide can be synthesized chemically using standard peptide synthesis techniques.

An "isolated" or "purified" polypeptide or biologically active portion thereof is substantially free of cellular material or other contaminating proteins from the cell or tissue source from which the TP-2 polypeptide is derived, or substantially free from chemical precursors or other chemicals when chemically synthesized. The language "substantially free of cellular material" includes preparations of TP-2 polypeptide in which the polypeptide is separated from cellular components of the cells from which it is isolated or recombinantly produced. In one embodiment, the language "substantially free of cellular material" includes preparations of TP-2 polypeptide having less than about 30% (by dry weight) of non-TP-2 polypeptide (also referred to herein as a "contaminating protein"), more preferably less than about 20% of non-TP-2 polypeptide, still more preferably less than about 10% of non-TP-2 polypeptide, and most preferably less than about 5% non-TP-2 polypeptide. When the TP-2 polypeptide or biologically active portion thereof is recombinantly produced, it is also preferably substantially free of culture medium, *i.e.*, culture medium represents less than about 20%, more preferably less than about 10%, and most preferably less than about 5% of the volume of the protein preparation.

The language "substantially free of chemical precursors or other chemicals" includes preparations of TP-2 polypeptide in which the polypeptide is separated from chemical precursors or other chemicals which are involved in the synthesis of the polypeptide. In one embodiment, the language "substantially free of chemical precursors or other chemicals" includes preparations of TP-2 polypeptide having less than about 30% (by dry weight) of chemical precursors or non-TP-2 chemicals, more preferably less than about 20% chemical precursors or non-TP-2 chemicals, still more preferably less than about 10% chemical precursors or non-TP-2 chemicals, and most preferably less than about 5% chemical precursors or non-TP-2 chemicals.

As used herein, a "biologically active portion" of a TP-2 polypeptide includes a fragment of a TP-2 polypeptide which participates in an interaction between a TP-2 molecule and a non-TP-2 molecule. Biologically active portions of a TP-2 polypeptide include peptides comprising amino acid sequences sufficiently identical to or derived from the amino acid sequence of the TP-2 polypeptide, *e.g.*, the amino acid sequence shown in SEQ ID NO:2, which include less amino acids than the full length TP-2 polypeptides, and exhibit at least one activity of a TP-2 polypeptide. Typically, biologically active portions comprise a domain or motif with at least one activity of the TP-2 polypeptide, *e.g.*, modulating transport mechanisms. A biologically active portion of a TP-2 polypeptide can be a polypeptide which is, for example, 25, 30, 35, 40, 45, 50, 75, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475 or

more amino acids in length. Biologically active portions of a TP-2 polypeptide can be used as targets for developing agents which modulate a TP-2 mediated activity, *e.g.*, modulating transport of biological molecules across membranes.

In one embodiment, a biologically active portion of a TP-2 polypeptide
5 comprises at least one transmembrane domain. It is to be understood that a preferred biologically active portion of a TP-2 polypeptide of the present invention comprises at least one or more of the following domains: a transmembrane domain, and/or a sugar transporter domain, and/or a LacY proton/sugar symporter domain. Moreover, other biologically active portions, in which other regions of the polypeptide are deleted, can
10 be prepared by recombinant techniques and evaluated for one or more of the functional activities of a native TP-2 polypeptide.

Another aspect of the invention features fragments of the polypeptide having the amino acid sequence of SEQ ID NO:2, for example, for use as immunogens. In one embodiment, a fragment comprises at least 5 amino acids (*e.g.*, contiguous or
15 consecutive amino acids) of the amino acid sequence of SEQ ID NO:2, or an amino acid sequence encoded by the DNA insert of the plasmid deposited with the ATCC as Accession Number _____. In another embodiment, a fragment comprises at least 10, 15, 20, 25, 30, 35, 40, 45, 50 or more amino acids (*e.g.*, contiguous or consecutive amino acids) of the amino acid sequence of SEQ ID NO:2, or an amino acid sequence
20 encoded by the DNA insert of the plasmid deposited with the ATCC as Accession Number _____.

In a preferred embodiment, a TP-2 polypeptide has an amino acid sequence shown in SEQ ID NO:2. In other embodiments, the TP-2 polypeptide is substantially identical to SEQ ID NO:2, and retains the functional activity of the polypeptide of SEQ
25 ID NO:2, yet differs in amino acid sequence due to natural allelic variation or mutagenesis, as described in detail in subsection I above. In another embodiment, the TP-2 polypeptide is a polypeptide which comprises an amino acid sequence at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more identical to SEQ ID NO:2.

In another embodiment, the invention features a TP-2 polypeptide which is encoded by a nucleic acid molecule consisting of a nucleotide sequence at least about
30 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more identical to a nucleotide sequence of SEQ ID NO:1 or SEQ ID NO:3, or a complement thereof. This invention further features a TP-2 polypeptide which is encoded by a
35 nucleic acid molecule consisting of a nucleotide sequence which hybridizes under stringent hybridization conditions to a complement of a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1 or SEQ ID NO:3, or a complement thereof.

To determine the percent identity of two amino acid sequences or of two nucleic acid sequences, the sequences are aligned for optimal comparison purposes (*e.g.*, gaps can be introduced in one or both of a first and a second amino acid or nucleic acid sequence for optimal alignment and non-identical sequences can be disregarded for comparison purposes). In a preferred embodiment, the length of a reference sequence aligned for comparison purposes is at least 30%, preferably at least 40%, more preferably at least 50%, even more preferably at least 60%, and even more preferably at least 70%, 80%, or 90% of the length of the reference sequence (*e.g.*, when aligning a second sequence to the TP-2 amino acid sequence of SEQ ID NO:2 having 474 amino acid residues, at least 142, preferably at least 189, more preferably at least 237, more preferably at least 284, even more preferably at least 331, and even more preferably at least 379 or 426 or more amino acid residues are aligned). The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are identical at that position (as used herein amino acid or nucleic acid "identity" is equivalent to amino acid or nucleic acid "homology"). The percent identity between the two sequences is a function of the number of identical positions shared by the sequences, taking into account the number of gaps, and the length of each gap, which need to be introduced for optimal alignment of the two sequences.

The comparison of sequences and determination of percent identity between two sequences can be accomplished using a mathematical algorithm. In a preferred embodiment, the percent identity between two amino acid sequences is determined using the Needleman and Wunsch (*J. Mol. Biol.* (48):444-453 (1970)) algorithm which has been incorporated into the GAP program in the GCG software package (available at <http://www.gcg.com>), using either a Blosum 62 matrix or a PAM250 matrix, and a gap weight of 16, 14, 12, 10, 8, 6, or 4 and a length weight of 1, 2, 3, 4, 5, or 6. In yet another preferred embodiment, the percent identity between two nucleotide sequences is determined using the GAP program in the GCG software package (available at <http://www.gcg.com>), using a NWSgapdna.CMP matrix and a gap weight of 40, 50, 60, 70, or 80 and a length weight of 1, 2, 3, 4, 5, or 6. A preferred, non-limiting example of parameters to be used in conjunction with the GAP program include a Blosum 62 scoring matrix with a gap penalty of 12, a gap extend penalty of 4, and a frameshift gap penalty of 5.

In another embodiment, the percent identity between two amino acid or nucleotide sequences is determined using the algorithm of E. Meyers and W. Miller (*Comput. Appl. Biosci.*, 4:11-17 (1988)) which has been incorporated into the ALIGN

program (version 2.0 or version 2.0U), using a PAM120 weight residue table, a gap length penalty of 12 and a gap penalty of 4.

The nucleic acid and polypeptide sequences of the present invention can further be used as a "query sequence" to perform a search against public databases to, for example, identify other family members or related sequences. Such searches can be performed using the NBLAST and XBLAST programs (version 2.0) of Altschul, *et al.* (1990) *J. Mol. Biol.* 215:403-10. BLAST nucleotide searches can be performed with the NBLAST program, score = 100, wordlength = 12 to obtain nucleotide sequences homologous to TP-2 nucleic acid molecules of the invention. BLAST protein searches can be performed with the XBLAST program, score = 100, wordlength = 3, and a Blosum62 matrix to obtain amino acid sequences homologous to TP-2 polypeptide 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(17):3389-3402. When utilizing BLAST and Gapped BLAST programs, the default parameters of the respective programs (*e.g.*, XBLAST and NBLAST) can be used. See <http://www.ncbi.nlm.nih.gov>.

The invention also provides TP-2 chimeric or fusion proteins. As used herein, a TP-2 "chimeric protein" or "fusion protein" comprises a TP-2 polypeptide operatively linked to a non-TP-2 polypeptide. An "TP-2 polypeptide" refers to a polypeptide having an amino acid sequence corresponding to TP-2, whereas a "non-TP-2 polypeptide" refers to a polypeptide having an amino acid sequence corresponding to a polypeptide which is not substantially homologous to the TP-2 polypeptide, *e.g.*, a polypeptide which is different from the TP-2 polypeptide and which is derived from the same or a different organism. Within a TP-2 fusion protein the TP-2 polypeptide can correspond to all or a portion of a TP-2 polypeptide. In a preferred embodiment, a TP-2 fusion protein comprises at least one biologically active portion of a TP-2 polypeptide. In another preferred embodiment, a TP-2 fusion protein comprises at least two biologically active portions of a TP-2 polypeptide. Within the fusion protein, the term "operatively linked" is intended to indicate that the TP-2 polypeptide and the non-TP-2 polypeptide are fused in-frame to each other. The non-TP-2 polypeptide can be fused to the N-terminus or C-terminus of the TP-2 polypeptide.

For example, in one embodiment, the fusion protein is a GST-TP-2 fusion protein in which the TP-2 sequences are fused to the C-terminus of the GST sequences. Such fusion proteins can facilitate the purification of recombinant TP-2.

In another embodiment, the fusion protein is a TP-2 polypeptide containing a heterologous signal sequence at its N-terminus. In certain host cells (*e.g.*, mammalian host cells), expression and/or secretion of TP-2 can be increased through the use of a heterologous signal sequence.

The TP-2 fusion proteins of the invention can be incorporated into pharmaceutical compositions and administered to a subject *in vivo*. The TP-2 fusion proteins can be used to affect the bioavailability of a TP-2 substrate. Use of TP-2 fusion proteins may be useful therapeutically for the treatment of disorders caused by, for example, (i) aberrant modification or mutation of a gene encoding a TP-2 polypeptide; (ii) mis-regulation of the TP-2 gene; and (iii) aberrant post-translational modification of a TP-2 polypeptide.

Moreover, the TP-2-fusion proteins of the invention can be used as immunogens to produce anti-TP-2 antibodies in a subject, to purify TP-2 ligands and in screening assays to identify molecules which inhibit the interaction of TP-2 with a TP-2 substrate.

Preferably, a TP-2 chimeric or fusion protein of the invention is produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional techniques, for example by employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers which give rise to complementary overhangs between two consecutive gene fragments which can subsequently be annealed and reamplified to generate a chimeric gene sequence (see, for example, *Current Protocols in Molecular Biology*, eds. Ausubel *et al.* John Wiley & Sons: 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (*e.g.*, a GST polypeptide). A TP-2-encoding nucleic acid can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the TP-2 polypeptide.

The present invention also pertains to variants of the TP-2 polypeptides which function as either TP-2 agonists (mimetics) or as TP-2 antagonists. Variants of the TP-2 polypeptides can be generated by mutagenesis, *e.g.*, discrete point mutation or truncation of a TP-2 polypeptide. An agonist of the TP-2 polypeptides can retain substantially the same, or a subset, of the biological activities of the naturally occurring form of a TP-2 polypeptide. An antagonist of a TP-2 polypeptide can inhibit one or more of the activities of the naturally occurring form of the TP-2 polypeptide by, for example, competitively modulating a TP-2-mediated activity of a TP-2 polypeptide. Thus, specific biological effects can be elicited by treatment with a variant of limited function. In one embodiment, treatment of a subject with a variant having a subset of the biological activities of the naturally occurring form of the polypeptide has fewer side

effects in a subject relative to treatment with the naturally occurring form of the TP-2 polypeptide.

In one embodiment, variants of a TP-2 polypeptide which function as either TP-2 agonists (mimetics) or as TP-2 antagonists can be identified by screening combinatorial libraries of mutants, *e.g.*, truncation mutants, of a TP-2 polypeptide for TP-2 polypeptide agonist or antagonist activity. In one embodiment, a variegated library of TP-2 variants is generated by combinatorial mutagenesis at the nucleic acid level and is encoded by a variegated gene library. A variegated library of TP-2 variants can be produced by, for example, enzymatically ligating a mixture of synthetic oligonucleotides into gene sequences such that a degenerate set of potential TP-2 sequences is expressible as individual polypeptides, or alternatively, as a set of larger fusion proteins (*e.g.*, for phage display) containing the set of TP-2 sequences therein. There are a variety of methods which can be used to produce libraries of potential TP-2 variants from a degenerate oligonucleotide sequence. Chemical synthesis of a degenerate gene sequence can be performed in an automatic DNA synthesizer, and the synthetic gene then ligated into an appropriate expression vector. Use of a degenerate set of genes allows for the provision, in one mixture, of all of the sequences encoding the desired set of potential TP-2 sequences. Methods for synthesizing degenerate oligonucleotides are known in the art (see, *e.g.*, Narang, S.A. (1983) *Tetrahedron* 39:3; Itakura *et al.* (1984) *Annu. Rev. Biochem.* 53:323; Itakura *et al.* (1984) *Science* 198:1056; Ike *et al.* (1983) *Nucleic Acid Res.* 11:477).

In addition, libraries of fragments of a TP-2 polypeptide coding sequence can be used to generate a variegated population of TP-2 fragments for screening and subsequent selection of variants of a TP-2 polypeptide. In one embodiment, a library of coding sequence fragments can be generated by treating a double stranded PCR fragment of a TP-2 coding sequence with a nuclease under conditions wherein nicking occurs only about once per molecule, denaturing the double stranded DNA, renaturing the DNA to form double stranded DNA which can include sense/antisense pairs from different nicked products, removing single stranded portions from reformed duplexes by treatment with S1 nuclease, and ligating the resulting fragment library into an expression vector. By this method, an expression library can be derived which encodes N-terminal, C-terminal and internal fragments of various sizes of the TP-2 polypeptide.

Several techniques are known in the art for screening gene products of combinatorial libraries made by point mutations or truncation, and for screening cDNA libraries for gene products having a selected property. Such techniques are adaptable for rapid screening of the gene libraries generated by the combinatorial mutagenesis of TP-2 polypeptides. The most widely used techniques, which are amenable to high throughput analysis, for screening large gene libraries typically include cloning the gene library

into replicable expression vectors, transforming appropriate cells with the resulting library of vectors, and expressing the combinatorial genes under conditions in which detection of a desired activity facilitates isolation of the vector encoding the gene whose product was detected. Recursive ensemble mutagenesis (REM), a new technique which
5 enhances the frequency of functional mutants in the libraries, can be used in combination with the screening assays to identify TP-2 variants (Arkin and Yourvan (1992) *Proc. Natl. Acad. Sci. USA* 89:7811-7815; Delgrave *et al.* (1993) *Protein Engineering* 6(3):327-331).

In one embodiment, cell based assays can be exploited to analyze a variegated
10 TP-2 library. For example, a library of expression vectors can be transfected into a cell line, *e.g.*, an endothelial cell line, which ordinarily responds to TP-2 in a particular TP-2 substrate-dependent manner. The transfected cells are then contacted with TP-2 and the effect of expression of the mutant on signaling by the TP-2 substrate can be detected, *e.g.*, by monitoring intra-cellular calcium, IP3, or diacylglycerol concentration,
15 phosphorylation profile of intra-cellular proteins, or the activity of a TP-2-regulated transcription factor. Plasmid DNA can then be recovered from the cells which score for inhibition, or alternatively, potentiation of signaling by the TP-2 substrate, and the individual clones further characterized.

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

Preferred epitopes encompassed by the antigenic peptide are regions of TP-2 that
30 are located on the surface of the polypeptide, *e.g.*, hydrophilic regions, as well as regions with high antigenicity (see, for example, Figure 2).

A TP-2 immunogen typically is used to prepare antibodies by immunizing a suitable subject, (*e.g.*, rabbit, goat, mouse or other mammal) with the immunogen. An appropriate immunogenic preparation can contain, for example, recombinantly
35 expressed TP-2 polypeptide or a chemically synthesized TP-2 polypeptide. The preparation can further include an adjuvant, such as Freund's complete or incomplete adjuvant, or similar immunostimulatory agent. Immunization of a suitable subject with an immunogenic TP-2 preparation induces a polyclonal anti-TP-2 antibody response.

Accordingly, another aspect of the invention pertains to anti-TP-2 antibodies. The term "antibody" as used herein refers to immunoglobulin molecules and immunologically active portions of immunoglobulin molecules, *i.e.*, molecules that contain an antigen binding site which specifically binds (immunoreacts with) an antigen, such as TP-2. Examples of immunologically active portions of immunoglobulin molecules include F(ab) and F(ab')₂ fragments which can be generated by treating the antibody with an enzyme such as pepsin. The invention provides polyclonal and monoclonal antibodies that bind TP-2. The term "monoclonal antibody" or "monoclonal antibody composition", as used herein, refers to a population of antibody molecules that contain only one species of an antigen binding site capable of immunoreacting with a particular epitope of TP-2. A monoclonal antibody composition thus typically displays a single binding affinity for a particular TP-2 polypeptide with which it immunoreacts.

Polyclonal anti-TP-2 antibodies can be prepared as described above by immunizing a suitable subject with a TP-2 immunogen. The anti-TP-2 antibody titer in the immunized subject can be monitored over time by standard techniques, such as with an enzyme linked immunosorbent assay (ELISA) using immobilized TP-2. If desired, the antibody molecules directed against TP-2 can be isolated from the mammal (*e.g.*, from the blood) and further purified by well known techniques, such as protein A chromatography to obtain the IgG fraction. At an appropriate time after immunization, *e.g.*, when the anti-TP-2 antibody titers are highest, antibody-producing cells can be obtained from the subject and used to prepare monoclonal antibodies by standard techniques, such as the hybridoma technique originally described by Kohler and Milstein (1975) *Nature* 256:495-497) (see also, Brown *et al.* (1981) *J. Immunol.* 127:539-46; Brown *et al.* (1980) *J. Biol. Chem.* 255:4980-83; Yeh *et al.* (1976) *Proc. Natl. Acad. Sci. USA* 76:2927-31; and Yeh *et al.* (1982) *Int. J. Cancer* 29:269-75), the more recent human B cell hybridoma technique (Kozbor *et al.* (1983) *Immunol Today* 4:72), the EBV-hybridoma technique (Cole *et al.* (1985), *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc., pp. 77-96) or trioma techniques. The technology for producing monoclonal antibody hybridomas is well known (see generally R. H. Kenneth, in *Monoclonal Antibodies: A New Dimension In Biological Analyses*, Plenum Publishing Corp., New York, New York (1980); E. A. Lerner (1981) *Yale J. Biol. Med.*, 54:387-402; M. L. Gefter *et al.* (1977) *Somatic Cell Genet.* 3:231-36). Briefly, an immortal cell line (typically a myeloma) is fused to lymphocytes (typically splenocytes) from a mammal immunized with a TP-2 immunogen as described above, and the culture supernatants of the resulting hybridoma cells are screened to identify a hybridoma producing a monoclonal antibody that binds TP-2.

Any of the many well known protocols used for fusing lymphocytes and immortalized cell lines can be applied for the purpose of generating an anti-TP-2 monoclonal antibody (see, *e.g.*, G. Galfré *et al.* (1977) *Nature* 266:55052; Geffer *et al. Somatic Cell Genet.*, cited *supra*; Lerner, *Yale J. Biol. Med.*, cited *supra*; Kenneth, *Monoclonal Antibodies*, cited *supra*). Moreover, the ordinarily skilled worker will appreciate that there are many variations of such methods which also would be useful. Typically, the immortal cell line (*e.g.*, a myeloma cell line) is derived from the same mammalian species as the lymphocytes. For example, murine hybridomas can be made by fusing lymphocytes from a mouse immunized with an immunogenic preparation of the present invention with an immortalized mouse cell line. Preferred immortal cell lines are mouse myeloma cell lines that are sensitive to culture medium containing hypoxanthine, aminopterin and thymidine ("HAT medium"). Any of a number of myeloma cell lines can be used as a fusion partner according to standard techniques, *e.g.*, the P3-NS1/1-Ag4-1, P3-x63-Ag8.653 or Sp2/O-Ag14 myeloma lines. These myeloma lines are available from ATCC. Typically, HAT-sensitive mouse myeloma cells are fused to mouse splenocytes using polyethylene glycol ("PEG"). Hybridoma cells resulting from the fusion are then selected using HAT medium, which kills unfused and unproductively fused myeloma cells (unfused splenocytes die after several days because they are not transformed). Hybridoma cells producing a monoclonal antibody of the invention are detected by screening the hybridoma culture supernatants for antibodies that bind TP-2, *e.g.*, using a standard ELISA assay.

Alternative to preparing monoclonal antibody-secreting hybridomas, a monoclonal anti-TP-2 antibody can be identified and isolated by screening a recombinant combinatorial immunoglobulin library (*e.g.*, an antibody phage display library) with TP-2 to thereby isolate immunoglobulin library members that bind TP-2. Kits for generating and screening phage display libraries are commercially available (*e.g.*, the Pharmacia *Recombinant Phage Antibody System*, Catalog No. 27-9400-01; and the Stratagene *SurfZAP™ Phage Display Kit*, Catalog No. 240612). Additionally, examples of methods and reagents particularly amenable for use in generating and screening antibody display library can be found in, for example, Ladner *et al.* U.S. Patent No. 5,223,409; Kang *et al.* PCT International Publication No. WO 92/18619; Dower *et al.* PCT International Publication No. WO 91/17271; Winter *et al.* PCT International Publication WO 92/20791; Markland *et al.* PCT International Publication No. WO 92/15679; Breitling *et al.* PCT International Publication WO 93/01288; McCafferty *et al.* PCT International Publication No. WO 92/01047; Garrard *et al.* PCT International Publication No. WO 92/09690; Ladner *et al.* PCT International Publication No. WO 90/02809; Fuchs *et al.* (1991) *Bio/Technology* 9:1370-1372; Hay *et al.* (1992) *Hum. Antibod. Hybridomas* 3:81-85; Huse *et al.* (1989) *Science* 246:1275-1281;

Griffiths *et al.* (1993) *EMBO J* 12:725-734; Hawkins *et al.* (1992) *J. Mol. Biol.* 226:889-896; Clarkson *et al.* (1991) *Nature* 352:624-628; Gram *et al.* (1992) *Proc. Natl. Acad. Sci. USA* 89:3576-3580; Garrad *et al.* (1991) *Bio/Technology* 9:1373-1377; Hoogenboom *et al.* (1991) *Nuc. Acid Res.* 19:4133-4137; Barbas *et al.* (1991) *Proc. Natl. Acad. Sci. USA* 88:7978-7982; and McCafferty *et al.* *Nature* (1990) 348:552-554.

Additionally, recombinant anti-TP-2 antibodies, such as chimeric and humanized monoclonal antibodies, comprising both human and non-human portions, which can be made using standard recombinant DNA techniques, are within the scope of the invention. Such chimeric and humanized monoclonal antibodies can be produced by recombinant DNA techniques known in the art, for example using methods described in Robinson *et al.* International Application No. PCT/US86/02269; Akira, *et al.* European Patent Application 184,187; Taniguchi, M., European Patent Application 171,496; Morrison *et al.* European Patent Application 173,494; Neuberger *et al.* PCT International Publication No. WO 86/01533; Cabilly *et al.* U.S. Patent No. 4,816,567; Cabilly *et al.* European Patent Application 125,023; Better *et al.* (1988) *Science* 240:1041-1043; Liu *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:3439-3443; Liu *et al.* (1987) *J. Immunol.* 139:3521-3526; Sun *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:214-218; Nishimura *et al.* (1987) *Canc. Res.* 47:999-1005; Wood *et al.* (1985) *Nature* 314:446-449; and Shaw *et al.* (1988) *J. Natl. Cancer Inst.* 80:1553-1559; Morrison, S. L. (1985) *Science* 229:1202-1207; Oi *et al.* (1986) *BioTechniques* 4:214; Winter U.S. Patent 5,225,539; Jones *et al.* (1986) *Nature* 321:552-525; Verhoeyan *et al.* (1988) *Science* 239:1534; and Beidler *et al.* (1988) *J. Immunol.* 141:4053-4060.

An anti-TP-2 antibody (*e.g.*, monoclonal antibody) can be used to isolate TP-2 by standard techniques, such as affinity chromatography or immunoprecipitation. An anti-TP-2 antibody can facilitate the purification of natural TP-2 from cells and of recombinantly produced TP-2 expressed in host cells. Moreover, an anti-TP-2 antibody can be used to detect TP-2 polypeptide (*e.g.*, in a cellular lysate or cell supernatant) in order to evaluate the abundance and pattern of expression of the TP-2 polypeptide. Anti-TP-2 antibodies can be used diagnostically to monitor polypeptide levels in tissue as part of a clinical testing procedure, *e.g.*, to, for example, determine the efficacy of a given treatment regimen. Detection can be facilitated by coupling (*i.e.*, physically linking) the antibody to a detectable substance. Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, and radioactive materials. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, β -galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine,

dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a luminescent material includes luminol; examples of bioluminescent materials include luciferase, luciferin, and aequorin, and examples of suitable radioactive material include ^{125}I , ^{131}I , ^{35}S or ^3H .

5

III. Recombinant Expression Vectors and Host Cells

Another aspect of the invention pertains to vectors, for example recombinant expression vectors, containing a nucleic acid containing a TP-2 nucleic acid molecule or vectors containing a nucleic acid molecule which encodes a TP-2 polypeptide (or a portion thereof). As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is a "plasmid", which refers to a circular double stranded DNA loop into which additional DNA segments can be ligated. Another type of vector is a viral vector, wherein additional DNA segments can be ligated into the viral genome. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (*e.g.*, bacterial vectors having a bacterial origin of replication and episomal mammalian vectors). Other vectors (*e.g.*, non-episomal mammalian vectors) are integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated along with the host genome. Moreover, certain vectors are capable of directing the expression of genes to which they are operatively linked. Such vectors are referred to herein as "expression vectors". In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids. In the present specification, "plasmid" and "vector" can be used interchangeably as the plasmid is the most commonly used form of vector. However, the invention is intended to include such other forms of expression vectors, such as viral vectors (*e.g.*, replication defective retroviruses, adenoviruses and adeno-associated viruses), which serve equivalent functions.

The recombinant expression vectors of the invention comprise a nucleic acid of the invention in a form suitable for expression of the nucleic acid in a host cell, which means that the recombinant expression vectors include one or more regulatory sequences, selected on the basis of the host cells to be used for expression, which is operatively linked to the nucleic acid sequence to be expressed. Within a recombinant expression vector, "operably linked" is intended to mean that the nucleotide sequence of interest is linked to the regulatory sequence(s) in a manner which allows for expression of the nucleotide sequence (*e.g.*, in an *in vitro* transcription/translation system or in a host cell when the vector is introduced into the host cell). The term "regulatory sequence" is intended to include promoters, enhancers and other expression control elements (*e.g.*, polyadenylation signals). Such regulatory sequences are described, for

example, in Goeddel; *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 cells and those which direct expression of the nucleotide sequence only in certain host cells (*e.g.*,
5 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 polypeptide desired, and the like. The expression vectors of the invention can be introduced into host cells to thereby produce proteins or peptides, including fusion proteins or peptides, encoded by nucleic
10 acids as described herein (*e.g.*, TP-2 polypeptides, mutant forms of TP-2 polypeptides, fusion proteins, and the like).

Accordingly, an exemplary embodiment provides a method for producing a polypeptide, preferably a TP-2 polypeptide, by culturing in a suitable medium a host cell of the invention (*e.g.*, a mammalian host cell such as a non-human mammalian cell)
15 containing a recombinant expression vector, such that the polypeptide is produced.

The recombinant expression vectors of the invention can be designed for expression of TP-2 polypeptides in prokaryotic or eukaryotic cells. For example, TP-2 polypeptides can be expressed in bacterial cells such as *E. coli*, insect cells (using baculovirus expression vectors) yeast cells or mammalian cells. Suitable host cells are
20 discussed further in Goeddel, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990). Alternatively, the recombinant expression vector can be transcribed and translated *in vitro*, for example using T7 promoter regulatory sequences and T7 polymerase.

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

Purified fusion proteins can be utilized in TP-2 activity assays, (e.g., direct assays or competitive assays described in detail below), or to generate antibodies specific for TP-2 polypeptides, for example. In a preferred embodiment, a TP-2 fusion protein expressed in a retroviral expression vector of the present invention can be
5 utilized to infect bone marrow cells which are subsequently transplanted into irradiated recipients. The pathology of the subject recipient is then examined after sufficient time has passed (e.g., six (6) weeks).

Examples of suitable inducible non-fusion *E. coli* expression vectors include pTrc (Amann *et al.*, (1988) *Gene* 69:301-315) and pET 11d (Studier *et al.*, *Gene*
10 *Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, California (1990) 60-89). Target gene expression from the pTrc vector relies on 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
15 polymerase is supplied by host strains BL21(DE3) or HMS174(DE3) from a resident 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 bacteria with an impaired capacity to proteolytically cleave the
20 recombinant protein (Gottesman, S., *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 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).
25 Such alteration of nucleic acid sequences of the invention can be carried out by standard DNA synthesis techniques.

In another embodiment, the TP-2 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 and Herskowitz, (1982) *Cell* 30:933-
30 943), pJRY88 (Schultz *et al.*, (1987) *Gene* 54:113-123), pYES2 (Invitrogen Corporation, San Diego, CA), and picZ (InVitrogen Corp, San Diego, CA).

Alternatively, TP-2 polypeptides can be expressed in insect cells using baculovirus expression vectors. 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.*
35 *Cell Biol.* 3:2156-2165) and the pVL series (Lucklow and Summers (1989) *Virology* 170:31-39).

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, B. (1987) *Nature* 329:840) and pMT2PC (Kaufman *et al.* (1987) *EMBO J.* 6:187-195). When used in mammalian cells, the expression vector's control functions are often provided by viral regulatory elements. For example, commonly used promoters are derived from polyoma, Adenovirus 2, cytomegalovirus and Simian Virus 40. For other suitable expression systems for both prokaryotic and eukaryotic cells see chapters 16 and 17 of Sambrook, J., Fritsch, E. F., and Maniatis, T. *Molecular Cloning: A Laboratory Manual. 2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.*

In another embodiment, the recombinant mammalian expression vector is capable of directing expression of the nucleic acid preferentially in a particular cell type (*e.g.*, tissue-specific regulatory elements are used to express the nucleic acid). 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 Dev.* 1:268-277), lymphoid-specific promoters (Calame and Eaton (1988) *Adv. Immunol.* 43:235-275), in particular promoters of T cell receptors (Winoto and Baltimore (1989) *EMBO J.* 8:729-733) and immunoglobulins (Banerji *et al.* (1983) *Cell* 33:729-740; Queen and Baltimore (1983) *Cell* 33:741-748), neuron-specific promoters (*e.g.*, the neurofilament promoter; Byrne and Ruddle (1989) *Proc. Natl. Acad. Sci. USA* 86:5473-5477), pancreas-specific promoters (Edlund *et al.* (1985) *Science* 230:912-916), and mammary gland-specific promoters (*e.g.*, milk whey promoter; U.S. 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 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 operatively linked to a regulatory sequence in a manner which allows for expression (by transcription of the DNA molecule) of an RNA molecule which is antisense to TP-2 mRNA. Regulatory sequences operatively linked to a nucleic acid cloned in the antisense orientation can be chosen which direct the continuous expression of the antisense RNA molecule in a variety of cell types, for instance viral promoters and/or enhancers, or regulatory sequences can be chosen which direct constitutive, tissue specific or cell type specific expression of antisense RNA. The antisense expression vector can be in the form of a recombinant plasmid, phagemid or attenuated virus in which antisense nucleic acids are produced under the control of a

high efficiency regulatory region, the activity of which can be determined by the cell type into which the vector is introduced. For a discussion of the regulation of gene expression using antisense genes see Weintraub, H. *et al.*, Antisense RNA as a molecular tool for genetic analysis, *Reviews - Trends in Genetics*, Vol. 1(1) 1986.

5 Another aspect of the invention pertains to host cells into which a TP-2 nucleic acid molecule of the invention is introduced, *e.g.*, a TP-2 nucleic acid molecule within a vector (*e.g.*, a recombinant expression vector) or a TP-2 nucleic acid molecule containing sequences which allow it to homologously recombine into a specific site of the host cell's genome. The terms "host cell" and "recombinant host cell" are used
10 interchangeably herein. It is understood that such terms refer not only to the particular subject cell but to the progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein.

15 A host cell can be any prokaryotic or eukaryotic cell. For example, a TP-2 polypeptide can be expressed in bacterial cells such as *E. coli*, insect cells, yeast or mammalian cells (such as Chinese hamster ovary cells (CHO) or COS cells). Other suitable host cells are known to those skilled in the art.

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

For stable transfection of mammalian cells, it is known that, depending upon the expression vector and transfection technique used, only a small fraction of cells may
30 integrate the foreign DNA into their genome. In order to identify and select these integrants, a gene that encodes a selectable marker (*e.g.*, resistance to antibiotics) is generally introduced into the host cells along with the gene of interest. Preferred selectable markers include those which confer resistance to drugs, such as G418, hygromycin and methotrexate. Nucleic acid encoding a selectable marker can be
35 introduced into a host cell on the same vector as that encoding a TP-2 polypeptide or can be introduced on a separate vector. 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 (*i.e.*, express) a TP-2 polypeptide. Accordingly, the invention further provides methods for producing a TP-2 polypeptide using the host cells of the invention. In one embodiment, the method comprises culturing the host cell
5 of the invention (into which a recombinant expression vector encoding a TP-2 polypeptide has been introduced) in a suitable medium such that a TP-2 polypeptide is produced. In another embodiment, the method further comprises isolating a TP-2 polypeptide from the medium or the host cell.

The host cells of the invention can also be used to produce non-human transgenic
10 animals. For example, in one embodiment, a host cell of the invention is a fertilized oocyte or an embryonic stem cell into which TP-2-coding sequences have been introduced. Such host cells can then be used to create non-human transgenic animals in which exogenous TP-2 sequences have been introduced into their genome or
15 homologous recombinant animals in which endogenous TP-2 sequences have been altered. Such animals are useful for studying the function and/or activity of a TP-2 and for identifying and/or evaluating modulators of TP-2 activity. As used herein, a "transgenic animal" is a non-human animal, preferably a mammal, more preferably a rodent such as a rat or mouse, in which one or more of the cells of the animal includes a transgene. Other examples of transgenic animals include non-human primates, sheep,
20 dogs, cows, goats, chickens, amphibians, and the like. A transgene is exogenous DNA which is integrated into the genome of a cell from which a transgenic animal develops and which remains in the genome of the mature animal, thereby directing the expression of an encoded gene product in one or more cell types or tissues of the transgenic animal. As used herein, a "homologous recombinant animal" is a non-human animal, preferably
25 a mammal, more preferably a mouse, in which an endogenous TP-2 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 a TP-2-
30 encoding nucleic acid into the male pronuclei of a fertilized oocyte, *e.g.*, by microinjection, retroviral infection, and allowing the oocyte to develop in a pseudopregnant female foster animal. The TP-2 cDNA sequence of SEQ ID NO:1 can be introduced as a transgene into the genome of a non-human animal. Alternatively, a nonhuman homologue of a human TP-2 gene, such as a mouse or rat TP-2 gene, can be
35 used as a transgene. Alternatively, a TP-2 gene homologue, such as another TP-2 family member, can be isolated based on hybridization to the TP-2 cDNA sequences of SEQ ID NO:1 or 3, or the DNA insert of the plasmid deposited with ATCC as Accession Number _____ (described further in subsection I above) and used as a transgene.

Intronic sequences and polyadenylation signals can also be included in the transgene to increase the efficiency of expression of the transgene. A tissue-specific regulatory sequence(s) can be operably linked to a TP-2 transgene to direct expression of a TP-2 polypeptide to particular cells. Methods for generating transgenic animals via embryo
5 manipulation and microinjection, particularly animals such as mice, have become conventional in the art and are described, for example, in U.S. Patent Nos. 4,736,866 and 4,870,009, both by Leder *et al.*, U.S. Patent No. 4,873,191 by Wagner *et al.* and in Hogan, B., *Manipulating the Mouse Embryo*, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1986). Similar methods are used for production of other
10 transgenic animals. A transgenic founder animal can be identified based upon the presence of a TP-2 transgene in its genome and/or expression of TP-2 mRNA in tissues or cells of the animals. A transgenic founder animal can then be used to breed additional animals carrying the transgene. Moreover, transgenic animals carrying a transgene encoding a TP-2 polypeptide can further be bred to other transgenic animals
15 carrying other transgenes.

To create a homologous recombinant animal, a vector is prepared which contains at least a portion of a TP-2 gene into which a deletion, addition or substitution has been introduced to thereby alter, *e.g.*, functionally disrupt, the TP-2 gene. The TP-2 gene can be a human gene (*e.g.*, the cDNA of SEQ ID NO:3), but more preferably, is a non-
20 human homologue of a human TP-2 gene (*e.g.*, a cDNA isolated by stringent hybridization with the nucleotide sequence of SEQ ID NO:1). For example, a mouse TP-2 gene can be used to construct a homologous recombination nucleic acid molecule, *e.g.*, a vector, suitable for altering an endogenous TP-2 gene in the mouse genome. In a preferred embodiment, the homologous recombination nucleic acid molecule is designed
25 such that, upon homologous recombination, the endogenous TP-2 gene is functionally disrupted (*i.e.*, no longer encodes a functional protein; also referred to as a "knock out" vector). Alternatively, the homologous recombination nucleic acid molecule can be designed such that, upon homologous recombination, the endogenous TP-2 gene is mutated or otherwise altered but still encodes functional polypeptide (*e.g.*, the upstream
30 regulatory region can be altered to thereby alter the expression of the endogenous TP-2 polypeptide). In the homologous recombination nucleic acid molecule, the altered portion of the TP-2 gene is flanked at its 5' and 3' ends by additional nucleic acid sequence of the TP-2 gene to allow for homologous recombination to occur between the exogenous TP-2 gene carried by the homologous recombination nucleic acid molecule
35 and an endogenous TP-2 gene in a cell, *e.g.*, an embryonic stem cell. The additional flanking TP-2 nucleic acid sequence is of sufficient length for successful homologous recombination with the endogenous gene. Typically, several kilobases of flanking DNA (both at the 5' and 3' ends) are included in the homologous recombination nucleic acid

molecule (see, *e.g.*, Thomas, K.R. and Capecchi, M. R. (1987) *Cell* 51:503 for a description of homologous recombination vectors). The homologous recombination nucleic acid molecule is introduced into a cell, *e.g.*, an embryonic stem cell line (*e.g.*, by electroporation) and cells in which the introduced TP-2 gene has homologously recombined with the endogenous TP-2 gene are selected (see *e.g.*, Li, E. *et al.* (1992) *Cell* 69:915). The selected cells can then be injected into a blastocyst of an animal (*e.g.*, a mouse) to form aggregation chimeras (see *e.g.*, Bradley, A. in *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach*, E.J. Robertson, ed. (IRL, Oxford, 1987) pp. 113-152). A chimeric embryo can then be implanted into a suitable pseudopregnant female foster animal and the embryo brought to term. Progeny harboring the homologously recombined DNA in their germ cells can be used to breed animals in which all cells of the animal contain the homologously recombined DNA by germline transmission of the transgene. Methods for constructing homologous recombination nucleic acid molecules, *e.g.*, vectors, or homologous recombinant animals are described further in Bradley, A. (1991) *Current Opinion in Biotechnology* 2:823-829 and in PCT International Publication Nos.: WO 90/11354 by Le Mouellec *et al.*; WO 91/01140 by Smithies *et al.*; WO 92/0968 by Zijlstra *et al.*; and WO 93/04169 by Berns *et al.*

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

Clones of the non-human transgenic animals described herein can also be produced according to the methods described in Wilmut, I. *et al.* (1997) *Nature* 385:810-813 and PCT International Publication Nos. WO 97/07668 and WO 97/07669. In brief, a cell, *e.g.*, a somatic cell, from the transgenic animal can be isolated and induced to exit the growth cycle and enter G₀ phase. The quiescent cell can then be fused, *e.g.*, through the use of electrical pulses, to an enucleated oocyte from an animal of the same species from which the quiescent cell is isolated. The reconstructed oocyte is then cultured such that it develops to morula or blastocyte and then transferred to

pseudopregnant female foster animal. The offspring borne of this female foster animal will be a clone of the animal from which the cell, *e.g.*, the somatic cell, is isolated.

IV. Pharmaceutical Compositions

5 The TP-2 nucleic acid molecules, fragments of TP-2 polypeptides, and anti-TP-2 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, polypeptide, or antibody and a pharmaceutically acceptable carrier. As used herein the language "pharmaceutically
10 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
15 in the compositions is contemplated. Supplementary active compounds can also be incorporated into the compositions.

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),
20 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
25 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
30 plastic.

Pharmaceutical compositions suitable for injectable use include sterile aqueous solutions (where water soluble) or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersion. For intravenous administration, suitable carriers include physiological saline, bacteriostatic
35 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
5 maintenance of the required particle size in the case of dispersion and by the use of surfactants. Prevention of the action of microorganisms can be achieved by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, ascorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars, polyalcohols such as manitol, sorbitol, sodium
10 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 fragment of a TP-2 polypeptide or an anti-TP-2 antibody) in the
15 required amount in an appropriate solvent with one or a combination of ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the active compound into a sterile vehicle which contains a basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions,
20 the preferred methods of preparation are vacuum drying and freeze-drying which yields a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

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

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

Systemic administration can also be by transmucosal or transdermal means. For
5 transmucosal or transdermal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration, detergents, bile salts, and fusidic acid derivatives. Transmucosal administration can be accomplished through the use of nasal sprays or suppositories. For transdermal administration, the active
10 compounds are formulated into ointments, salves, gels, or creams as 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, biocompatible polymers can be used, such as ethylene vinyl acetate, polyanhydrides, polyglycolic acid, collagen, polyorthoesters, and polylactic acid.
20 Methods for preparation of such formulations will be apparent to those skilled in the art. The materials can also be obtained commercially from Alza Corporation and Nova Pharmaceuticals, Inc. Liposomal suspensions (including liposomes targeted to infected cells with monoclonal antibodies to viral antigens) can also be used as pharmaceutically acceptable carriers. These can be prepared according to methods known to those skilled
25 in the art, for example, as described in U.S. Patent No. 4,522,811.

It is especially advantageous to formulate oral or parenteral compositions in dosage unit form for ease of administration and uniformity of dosage. Dosage unit form as used herein refers to physically discrete units suited as unitary dosages for the subject to be treated; each unit containing a predetermined quantity of active compound
30 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.

35 Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, *e.g.*, for determining the LD₅₀ (the dose lethal to 50% of the population) and the ED₅₀ (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and

therapeutic effects is the therapeutic index and it can be expressed as the ratio LD50/ED50. Compounds which exhibit large therapeutic indices are preferred. While compounds that exhibit toxic side effects may be used, care should be taken to design a delivery system that targets such compounds to the site of affected tissue in order to minimize potential damage to uninfected cells and, thereby, reduce side effects.

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

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

In a preferred example, a subject is treated with antibody or polypeptide in the range of between about 0.1 to 20 mg/kg body weight, one time per week for between about 1 to 10 weeks, preferably between 2 to 8 weeks, more preferably between about 3 to 7 weeks, and even more preferably for about 4, 5, or 6 weeks. It will also be appreciated that the effective dosage of antibody or polypeptide used for treatment may increase or decrease over the course of a particular treatment. Changes in dosage may result and become apparent from the results of diagnostic assays as described herein.

The present invention encompasses agents which modulate expression or activity. An agent may, for example, be a small molecule. For example, such small molecules include, but are not limited to, peptides, peptidomimetics, amino acids, amino acid analogs, polynucleotides, polynucleotide analogs, nucleotides, nucleotide analogs,

organic or inorganic compounds (*i.e.*, including heteroorganic and organometallic compounds) having a molecular weight less than about 10,000 grams per mole, organic or inorganic compounds having a molecular weight less than about 5,000 grams per mole, organic or inorganic compounds having a molecular weight less than about 1,000
5 grams per mole, organic or inorganic compounds having a molecular weight less than about 500 grams per mole, and salts, esters, and other pharmaceutically acceptable forms of such compounds. It is understood that appropriate doses of small molecule agents depends upon a number of factors within the ken of the ordinarily skilled physician, veterinarian, or researcher. The dose(s) of the small molecule will vary, for
10 example, depending upon the identity, size, and condition of the subject or sample being treated, further depending upon the route by which the composition is to be administered, if applicable, and the effect which the practitioner desires the small molecule to have upon the nucleic acid or polypeptide of the invention.

Exemplary doses include milligram or microgram amounts of the small molecule
15 per kilogram of subject or sample weight (*e.g.*, about 1 microgram per kilogram to about 500 milligrams per kilogram, about 100 micrograms per kilogram to about 5 milligrams per kilogram, or about 1 microgram per kilogram to about 50 micrograms per kilogram. It is

furthermore understood that appropriate doses of a small molecule depend upon the
20 potency of the small molecule with respect to the expression or activity to be modulated. Such appropriate doses may be determined using the assays described herein. When one or more of these small molecules is to be administered to an animal (*e.g.*, a human) in order to modulate expression or activity of a polypeptide or nucleic acid of the invention, a physician, veterinarian, or researcher may, for example, prescribe a
25 relatively low dose at first, subsequently increasing the dose until an appropriate response is obtained. In addition, it is understood that the specific dose level for any particular animal subject will depend upon a variety of factors including the activity of the specific compound employed, the age, body weight, general health, gender, and diet of the subject, the time of administration, the route of administration, the rate of
30 excretion, any drug combination, and the degree of expression or activity to be modulated.

Further, an antibody (or fragment thereof) may be conjugated to a therapeutic moiety such as a cytotoxin, a therapeutic agent or a radioactive metal ion. A cytotoxin or cytotoxic agent includes any agent that is detrimental to cells. Examples include
35 taxol, cytochalasin B, gramicidin D, ethidium bromide, emetine, mitomycin, etoposide, tenoposide, vincristine, vinblastine, colchicin, doxorubicin, daunorubicin, dihydroxy anthracin dione, mitoxantrone, mithramycin, actinomycin D, 1-dehydrotestosterone, glucocorticoids, procaine, tetracaine, lidocaine, propranolol, and puromycin and analogs

or homologues thereof. Therapeutic agents include, but are not limited to, antimetabolites (*e.g.*, methotrexate, 6-mercaptopurine, 6-thioguanine, cytarabine, 5-fluorouracil decarbazine), alkylating agents (*e.g.*, mechlorethamine, thioepa chlorambucil, melphalan, carmustine (BSNU) and lomustine (CCNU),
5 cyclophosphamide, busulfan, dibromomannitol, streptozotocin, mitomycin C, and cis-dichlorodiamine platinum (II) (DDP) cisplatin), anthracyclines (*e.g.*, daunorubicin (formerly daunomycin) and doxorubicin), antibiotics (*e.g.*, dactinomycin (formerly actinomycin), bleomycin, mithramycin, and anthramycin (AMC)), and anti-mitotic agents (*e.g.*, vincristine and vinblastine).

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

20 Techniques for conjugating such therapeutic moiety to antibodies are well known, see, *e.g.*, Arnon *et al.*, "Monoclonal Antibodies For Immunotargeting Of Drugs In Cancer Therapy", in *Monoclonal Antibodies And Cancer Therapy*, Reisfeld *et al.* (eds.), pp. 243-56 (Alan R. Liss, Inc. 1985); Hellstrom *et al.*, "Antibodies For Drug Delivery", in *Controlled Drug Delivery* (2nd Ed.), Robinson *et al.* (eds.), pp. 623-53
25 (Marcel Dekker, Inc. 1987); Thorpe, "Antibody Carriers Of Cytotoxic Agents In Cancer Therapy: A Review", in *Monoclonal Antibodies '84: Biological And Clinical Applications*, Pinchera *et al.* (eds.), pp. 475-506 (1985); "Analysis, Results, And Future Prospective Of The Therapeutic Use Of Radiolabeled Antibody In Cancer Therapy", in *Monoclonal Antibodies For Cancer Detection And Therapy*, Baldwin *et al.* (eds.), pp.
30 303-16 (Academic Press 1985), and Thorpe *et al.*, "The Preparation And Cytotoxic Properties Of Antibody-Toxin Conjugates", *Immunol. Rev.*, 62:119-58 (1982). Alternatively, an antibody can be conjugated to a second antibody to form an antibody heteroconjugate as described by Segal in U.S. Patent No. 4,676,980.

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

therapy vector in an acceptable diluent, or can comprise a slow release matrix in which the gene delivery vehicle is imbedded. Alternatively, where the complete gene delivery vector can be produced intact from recombinant cells, *e.g.*, retroviral vectors, the pharmaceutical preparation can include one or more cells which produce the gene
5 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) predictive medicine (*e.g.*, diagnostic assays, prognostic assays, monitoring clinical trials, and pharmacogenetics); and c) methods of treatment (*e.g.*, therapeutic and prophylactic). As described herein, a TP-2 polypeptide of the invention has one or more
15 of the following activities: (1) modulate the import and export of molecules, *e.g.*, hormones, ions, cytokines, neurotransmitters, monosaccharides, and metabolites, from cells, 2) modulate intra- or inter-cellular signaling, 3) modulate removal of potentially harmful compounds from the cell, or facilitate the compartmentalization of these molecules into a sequestered intra-cellular space (*e.g.*, the peroxisome), and 4) modulate
20 transport of biological molecules across membranes, *e.g.*, the plasma membrane, or the membrane of the mitochondrion, the peroxisome, the lysosome, the endoplasmic reticulum, the nucleus, or the vacuole.

The isolated nucleic acid molecules of the invention can be used, for example, to express TP-2 polypeptides (*e.g.*, via a recombinant expression vector in a host cell in
25 gene therapy applications), to detect TP-2 mRNA (*e.g.*, in a biological sample) or a genetic alteration in a TP-2 gene, and to modulate TP-2 activity, as described further below. The TP-2 polypeptides can be used to treat disorders characterized by insufficient or excessive production of a TP-2 substrate or production of TP-2 inhibitors. In addition, the TP-2 polypeptides can be used to screen for naturally occurring TP-2
30 substrates, to screen for drugs or compounds which modulate TP-2 activity, as well as to treat disorders characterized by insufficient or excessive production of TP-2 polypeptide or production of TP-2 polypeptide forms which have decreased, aberrant or unwanted activity compared to TP-2 wild type polypeptide (*e.g.*, transporter-associated disorders). Moreover, the anti-TP-2 antibodies of the invention can be used to detect and isolate TP-
35 2 polypeptides, to regulate the bioavailability of TP-2 polypeptides, and modulate TP-2 activity.

A. Screening Assays

The invention provides a method (also referred to herein as a "screening assay") for identifying modulators, *i.e.*, candidate or test compounds or agents (*e.g.*, peptides, peptidomimetics, small molecules or other drugs) which bind to TP-2 polypeptides, 5 have a stimulatory or inhibitory effect on, for example, TP-2 expression or TP-2 activity, or have a stimulatory or inhibitory effect on, for example, the expression or activity of TP-2 substrate.

In one embodiment, the invention provides assays for screening candidate or test compounds which are substrates of a TP-2 polypeptide or polypeptide or biologically 10 active portion thereof. In another embodiment, the invention provides assays for screening candidate or test compounds which bind to or modulate the activity of a TP-2 polypeptide or polypeptide or biologically active portion thereof. The test compounds of the present invention can be obtained using any of the numerous approaches in combinatorial library methods known in the art, including: biological libraries; spatially 15 addressable parallel solid phase or solution phase libraries; synthetic library methods requiring deconvolution; the 'one-bead one-compound' library method; and synthetic library methods using affinity chromatography selection. The biological library approach is limited to peptide libraries, while the other four approaches are applicable to peptide, non-peptide oligomer or small molecule libraries of compounds (Lam, K.S. 20 (1997) *Anticancer Drug Des.* 12:145).

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

Libraries of compounds may be presented in solution (*e.g.*, Houghten (1992) *Biotechniques* 13:412-421), or on beads (Lam (1991) *Nature* 354:82-84), chips (Fodor (1993) *Nature* 364:555-556), bacteria (Ladner USP 5,223,409), spores (Ladner USP 30 '409), plasmids (Cull *et al.* (1992) *Proc Natl Acad Sci USA* 89:1865-1869) or on phage (Scott and Smith (1990) *Science* 249:386-390); (Devlin (1990) *Science* 249:404-406); (Cwirla *et al.* (1990) *Proc. Natl. Acad. Sci.* 87:6378-6382); (Felici (1991) *J. Mol. Biol.* 222:301-310); (Ladner *supra.*).

In one embodiment, an assay is a cell-based assay in which a cell which 35 expresses a TP-2 polypeptide or biologically active portion thereof is contacted with a test compound and the ability of the test compound to modulate TP-2 activity is determined. Determining the ability of the test compound to modulate TP-2 activity can be accomplished by monitoring, for example, intra- or extra-cellular D-glucose, D-

fructose or D-galactose concentration, or insulin or glucagon secretion. The cell, for example, can be of mammalian origin, *e.g.*, a liver cell, fat cell, muscle cell, or a blood cell, such as an erythrocyte.

5 The ability of the test compound to modulate TP-2 binding to a substrate or to bind to TP-2 can also be determined. Determining the ability of the test compound to modulate TP-2 binding to a substrate can be accomplished, for example, by coupling the TP-2 substrate with a radioisotope or enzymatic label such that binding of the TP-2 substrate to TP-2 can be determined by detecting the labeled TP-2 substrate in a complex. Alternatively, TP-2 could be coupled with a radioisotope or enzymatic label
10 to monitor the ability of a test compound to modulate TP-2 binding to a TP-2 substrate in a complex. Determining the ability of the test compound to bind TP-2 can be accomplished, for example, by coupling the compound with a radioisotope or enzymatic label such that binding of the compound to TP-2 can be determined by detecting the labeled TP-2 compound in a complex. For example, compounds (*e.g.*, TP-2 substrates)
15 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. Alternatively, compounds can be enzymatically labeled with, for example, horseradish peroxidase, alkaline phosphatase, or luciferase, and the enzymatic label detected by determination of conversion of an appropriate substrate to product.

20 It is also within the scope of this invention to determine the ability of a compound (*e.g.*, a TP-2 substrate) to interact with TP-2 without the labeling of any of the interactants. For example, a microphysiometer can be used to detect the interaction of a compound with TP-2 without the labeling of either the compound or the TP-2. McConnell, H. M. *et al.* (1992) *Science* 257:1906-1912. As used herein, a
25 "microphysiometer" (*e.g.*, Cytosensor) is an analytical instrument that measures the rate at which a cell acidifies its environment using a light-addressable potentiometric sensor (LAPS). Changes in this acidification rate can be used as an indicator of the interaction between a compound and TP-2.

In another embodiment, an assay is a cell-based assay comprising contacting a
30 cell expressing a TP-2 target molecule (*e.g.*, a TP-2 substrate) with a test compound and determining the ability of the test compound to modulate (*e.g.*, stimulate or inhibit) the activity of the TP-2 target molecule. Determining the ability of the test compound to modulate the activity of a TP-2 target molecule can be accomplished, for example, by determining the ability of the TP-2 polypeptide to bind to or interact with the TP-2 target
35 molecule.

Determining the ability of the TP-2 polypeptide, or a biologically active fragment thereof, to bind to or interact with a TP-2 target molecule can be accomplished by one of the methods described above for determining direct binding. In a preferred

embodiment, determining the ability of the TP-2 polypeptide to bind to or interact with a TP-2 target molecule can be accomplished by determining the activity of the target molecule. For example, the activity of the target molecule can be determined by detecting induction of a cellular second messenger of the target (*i.e.*, intra-cellular Ca^{2+} , diacylglycerol, IP_3 , and the like), detecting catalytic/enzymatic activity of the target using an appropriate substrate, detecting the induction of a reporter gene (comprising a target-responsive regulatory element operatively linked to a nucleic acid encoding a detectable marker, *e.g.*, luciferase), or detecting a target-regulated cellular response.

In yet another embodiment, an assay of the present invention is a cell-free assay in which a TP-2 polypeptide or biologically active portion thereof is contacted with a test compound and the ability of the test compound to bind to the TP-2 polypeptide or biologically active portion thereof is determined. Preferred biologically active portions of the TP-2 polypeptides to be used in assays of the present invention include fragments which participate in interactions with non-TP-2 molecules, *e.g.*, fragments with high surface probability scores (see, for example, Figure 2). Binding of the test compound to the TP-2 polypeptide can be determined either directly or indirectly as described above. In a preferred embodiment, the assay includes contacting the TP-2 polypeptide or biologically active portion thereof with a known compound which binds TP-2 to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with a TP-2 polypeptide, wherein determining the ability of the test compound to interact with a TP-2 polypeptide comprises determining the ability of the test compound to preferentially bind to TP-2 or biologically active portion thereof as compared to the known compound.

In another embodiment, the assay is a cell-free assay in which a TP-2 polypeptide or biologically active portion thereof is contacted with a test compound and the ability of the test compound to modulate (*e.g.*, stimulate or inhibit) the activity of the TP-2 polypeptide or biologically active portion thereof is determined. Determining the ability of the test compound to modulate the activity of a TP-2 polypeptide can be accomplished, for example, by determining the ability of the TP-2 polypeptide to bind to a TP-2 target molecule by one of the methods described above for determining direct binding. Determining the ability of the TP-2 polypeptide to bind to a TP-2 target molecule can also be accomplished using a technology such as real-time Biomolecular Interaction Analysis (BIA). Sjolander, S. and Urbaniczky, C. (1991) *Anal. Chem.* 63:2338-2345 and Szabo *et al.* (1995) *Curr. Opin. Struct. Biol.* 5:699-705. As used herein, "BIA" is a technology for studying biospecific interactions in real time, without labeling any of the interactants (*e.g.*, BIAcore). Changes in the optical phenomenon of surface plasmon resonance (SPR) can be used as an indication of real-time reactions between biological molecules.

In an alternative embodiment, determining the ability of the test compound to modulate the activity of a TP-2 polypeptide can be accomplished by determining the ability of the TP-2 polypeptide to further modulate the activity of a downstream effector of a TP-2 target molecule. For example, the activity of the effector molecule on an appropriate target can be determined or the binding of the effector to an appropriate target can be determined as previously described.

In yet another embodiment, the cell-free assay involves contacting a TP-2 polypeptide or biologically active portion thereof with a known compound which binds the TP-2 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 TP-2 polypeptide, wherein determining the ability of the test compound to interact with the TP-2 polypeptide comprises determining the ability of the TP-2 polypeptide to preferentially bind to or modulate the activity of a TP-2 target molecule.

In more than one embodiment of the above assay methods of the present invention, it may be desirable to immobilize either TP-2 or its target molecule to facilitate separation of complexed from uncomplexed forms of one or both of the proteins, as well as to accommodate automation of the assay. Binding of a test compound to a TP-2 polypeptide, or interaction of a TP-2 polypeptide with a target molecule in the presence and absence of a candidate compound, can be accomplished in any vessel suitable for containing the reactants. Examples of such vessels include microtiter plates, test tubes, and micro-centrifuge tubes. In one embodiment, a fusion protein can be provided which adds a domain that allows one or both of the proteins to be bound to a matrix. For example, glutathione-S-transferase/TP-2 fusion proteins or glutathione-S-transferase/target fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, MO) or glutathione derivatized micrometer plates, which are then combined with the test compound or the test compound and either the non-adsorbed target protein or TP-2 polypeptide, and the mixture incubated under conditions conducive to complex formation (*e.g.*, at physiological conditions for salt and pH). Following incubation, the beads or micrometer plate wells are washed to remove any unbound components, the matrix immobilized in the case of beads, complex determined either directly or indirectly, for example, as described above. Alternatively, the complexes can be dissociated from the matrix, and the level of TP-2 binding or activity determined using standard techniques.

Other techniques for immobilizing proteins on matrices can also be used in the screening assays of the invention. For example, either a TP-2 polypeptide or a TP-2 target molecule can be immobilized utilizing conjugation of biotin and streptavidin. Biotinylated TP-2 polypeptide or target molecules can be prepared from biotin-NHS (N-hydroxy-succinimide) using techniques known in the art (*e.g.*, biotinylation kit, Pierce

Chemicals, Rockford, IL), and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical). Alternatively, antibodies reactive with TP-2 polypeptide or target molecules but which do not interfere with binding of the TP-2 polypeptide to its target molecule can be derivatized to the wells of the plate, and unbound target or TP-2 polypeptide 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 TP-2 polypeptide or target molecule, as well as enzyme-linked assays which rely on detecting an enzymatic activity associated with the TP-2 polypeptide or target molecule.

10 In another embodiment, modulators of TP-2 expression are identified in a method wherein a cell is contacted with a candidate compound and the expression of TP-2 mRNA or polypeptide in the cell is determined. The level of expression of TP-2 mRNA or polypeptide in the presence of the candidate compound is compared to the level of expression of TP-2 mRNA or polypeptide in the absence of the candidate compound. The candidate compound can then be identified as a modulator of TP-2 expression based on this comparison. For example, when expression of TP-2 mRNA or polypeptide 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 TP-2 mRNA or polypeptide expression. Alternatively, when expression of TP-2 mRNA or polypeptide 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 TP-2 mRNA or polypeptide expression. The level of TP-2 mRNA or polypeptide expression in the cells can be determined by methods described herein for detecting TP-2 mRNA or polypeptide.

25 In yet another aspect of the invention, the TP-2 polypeptides can be used as "bait proteins" in a two-hybrid assay or three-hybrid assay (see, *e.g.*, U.S. Patent No. 5,283,317; Zervos *et al.* (1993) *Cell* 72:223-232; Madura *et al.* (1993) *J. Biol. Chem.* 268:12046-12054; Bartel *et al.* (1993) *Biotechniques* 14:920-924; Iwabuchi *et al.* (1993) *Oncogene* 8:1693-1696; and Brent WO94/10300), to identify other proteins, which bind to or interact with TP-2 ("TP-2-binding proteins" or "TP-2-bp") and are involved in TP-2 activity. Such TP-2-binding proteins are also likely to be involved in the propagation of signals by the TP-2 polypeptides or TP-2 targets as, for example, downstream elements of a TP-2-mediated signaling pathway. Alternatively, such TP-2-binding proteins are likely to be TP-2 inhibitors.

35 The two-hybrid system is based on the modular nature of most transcription factors, which consist of separable DNA-binding and activation domains. Briefly, the assay utilizes two different DNA constructs. In one construct, the gene that codes for a TP-2 polypeptide is fused to a gene encoding the DNA binding domain of a known

transcription factor (*e.g.*, GAL-4). In the other construct, a DNA sequence, from a library of DNA sequences, that encodes an unidentified protein ("prey" or "sample") is fused to a gene that codes for the activation domain of the known transcription factor. If the "bait" and the "prey" proteins are able to interact, *in vivo*, forming a TP-2-dependent complex, the DNA-binding and activation domains of the transcription factor are brought into close proximity. This proximity allows transcription of a reporter gene (*e.g.*, LacZ) which is operably linked to a transcriptional regulatory site responsive to the transcription factor. Expression of the reporter gene can be detected and cell colonies containing the functional transcription factor can be isolated and used to obtain the cloned gene which encodes the protein which interacts with the TP-2 polypeptide.

In another aspect, the invention pertains to a combination of two or more of the assays described herein. For example, a modulating agent can be identified using a cell-based or a cell free assay, and the ability of the agent to modulate the activity of a TP-2 polypeptide can be confirmed *in vivo*, *e.g.*, in an animal such as an animal model for cellular transformation and/or tumorigenesis.

This invention further pertains to novel agents identified by the above-described screening assays. Accordingly, it is within the scope of this invention to further use an agent identified as described herein in an appropriate animal model. For example, an agent identified as described herein (*e.g.*, a TP-2 modulating agent, an antisense TP-2 nucleic acid molecule, a TP-2-specific antibody, or a TP-2-binding partner) can be used in an animal model to determine the efficacy, toxicity, or side effects of treatment with such an agent. Alternatively, an agent identified as described herein can be used in an animal model to determine the mechanism of action of such an agent. Furthermore, this invention pertains to uses of novel agents identified by the above-described screening assays for treatments as described herein.

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

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. This process is

called chromosome mapping. Accordingly, portions or fragments of the TP-2 nucleotide sequences, described herein, can be used to map the location of the TP-2 genes on a chromosome. The mapping of the TP-2 sequences to chromosomes is an important first step in correlating these sequences with genes associated with disease.

5 Briefly, TP-2 genes can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp in length) from the TP-2 nucleotide sequences. Computer analysis of the TP-2 sequences can be used to predict primers that do not span more 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
10 chromosomes. Only those hybrids containing the human gene corresponding to the TP-2 sequences will yield an amplified fragment.

Somatic cell hybrids are prepared by fusing somatic cells from different mammals (*e.g.*, human and mouse cells). As hybrids of human and mouse cells grow and divide, they gradually lose human chromosomes in random order, but retain the
15 mouse chromosomes. By using media in which mouse cells cannot grow, because they lack a particular enzyme, but human cells can, the one human chromosome that contains the gene encoding the needed enzyme, will be retained. By using various media, panels of hybrid cell lines can be established. Each cell line in a panel contains either a single human chromosome or a small number of human chromosomes, and a full set of mouse
20 chromosomes, allowing easy mapping of individual genes to specific human chromosomes (D'Eustachio P. *et al.* (1983) *Science* 220:919-924). Somatic cell hybrids containing only fragments of human chromosomes can also be produced by using human chromosomes with translocations and deletions.

PCR mapping of somatic cell hybrids is a rapid procedure for assigning a
25 particular sequence to a particular chromosome. Three or more sequences can be assigned per day using a single thermal cycler. Using the TP-2 nucleotide sequences to design oligonucleotide primers, sublocalization can be achieved with panels of fragments from specific chromosomes. Other mapping strategies which can similarly be used to map a TP-2 sequence to its chromosome include *in situ* hybridization (described
30 in Fan, Y. *et al.* (1990) *Proc. Natl. Acad. Sci. USA*, 87:6223-27), pre-screening with labeled flow-sorted chromosomes, and pre-selection by hybridization to chromosome specific cDNA libraries.

Fluorescence *in situ* hybridization (FISH) of a DNA sequence to a metaphase
35 chromosomal spread can further be used to provide a precise chromosomal location in one step. Chromosome spreads can be made using cells whose division has been blocked in metaphase by a chemical such as colcemid that disrupts the mitotic spindle. The chromosomes can be treated briefly with trypsin, and then stained with Giemsa. A pattern of light and dark bands develops on each chromosome, so that the chromosomes

can be identified individually. The FISH technique can be used with a DNA sequence as short as 500 or 600 bases. However, clones larger than 1,000 bases have a higher likelihood of binding to a unique chromosomal location with sufficient signal intensity for simple detection. Preferably 1,000 bases, and more preferably 2,000 bases will suffice to get good results at a reasonable amount of time. For a review of this technique, see Verma *et al.*, 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 marking multiple sites and/or multiple chromosomes. Reagents corresponding to noncoding regions of the genes actually are preferred for mapping purposes. Coding sequences are more likely to be conserved within gene families, thus increasing the chance of cross hybridizations during chromosomal mapping.

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. McKusick, Mendelian Inheritance in Man, available on-line through Johns Hopkins University Welch Medical Library). The relationship between a gene and a disease, mapped to the same chromosomal region, can then be identified through linkage analysis (co-inheritance of physically adjacent genes), described in, for example, Egeland, J. *et al.* (1987) *Nature*, 325:783-787.

Moreover, differences in the DNA sequences between individuals affected and unaffected with a disease associated with the TP-2 gene, can be determined. If a mutation is observed in some or all of the affected individuals but not in any unaffected individuals, then the mutation is likely to be the causative agent of the particular disease. Comparison of affected and unaffected individuals generally involves first looking for structural alterations in the chromosomes, such as deletions or translocations that are visible from chromosome spreads or detectable using PCR based on that DNA sequence. Ultimately, complete sequencing of genes from several individuals can be performed to confirm the presence of a mutation and to distinguish mutations from polymorphisms.

2. Tissue Typing

The TP-2 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 TP-2 nucleotide sequences described herein can be used to prepare two PCR primers from the 5' and 3' ends of the sequences. These primers can then be used to amplify an individual's DNA and subsequently sequence it.

Panels of corresponding DNA sequences from individuals, prepared in this manner, can provide unique individual identifications, as each individual will have a unique set of such DNA sequences due to allelic differences. The sequences of the present invention can be used to obtain such identification sequences from individuals and from tissue. The TP-2 nucleotide 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 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 are used, a more appropriate number of primers for positive individual identification would be 500-2,000.

If a panel of reagents from TP-2 nucleotide sequences described herein is used to generate a unique identification database for an individual, those same reagents can later be used to identify tissue from that individual. Using the unique identification database, positive identification of the individual, living or dead, can be made from extremely small tissue samples.

3. Use of TP-2 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 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
5 enhance the reliability of DNA-based forensic identifications by, for example, providing another "identification marker" (*i.e.* another DNA sequence that is unique to a particular individual). As mentioned above, actual base sequence information can be used for identification as an accurate alternative to patterns formed by restriction enzyme generated fragments. Sequences targeted to noncoding regions of SEQ ID NO:1 are
10 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 TP-2 nucleotide sequences or portions thereof, *e.g.*, fragments derived from the noncoding regions of SEQ ID NO:1 having a length of at least 20 bases, preferably at least 30 bases.

15 The TP-2 nucleotide sequences described herein can further be used to provide polynucleotide reagents, *e.g.*, labeled or labelable probes which can be used in, for example, an *in situ* hybridization technique, to identify a specific tissue, *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 TP-2 probes can be used to identify tissue by
20 species and/or by organ type.

In a similar fashion, these reagents, *e.g.*, TP-2 primers or probes can be used to screen tissue culture for contamination (*i.e.* screen for the presence of a mixture of different types of cells in a culture).

25 C. Predictive Medicine:

The present invention also pertains to the field of predictive medicine in which diagnostic assays, prognostic assays, and monitoring clinical trials are used for prognostic (predictive) purposes to thereby treat an individual prophylactically. Accordingly, one aspect of the present invention relates to diagnostic assays for
30 determining TP-2 polypeptide and/or nucleic acid expression as well as TP-2 activity, 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 or unwanted TP-2 expression or activity. The invention also provides for prognostic (or predictive) assays for determining whether an
35 individual is at risk of developing a disorder associated with TP-2 polypeptide, nucleic acid expression or activity. For example, mutations in a TP-2 gene 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 with TP-2 polypeptide, nucleic acid expression or activity.

Another aspect of the invention pertains to monitoring the influence of agents (*e.g.*, drugs, compounds) on the expression or activity of TP-2 in clinical trials.

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

1. Diagnostic Assays

An exemplary method for detecting the presence or absence of TP-2 polypeptide or nucleic acid in a biological sample involves obtaining a biological sample from a test
10 subject and contacting the biological sample with a compound or an agent capable of detecting TP-2 polypeptide or nucleic acid (*e.g.*, mRNA, or genomic DNA) that encodes TP-2 polypeptide such that the presence of TP-2 polypeptide or nucleic acid is detected in the biological sample. In another aspect, the present invention provides a method for detecting the presence of TP-2 activity in a biological sample by contacting the
15 biological sample with an agent capable of detecting an indicator of TP-2 activity such that the presence of TP-2 activity is detected in the biological sample. A preferred agent for detecting TP-2 mRNA or genomic DNA is a labeled nucleic acid probe capable of hybridizing to TP-2 mRNA or genomic DNA. The nucleic acid probe can be, for example, the TP-2 nucleic acid set forth in SEQ ID NO:1 or 3, or the DNA insert of the
20 plasmid deposited with ATCC as Accession Number _____, 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 TP-2 mRNA or genomic DNA. Other suitable probes for use in the diagnostic assays of the invention are described herein.

25 A preferred agent for detecting TP-2 polypeptide is an antibody capable of binding to TP-2 polypeptide, preferably an antibody with a detectable label. Antibodies can be polyclonal, or more preferably, monoclonal. An intact antibody, or a fragment thereof (*e.g.*, Fab or F(ab')₂) can be used. The term "labeled", with regard to the probe or antibody, is intended to encompass direct labeling of the probe or antibody by
30 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.
35 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 TP-2 mRNA, polypeptide, or genomic DNA in a biological sample *in vitro* as well as *in vivo*. For example, *in vitro*

techniques for detection of TP-2 mRNA include Northern hybridizations and *in situ* hybridizations. *In vitro* techniques for detection of TP-2 polypeptide include enzyme linked immunosorbent assays (ELISAs), Western blots, immunoprecipitations and immunofluorescence. *In vitro* techniques for detection of TP-2 genomic DNA include
5 Southern hybridizations. Furthermore, *in vivo* techniques for detection of TP-2 polypeptide include introducing into a subject a labeled anti-TP-2 antibody. For example, the antibody can be labeled with a radioactive marker whose presence and location in a subject can be detected by standard imaging techniques.

The present invention also provides diagnostic assays for identifying the
10 presence or absence of a genetic alteration characterized by at least one of (i) aberrant modification or mutation of a gene encoding a TP-2 polypeptide; (ii) aberrant expression of a gene encoding a TP-2 polypeptide; (iii) mis-regulation of the gene; and (iii) aberrant post-translational modification of a TP-2 polypeptide, wherein a wild-type form of the gene encodes a polypeptide with a TP-2 activity. "Misexpression or aberrant
15 expression", as used herein, refers to a non-wild type pattern of gene expression, at the RNA or protein level. It includes, but is not limited to, expression at non-wild type levels (*e.g.*, over or under expression); a pattern of expression that differs from wild type in terms of the time or stage at which the gene is expressed (*e.g.*, increased or decreased expression (as compared with wild type) at a predetermined developmental period or
20 stage); a pattern of expression that differs from wild type in terms of decreased expression (as compared with wild type) in a predetermined cell type or tissue type; a pattern of expression that differs from wild type in terms of the splicing size, amino acid sequence, post-translational modification, or biological activity of the expressed polypeptide; a pattern of expression that differs from wild type in terms of the effect of
25 an environmental stimulus or extracellular stimulus on expression of the gene (*e.g.*, a pattern of increased or decreased expression (as compared with wild type) in the presence of an increase or decrease in the strength of the stimulus).

In one embodiment, the biological sample contains protein molecules from the test subject. Alternatively, the biological sample can contain mRNA molecules from the
30 test subject or genomic DNA molecules from the test subject. A preferred biological sample is a serum sample isolated by conventional means 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 TP-2 polypeptide, mRNA, or genomic DNA,
35 such that the presence of TP-2 polypeptide, mRNA or genomic DNA is detected in the biological sample, and comparing the presence of TP-2 polypeptide, mRNA or genomic DNA in the control sample with the presence of TP-2 polypeptide, mRNA or genomic DNA in the test sample.

The invention also encompasses kits for detecting the presence of TP-2 in a biological sample. For example, the kit can comprise a labeled compound or agent capable of detecting TP-2 polypeptide or mRNA in a biological sample; means for determining the amount of TP-2 in the sample; and means for comparing the amount of TP-2 in the sample with a standard. The compound or agent can be packaged in a suitable container. The kit can further comprise instructions for using the kit to detect TP-2 polypeptide or nucleic acid.

2. Prognostic Assays

The diagnostic methods described herein can furthermore be utilized to identify subjects having or at risk of developing a disease or disorder associated with aberrant or unwanted TP-2 expression or activity. As used herein, the term "aberrant" includes a TP-2 expression or activity which deviates from the wild type TP-2 expression or activity. Aberrant expression or activity includes increased or decreased expression or activity, as well as expression or activity which does not follow the wild type developmental pattern of expression or the subcellular pattern of expression. For example, aberrant TP-2 expression or activity is intended to include the cases in which a mutation in the TP-2 gene causes the TP-2 gene to be under-expressed or over-expressed and situations in which such mutations result in a non-functional TP-2 polypeptide or a polypeptide which does not function in a wild-type fashion, *e.g.*, a polypeptide which does not interact with a TP-2 substrate, *e.g.*, a transporter subunit or ligand, or one which interacts with a non-TP-2 substrate, *e.g.* a non- transporter subunit or ligand. As used herein, the term "unwanted" includes an unwanted phenomenon involved in a biological response, such as cellular proliferation. For example, the term unwanted includes a TP-2 expression or activity which is undesirable in a subject.

The 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 a misregulation in TP-2 polypeptide activity or nucleic acid expression, such as a transporter-associated disorder. Alternatively, the prognostic assays can be utilized to identify a subject having or at risk for developing a disorder associated with a misregulation in TP-2 polypeptide activity or nucleic acid expression, such as a transporter-associated disorder. Thus, the present invention provides a method for identifying a disease or disorder associated with aberrant or unwanted TP-2 expression or activity in which a test sample is obtained from a subject and TP-2 polypeptide or nucleic acid (*e.g.*, mRNA or genomic DNA) is detected, wherein the presence of TP-2 polypeptide or nucleic acid is diagnostic for a subject having or at risk of developing a disease or disorder associated with aberrant or unwanted TP-2 expression or activity. 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 herein can be used to determine whether a subject can be administered an agent (e.g., an agonist, antagonist, peptidomimetic, protein, peptide, nucleic acid, small molecule, or other drug candidate) to treat a disease or disorder associated with aberrant or unwanted TP-2 expression or activity. For example, such methods can be used to determine whether a subject can be effectively treated with an agent for a transporter-associated disorder. Thus, the present invention provides methods for determining whether a subject can be effectively treated with an agent for a disorder associated with aberrant or unwanted TP-2 expression or activity in which a test sample is obtained and TP-2 polypeptide or nucleic acid expression or activity is detected (e.g., wherein the abundance of TP-2 polypeptide or nucleic acid expression or activity is diagnostic for a subject that can be administered the agent to treat a disorder associated with aberrant or unwanted TP-2 expression or activity).

The methods of the invention can also be used to detect genetic alterations in a TP-2 gene, thereby determining if a subject with the altered gene is at risk for a disorder characterized by misregulation in TP-2 polypeptide activity or nucleic acid expression, such as a transporter-associated disorder. In preferred embodiments, the methods include detecting, in a sample of cells from the subject, the presence or absence of a genetic alteration characterized by at least one of an alteration affecting the integrity of a gene encoding a TP-2 -polypeptide, or the mis-expression of the TP-2 gene. For example, such genetic alterations can be detected by ascertaining the existence of at least one of 1) a deletion of one or more nucleotides from a TP-2 gene; 2) an addition of one or more nucleotides to a TP-2 gene; 3) a substitution of one or more nucleotides of a TP-2 gene, 4) a chromosomal rearrangement of a TP-2 gene; 5) an alteration in the level of a messenger RNA transcript of a TP-2 gene, 6) aberrant modification of a TP-2 gene, such as of the methylation pattern of the genomic DNA, 7) the presence of a non-wild type splicing pattern of a messenger RNA transcript of a TP-2 gene, 8) a non-wild type level of a TP-2-polypeptide, 9) allelic loss of a TP-2 gene, and 10) inappropriate post-translational modification of a TP-2-polypeptide. As described herein, there are a large number of assays known in the art which can be used for detecting alterations in a TP-2 gene. A preferred biological sample is a tissue or serum sample isolated by conventional means from a subject.

In certain embodiments, detection of the alteration involves the use of a probe/primer in a polymerase chain reaction (PCR) (see, e.g., U.S. Patent Nos. 4,683,195 and 4,683,202), such as anchor PCR or RACE PCR, or, alternatively, in a ligation chain reaction (LCR) (see, e.g., Landegran *et al.* (1988) *Science* 241:1077-1080;

and Nakazawa *et al.* (1994) *Proc. Natl. Acad. Sci. USA* 91:360-364), the latter of which can be particularly useful for detecting point mutations in the TP-2-gene (see Abravaya *et al.* (1995) *Nucleic Acids Res.* 23:675-682). This method can include the steps of collecting a sample of cells from a subject, isolating nucleic acid (*e.g.*, genomic, mRNA or both) from the cells of the sample, contacting the nucleic acid sample with one or more primers which specifically hybridize to a TP-2 gene under conditions such that hybridization and amplification of the TP-2-gene (if present) occurs, and detecting the presence or absence of an amplification product, or detecting the size of the amplification product and comparing the length to a control sample. It is anticipated that PCR and/or LCR may be desirable to use as a preliminary amplification step in conjunction with any of the techniques used for detecting mutations described herein.

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

In an alternative embodiment, mutations in a TP-2 gene from a sample cell can be identified by alterations in restriction enzyme cleavage patterns. For example, sample and control DNA is isolated, amplified (optionally), digested with one or more restriction endonucleases, and fragment length sizes are determined by gel electrophoresis and compared. Differences in fragment length sizes between sample and control DNA indicates mutations in the sample DNA. Moreover, the use of sequence specific ribozymes (see, for example, U.S. Patent No. 5,498,531) can be used to score for the presence of specific mutations by development or loss of a ribozyme cleavage site.

In other embodiments, genetic mutations in TP-2 can be identified by hybridizing a sample and control nucleic acids, *e.g.*, DNA or RNA, to high density arrays containing hundreds or thousands of oligonucleotide probes (Cronin, M.T. *et al.* (1996) *Human Mutation* 7: 244-255; Kozal, M.J. *et al.* (1996) *Nature Medicine* 2: 753-759). For example, genetic mutations in TP-2 can be identified in two dimensional arrays containing light-generated DNA probes as described in Cronin, M.T. *et al. supra.* Briefly, a first hybridization array of probes can be used to scan through long stretches of DNA in a sample and control to identify base changes between the sequences by making linear arrays of sequential overlapping probes. This step allows the identification of point mutations. This step is followed by a second hybridization array

that allows the characterization of specific mutations by using smaller, specialized probe arrays complementary to all variants or mutations detected. Each mutation array is composed of parallel probe sets, one complementary to the wild-type gene and the other complementary to the mutant gene.

5 In yet another embodiment, any of a variety of sequencing reactions known in the art can be used to directly sequence the TP-2 gene and detect mutations by comparing the sequence of the sample TP-2 with the corresponding wild-type (control) sequence. Examples of sequencing reactions include those based on techniques developed by Maxam and Gilbert ((1977) *Proc. Natl. Acad. Sci. USA* 74:560) or Sanger
10 ((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) *Biotechniques* 19:448), including sequencing by mass spectrometry (see, e.g., PCT International Publication No. WO 94/16101; Cohen *et al.* (1996) *Adv. Chromatogr.* 36:127-162; and Griffin *et al.* (1993) *Appl. Biochem.*
15 *Biotechnol.* 38:147-159).

Other methods for detecting mutations in the TP-2 gene include methods in which protection from cleavage agents is used to detect mismatched bases in RNA/RNA or RNA/DNA heteroduplexes (Myers *et al.* (1985) *Science* 230:1242). In general, the art technique of "mismatch cleavage" starts by providing heteroduplexes of formed by
20 hybridizing (labeled) RNA or DNA containing the wild-type TP-2 sequence with potentially mutant RNA or DNA obtained from a tissue sample. The double-stranded duplexes are treated with an agent which cleaves single-stranded regions of the duplex such as which will exist due to basepair mismatches between the control and sample strands. For instance, RNA/DNA duplexes can be treated with RNase and DNA/DNA
25 hybrids treated with S1 nuclease to enzymatically digesting the mismatched regions. In other embodiments, either DNA/DNA or RNA/DNA duplexes can be treated with hydroxylamine or osmium tetroxide and with piperidine in order to digest mismatched regions. After digestion of the mismatched regions, the resulting material is then separated by size on denaturing polyacrylamide gels to determine the site of mutation.
30 See, for example, Cotton *et al.* (1988) *Proc. Natl Acad Sci 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 reaction employs one or more proteins that recognize mismatched base pairs in double-stranded DNA (so called
35 "DNA mismatch repair" enzymes) in defined systems for detecting and mapping point mutations in TP-2 cDNAs obtained from samples of cells. For example, the mutY enzyme of *E. coli* cleaves A at G/A mismatches and the thymidine DNA glycosylase from HeLa cells cleaves T at G/T mismatches (Hsu *et al.* (1994) *Carcinogenesis*

15:1657-1662). According to an exemplary embodiment, a probe based on a TP-2 sequence, *e.g.*, a wild-type TP-2 sequence, is hybridized to a cDNA or other DNA product from a test cell(s). The duplex is treated with a DNA mismatch repair enzyme, and the cleavage products, if any, can be detected from electrophoresis protocols or the like. See, for example, U.S. Patent No. 5,459,039.

In other embodiments, alterations in electrophoretic mobility will be used to identify mutations in TP-2 genes. For example, single strand conformation polymorphism (SSCP) may be used to detect differences in electrophoretic mobility between mutant and wild type nucleic acids (Orita *et al.* (1989) *Proc Natl. Acad. Sci USA*: 86:2766, see also Cotton (1993) *Mutat. Res.* 285:125-144; and Hayashi (1992) *Genet. Anal. Tech. Appl.* 9:73-79). Single-stranded DNA fragments of sample and control TP-2 nucleic acids will be denatured and allowed to renature. The secondary structure of single-stranded nucleic acids varies according to sequence, the resulting alteration in electrophoretic mobility enables the detection of even a single base change. The DNA fragments may be labeled or detected with labeled probes. The sensitivity of the assay may be enhanced by using RNA (rather than DNA), in which the secondary structure is more sensitive to a change in sequence. In a preferred embodiment, the subject method utilizes heteroduplex analysis to separate double stranded heteroduplex molecules on the basis of changes in electrophoretic mobility (Keen *et al.* (1991) *Trends Genet* 7:5).

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

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

Alternatively, allele specific amplification technology which depends on selective PCR amplification may be used in conjunction with the instant invention.

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

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 TP-2 gene.

Furthermore, any cell type or tissue in which TP-2 is expressed may be utilized in the prognostic assays described herein.

3. Monitoring of Effects During Clinical Trials

Monitoring the influence of agents (*e.g.*, drugs) on the expression or activity of a TP-2 polypeptide (*e.g.*, the modulation of transport of biological molecules across membranes) can be applied not only in basic drug screening, but also in clinical trials. For example, the effectiveness of an agent determined by a screening assay as described herein to increase TP-2 gene expression, polypeptide levels, or upregulate TP-2 activity, can be monitored in clinical trials of subjects exhibiting decreased TP-2 gene expression, polypeptide levels, or downregulated TP-2 activity. Alternatively, the effectiveness of an agent determined by a screening assay to decrease TP-2 gene expression, polypeptide levels, or downregulate TP-2 activity, can be monitored in clinical trials of subjects exhibiting increased TP-2 gene expression, polypeptide levels, or upregulated TP-2 activity. In such clinical trials, the expression or activity of a TP-2 gene, and preferably, other genes that have been implicated in, for example, a TP-2-associated disorder can be used as a "read out" or markers of the phenotype of a particular cell.

For example, and not by way of limitation, genes, including TP-2, that are modulated in cells by treatment with an agent (*e.g.*, compound, drug or small molecule) which modulates TP-2 activity (*e.g.*, identified in a screening assay as described herein) can be identified. Thus, to study the effect of agents on transporter-associated disorders, for example, in a clinical trial, cells can be isolated and RNA prepared and analyzed for the levels of expression of TP-2 and other genes implicated in the transporter-associated disorder, respectively. The levels of gene expression (*e.g.*, a gene expression pattern) can be quantified by northern blot analysis or RT-PCR, as described herein, or alternatively by measuring the amount of polypeptide produced, by one of the methods as described herein, or by measuring the levels of activity of TP-2 or other genes. In this way, the gene expression pattern can serve as a marker, indicative of the physiological response of the cells to the agent. Accordingly, this response state may be determined before, and at various points during treatment of the individual with the agent.

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 by the screening assays described herein) including the steps of (i) obtaining a pre-administration sample from a subject prior to administration of the agent; (ii) detecting the level of expression of a TP-2 polypeptide, mRNA, or genomic DNA in the preadministration sample; (iii) obtaining one or more post-administration samples from the subject; (iv) detecting the level of expression or activity of the TP-2 polypeptide, mRNA, or genomic DNA in the post-administration samples; (v) comparing the level of expression or activity of the TP-2 polypeptide, mRNA, or genomic DNA in the pre-administration sample with the TP-2 polypeptide, mRNA, or genomic DNA in the post administration sample or samples; and (vi) altering the administration of the agent to the subject accordingly. For example, increased administration of the agent may be desirable to increase the expression or activity of TP-2 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 TP-2 to lower levels than detected, *i.e.* to decrease the effectiveness of the agent. According to such an embodiment, TP-2 expression or activity may be used as an indicator of the effectiveness of an agent, even in the absence of an observable phenotypic response.

35 D. 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 or unwanted TP-2 expression or activity, *e.g.* a transporter-associated

disorder. "Treatment", or "treating" as used herein, is defined as the application or administration of a therapeutic agent to a patient, or application or administration of a therapeutic agent to an isolated tissue or cell line from a patient, who has a disease or disorder, a symptom of disease or disorder or a predisposition toward a disease or disorder, with the purpose to cure, heal, alleviate, relieve, alter, remedy, ameliorate, improve or affect the disease or disorder, the symptoms of the disease or disorder, or the predisposition toward disease. A therapeutic agent includes, but is not limited to, small molecules, peptides, antibodies, ribozymes and antisense oligonucleotides. With regards to both prophylactic and therapeutic methods of treatment, such treatments may be specifically tailored or modified, based on knowledge obtained from the field of pharmacogenomics. "Pharmacogenomics", as used herein, refers to the application of genomics technologies such as gene sequencing, statistical genetics, and gene expression analysis to drugs in clinical development and on the market. More specifically, the term refers the study of how a patient's genes determine his or her response to a drug (*e.g.*, a patient's "drug response phenotype", or "drug response genotype"). Thus, another aspect of the invention provides methods for tailoring an individual's prophylactic or therapeutic treatment with either the TP-2 molecules of the present invention or TP-2 modulators according to that individual's drug response genotype. Pharmacogenomics allows a clinician or physician to target prophylactic or therapeutic treatments to patients who will most benefit from the treatment and to avoid treatment of patients who will experience toxic drug-related side effects.

1. Prophylactic Methods

In one aspect, the invention provides a method for preventing in a subject, a disease or condition associated with an aberrant or unwanted TP-2 expression or activity, by administering to the subject a TP-2 or an agent which modulates TP-2 expression or at least one TP-2 activity. Subjects at risk for a disease which is caused or contributed to by aberrant or unwanted TP-2 expression or activity can be identified by, for example, any or a combination of diagnostic or prognostic assays as described herein. Administration of a prophylactic agent can occur prior to the manifestation of symptoms characteristic of the TP-2 aberrancy, such that a disease or disorder is prevented or, alternatively, delayed in its progression. Depending on the type of TP-2 aberrancy, for example, a TP-2, TP-2 agonist or TP-2 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 methods of modulating TP-2 expression or activity for therapeutic purposes. Accordingly, in an exemplary embodiment, the modulatory method of the invention involves contacting a cell capable of expressing TP-2 with an agent that modulates one or more of the activities of TP-2 polypeptide activity associated with the cell, such that TP-2 activity in the cell is modulated. An agent that modulates TP-2 polypeptide activity can be an agent as described herein, such as a nucleic acid or a polypeptide, a naturally-occurring target molecule of a TP-2 polypeptide (*e.g.*, a TP-2 substrate), a TP-2 antibody, a TP-2 agonist or antagonist, a peptidomimetic of a TP-2 agonist or antagonist, or other small molecule. In one embodiment, the agent stimulates one or more TP-2 activities. Examples of such stimulatory agents include active TP-2 polypeptide and a nucleic acid molecule encoding TP-2 that has been introduced into the cell. In another embodiment, the agent inhibits one or more TP-2 activities. Examples of such inhibitory agents include antisense TP-2 nucleic acid molecules, anti-TP-2 antibodies, and TP-2 inhibitors. These modulatory methods can be performed *in vitro* (*e.g.*, by culturing the cell with the agent) or, alternatively, *in vivo* (*e.g.*, by administering the agent to a subject). As such, the present invention provides methods of treating an individual afflicted with a disease or disorder characterized by aberrant or unwanted expression or activity of a TP-2 polypeptide or nucleic acid molecule. In one embodiment, the method involves administering an agent (*e.g.*, an agent identified by a screening assay described herein), or combination of agents that modulates (*e.g.*, upregulates or downregulates) TP-2 expression or activity. In another embodiment, the method involves administering a TP-2 polypeptide or nucleic acid molecule as therapy to compensate for reduced, aberrant, or unwanted TP-2 expression or activity.

Stimulation of TP-2 activity is desirable in situations in which TP-2 is abnormally downregulated and/or in which increased TP-2 activity is likely to have a beneficial effect. Likewise, inhibition of TP-2 activity is desirable in situations in which TP-2 is abnormally upregulated and/or in which decreased TP-2 activity is likely to have a beneficial effect.

3. Pharmacogenomics

The TP-2 molecules of the present invention, as well as agents, or modulators which have a stimulatory or inhibitory effect on TP-2 activity (*e.g.*, TP-2 gene expression) as identified by a screening assay described herein can be administered to individuals to treat (prophylactically or therapeutically) transporter-associated disorders (*e.g.*, proliferative disorders) associated with aberrant or unwanted TP-2 activity. In conjunction with such treatment, pharmacogenomics (*i.e.*, the study of the relationship

between an individual's genotype and that individual's response to a foreign compound or drug) may be considered. Differences in metabolism of therapeutics can lead to severe toxicity or therapeutic failure by altering the relation between dose and blood concentration of the pharmacologically active drug. Thus, a physician or clinician may consider applying knowledge obtained in relevant pharmacogenomics studies in determining whether to administer a TP-2 molecule or TP-2 modulator as well as tailoring the dosage and/or therapeutic regimen of treatment with a TP-2 molecule or TP-2 modulator.

Pharmacogenomics deals with clinically significant hereditary variations in the response to drugs due to altered drug disposition and abnormal action in affected persons. See, for example, Eichelbaum, M. *et al.* (1996) *Clin. Exp. Pharmacol. Physiol.* 23(10-11): 983-985 and Linder, M.W. *et al.* (1997) *Clin. Chem.* 43(2):254-266. In general, two types of pharmacogenetic conditions can be differentiated. Genetic conditions transmitted as a single factor altering the way drugs act on the body (altered drug action) or genetic conditions transmitted as single factors altering the way the body acts on drugs (altered drug metabolism). These pharmacogenetic conditions can occur either as rare genetic defects or as naturally-occurring polymorphisms. For example, glucose-6-phosphate dehydrogenase deficiency (G6PD) is a common inherited enzymopathy in which the main clinical complication is haemolysis after ingestion of oxidant drugs (anti-malarials, sulfonamides, analgesics, nitrofurans) and consumption of fava beans.

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

genetically similar individuals, taking into account traits that may be common among such genetically similar individuals.

Alternatively, a method termed the "candidate gene approach", can be utilized to identify genes that predict drug response. According to this method, if a gene that encodes a drug target is known (*e.g.*, a TP-2 polypeptide of the present invention), all common variants of that gene can be fairly easily identified in the population and it can be determined if having one version of the gene versus another is associated with a particular drug response.

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

Alternatively, a method termed the "gene expression profiling", can be utilized to identify genes that predict drug response. For example, the gene expression of an animal dosed with a drug (*e.g.*, a TP-2 molecule or TP-2 modulator of the present invention) can give an indication whether gene pathways related to toxicity have been turned on.

Information generated from more than one of the above pharmacogenomics approaches can be used to determine appropriate dosage and treatment regimens for prophylactic or therapeutic treatment an individual. This knowledge, when applied to dosing or drug selection, can avoid adverse reactions or therapeutic failure and thus enhance therapeutic or prophylactic efficiency when treating a subject with a TP-2 molecule or TP-2 modulator, such as a modulator identified by one of the exemplary screening assays described herein.

4. Use of TP-2 Molecules as Surrogate Markers

The TP-2 molecules of the invention are also useful as markers of disorders or disease states, as markers for precursors of disease states, as markers for predisposition of disease states, as markers of drug activity, or as markers of the pharmacogenomic profile of a subject. Using the methods described herein, the presence, absence and/or quantity of the TP-2 molecules of the invention may be detected, and may be correlated with one or more biological states *in vivo*. For example, the TP-2 molecules of the invention may serve as surrogate markers for one or more disorders or disease states or for conditions leading up to disease states. As used herein, a “surrogate marker” is an objective biochemical marker which correlates with the absence or presence of a disease or disorder, or with the progression of a disease or disorder (*e.g.*, with the presence or absence of a tumor). The presence or quantity of such markers is independent of the disease. Therefore, these markers may serve to indicate whether a particular course of treatment is effective in lessening a disease state or disorder. Surrogate markers are of particular use when the presence or extent of a disease state or disorder is difficult to assess through standard methodologies (*e.g.*, early stage tumors), or when an assessment of disease progression is desired before a potentially dangerous clinical endpoint is reached (*e.g.*, an assessment of cardiovascular disease may be made using cholesterol levels as a surrogate marker, and an analysis of HIV infection may be made using HIV RNA levels as a surrogate marker, well in advance of the undesirable clinical outcomes of myocardial infarction or fully-developed AIDS). Examples of the use of surrogate markers in the art include: Koomen *et al.* (2000) *J. Mass. Spectrom.* 35: 258-264; and James (1994) *AIDS Treatment News Archive* 209.

The TP-2 molecules of the invention are also useful as pharmacodynamic markers. As used herein, a “pharmacodynamic marker” is an objective biochemical marker which correlates specifically with drug effects. The presence or quantity of a pharmacodynamic marker is not related to the disease state or disorder for which the drug is being administered; therefore, the presence or quantity of the marker is indicative of the presence or activity of the drug in a subject. For example, a pharmacodynamic marker may be indicative of the concentration of the drug in a biological tissue, in that the marker is either expressed or transcribed or not expressed or transcribed in that tissue in relationship to the level of the drug. In this fashion, the distribution or uptake of the drug may be monitored by the pharmacodynamic marker. Similarly, the presence or quantity of the pharmacodynamic marker may be related to the presence or quantity of the metabolic product of a drug, such that the presence or quantity of the marker is indicative of the relative breakdown rate of the drug *in vivo*. Pharmacodynamic markers are of particular use in increasing the sensitivity of detection of drug effects, particularly when the drug is administered in low doses. Since even a

small amount of a drug may be sufficient to activate multiple rounds of marker (*e.g.*, a TP-2 marker) transcription or expression, the amplified marker may be in a quantity which is more readily detectable than the drug itself. Also, the marker may be more easily detected due to the nature of the marker itself; for example, using the methods
5 described herein, anti-TP-2 antibodies may be employed in an immune-based detection system for a TP-2 polypeptide marker, or TP-2-specific radiolabeled probes may be used to detect a TP-2 mRNA marker. Furthermore, the use of a pharmacodynamic marker may offer mechanism-based prediction of risk due to drug treatment beyond the range of possible direct observations. Examples of the use of pharmacodynamic markers in the
10 art include: Matsuda *et al.* US 6,033,862; Hattis *et al.* (1991) *Env. Health Perspect.* 90: 229-238; Schentag (1999) *Am. J. Health-Syst. Pharm.* 56 Suppl. 3: S21-S24; and Nicolau (1999) *Am. J. Health-Syst. Pharm.* 56 Suppl. 3: S16-S20.

The TP-2 molecules of the invention are also useful as pharmacogenomic markers. As used herein, a “pharmacogenomic marker” is an objective biochemical
15 marker which correlates with a specific clinical drug response or susceptibility in a subject (see, *e.g.*, McLeod *et al.* (1999) *Eur. J. Cancer* 35(12): 1650-1652). The presence or quantity of the pharmacogenomic marker is related to the predicted response of the subject to a specific drug or class of drugs prior to administration of the drug. By assessing the presence or quantity of one or more pharmacogenomic markers in a
20 subject, a drug therapy which is most appropriate for the subject, or which is predicted to have a greater degree of success, may be selected. For example, based on the presence or quantity of RNA, or polypeptide (*e.g.*, TP-2 polypeptide or RNA) for specific tumor markers in a subject, a drug or course of treatment may be selected that is optimized for the treatment of the specific tumor likely to be present in the subject.
25 Similarly, the presence or absence of a specific sequence mutation in TP-2 DNA may correlate TP-2 drug response. The use of pharmacogenomic markers therefore permits the application of the most appropriate treatment for each subject without having to administer the therapy.

30 VI. Electronic Apparatus Readable Media and Arrays

Electronic apparatus readable media comprising TP-2 sequence information is also provided. As used herein, “TP-2 sequence information” refers to any nucleotide and/or amino acid sequence information particular to the TP-2 molecules of the present invention, including but not limited to full-length nucleotide and/or amino acid
35 sequences, partial nucleotide and/or amino acid sequences, polymorphic sequences including single nucleotide polymorphisms (SNPs), epitope sequences, and the like. Moreover, information “related to” said TP-2 sequence information includes detection of the presence or absence of a sequence (*e.g.*, detection of expression of a sequence,

fragment, polymorphism, etc.), determination of the level of a sequence (*e.g.*, detection of a level of expression, for example, a quantitative detection), detection of a reactivity to a sequence (*e.g.*, detection of protein expression and/or levels, for example, using a sequence-specific antibody), and the like. As used herein, “electronic apparatus
5 readable media” refers to any suitable medium for storing, holding or containing data or information that can be read and accessed directly by an electronic apparatus. Such media can include, but are not limited to: magnetic storage media, such as floppy discs, hard disc storage medium, and magnetic tape; optical storage media such as compact disc; electronic storage media such as RAM, ROM, EPROM, EEPROM and the like;
10 general hard disks and hybrids of these categories such as magnetic/optical storage media. The medium is adapted or configured for having recorded thereon TP-2 sequence information of the present invention.

As used herein, the term “electronic apparatus” is intended to include any suitable computing or processing apparatus or other device configured or adapted for
15 storing data or information. Examples of electronic apparatus suitable for use with the present invention include stand-alone computing apparatus; networks, including a local area network (LAN), a wide area network (WAN) Internet, Intranet, and Extranet; electronic appliances such as a personal digital assistants (PDAs), cellular phone, pager and the like; and local and distributed processing systems.

20 As used herein, “recorded” refers to a process for storing or encoding information on the electronic apparatus readable medium. Those skilled in the art can readily adopt any of the presently known methods for recording information on known media to generate manufactures comprising the TP-2 sequence information.

A variety of software programs and formats can be used to store the sequence
25 information on the electronic apparatus readable medium. For example, the sequence information can be represented in a word processing text file, formatted in commercially-available software such as WordPerfect and MicroSoft Word, or represented in the form of an ASCII file, stored in a database application, such as DB2, Sybase, Oracle, or the like, as well as in other forms. Any number of data processor
30 structuring formats (*e.g.*, text file or database) may be employed in order to obtain or create a medium having recorded thereon the TP-2 sequence information.

By providing TP-2 sequence information in readable form, one can routinely access the sequence information for a variety of purposes. For example, one skilled in the art can use the sequence information in readable form to compare a target sequence
35 or target structural motif with the sequence information stored within the data storage means. Search means are used to identify fragments or regions of the sequences of the invention which match a particular target sequence or target motif.

The present invention therefore provides a medium for holding instructions for performing a method for determining whether a subject has a TP-2- associated disease or disorder or a pre-disposition to a TP-2-associated disease or disorder, wherein the method comprises the steps of determining TP-2 sequence information associated with the subject and based on the TP-2 sequence information, determining whether the subject has a TP-2 -associated disease or disorder or a pre-disposition to a TP-2-associated disease or disorder and/or recommending a particular treatment for the disease, disorder or pre-disease condition.

The present invention further provides in an electronic system and/or in a network, a method for determining whether a subject has a TP-2-associated disease or disorder or a pre-disposition to a disease associated with a TP-2 wherein the method comprises the steps of determining TP-2 sequence information associated with the subject, and based on the TP-2 sequence information, determining whether the subject has a TP-2 -associated disease or disorder or a pre-disposition to a TP-2-associated disease or disorder, and/or recommending a particular treatment for the disease, disorder or pre-disease condition. The method may further comprise the step of receiving phenotypic information associated with the subject and/or acquiring from a network phenotypic information associated with the subject.

The present invention also provides in a network, a method for determining whether a subject has a TP-2 -associated disease or disorder or a pre-disposition to a TP-2 -associated disease or disorder associated with TP-2, said method comprising the steps of receiving TP-2 sequence information from the subject and/or information related thereto, receiving phenotypic information associated with the subject, acquiring information from the network corresponding to TP-2 and/or a TP-2-associated disease or disorder, and based on one or more of the phenotypic information, the TP-2 information (e.g., sequence information and/or information related thereto), and the acquired information, determining whether the subject has a TP-2-associated disease or disorder or a pre-disposition to a TP-2-associated disease or disorder. The method may further comprise the step of recommending a particular treatment for the disease, disorder or pre-disease condition.

The present invention also provides a business method for determining whether a subject has a TP-2-associated disease or disorder or a pre-disposition to a TP-2-associated disease or disorder, said method comprising the steps of receiving information related to TP-2 (e.g., sequence information and/or information related thereto), receiving phenotypic information associated with the subject, acquiring information from the network related to TP-2 and/or related to a TP-2-associated disease or disorder, and based on one or more of the phenotypic information, the TP-2 information, and the acquired information, determining whether the subject has a TP-2-

associated disease or disorder or a pre-disposition to a TP-2-associated disease or disorder. The method may further comprise the step of recommending a particular treatment for the disease, disorder or pre-disease condition.

5 The invention also includes an array comprising a TP-2 sequence of the present invention. The array can be used to assay expression of one or more genes in the array. In one embodiment, the array can be used to assay gene expression in a tissue to ascertain tissue specificity of genes in the array. In this manner, up to about 7600 genes can be simultaneously assayed for expression, one of which can be TP-2. This allows a profile to be developed showing a battery of genes specifically expressed in one or more
10 tissues.

In addition to such qualitative determination, the invention allows the quantitation of gene expression. Thus, not only tissue specificity, but also the level of expression of a battery of genes in the tissue is ascertainable. Thus, genes can be grouped on the basis of their tissue expression *per se* and level of expression in that
15 tissue. This is useful, for example, in ascertaining the relationship of gene expression between or among tissues. Thus, one tissue can be perturbed and the effect on gene expression in a second tissue can be determined. In this context, the effect of one cell type on another cell type in response to a biological stimulus can be determined. Such a determination is useful, for example, to know the effect of cell-cell interaction at the
20 level of gene expression. If an agent is administered therapeutically to treat one cell type but has an undesirable effect on another cell type, the invention provides an assay to determine the molecular basis of the undesirable effect and thus provides the opportunity to co-administer a counteracting agent or otherwise treat the undesired effect. Similarly, even within a single cell type, undesirable biological effects can be
25 determined at the molecular level. Thus, the effects of an agent on expression of other than the target gene can be ascertained and counteracted.

In another embodiment, the array can be used to monitor the time course of expression of one or more genes in the array. This can occur in various biological contexts, as disclosed herein, for example development of a TP-2-associated disease or
30 disorder, progression of TP-2-associated disease or disorder, and processes, such a cellular transformation associated with the TP-2-associated disease or disorder.

The array is also useful for ascertaining the effect of the expression of a gene on the expression of other genes in the same cell or in different cells (*e.g.*, ascertaining the effect of TP-2 expression on the expression of other genes). This provides, for example,
35 for a selection of alternate molecular targets for therapeutic intervention if the ultimate or downstream target cannot be regulated.

The array is also useful for ascertaining differential expression patterns of one or more genes in normal and abnormal cells. This provides a battery of genes (*e.g.*, including TP-2) that could serve as a molecular target for diagnosis or therapeutic intervention.

5

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

10

EXAMPLES

EXAMPLE 1: IDENTIFICATION AND CHARACTERIZATION OF HUMAN TP-2 cDNA

15

In this example, the identification and characterization of the gene encoding human TP-2 (clone 63760) is described.

Isolation of the human TP-2 cDNA

The invention is based, at least in part, on the discovery of a human gene encoding a novel polypeptide, referred to herein as human TP-2. The entire sequence of the human clone 63760 was determined and found to contain an open reading frame termed human "TP-2." The nucleotide sequence of the human TP-2 gene is set forth in Figures 1A-B and in the Sequence Listing as SEQ ID NO:1. The amino acid sequence of the human TP-2 expression product is set forth in Figures 1A-B and in the Sequence Listing as SEQ ID NO: 2. The TP-2 polypeptide comprises 474 amino acids. The coding region (open reading frame) of SEQ ID NO:1 is set forth as SEQ ID NO:3. Clone 63760, comprising the coding region of human TP-2, was deposited with the American Type Culture Collection (ATCC[®]), 10801 University Boulevard, Manassas, VA 20110-2209, on _____, and assigned Accession No. _____.

Analysis of the Human TP-2 Molecules

The human TP-2 amino acid sequence was aligned with the amino acid sequence of the tetracycline-6-hydroxylase/oxygenase homolog gene from *Salmonella typhi* (SEQ ID NO:4) using the CLUSTAL W (1.74) multiple sequence alignment program. The results of the alignment are set forth in Figure 5.

A search using the polypeptide sequence of SEQ ID NO:2 was performed against the HMM database in PFAM (Figures 3A-C) resulting in the identification of a potential sugar transporter domain in the amino acid sequence of human TP-2 at about residues 37-454 of SEQ ID NO:2 (score = -101.1), a potential LacY proton/sugar symporter domain in the amino acid sequence of human TP-2 at about residues 39-383 of SEQ ID NO:2 (score = -336.7), a potential glutamine amidotransferases class-II domain in the amino acid sequence of human TP-2 at about residues 165-170 of SEQ ID NO:2 (score = 1.2), and a potential MCT domain in the amino acid sequence of human TP-2 at about residues 33-458 of SEQ ID NO:2 (score = -167.8).

10 The amino acid sequence of human TP-2 was analyzed using the program PSORT (<http://www.psort.nibb.ac.jp>) to predict the localization of the proteins within the cell. This program assesses the presence of different targeting and localization amino acid sequences within the query sequence. The results of this analysis show that human TP-2 may be localized to the endoplasmic reticulum, secretory vesicles, or
15 mitochondria.

Searches of the amino acid sequence of human TP-2 were further performed against the Prosite database. These searches resulted in the identification in the amino acid sequence of human TP-2 of two potential glycosaminoglycan attachment sites at about residues 176-179 and 464-467 of SEQ ID NO:2, two potential cAMP- and cGMP-
20 dependent protein kinase phosphorylation sites at about residues 108-111 and 460-463 of SEQ ID NO:2, a number of potential protein kinase C phosphorylation sites at about residues 228-230, 253-255, and 260-262 of SEQ ID NO:2, a number of potential casein kinase II phosphorylation sites at about residues 28-31, 191-194, 247-250, and 463-466 of SEQ ID NO:2, a number of potential N-myristoylation sites at about residues 38-43,
25 75-80, 82-87, 127-132, 187-192, 332-337, 403-408, 409-414, 415-420, and 445-450 of SEQ ID NO:2, one potential amidation site at about residues 106-109 of SEQ ID NO:2, and a potential prokaryotic membrane lipoprotein lipid attachment site at about residues 99-114 of SEQ ID NO:2.

A MEMSAT analysis of the polypeptide sequence of SEQ ID NO:2 was also
30 performed (Figure 4), predicting eleven potential transmembrane domains in the amino acid sequence of human TR-2 (SEQ ID NO:2) at about residues 45-69, 80-102, 112-136, 167-190, 197-218, 288-310, 323-343, 352-368, 375-391, 409-433, and 422-458. However, a structural, hydrophobicity, and antigenicity analysis (Figure 2) resulted in the identification of twelve transmembrane domains. Accordingly, the TP-2 protein of
35 SEQ ID NO:2 is predicted to have at least 12 transmembrane domains, which are identified in Figures 2 and 4 as transmembrane (TM) domains 1 through 12. TM1 is at about residues 45-69, TM2 is at about residues 80-102, TM3 is at about residues 112-126, TM4 is at about residues 133-156, TM5 is at about residues 167-190, TM6 is at

about residues 197-218, TM7 is at about residues 288-310, TM8 is at about residues 323-343, TM9 is at about residues 352-368, TM10 is at about residues 375-391, TM11 is at about residues 409-433, and TM12 is at about residues 442-458.

A search of the amino acid sequence of human TP-2 was also performed against the ProDom database. These searches resulted in the identification of a “kinase activity integral membrane domain” at about amino acid residues 36-235, a “transport integral membrane” at about amino acid residues 41-190, a “YFKF transporter MFS transmembrane domain” at about amino acid residues 45-229, a “multidrug transmembrane domain” at about amino acid residues 130-250, a “transporter-like polyspecific organic subtransferable suppressing membrane tumor domain” at about amino acid residues 133-214, a “transport membrane domain” at about amino acid residues 163-244, a “NORA domain” at about amino acid residues 190-462, and a “family C2-domain” at about amino acid residues 399-466 in the amino acid protein sequence of TP-2 (SEQ ID NO:2).

15

EXAMPLE 2: EXPRESSION OF RECOMBINANT TP-2 POLYPEPTIDE IN BACTERIAL CELLS

In this example, human TP-2 is expressed as a recombinant glutathione-S-transferase (GST) fusion polypeptide in *E. coli* and the fusion polypeptide is isolated and characterized. Specifically, TP-2 is fused to GST and this fusion polypeptide is expressed in *E. coli*, e.g., strain PEB199. Expression of the GST-TP-2 fusion polypeptide in PEB199 is induced with IPTG. The recombinant fusion polypeptide is purified from crude bacterial lysates of the induced PEB199 strain by affinity chromatography on glutathione beads. Using polyacrylamide gel electrophoretic analysis of the polypeptide purified from the bacterial lysates, the molecular weight of the resultant fusion polypeptide is determined.

25

EXAMPLE 3: EXPRESSION OF RECOMBINANT TP-2 POLYPEPTIDE IN COS CELLS

To express the human TP-2 gene in COS cells, the pcDNA/Amp vector by Invitrogen Corporation (San Diego, CA) is used. This vector contains an SV40 origin of replication, an ampicillin resistance gene, an *E. coli* replication origin, a CMV promoter followed by a polylinker region, and an SV40 intron and polyadenylation site. A DNA fragment encoding the entire TP-2 polypeptide and an HA tag (Wilson *et al.* (1984) *Cell* 37:767) or a FLAG tag fused in-frame to its 3' end of the fragment is cloned into the

35

polylinker region of the vector, thereby placing the expression of the recombinant polypeptide under the control of the CMV promoter.

To construct the plasmid, the human TP-2 DNA sequence is amplified by PCR using two primers. The 5' primer contains the restriction site of interest followed by approximately twenty nucleotides of the TP-2 coding sequence starting from the initiation codon; the 3' end sequence contains complementary sequences to the other restriction site of interest, a translation stop codon, the HA tag or FLAG tag and the last 20 nucleotides of the TP-2 coding sequence. The PCR amplified fragment and the pCDNA/Amp vector are digested with the appropriate restriction enzymes and the vector is dephosphorylated using the CIAP enzyme (New England Biolabs, Beverly, MA). Preferably the two restriction sites chosen are different so that the TP-2 gene is inserted in the correct orientation. The ligation mixture is transformed into *E. coli* cells (strains HB101, DH5 α , SURE, available from Stratagene Cloning Systems, La Jolla, CA, can be used), the transformed culture is plated on ampicillin media plates, and resistant colonies are selected. Plasmid DNA is isolated from transformants and examined by restriction analysis for the presence of the correct fragment.

COS cells are subsequently transfected with the human TP-2-pcDNA/Amp plasmid DNA using the calcium phosphate or calcium chloride co-precipitation methods, DEAE-dextran-mediated transfection, lipofection, or electroporation. Other suitable methods for transfecting host cells can be found in Sambrook, J., Fritsh, E. F., and Maniatis, T. *Molecular Cloning: A Laboratory Manual*. 2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989. The expression of the IC54420 polypeptide is detected by radiolabelling (^{35}S -methionine or ^{35}S -cysteine available from NEN, Boston, MA, can be used) and immunoprecipitation (Harlow, E. and Lane, D. *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1988) using an HA specific monoclonal antibody. Briefly, the cells are labeled for 8 hours with ^{35}S -methionine (or ^{35}S -cysteine). The culture media are then collected and the cells are lysed using detergents (RIPA buffer, 150 mM NaCl, 1% NP-40, 0.1% SDS, 0.5% DOC, 50 mM Tris, pH 7.5). Both the cell lysate and the culture media are precipitated with an HA specific monoclonal antibody. Precipitated polypeptides are then analyzed by SDS-PAGE.

Alternatively, DNA containing the human TP-2 coding sequence is cloned directly into the polylinker of the pCDNA/Amp vector using the appropriate restriction sites. The resulting plasmid is transfected into COS cells in the manner described above, and the expression of the TP-2 polypeptide is detected by radiolabelling and immunoprecipitation using a TP-2-specific monoclonal antibody.

Equivalents

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

What is claimed:

1. An isolated nucleic acid molecule selected from the group consisting of:
 - 5 (a) a nucleic acid molecule comprising the nucleotide sequence set forth in SEQ ID NO:1; and
 - (b) a nucleic acid molecule comprising the nucleotide sequence set forth in SEQ ID NO:3.
- 10 2. An isolated nucleic acid molecule which encodes a polypeptide comprising the amino acid sequence set forth in SEQ ID NO:2.
3. An isolated nucleic acid molecule comprising the nucleotide sequence contained in the plasmid deposited with ATCC® as Accession Number _____.
15
4. An isolated nucleic acid molecule which encodes a naturally occurring allelic variant of a polypeptide comprising the amino acid sequence set forth in SEQ ID NO:2.
- 20 5. An isolated nucleic acid molecule selected from the group consisting of:
 - a) a nucleic acid molecule comprising a nucleotide sequence which is at least 60% identical to the nucleotide sequence of SEQ ID NO:1 or 3, or a complement thereof;
 - 25 b) a nucleic acid molecule comprising a fragment of at least 30 nucleotides of a nucleic acid comprising the nucleotide sequence of SEQ ID NO:1 or 3, or a complement thereof;
 - c) a nucleic acid molecule which encodes a polypeptide comprising an amino acid sequence at least about 60% identical to the amino acid sequence of SEQ
30 ID NO:2; and
 - d) a nucleic acid molecule which encodes a fragment of a polypeptide comprising the amino acid sequence of SEQ ID NO:2, wherein the fragment comprises at least 10 contiguous amino acid residues of the amino acid sequence of SEQ
35 ID NO:2.
6. An isolated nucleic acid molecule which hybridizes to a complement of the nucleic acid molecule of any one of claims 1, 2, 3, 4, or 5 under stringent conditions.

7. An isolated nucleic acid molecule comprising a nucleotide sequence which is complementary to the nucleotide sequence of the nucleic acid molecule of any one of claims 1, 2, 3, 4, or 5.
- 5 8. An isolated nucleic acid molecule comprising the nucleic acid molecule of any one of claims 1, 2, 3, 4, or 5, and a nucleotide sequence encoding a heterologous polypeptide.
9. A vector comprising the nucleic acid molecule of any one of claims 1, 2,
10 3, 4, or 5.
10. The vector of claim 9, which is an expression vector.
11. A host cell transfected with the expression vector of claim 10.
15
12. A method of producing a polypeptide comprising culturing the host cell of claim 11 in an appropriate culture medium to, thereby, produce the polypeptide.
13. An isolated polypeptide selected from the group consisting of:
20
- a) a fragment of a polypeptide comprising the amino acid sequence of SEQ ID NO:2, wherein the fragment comprises at least 10 contiguous amino acids of SEQ ID NO:2;
 - b) a naturally occurring allelic variant of a polypeptide comprising
25 the amino acid sequence of SEQ ID NO:2, wherein the polypeptide is encoded by a nucleic acid molecule which hybridizes to a complement of a nucleic acid molecule consisting of SEQ ID NO:1 or 3 under stringent conditions;
 - c) a polypeptide which is encoded by a nucleic acid molecule comprising a nucleotide sequence which is at least 60% identical to a nucleic acid
30 comprising the nucleotide sequence of SEQ ID NO:1 or 3; and
 - d) a polypeptide comprising an amino acid sequence which is at least 60% identical to the amino acid sequence of SEQ ID NO:2.
14. The isolated polypeptide of claim 13 comprising the amino acid sequence
35 of SEQ ID NO:2.
15. The polypeptide of claim 13, further comprising heterologous amino acid sequences.

16. An antibody which selectively binds to a polypeptide of claim 13.
17. A method for detecting the presence of a polypeptide of claim 13 in a
5 sample comprising:
a) contacting the sample with a compound which selectively binds to
the polypeptide; and
b) determining whether the compound binds to the polypeptide in
the sample to thereby detect the presence of a polypeptide of claim 13 in the sample.
10
18. The method of claim 17, wherein the compound which binds to the
polypeptide is an antibody.
19. A kit comprising a compound which selectively binds to a polypeptide of
15 claim 13 and instructions for use.
20. A method for detecting the presence of a nucleic acid molecule of any
one of claims 1, 2, 3, 4, or 5 in a sample comprising:
a) contacting the sample with a nucleic acid probe or primer which
20 selectively hybridizes to a complement of the nucleic acid molecule; and
b) determining whether the nucleic acid probe or primer binds to the
complement of the nucleic acid molecule in the sample to thereby detect the presence of
the nucleic acid molecule of any one of claims 1, 2, 3, 4, or 5 in the sample.
21. The method of claim 20, wherein the sample comprises mRNA
25 molecules and is contacted with a nucleic acid probe.
22. A kit comprising a compound which selectively hybridizes to a
complement of the nucleic acid molecule of any one of claims 1, 2, 3, 4, or 5 and
30 instructions for use.
23. A method for identifying a compound which binds to a polypeptide of
claim 13 comprising:
a) contacting the polypeptide, or a cell expressing the polypeptide
35 with a test compound; and
b) determining whether the polypeptide binds to the test compound.

24. The method of claim 23, 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 detection of test compound/polypeptide binding;
- 5 b) detection of binding using a competition binding assay; and
- c) detection of binding using an assay for TP-2 activity.

25. A method for modulating the activity of a polypeptide of claim 13 comprising contacting the polypeptide or a cell expressing the polypeptide with a
10 compound which binds to the polypeptide in a sufficient concentration to modulate the activity of the polypeptide.

26. A method for identifying a compound which modulates the activity of a polypeptide of claim 13 comprising:
15 a) contacting a polypeptide of claim 13 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.

20

Input file Fbh63760FL.seq; Output File Fbh63760FL.tra
 Sequence length 1963

GTCGACCCACGCGTCCGGGGTAGGTCCGCGGCTGCTGCTGCGTCCGACCCCGGCCGGCGCGGGT	M E L	3
ATG GAG CTT		9
G G H W D M N S A P R L V S E T A E R K		23
GGG GGT CAC TGG GAC ATG AAC TCG GCC CCG AGG CTG GTC TCG GAG ACC GCA GAG CGG AAA		69
Q E Q K T G T E A E A A D S G A V G A R		43
CAG GAG CAG AAG ACA GGA ACC GAG GCG GAG GCT GCC GAC TCC GGT GCC GTC GGA GCC CGC		129
R F L L C L Y L V G F L D L F G V S M V		63
CGC TTC CTG CTC TGT CTC TAC TTG GTG GGC TTC TTG GAT TTG TTT GGT GTC AGC ATG GTT		189
V P L L S L H V K S L G A S P T V A G I		83
GTG CCT TTA TTG AGC CTT CAT GTC AAG TCC CTT GGA GCA AGT CCA ACA GTT GCT GGA ATA		249
V G S S Y G I L Q L F S S T L V G C W S		103
GTA GGC TCC TCC TAT GGC ATT TTG CAA CTC TTT TCT AGC ACA TTG GTG GGC TGC TGG AGC		309
D V V G R R S S L L A C I L L S A L G Y		123
GAT GTA GTG GGA AGA CGG TCT TCC TTG CTG GCA TGC ATT CTA CTC AGT GCT CTG GGC TAT		369
L L L G A A T N V F L F V L A R V P A G		143
CTC CTT CTC GGA GCA GCC ACC AAT GTG TTT CTG TTT GTC CTG GCT AGA GTC CCG GCA GGT		429
I F K H T L S I S R A L L S D V V P E K		163
ATT TTT AAA CAC ACT CTC TCC ATC TCA AGG GCT CTA CTT TCT GAT GTG GTT CCA GAG AAG		489
E R P L V I G H F N T A S G V G F I L G		183
GAA CGG CCG CTT GTA ATC GGA CAC TTC AAC ACA GCC TCC GGT GTG GGC TTC ATC TTG GGC		549
P V V G G Y L T E L E D G F Y L T A F I		203
CCC GTG GTC GGT GGC TAT CTC ACT GAA TTA GAG GAT GGG TTT TAT CTC ACA GCC TTC ATC		609
C F L V F I L N A G L V W F F P R R E A		223
TGC TTT TTG GTC TTC ATT CTC AAT GCT GGT CTC GTT TGG TTC TTT CCA CGG AGG GAA GCA		669
K P G S T E K G L P L R K T H V L L G R		243
AAA CCG GGC AGT ACA GAG AAG GGC CTG CCA TTG CGA AAG ACC CAT GTG CTG TTG GGA AGG		729
S H D T V Q E A A T S R R A R A S K K T		263
AGC CAT GAC ACA GTG CAG GAG GCA GCC ACC AGC CGC AGA GCC AGG GCC AGC AAG AAG ACT		789
A Q P W V E V V L A L R N M K N L L F S		283
GCC CAG CCC TGG GTC GAA GTA GTG TTG GCC TTG CGG AAC ATG AAG AAC CTG CTG TTT TCC		849
E M W D I F L V R L L M A M A V M L Y Y		303
GAA ATG TGG GAC ATA TTT CTG GTG CGC TTG CTG ATG GCC ATG GCA GTC ATG CTG TAC TAC		909
S N F V L A L E E R F G V R P K V T G Y		323
AGT AAC TTT GTC CTG GCC CTG GAG GAG CGC TTT GGG GTG CGG CCC AAG GTG ACA GGC TAC		969
L I S Y S S M L G A V A G L A L G P I L		343
CTC ATC AGT TAC AGC AGC ATG CTG GGG GCC GTG GCC GGC CTT GCC CTG GGG CCA ATC CTA		1029

Fig. 1A

R	L	Y	K	H	N	S	Q	A	L	L	L	H	S	S	I	L	T	C	T	363
CGG	CTG	TAC	AAG	CAC	AAC	TCG	CAG	GCA	CTG	CTG	CTG	CAT	TCC	AGC	ATA	CTC	ACC	TGC	ACA	1089
L	L	L	L	Y	S	L	A	P	T	M	G	A	V	V	L	S	S	T	L	383
CTG	CTG	CTG	CTC	TAC	TCC	TTG	GCC	CCC	ACC	ATG	GGT	GCA	GTT	GTC	CTC	TCC	TCC	ACT	CTC	1149
L	S	F	S	T	A	I	G	R	T	C	I	T	D	L	Q	L	T	V	G	403
CTG	TCC	TTC	TCC	ACT	GCC	ATT	GGC	AGG	ACG	TGC	ATC	ACG	GAC	CTC	CAG	CTG	ACT	GTG	GGC	1209
G	A	Q	A	S	G	T	L	I	G	V	G	Q	S	V	T	A	V	G	R	423
GGG	GCC	CAG	GCC	AGC	GGC	ACC	CTT	ATT	GGC	GTG	GGG	CAG	TCT	GTG	ACT	GCA	GTG	GGC	CGC	1269
I	I	A	P	L	L	S	G	V	A	Q	E	V	S	P	C	G	P	P	S	443
ATC	ATC	GCC	CCT	CTC	CTC	TCG	GGG	GTT	GCC	CAG	GAG	GTC	AGC	CCT	TGC	GGC	CCC	CCC	AGC	1329
L	G	A	V	L	A	L	V	A	I	F	I	M	S	L	N	K	R	H	S	463
CTG	GGT	GCT	GTG	TTA	GCC	TTA	GTG	GCC	ATT	TTC	ATA	ATG	TCT	CTA	AAC	AAG	CGA	CAC	TCT	1389
S	G	D	G	N	S	K	L	K	S	E	*									475
AGT	GGT	GAT	GGG	AAT	AGT	AAA	TTA	AAA	AGT	GAG	TAG									1425

ATGGATTTGGACAACATAAAGCAACAAAATTTGAGATGGTTGAATGAGGGCCGGAGGCCATGATGAAAAGGGCACTTTG
GAAAGGGTTGGGGTGAAGGGAAATATTTCCGGGTGGGTGTGAGCTGTTGGGCTTCCAGGTCAGCTCTTGGCCATGCAG
CCATGCCTGCAGGATGATCAGAAGTCACGGCACCTCATGGGAAGGTTAAGACTGGAGCAAAGCTTTTCCAAGTGTGACC
CAGAGACCAGTAGCATCAGCATCACCTGAGACCTGTGAACACTGAAGCTCCAGCTCAGACATGTTGGGGACCATTTTAA
TAAGATACCTAGCTGATTTTTCACAGCAAACCTTTGAAAACCCCTGGTCTAAGGGGTAGTATTTGTATCACTTATGGA
ATATAATCTCAGGGAAATTAATCTGCTCAATTGACAAAAAAGGGCGCCGC

Fig. 1B

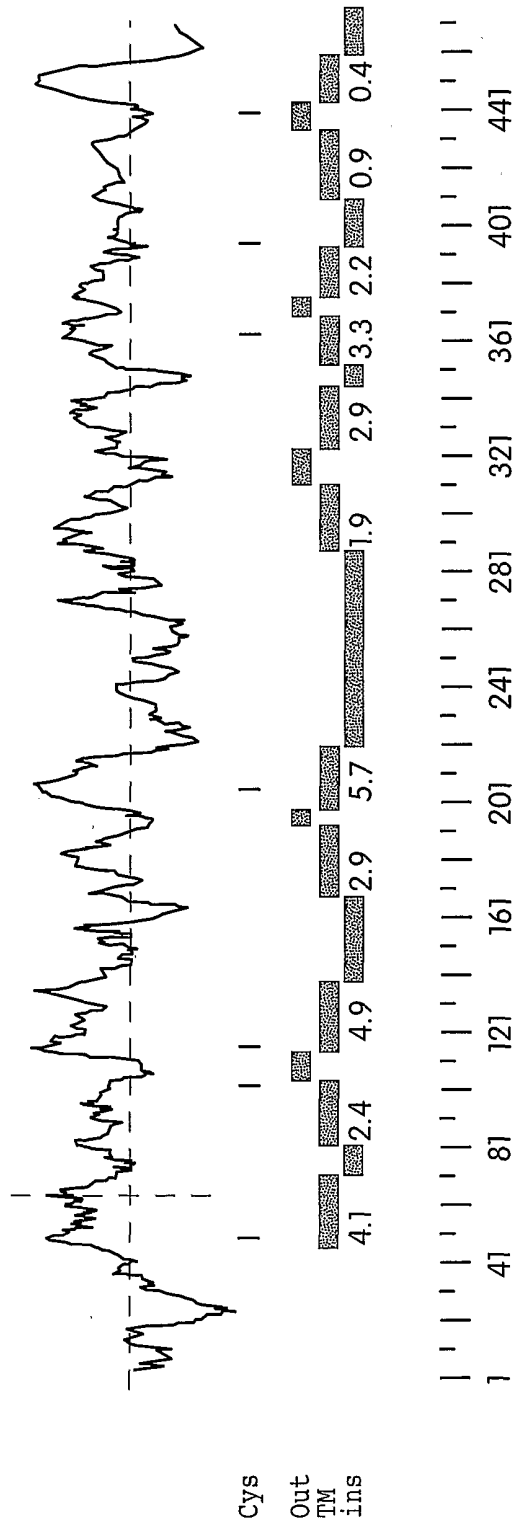


Fig. 2

Protein Family / Domain Matches, Hmmer version 2

Searching for complete domains in PFAM
 hmmpfam - search a single seq against HMM database
 HMMER 2.1.1 (Dec 1998)
 Copyright (C) 1992-1998 Washington University School of Medicine
 HMMER is freely distributed under the GNU General Public License (GPL).

HMM file: /prod/ddm/seqanal/PFAM/pfam6.4/Pfam
 Sequence file: /prod/ddm/wspace/orfanal/oa-script.27045.seq

Query: 63760

Scores for sequence family classification (score includes all domains):

Model	Description	Score	E-value	N
GATase_2	Glutamine amidotransferases class-II	1.2	8.6	1
sugar_tr	Sugar (and other) transporter	-103.0	0.0041	1
MCT	Monocarboxylate transporter	-167.8	0.014	1
LacY_symp	LacY proton/sugar symporter	-336.7	5.6	1

Parsed for domains:

Model	Domain	seq-f	seq-t	hmm-f	hmm-t	score	E-value
GATase_2	1/1	165	170	..	207 212 .]	1.2	8.6
LacY_symp	1/1	39	383	..	1 415 []	-336.7	5.6
sugar_tr	1/1	37	454	..	1 488 []	-103.1	0.0041
MCT	1/1	33	458	..	1 611 []	-167.8	0.014

Alignments of top-scoring domains:

GATase_2: domain 1 of 1, from 165 to 170: score 1.2, E = 8.6
 ->RPLviG<-
 RPLviG
 63760 165 RPLVIG 170

LacY_symp: domain 1 of 1, from 39 to 383: score -336.7, E = 5.6
 *->lipfkNtnFwrFglFFFFYFFISMSayFPFfPiWLkeVngLkTetG
 + ++ F +F +SM P++ k +g + T G
 63760 39 AVGARRFLCLYLVGFLDLFGVSMV--PLLSLHVKS-LGASPTVAG 82

Iv...FScisLFsIlFqPlfGlisDKLGLkKhLiWcI..sll..LVlFAP
 Iv++++ ++LFS l G sD G + L+ cI s l+ L l A
 63760 83 IVgssYGILQLFS---STLVGCWSDVVGRRSLLACI11SALgyLLLGAA 129

FFIYVFePLLQlNilaGalvGGvflglvysAGaGaiEAYIEkvSRnshFE
 ++ F + G+f ++a + + Ek
 63760 130 TNVFLF-----VLARVPAGIFKHTLSISRALLSDVVPEKERPLVIGH 171

YGkARmFGCvGwALCasiAGILFsIdPhivFWl.gSGfAlilLl...LLl
 + +A +G vG+ L+ G l F l++ + l++ l+ +L +
 63760 172 FNtA--SG-VGFILGPVGGYLTELEDG--FYLtAFICFLVFiLnaGLVW 216

lsKpdkshsAiVaDALGanksaF..SlrlaieLFkmrkfWvFvlyvGVa
 ++ ++a ++S+ + L k +
 63760 217 FFP-----RREAKpgSTEKGLPLRkTHVL----- 240

svYDVFdqQLFAvFFagFFespqvGTrvFGyvTTFGelLnAlIMFcaPFI
 + s++ ++ +T+ A+ P++
 63760 241 -----LGRSHDTVQEA---ATSRRA--RASKKTAQPWV 268

vNRIGaKNALLiAGvIMsvRilGsafattaleVvILKlLHafEvPFLlVG
 v+ ++R + ++ ++ + ++ lL a+ v L +
 63760 269 -----EVVLALRNmKNLLFSEMWDIFLVRLlMAMAV-MLYYS 304

Fig. 3A

```

vFKYitsnFdkRLS...ATifLIGFqfskqLaiv..lLStlaGkLYDhv
F   + + + R+   +++ T +LI   s+ L++v +l   +   +LY h
63760  305 NF---VLALEERFGvrpkVTGYLISY--SSMLGAVagLALGPILRLYKHN 349

GFQtaYlvLGiIvlsftlISiFTLSgsreqivlptpe<-*
+a l + I l+ tI + L   + vl ++
63760  350 --SQALLLHSSI-LTCTLLLLLYSLAPTMGAVVLSSTL 383

sugar_tr: domain 1 of 1, from 37 to 454: score -103.1, E = 0.0041
*->valvaalgGgflfGyDtgvigggflalidflfrfglltssgalaelvg
v a + +l+ y g++ f + + l+ + ++l++
63760  37 SGAVGARRF-LLCLYLVGFLDLFGVSMV---VPLLSLHVKSLGA--- 76

ystvltglvvsiffIgrliGslfaGklqdrfGRkksllialvlfviGall
s g+v s++ + l++s ++G ++d GR++sll++++l++ G ll
63760  77 -SPTVAGIVGSSYGILQLFSSTLVGCWSDVVGRRSLLACILLSALGYLL 125

sgaapgytTiGlwafyllivGRvIvGlgvGgasvIvPmYisEiAPkalRG
gaa+ ++ l+ ++Rv +G++ s +++ s + P++ R
63760  126 LGAAT-----NVFLFVLARVPAGIFKHLSISRALL-SDVVPEKERP 166

algslyqlaitiGilvAaiiglgnktnndsalnswgWRiplglqlvpal
++ +++a +G++++ +g g + + +g+ + +++++ +
63760  167 LVIGHFNTASGVGFILGPVVG-GYLTELE-----DGFYLTAIFICFLVFI 209

llligllflPESPRwLvekgkle.eArevLaklrgv...edvdqeigeik
l + +++f P ++ + ++e lr+++ + + +++
63760  210 LNAGLVWFFP-----RREAKpGSTEKGLPLRKThvllGRSHDTVQEA 251

aeleatvseekagkaswgelfrgrtrpkvrqrllmgvmlqafqQltGiNa
a++ +++ ++ ++ + ++ v+l++
63760  252 ATSRRARASK-----KTA-----QPWVEVVLALRN----- 276

ifYYsptifksvGvsdsvasllvtiivg.vvNfvfTfvaLiflvDrfGRR
+k++ +s+ + +lv ++++ +v+ + ++l +rfG R
63760  277 -----MKNLLFSEMWDIFLVRLLMamAVMLYYSNFV-LALEERFGVR 317

...pllllGaagmaicflilgasigvallllnkpk.dpsskaagivaivf
++ + l++ + m ++ + l+ +g+ l l + +++++ +i+ ++
63760  318 pkvTGYLISYSSMLGAVAGLA--LGPILRLYKH-NsQALLLHSSILTCTL 364

illfiafFalgwGpipwvilsElFPtkvRskalalataanwlanfiigfl
+ll+ a++ G + v++s l + ++ + + + t ++ + + + +
63760  365 LLLY--SLAPTMG--AVVLSSTLLSFST-AIGRTCITDLQLTVGGAQASG 409

fpyitgaiglalggyvflvfagllvlfilfvfffvPETkGrtLEeieelf
+ + g++ v+ ++a ll v + P G L ++ ++f
63760  410 TLIGVGQSV----TAVGRIIAPLLSGVAQEVSPCGPPSLGAVL-ALVAIF 454

<-*
63760 - -

```

```

MCT: domain 1 of 1, from 33 to 458: score -167.8, E = 0.014
*->kpPDGGwGWvVVFasFlingfvdGfiksfgvffse....llqeetlf
++ D G +V Fl Gf fGv + + +l +++
63760  33 EAADSG---AVGARRFLCLYLVGFLDLFGVSMVvllsLHVKS--L 74

nesksdvdAwIgsimlavllfsGPlsSilvnrfGcRivmiaGgllagaG
++s + +++++gS + + lfs+ l++ ++ G R + +a ll+++G
63760  75 GASPTV--AGIVGSSYGILQLFSSTLVGCWSDVVGRRSLLACILLSALG 122

```

Fig. 3B

```

lllasFstniwelyltfGvitGl...GfgfifqPaivilgqYFeKrRslA
+ll +tn+++++l+ v G+ ++ +++++ + +++ ek R+l+

63760 123 YLLLGAATNVFLFVLAR-VPAGIfkhTLSISRALLSDVVP---EKERPLV 168

tGiAvaGsGvGtvvfppllqflidnyGsDWrgallilggillncvicGal
G + sGvG+++ p + +l + ++ a++ + +++ln +G++
63760 169 IGHFNTASGVGFILGPVVGGYLLEDEGDFYLTAFICFLVFILN---AGLV 215

llRPlepsvpqdekdketlkeakkkkndtettkeetplkslpkasi
P+ +p +++++ +l++ ++ + + +t e +
63760 216 WFFPRREAKPGSTEKG--LPLRKTHVLLGRSHDTVQEAA----- 252

lkledakaersvdsLlsskSvgerdksqlsekqksqasgrpsssatavql

63760 - -----

vllrsrlekadlplkrvrsvrrrvlSkVSaeSgtdgersSgylNrkdvFY
+sr+ r+s+
63760 250 ---TSRR-----ARASK----- 261

tGsisNvaefkedpdkYrssslhgtrttvгнаesqstlrldsresgdg

63760 - -----

sssedlsektrgdggkessskeiretikllDfs.vlknrtFllyaisn
++ ++ e l +++++l+ ++++
63760 262 -----KTAQPWVEVVLALRNMKnLLFSEMWDIFLVR 293

lfaslGffvPlvfLvsYaikslgldekeAsfllsiiGvsnivGRpifGlv
l a+ ++ + f+ ++ + g+ k++ +L+s+ ++ v+ +++G +
63760 294 LMAMAVMLYYSNFVLALE-ERFGVRPKVTGYLISYSSMLGAVAGLALGPI 342

ADkkgvrpTarhivyifnslallGlttlacPlatsfwgLvvy CilFGfs
+++ + +++ ++s ++ ++l+ la + + v + l +fs
63760 343 LRLYKHNS-----QALLLHSSILTCTLLLLYSLAPTMGAVVLSSTLLSFS 387

iGsygaLtfvvLvdLvgWlekfsnAfGllllfeGvavLvGPPiaGllvDa
+ ++++ g + + +G+ + +v+ ++ P ++G +
63760 388 TAIGRTCITDLQlTVGG-AQASGTLIGVGQSVTAVGRIIAPLLSGVAQEV 436

kttgdYtvaFyfsGillllsgl<-*
g+ + ++++ +++++ +l
63760 437 SPCGPPSLGAVLALVAIFIMSL 458

```

Fig. 3C

Transmembrane Segments Predicted by MEMSAT

START	END	ORIENT	SCORE		
45	69	out-->ins	4.1	TM 1	
80	102	ins-->out	2.4	TM 2	TM 3 = 112-126
112	136	out-->ins	4.9	ND	TM 4 = 133-156
167	190	ins-->out	2.9	TM 5	
197	218	out-->ins	5.7	TM 6	
288	310	ins-->out	1.9	TM 7	
323	343	out-->ins	2.9	TM 8	
352	368	ins-->out	3.3	TM 9	
375	391	out-->ins	2.2	TM 10	
409	433	ins-->out	0.6	TM 11	
442	458	out-->ins	3.4	TM 12	

Fig. 4

LUSTAL W (1.74) multiple sequence alignment

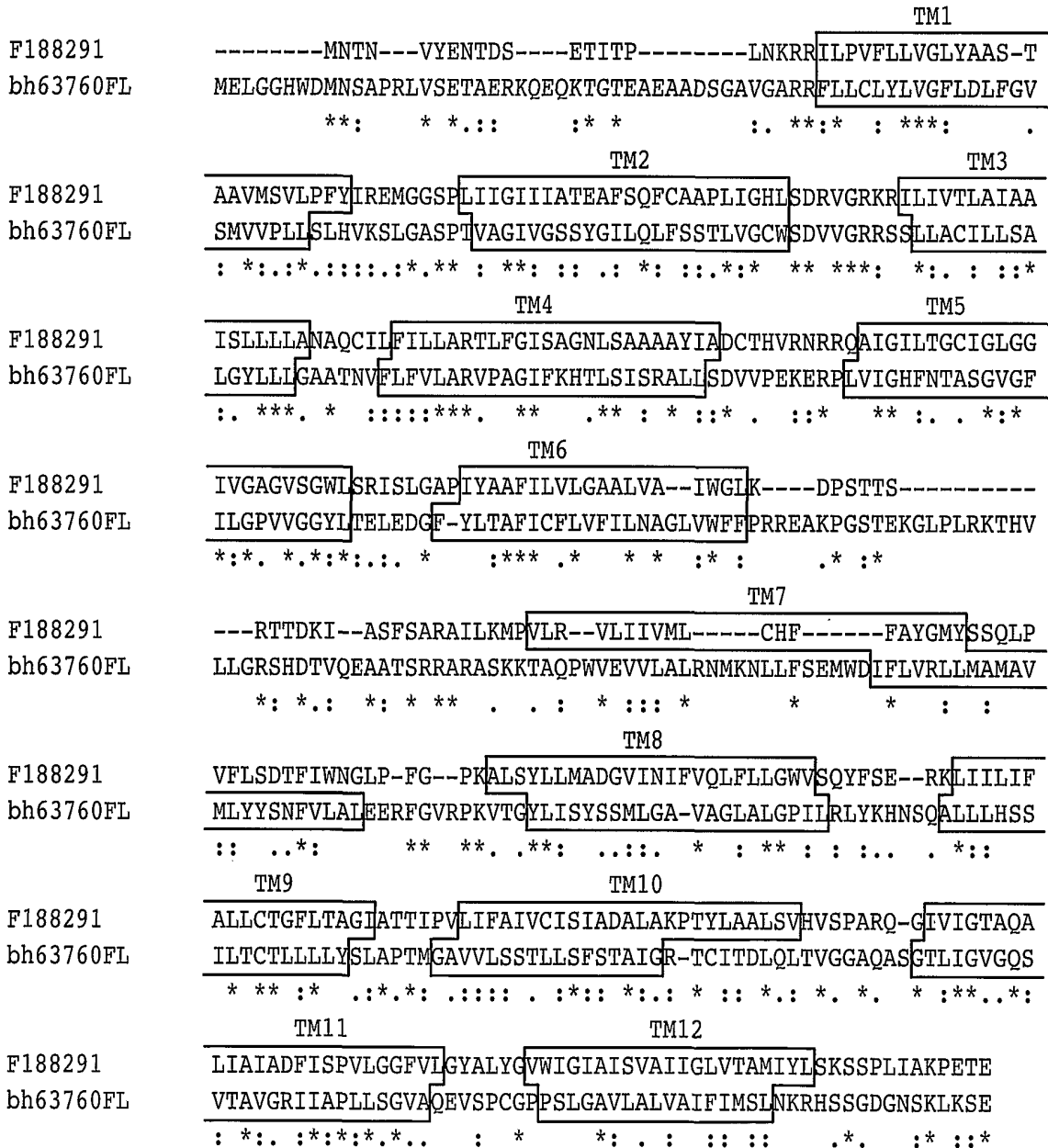


Fig. 5