



(86) Date de dépôt PCT/PCT Filing Date: 2001/12/05
(87) Date publication PCT/PCT Publication Date: 2002/06/13
(85) Entrée phase nationale/National Entry: 2003/06/06
(86) N° demande PCT/PCT Application No.: US 2001/046964
(87) N° publication PCT/PCT Publication No.: 2002/046383
(30) Priorités/Priorities: 2000/12/08 (60/254,399) US;
2000/12/21 (60/257,803) US; 2001/01/05 (60/260,110) US;
2001/01/19 (60/262,851) US; 2001/01/25 (60/264,623) US

(51) Cl.Int.⁷/Int.Cl.⁷ C12N 15/12, A61K 38/17, A01K 67/027,
C12N 15/10, C07K 16/18, C12Q 1/68, C07K 14/47,
G01N 33/577, G01N 33/68

(71) Demandeur/Applicant:
INCYTE GENOMICS, INC., US

(72) Inventeurs/Inventors:
YUE, HENRY, US;
AZIMZAI, YALDA, US;
KALLICK, DEBORAH A., US;
BAUGHN, MARIAH R., US;
GRIFFIN, JENNIFER A., US;
...

(74) Agent: SMART & BIGGAR

(54) Titre : MOLECULES DE MODIFICATION ET DE MAINTENANCE PROTEIQUES
(54) Title: PROTEIN MODIFICATION AND MAINTENANCE MOLECULES

(57) **Abrégé/Abstract:**

The invention provides human protein modification and maintenance molecules (PMMM) and polynucleotides which identify and encode PMMM. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating, or preventing disorders associated with aberrant expression of PMMM.



(72) **Inventeurs(suite)/Inventors(continued)**: SWARNAKAR, ANITA, US; LAL, PREETI G., US; CHAWLA, NARINDER K., US; HAFALIA, APRIL J. A., US; GANDHI, AMEENA R., US; AU-YOUNG, JANICE, US; ELLIOTT, VICKI S., US; RAMKUMAR, JAYALAXMI, US; THANGAVELU, KAVITHA, US; LU, YAN, US; WARREN, BRIDGET A., US; LU, DYUNG AINA M., US; LEE, ERNESTINE A., US; TRIBOULEY, CATHERINE M., US; ARVIZU, CHANDRA, US; DELEGEANE, ANGELO M., US; YAO, MONIQUE G., US; KHAN, FARRAH A., US; SANJANWALA, MADHUSUDAN M., US

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

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
13 June 2002 (13.06.2002)

PCT

(10) International Publication Number
WO 02/046383 A3

(51) International Patent Classification⁷: **C12N 15/12**,
15/10, C07K 14/47, 16/18, C12Q 1/68, G01N 33/577,
33/68, A61K 38/17, A01K 67/027

(21) International Application Number: PCT/US01/46964

(22) International Filing Date: 5 December 2001 (05.12.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/254,399	8 December 2000 (08.12.2000)	US
60/257,803	21 December 2000 (21.12.2000)	US
60/260,110	5 January 2001 (05.01.2001)	US
60/262,851	19 January 2001 (19.01.2001)	US
60/264,623	25 January 2001 (25.01.2001)	US

(71) Applicant (for all designated States except US): **INCYTE GENOMICS, INC.** [US/US]; 3160 Porter Drive, Palo Alto, CA 94304 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **YUE, Henry** [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). **AZIMZAI, Yalda** [US/US]; 5518 Boulder Canyon Drive, Castro Valley, CA 94552 (US). **KALLICK, Deborah, A.** [US/US]; 900 Olive Street, Menlo Park, CA 94025 (US). **BAUGHN, Mariah, R.** [US/US]; 14244 Santiago Road, San Leandro, CA 94577 (US). **GRIFFIN, Jennifer, A.** [US/US]; 33691 Mello Way, Fremont, CA 94555 (US). **SWARNAKAR, Anita** [CA/US]; 8 Locksley Avenue #5D, San Francisco, CA 94122 (US). **LAL, Preeti, G.** [IN/US]; P.O. Box 5142, Santa Clara, CA 95056 (US). **WALIA, Narinder, K.** [US/US]; 890 Davis Street #205, San Leandro, CA 94577 (US). **HAFALIA, April, J., A.** [US/US]; 2227 Calle de Primavera, Santa Clara, CA 95054 (US). **GANDHI, Ameena, R.** [US/US]; 705 5th Avenue, San Francisco, CA 94118 (US). **AU-YOUNG, Janice** [US/US]; 233 Golden Eagle Lane, Brisbane, CA 94005 (US). **ELLIOTT, Vicki, S.** [US/US]; 3770 Polton Place Way, San Jose, CA 95121 (US). **RAMKUMAR, Jayalaxmi** [IN/US]; 34359 Maybird Circle, Fremont, CA 94555 (US). **THANGAVELU, Kavitha** [IN/US]; 1950

Montecito Avenue #23, Mountain View, CA 94043 (US). **LU, Yan** [US/US]; 3885 Corrina Way, Palo Alto, CA 94303 (US). **WARREN, Bridget, A.** [US/US]; 10130 Parkwood Drive #2, Cupertino, CA 95014 (US). **LU, Dyung, Aina, M.** [US/US]; 233 Coy Drive, San Jose, CA 95123 (US). **LEE, Ernestine, A.** [US/US]; 624 Kains Street, Albany, CA 94706 (US). **TRIBOULEY, Catherine, M.** [FR/US]; 1121 Tennessee Street #5, San Francisco, CA 94107 (US). **ARVIZU, Chandra** [US/US]; 490 Sherwood Way #1, Menlo Park, CA 94025 (US). **DELEGEANE, Angelo, M.** [US/US]; 594 Angus Drive, Milpitas, CA 95035 (US). **YAO, Monique, G.** [US/US]; 1189 Woodgate Drive, Carmel, IN 46033 (US). **KHAN, Farrah, A.** [IN/US]; 9445 Harrison Street, Des Plaines, IL 60016 (US). **SANJANWALA, Madhusudan, M.** [US/US]; 210 Silvia Court, Los Altos, CA 94024 (US).

(74) Agents: **HAMLET-COX, Diana** et al.; Incyte Genomics, Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).

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

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

Published:

— with international search report

(88) Date of publication of the international search report:
24 April 2003

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

(54) Title: PROTEIN MODIFICATION AND MAINTENANCE MOLECULES

(57) Abstract: The invention provides human protein modification and maintenance molecules (PMMM) and polynucleotides which identify and encode PMMM. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating, or preventing disorders associated with aberrant expression of PMMM.



WO 02/046383 A3

PROTEIN MODIFICATION AND MAINTENANCE MOLECULES

TECHNICAL FIELD

This invention relates to nucleic acid and amino acid sequences of protein modification and maintenance molecules and to the use of these sequences in the diagnosis, treatment, and prevention of gastrointestinal, cardiovascular, autoimmune/inflammatory, cell proliferative, developmental, epithelial, neurological, and reproductive disorders, and in the assessment of the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of protein modification and maintenance molecules.

10

BACKGROUND OF THE INVENTION

The cellular processes regulating modification and maintenance of protein molecules coordinate their function, conformation, stabilization, and degradation. Each of these processes is mediated by key enzymes or proteins such as kinases, phosphatases, proteases, protease inhibitors, isomerases, transferases, and molecular chaperones.

Kinases

Kinases catalyze the transfer of high-energy phosphate groups from adenosine triphosphate (ATP) to target proteins on the hydroxyamino acid residues serine, threonine, or tyrosine. Addition of a phosphate group alters the local charge on the acceptor molecule, causing internal conformational changes and potentially influencing intermolecular contacts. Reversible protein phosphorylation is the ubiquitous strategy used to control many of the intracellular events in eukaryotic cells. It is estimated that more than ten percent of proteins active in a typical mammalian cell are phosphorylated. Extracellular signals including hormones, neurotransmitters, and growth and differentiation factors can activate kinases, which can occur as cell surface receptors or as the activators of the final effector protein, as well as elsewhere along the signal transduction pathway. Kinases are involved in all aspects of a cell's function, from basic metabolic processes, such as glycolysis, to cell-cycle regulation, differentiation, and communication with the extracellular environment through signal transduction cascades. Inappropriate phosphorylation of proteins in cells has been linked to changes in cell cycle progression and cell differentiation. Changes in the cell cycle have been linked to induction of apoptosis or cancer. Changes in cell differentiation have been linked to diseases and disorders of the reproductive system, immune system, and skeletal muscle.

There are two classes of protein kinases. One class, protein tyrosine kinases (PTKs), phosphorylates tyrosine residues, and the other class, protein serine/threonine kinases (STKs), phosphorylates serine and threonine residues. Some PTKs and STKs possess structural

characteristics of both families and have dual specificity for both tyrosine and serine/threonine residues. Almost all kinases contain a conserved 250-300 amino acid catalytic domain containing specific residues and sequence motifs characteristic of the kinase family. (Reviewed in Hardie, G. and Hanks, S. (1995) The Protein Kinase Facts Book, Vol I p.p. 17-20 Academic Press, San Diego, CA.).

Phosphatases

Phosphatases hydrolytically remove phosphate groups from proteins. Phosphatases are essential in determining the extent of phosphorylation in the cell and, together with kinases, regulate key cellular processes such as metabolic enzyme activity, proliferation, cell growth and differentiation, cell adhesion, and cell cycle progression. Protein phosphatases are characterized as either serine/threonine- or tyrosine-specific based on their preferred phospho-amino acid substrate. Some phosphatases (DSPs, for dual specificity phosphatases) can act on phosphorylated tyrosine, serine, or threonine residues. The protein serine/threonine phosphatases (PSPs) are important regulators of many cAMP-mediated hormone responses in cells. Protein tyrosine phosphatases (PTPs) play a significant role in cell cycle and cell signaling processes.

Proteases

Proteases cleave proteins and peptides at the peptide bond that forms the backbone of the protein or peptide chain. Proteolysis is one of the most important and frequent enzymatic reactions that occurs both within and outside of cells. Proteolysis is responsible for the activation and maturation of nascent polypeptides, the degradation of misfolded and damaged proteins, and the controlled turnover of peptides within the cell. Proteases participate in digestion, endocrine function, tissue remodeling during embryonic development, wound healing, and normal growth. Proteases can play a role in regulatory processes by affecting the half life of regulatory proteins. Proteases are involved in the etiology or progression of disease states such as inflammation, angiogenesis, tumor dispersion and metastasis, cardiovascular disease, neurological disease, and bacterial, parasitic, and viral infections.

Proteases can be categorized on the basis of where they cleave their substrates. Exopeptidases, which include aminopeptidases, dipeptidyl peptidases, tripeptidases, carboxypeptidases, peptidyl-di-peptidases, dipeptidases, and omega peptidases, cleave residues at the termini of their substrates. Endopeptidases, including serine proteases, cysteine proteases, and metalloproteases, cleave at residues within the peptide. Four principal categories of mammalian proteases have been identified based on active site structure, mechanism of action, and overall three-dimensional structure. (See Beynon, R.J. and J.S. Bond (1994) Proteolytic Enzymes: A Practical Approach, Oxford University Press, New York NY, pp. 1-5.)

Serine Proteases

The serine proteases (SPs) are a large, widespread family of proteolytic enzymes that include the digestive enzymes trypsin and chymotrypsin, components of the complement and blood-clotting cascades, and enzymes that control the degradation and turnover of macromolecules within the cell and in the extracellular matrix. Most of the more than 20 subfamilies can be grouped into
 5 six clans, each with a common ancestor. These six clans are hypothesized to have descended from at least four evolutionarily distinct ancestors. SPs are named for the presence of a serine residue found in the active catalytic site of most families. The active site is defined by the catalytic triad, a set of conserved asparagine, histidine, and serine residues critical for catalysis. These residues form a charge relay network that facilitates substrate binding. Other residues outside the active site form
 10 an oxyanion hole that stabilizes the tetrahedral transition intermediate formed during catalysis. SPs have a wide range of substrates and can be subdivided into subfamilies on the basis of their substrate specificity. The main subfamilies are named for the residue(s) after which they cleave: trypases (after arginine or lysine), aspases (after aspartate), chymases (after phenylalanine or leucine), metases (methionine), and serases (after serine) (Rawlings, N.D. and A.J. Barrett (1994)
 15 *Methods Enzymol.* 244:19-61).

Most mammalian serine proteases are synthesized as zymogens, inactive precursors that are activated by proteolysis. For example, trypsinogen is converted to its active form, trypsin, by enteropeptidase. Enteropeptidase is an intestinal protease that removes an N-terminal fragment from trypsinogen. The remaining active fragment is trypsin, which in turn activates the precursors
 20 of the other pancreatic enzymes. Likewise, proteolysis of prothrombin, the precursor of thrombin, generates three separate polypeptide fragments. The N-terminal fragment is released while the other two fragments, which comprise active thrombin, remain associated through disulfide bonds.

The two largest SP subfamilies are the chymotrypsin (S1) and subtilisin (S8) families. Some members of the chymotrypsin family contain two structural domains unique to this family.
 25 Kringle domains are triple-looped, disulfide cross-linked domains found in varying copy number. Kringle domains are thought to play a role in binding mediators such as membranes, other proteins or phospholipids, and in the regulation of proteolytic activity (PROSITE PDOC00020). Apple domains are 90 amino-acid repeated domains, each containing six conserved cysteines. Three disulfide bonds link the first and sixth, second and fifth, and third and fourth cysteines (PROSITE
 30 PDOC00376). Apple domains are involved in protein-protein interactions. S1 family members include trypsin, chymotrypsin, coagulation factors IX-XII, complement factors B, C, and D, granzymes, kallikrein, and tissue- and urokinase-plasminogen activators. The subtilisin family has members found in the eubacteria, archaebacteria, eukaryotes, and viruses. Subtilisins include the proprotein-processing endopeptidases kexin and furin and the pituitary prohormone convertases
 35 PC1, PC2, PC3, PC6, and PACE4 (Rawlings and Barrett, *supra*).

SPs have functions in many normal processes and some have been implicated in the etiology or treatment of disease. Enterokinase, the initiator of intestinal digestion, is found in the intestinal brush border, where it cleaves the acidic propeptide from trypsinogen to yield active trypsin (Kitamoto, Y. et al. (1994) *Proc. Natl. Acad. Sci. USA* 91:7588-7592).

- 5 Prolylcarboxypeptidase, a lysosomal serine peptidase that cleaves peptides such as angiotensin II and III and [des-Arg9] bradykinin, shares sequence homology with members of both the serine carboxypeptidase and prolylendopeptidase families (Tan, F. et al. (1993) *J. Biol. Chem.* 268:16631-16638). The protease neuropsin may influence synapse formation and neuronal connectivity in the hippocampus in response to neural signaling (Chen, Z.-L. et al. (1995) *J. Neurosci.* 15:5088-5097).
- 10 Tissue plasminogen activator is useful for acute management of stroke (Zivin, J.A. (1999) *Neurology* 53:14-19) and myocardial infarction (Ross, A.M. (1999) *Clin. Cardiol.* 22:165-171). Some receptors (PAR, for proteinase-activated receptor), highly expressed throughout the digestive tract, are activated by proteolytic cleavage of an extracellular domain. The major agonists for PARs, thrombin, trypsin, and mast cell tryptase, are released in allergy and inflammatory
- 15 conditions. Control of PAR activation by proteases has been suggested as a promising therapeutic target (Vergnolle, N. (2000) *Aliment. Pharmacol. Ther.* 14:257-266; Rice, K.D. et al. (1998) *Curr. Pharm. Des.* 4:381-396). Prostate-specific antigen (PSA) is a kallikrein-like serine protease synthesized and secreted exclusively by epithelial cells in the prostate gland. Serum PSA is elevated in prostate cancer and is the most sensitive physiological marker for monitoring cancer
- 20 progression and response to therapy. PSA can also identify the prostate as the origin of a metastatic tumor (Brawer, M.K. and P.H. Lange (1989) *Urology* 33:11-16).

The signal peptidase is a specialized class of SP found in all prokaryotic and eukaryotic cell types that serves in the processing of signal peptides from certain proteins. Signal peptides are amino-terminal domains of a protein which direct the protein from its ribosomal assembly site to a

25 particular cellular or extracellular location. Once the protein has been exported, removal of the signal sequence by a signal peptidase and posttranslational processing, e.g., glycosylation or phosphorylation, activate the protein. Signal peptidases exist as multi-subunit complexes in both yeast and mammals. The canine signal peptidase complex is composed of five subunits, all associated with the microsomal membrane and containing hydrophobic regions that span the

30 membrane one or more times (Shelness, G.S. and G. Blobel (1990) *J. Biol. Chem.* 265:9512-9519). Some of these subunits serve to fix the complex in its proper position on the membrane while others contain the actual catalytic activity.

Thrombin is a serine protease with an essential role in the process of blood coagulation. Prothrombin, synthesized in the liver, is converted to active thrombin by Factor Xa. Activated

35 thrombin then cleaves soluble fibrinogen to polymer-forming fibrin, a primary component of blood

clots. In addition, thrombin activates Factor XIIIa, which plays a role in cross-linking fibrin.

Thrombin also stimulates platelet aggregation through proteolytic processing of a 41-residue amino-terminal peptide from protease-activated receptor 1 (PAR-1), formerly known as the thrombin receptor. The cleavage of the amino-terminal peptide exposes a new amino terminus and
5 may also be associated with PAR-1 internalization (Stubbs, M.T. and Bode, W. (1994) *Current Opinion in Structural Biology* 4:823-832 and Ofoso, F.A. et al. (1998) *Biochem. J.* 336:283-285). In addition to stimulating platelet activation through cleavage of the PAR-1 receptor, thrombin also induces platelet aggregation following cleavage of glycoprotein V, also on the surface of platelets. Glycoprotein V appears to be the major thrombin substrate on intact platelets. Platelets deficient for
10 glycoprotein V are hypersensitive to thrombin, which is still required to cleave PAR-1. While platelet aggregation is required for normal hemostasis in mammals, excessive platelet aggregation can result in arterial thrombosis, atherosclerotic arteries, acute myocardial infarction, and stroke (Ramakrishnan, V. et al. (1999) *Proc. Natl. Acad. Sci. U.S.A.* 96:13336-41 and references within).

Proteases in another family have a serine in their active site and are dependent on the
15 hydrolysis of ATP for their activity. These proteases contain proteolytic core domains and regulatory ATPase domains which can be identified by the presence of the P-loop, an ATP/GTP-binding motif (PROSITE PDOC00803). Members of this family include the eukaryotic mitochondrial matrix proteases, Clp protease and the proteasome. Clp protease was originally found in plant chloroplasts but is believed to be widespread in both prokaryotic and eukaryotic cells.
20 The gene for early-onset torsion dystonia encodes a protein related to Clp protease (Ozelius, L.J. et al. (1998) *Adv. Neurol.* 78:93-105).

The proteasome is an intracellular protease complex found in some bacteria and in all eukaryotic cells, and plays an important role in cellular physiology. The proteasome is a large (~2000 kDa) multisubunit complex composed of a central catalytic core containing a variety of
25 proteases arranged in four seven-membered rings with the active sites facing inwards into the central cavity, and terminal ATPase subunits covering the outer port of the cavity and regulating substrate entry (for review, see Schmidt, M. et al. (1999) *Curr. Opin. Chem. Biol.* 3:584-591). Proteasomes are associated with the ubiquitin conjugation system (UCS), a major pathway for the degradation of cellular proteins of all types, including proteins that function to activate or repress
30 cellular processes such as transcription and cell cycle progression (Ciechanover, A. (1994) *Cell* 79:13-21). In the UCS pathway, proteins targeted for degradation are conjugated to ubiquitin, a small heat stable protein. The ubiquitinated protein is then recognized and degraded by the proteasome. The resultant ubiquitin-peptide complex is hydrolyzed by a ubiquitin carboxyl terminal hydrolase, and free ubiquitin is released for reutilization by the UCS. Ubiquitin-proteasome
35 systems are implicated in the degradation of mitotic cyclic kinases, oncoproteins, tumor suppressor

genes (p53), cell surface receptors associated with signal transduction, transcriptional regulators, and mutated or damaged proteins (Ciechanover, *supra*). This pathway has been implicated in a number of diseases, including cystic fibrosis, Angelman's syndrome, and Liddle syndrome (reviewed in Schwartz, A.L. and A. Ciechanover (1999) *Annu. Rev. Med.* 50:57-74). A murine

5 proto-oncogene, Unp, encodes a nuclear ubiquitin protease whose overexpression leads to oncogenic transformation of NIH3T3 cells. The human homolog of this gene is consistently elevated in small cell tumors and adenocarcinomas of the lung (Gray, D.A. (1995) *Oncogene* 10:2179-2183). Ubiquitin carboxyl terminal hydrolase is involved in the differentiation of a lymphoblastic leukemia cell line to a non-dividing mature state (Maki, A. et al. (1996)

10 Differentiation 60:59-66). In neurons, ubiquitin carboxyl terminal hydrolase (PGP 9.5) expression is strong in the abnormal structures that occur in human neurodegenerative diseases (Lowe, J. et al. (1990) *J. Pathol.* 161:153-160).

Cysteine Proteases

Cysteine proteases (CPs) are involved in diverse cellular processes ranging from the

15 processing of precursor proteins to intracellular degradation. Nearly half of the CPs known are present only in viruses. CPs have a cysteine as the major catalytic residue at the active site where catalysis proceeds via a thioester intermediate and is facilitated by nearby histidine and asparagine residues. A glutamine residue is also important, as it helps to form an oxyanion hole. Two important CP families include the papain-like enzymes (C1) and the calpains (C2). Papain-like

20 family members are generally lysosomal or secreted and therefore are synthesized with signal peptides as well as propeptides. Most members bear a conserved motif in the propeptide that may have structural significance (Karrer, K.M. et al. (1993) *Proc. Natl. Acad. Sci. USA* 90:3063-3067). Three-dimensional structures of papain family members show a bilobed molecule with the catalytic site located between the two lobes. Papains include cathepsins B, C, H, L, and S, certain plant

25 allergens and dipeptidyl peptidase (for a review, see Rawlings, N.D. and A.J. Barrett (1994) *Methods Enzymol.* 244:461-486).

Some CPs are expressed ubiquitously, while others are produced only by cells of the immune system. Of particular note, CPs are produced by monocytes, macrophages and other cells which migrate to sites of inflammation and secrete molecules involved in tissue repair.

30 Overabundance of these repair molecules plays a role in certain disorders. In autoimmune diseases such as rheumatoid arthritis, secretion of the cysteine peptidase cathepsin C degrades collagen, laminin, elastin and other structural proteins found in the extracellular matrix of bones. Bone weakened by such degradation is also more susceptible to tumor invasion and metastasis. Cathepsin L expression may also contribute to the influx of mononuclear cells which exacerbates the

35 destruction of the rheumatoid synovium (Keyszer, G.M. (1995) *Arthritis Rheum.* 38:976-984).

Calpains are calcium-dependent cytosolic endopeptidases which contain both an N-terminal catalytic domain and a C-terminal calcium-binding domain. Calpain is expressed as a proenzyme heterodimer consisting of a catalytic subunit unique to each isoform and a regulatory subunit common to different isoforms. Each subunit bears a calcium-binding EF-hand domain. The
5 regulatory subunit also contains a hydrophobic glycine-rich domain that allows the enzyme to associate with cell membranes. Calpains are activated by increased intracellular calcium concentration, which induces a change in conformation and limited autolysis. The resultant active molecule requires a lower calcium concentration for its activity (Chan, S.L. and M.P. Mattson (1999) J. Neurosci. Res. 58:167-190). Calpain expression is predominantly neuronal, although it is
10 present in other tissues. Several chronic neurodegenerative disorders, including ALS, Parkinson's disease and Alzheimer's disease are associated with increased calpain expression (Chan and Mattson, supra). Calpain-mediated breakdown of the cytoskeleton has been proposed to contribute to brain damage resulting from head injury (McCracken, E. et al. (1999) J. Neurotrauma 16:749-761). Calpain-3 is predominantly expressed in skeletal muscle, and is responsible for limb-girdle
15 muscular dystrophy type 2A (Minami, N. et al. (1999) J. Neurol. Sci. 171:31-37).

Another family of thiol proteases is the caspases, which are involved in the initiation and execution phases of apoptosis. A pro-apoptotic signal can activate initiator caspases that trigger a proteolytic caspase cascade, leading to the hydrolysis of target proteins and the classic apoptotic death of the cell. Two active site residues, a cysteine and a histidine, have been implicated in the
20 catalytic mechanism. Caspases are among the most specific endopeptidases, cleaving after aspartate residues. Caspases are synthesized as inactive zymogens consisting of one large (p20) and one small (p10) subunit separated by a small spacer region, and a variable N-terminal prodomain. This prodomain interacts with cofactors that can positively or negatively affect apoptosis. An activating signal causes autoproteolytic cleavage of a specific aspartate residue (D297 in the caspase-1
25 numbering convention) and removal of the spacer and prodomain, leaving a p10/p20 heterodimer. Two of these heterodimers interact via their small subunits to form the catalytically active tetramer. The long prodomains of some caspase family members have been shown to promote dimerization and auto-processing of procaspases. Some caspases contain a "death effector domain" in their prodomain by which they can be recruited into self-activating complexes with other caspases and
30 FADD protein associated death receptors or the TNF receptor complex. In addition, two dimers from different caspase family members can associate, changing the substrate specificity of the resultant tetramer. Endogenous caspase inhibitors (inhibitor of apoptosis proteins, or IAPs) also exist. All these interactions have clear effects on the control of apoptosis (reviewed in Chan and Mattson, supra; Salveson, G.S. and V.M. Dixit (1999) Proc. Natl. Acad. Sci. USA 96:10964-10967).

Caspases have been implicated in a number of diseases. Mice lacking some caspases have severe nervous system defects due to failed apoptosis in the neuroepithelium and suffer early lethality. Others show severe defects in the inflammatory response, as caspases are responsible for processing IL-1b and possibly other inflammatory cytokines (Chan and Mattson, *supra*). Cowpox virus and baculoviruses target caspases to avoid the death of their host cell and promote successful infection. In addition, increases in inappropriate apoptosis have been reported in AIDS, neurodegenerative diseases and ischemic injury, while a decrease in cell death is associated with cancer (Salveson and Dixit, *supra*; Thompson, C.B. (1995) Science 267:1456-1462).

Aspartyl proteases

Aspartyl proteases (APs) include the lysosomal proteases cathepsins D and E, as well as chymosin, renin, and the gastric pepsins. Most retroviruses encode an AP, usually as part of the pol polyprotein. APs, also called acid proteases, are monomeric enzymes consisting of two domains, each domain containing one half of the active site with its own catalytic aspartic acid residue. APs are most active in the range of pH 2–3, at which one of the aspartate residues is ionized and the other neutral. The pepsin family of APs contains many secreted enzymes, and all are likely to be synthesized with signal peptides and propeptides. Most family members have three disulfide loops, the first ~5 residue loop following the first aspartate, the second 5-6 residue loop preceding the second aspartate, and the third and largest loop occurring toward the C terminus. Retropepsins, on the other hand, are analogous to a single domain of pepsin, and become active as homodimers with each retropepsin monomer contributing one half of the active site. Retropepsins are required for processing the viral polyproteins.

APs have roles in various tissues, and some have been associated with disease. Renin mediates the first step in processing the hormone angiotensin, which is responsible for regulating electrolyte balance and blood pressure (reviewed in Crews, D.E. and S.R. Williams (1999) Hum. Biol. 71:475-503). Abnormal regulation and expression of cathepsins are evident in various inflammatory disease states. Expression of cathepsin D is elevated in synovial tissues from patients with rheumatoid arthritis and osteoarthritis. The increased expression and differential regulation of the cathepsins are linked to the metastatic potential of a variety of cancers (Chambers, A.F. et al. (1993) Crit. Rev. Oncol. 4:95-114).

30 Metalloproteases

Metalloproteases require a metal ion for activity, usually manganese or zinc. Examples of manganese metalloenzymes include aminopeptidase P and human proline dipeptidase (PEPD). Aminopeptidase P can degrade bradykinin, a nonapeptide activated in a variety of inflammatory responses. Aminopeptidase P has been implicated in coronary ischemia/reperfusion injury.

Administration of aminopeptidase P inhibitors has been shown to have a cardioprotective effect in rats (Ersahin, C. et al (1999) *J. Cardiovasc. Pharmacol.* 34:604-611).

Most zinc-dependent metalloproteases share a common sequence in the zinc-binding domain. The active site is made up of two histidines which act as zinc ligands and a catalytic glutamic acid C-terminal to the first histidine. Proteins containing this signature sequence are known as the metzincins and include aminopeptidase N, angiotensin-converting enzyme, neurolysin, the matrix metalloproteases and the adamalysins (ADAMS). An alternate sequence is found in the zinc carboxypeptidases, in which all three conserved residues – two histidines and a glutamic acid – are involved in zinc binding.

A number of the neutral metalloendopeptidases, including angiotensin converting enzyme and the aminopeptidases, are involved in the metabolism of peptide hormones. High aminopeptidase B activity, for example, is found in the adrenal glands and neurohypophyses of hypertensive rats (Prieto, I. et al. (1998) *Horm. Metab. Res.* 30:246-248). Oligopeptidase M/neurolysin can hydrolyze bradykinin as well as neurotensin (Serizawa, A. et al. (1995) *J. Biol. Chem.* 270:2092-2098). Neurotensin is a vasoactive peptide that can act as a neurotransmitter in the brain, where it has been implicated in limiting food intake (Tritos, N.A. et al. (1999) *Neuropeptides* 33:339-349).

The matrix metalloproteases (MMPs) are a family of at least 23 enzymes that can degrade components of the extracellular matrix (ECM). They are Zn^{2+} endopeptidases with an N-terminal catalytic domain. Nearly all members of the family have a hinge peptide and a C-terminal domain which can bind to substrate molecules in the ECM or to inhibitors produced by the tissue (TIMPs, for tissue inhibitor of metalloprotease; Campbell, I.L. et al. (1999) *Trends Neurosci.* 22:285). The presence of fibronectin-like repeats, transmembrane domains, or C-terminal hemopexinase-like domains can be used to separate MMPs into collagenase, gelatinase, stromelysin and membrane-type MMP subfamilies. In the inactive form, the Zn^{2+} ion in the active site interacts with a cysteine in the pro-sequence. Activating factors disrupt the Zn^{2+} -cysteine interaction, or "cysteine switch," exposing the active site. This partially activates the enzyme, which then cleaves off its propeptide and becomes fully active. MMPs are often activated by the serine proteases plasmin and furin. MMPs are often regulated by stoichiometric, noncovalent interactions with inhibitors; the balance of protease to inhibitor, then, is very important in tissue homeostasis (reviewed in Yong, V.W. et al. (1998) *Trends Neurosci.* 21:75).

MMPs are implicated in a number of diseases including osteoarthritis (Mitchell, P. et al. (1996) *J. Clin. Invest.* 97:761), atherosclerotic plaque rupture (Sukhova, G.K. et al. (1999) *Circulation* 99:2503), aortic aneurysm (Schneiderman, J. et al. (1998) *Am. J. Path.* 152:703), non-healing wounds (Saarialho-Kere, U.K. et al. (1994) *J. Clin. Invest.* 94:79), bone resorption

(Blavier, L. and J.M. Delaisse (1995) J. Cell Sci. 108:3649), age-related macular degeneration (Steen, B. et al. (1998) Invest. Ophthalmol. Vis. Sci. 39:2194), emphysema (Finlay, G.A. et al. (1997) Thorax 52:502), myocardial infarction (Rohde, L.E. et al. (1999) Circulation 99:3063) and dilated cardiomyopathy (Thomas, C.V. et al. (1998) Circulation 97:1708). MMP inhibitors prevent
 5 metastasis of mammary carcinoma and experimental tumors in rat, and Lewis lung carcinoma, hemangioma, and human ovarian carcinoma xenografts in mice (Eccles, S.A. et al. (1996) Cancer Res. 56:2815; Anderson et al. (1996) Cancer Res. 56:715-718; Volpert, O.V. et al. (1996) J. Clin. Invest. 98:671; Taraboletti, G. et al. (1995) J. NCI 87:293; Davies, B. et al. (1993) Cancer Res. 53:2087). MMPs may be active in Alzheimer's disease. A number of MMPs are implicated in
 10 multiple sclerosis, and administration of MMP inhibitors can relieve some of its symptoms (reviewed in Yong, supra).

Another family of metalloproteases is the ADAMs, for A Disintegrin and Metalloprotease Domain, which they share with their close relatives the adamalysins, snake venom metalloproteases (SVMPs). ADAMs combine features of both cell surface adhesion molecules and proteases,
 15 containing a prodomain, a protease domain, a disintegrin domain, a cysteine rich domain, an epidermal growth factor repeat, a transmembrane domain, and a cytoplasmic tail. The first three domains listed above are also found in the SVMPs. The ADAMs possess four potential functions: proteolysis, adhesion, signaling and fusion. The ADAMs share the metzincin zinc binding sequence and are inhibited by some MMP antagonists such as TIMP-1.

20 ADAMs are implicated in such processes as sperm-egg binding and fusion, myoblast fusion, and protein-ectodomain processing or shedding of cytokines, cytokine receptors, adhesion proteins and other extracellular protein domains (Schlöndorff, J. and C.P. Blobel (1999) J. Cell. Sci. 112:3603-3617). The Kuzbanian protein cleaves a substrate in the NOTCH pathway (possibly NOTCH itself), activating the program for lateral inhibition in Drosophila neural development.
 25 Two ADAMs, TACE (ADAM 17) and ADAM 10, are proposed to have analogous roles in the processing of amyloid precursor protein in the brain (Schlöndorff and Blobel, supra). TACE has also been identified as the TNF activating enzyme (Black, R.A. et al. (1997) Nature 385:729). TNF is a pleiotropic cytokine that is important in mobilizing host defenses in response to infection or trauma, but can cause severe damage in excess and is often overproduced in autoimmune disease.
 30 TACE cleaves membrane-bound pro-TNF to release a soluble form. Other ADAMs may be involved in a similar type of processing of other membrane-bound molecules.

Proteins of the ADAMTS sub-family have all of the features of ADAM family metalloproteases and contain an additional thrombospondin domain (TS). The prototypic ADAMTS was identified in mouse, and found to be expressed in heart and kidney and upregulated
 35 by proinflammatory stimuli (Kuno, K. et al. (1997) J. Biol. Chem. 272:556-562). To date eleven

members are recognized by the Human Genome Organization (HUGO; <http://www.gene.ucl.ac.uk/users/hester/adamts.html#Approved>). Members of this family have the ability to degrade aggrecan, a high molecular weight proteoglycan which provides cartilage with important mechanical properties including compressibility, and which is lost during the

5 development of arthritis. Enzymes which degrade aggrecan are thus considered attractive targets to prevent and slow the degradation of articular cartilage (See, e.g., Tortorella, M.D. (1999) *Science* 284:1664; Abbaszade, I. (1999) *J. Biol. Chem.* 274:23443). Other members are reported to have antiangiogenic potential (Kuno et al., *supra*) and/or procollagen processing (Colige, A. et al. (1997) *Proc. Natl. Acad. Sci. USA* 94:2374).

10 Protease inhibitors

Protease inhibitors and other regulators of protease activity control the activity and effects of proteases. Protease inhibitors have been shown to control pathogenesis in animal models of proteolytic disorders (Murphy, G. (1991) *Agents Actions Suppl.* 35:69-76). Low levels of the cystatins, low molecular weight inhibitors of the cysteine proteases, correlate with malignant

15 progression of tumors (Calkins, C. et al. (1995) *Biol. Biochem. Hoppe Seyler* 376:71-80). The cystatin superfamily of protease inhibitors is characterized by a particular pattern of linearly arranged and tandemly repeated disulfide loops (Kellermann, J. et al. (1989) *J. Biol. Chem.* 264:14121-14128). Serpins are inhibitors of mammalian plasma serine proteases. Many serpins serve to regulate the blood clotting cascade and/or the complement cascade in mammals. Sp32 is a

20 positive regulator of the mammalian acrosomal protease, acrosin, that binds the proenzyme, proacrosin, and thereby aides in packaging the enzyme into the acrosomal matrix (Baba, T. et al. (1994) *J. Biol. Chem.* 269:10133-10140). The Kunitz family of serine protease inhibitors are characterized by one or more "Kunitz domains" containing a series of cysteine residues that are regularly spaced over approximately 50 amino acid residues and form three intrachain disulfide

25 bonds. Members of this family include aprotinin, tissue factor pathway inhibitor (TFPI-1 and TFPI-2), inter- α -trypsin inhibitor, and bikunin (Marlor, C.W. et al. (1997) *J. Biol. Chem.* 272:12202-12208). Members of this family are potent inhibitors (in the nanomolar range) against serine proteases such as kallikrein and plasmin. Aprotinin has clinical utility in reduction of perioperative blood loss.

30 A major portion of all proteins synthesized in eukaryotic cells are synthesized on the cytosolic surface of the endoplasmic reticulum (ER). Before these immature proteins are distributed to other organelles in the cell or are secreted, they must be transported into the interior lumen of the ER where post-translational modifications are performed. These modifications include protein folding and the formation of disulfide bonds, and N-linked glycosylations.

35 **Protein Isomerases**

Protein folding in the ER is aided by two principal types of protein isomerases, protein disulfide isomerase (PDI), and peptidyl-prolyl isomerase (PPI). PDI catalyzes the oxidation of free sulfhydryl groups in cysteine residues to form intramolecular disulfide bonds in proteins. PPI, an enzyme that catalyzes the isomerization of certain proline imidic bonds in oligopeptides and
5 proteins, is considered to govern one of the rate limiting steps in the folding of many proteins to their final functional conformation. The cyclophilins represent a major class of PPI that was originally identified as the major receptor for the immunosuppressive drug cyclosporin A (Handschumacher, R.E. et al. (1984) Science 226: 544-547).

Protein Glycosylation

10 The glycosylation of most soluble secreted and membrane-bound proteins by oligosaccharides linked to asparagine residues in proteins is also performed in the ER. This reaction is catalyzed by a membrane-bound enzyme, oligosaccharyl transferase. Although the exact purpose of this "N-linked" glycosylation is unknown, the presence of oligosaccharides tends to make a glycoprotein resistant to protease digestion. In addition, oligosaccharides attached to cell-
15 surface proteins called selectins are known to function in cell-cell adhesion processes (Alberts, B. et al. (1994) Molecular Biology of the Cell Garland Publishing Co., New York, NY. p.608). "O-linked" glycosylation of proteins also occurs in the ER by the addition of N-acetylgalactosamine to the hydroxyl group of a serine or threonine residue followed by the sequential addition of other sugar residues to the first. This process is catalyzed by a series of glycosyltransferases, each
20 specific for a particular donor sugar nucleotide and acceptor molecule (Lodish, H. et al. (1995) Molecular Cell Biology, W. H. Freeman and Co., New York, NY pp.700-708). In many cases, both N- and O-linked oligosaccharides appear to be required for the secretion of proteins or the movement of plasma membrane glycoproteins to the cell surface.

An additional glycosylation mechanism operates in the ER specifically to target lysosomal
25 enzymes to lysosomes and prevent their secretion. Lysosomal enzymes in the ER receive an N-linked oligosaccharide, like plasma membrane and secreted proteins, but are then phosphorylated on one or two mannose residues. The phosphorylation of mannose residues occurs in two steps, the first step being the addition of an N-acetylglucosamine phosphate residue by N-acetylglucosamine phosphotransferase, and the second the removal of the N-acetylglucosamine group by
30 phosphodiesterase. The phosphorylated mannose residue then targets the lysosomal enzyme to a mannose 6-phosphate receptor which transports it to a lysosome vesicle (Lodish et al. supra, pp. 708-711).

Chaperones

Molecular chaperones are proteins that aid in the proper folding of immature proteins and
35 refolding of improperly folded ones, the assembly of protein subunits, and in the transport of

unfolded proteins across membranes. Chaperones are also called heat-shock proteins (hsp) because of their tendency to be expressed in dramatically increased amounts following brief exposure of cells to elevated temperatures. This latter property most likely reflects their need in the refolding of proteins that have become denatured by the high temperatures. Chaperones may be divided into

5 several classes according to their location, function, and molecular weight, and include hsp60, TCP1, hsp70, hsp40 (also called DnaJ), and hsp90. For example, hsp90 binds to steroid hormone receptors, represses transcription in the absence of the ligand, and provides proper folding of the ligand-binding domain of the receptor in the presence of the hormone (Burston, S.G. and A.R. Clarke (1995) *Essays Biochem.* 29:125-136). Hsp60 and hsp70 chaperones aid in the transport and

10 folding of newly synthesized proteins. Hsp70 acts early in protein folding, binding a newly synthesized protein before it leaves the ribosome and transporting the protein to the mitochondria or ER before releasing the folded protein. Hsp60, along with hsp10, binds misfolded proteins and gives them the opportunity to refold correctly. All chaperones share an affinity for hydrophobic patches on incompletely folded proteins and the ability to hydrolyze ATP. The energy of ATP

15 hydrolysis is used to release the hsp-bound protein in its properly folded state (Alberts, B. et al.supra, pp 214, 571-572).

The discovery of new protein modification and maintenance molecules, and the polynucleotides encoding them, satisfies a need in the art by providing new compositions which are useful in the diagnosis, prevention, and treatment of gastrointestinal, cardiovascular,

20 autoimmune/inflammatory, cell proliferative, developmental, epithelial, neurological, and reproductive disorders, and in the assessment of the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of protein modification and maintenance molecules.

25

SUMMARY OF THE INVENTION

The invention features purified polypeptides, protein modification and maintenance molecules, referred to collectively as "PMMM" and individually as "PMMM-1," "PMMM-2," "PMMM-3," "PMMM-4," "PMMM-5," "PMMM-6," "PMMM-7," "PMMM-8," "PMMM-9," "PMMM-10," "PMMM-11," "PMMM-12," "PMMM-13," "PMMM-14," "PMMM-15," "PMMM-

30 16," "PMMM-17," and "PMMM-18." In one aspect, the invention provides an isolated polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-18, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-18, c) a biologically active fragment of a polypeptide having an amino

35 acid sequence selected from the group consisting of SEQ ID NO:1-18, and d) an immunogenic

fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-18. In one alternative, the invention provides an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:1-18.

The invention further provides an isolated polynucleotide encoding a polypeptide selected
5 from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-18, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-18, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-18, and d) an immunogenic fragment of a
10 polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-18. In one alternative, the polynucleotide encodes a polypeptide selected from the group consisting of SEQ ID NO:1-18. In another alternative, the polynucleotide is selected from the group consisting of SEQ ID NO:19-36.

Additionally, the invention provides a recombinant polynucleotide comprising a promoter
15 sequence operably linked to a polynucleotide encoding a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-18, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-18, c) a biologically active fragment of a polypeptide having an amino acid sequence
20 selected from the group consisting of SEQ ID NO:1-18, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-18. In one alternative, the invention provides a cell transformed with the recombinant polynucleotide. In another alternative, the invention provides a transgenic organism comprising the recombinant polynucleotide.

25 The invention also provides a method for producing a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-18, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-18, c) a biologically active fragment of a polypeptide having an amino acid sequence
30 selected from the group consisting of SEQ ID NO:1-18, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-18. The method comprises a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding the polypeptide, and b) recovering
35 the polypeptide so expressed.

Additionally, the invention provides an isolated antibody which specifically binds to a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-18, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected
5 from the group consisting of SEQ ID NO:1-18, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-18, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-18.

The invention further provides an isolated polynucleotide selected from the group
10 consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:19-36, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:19-36, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-
15 d). In one alternative, the polynucleotide comprises at least 60 contiguous nucleotides.

Additionally, the invention provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:19-36, b) a polynucleotide comprising a naturally occurring
20 polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:19-36, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). The method comprises a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and
25 which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and b) detecting the presence or absence of said hybridization complex, and optionally, if present, the amount thereof. In one alternative, the probe comprises at least 60 contiguous nucleotides.

30 The invention further provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:19-36, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID
35 NO:19-36, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide

complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). The method comprises a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.

5 The invention further provides a composition comprising an effective amount of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-18, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-18, c) a biologically active fragment of a polypeptide
10 having an amino acid sequence selected from the group consisting of SEQ ID NO:1-18, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-18, and a pharmaceutically acceptable excipient. In one embodiment, the composition comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-18. The invention additionally provides a method of treating a disease or condition associated
15 with decreased expression of functional PMMM, comprising administering to a patient in need of such treatment the composition.

 The invention also provides a method for screening a compound for effectiveness as an agonist of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-18, b) a polypeptide comprising
20 a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-18, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-18, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-18. The method comprises a) exposing a sample comprising the
25 polypeptide to a compound, and b) detecting agonist activity in the sample. In one alternative, the invention provides a composition comprising an agonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with decreased expression of functional PMMM, comprising administering to a patient in need of such treatment the composition.

30 Additionally, the invention provides a method for screening a compound for effectiveness as an antagonist of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-18, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-18, c) a biologically active fragment
35 of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID

NO:1-18, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-18. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting antagonist activity in the sample. In one alternative, the invention provides a composition comprising an antagonist compound identified
5 by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with overexpression of functional PMMM, comprising administering to a patient in need of such treatment the composition.

The invention further provides a method of screening for a compound that specifically binds to a polypeptide selected from the group consisting of a) a polypeptide comprising an amino
10 acid sequence selected from the group consisting of SEQ ID NO:1-18, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-18, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-18, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group
15 consisting of SEQ ID NO:1-18. The method comprises a) combining the polypeptide with at least one test compound under suitable conditions, and b) detecting binding of the polypeptide to the test compound, thereby identifying a compound that specifically binds to the polypeptide.

The invention further provides a method of screening for a compound that modulates the activity of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino
20 acid sequence selected from the group consisting of SEQ ID NO:1-18, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-18, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-18, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group
25 consisting of SEQ ID NO:1-18. The method comprises a) combining the polypeptide with at least one test compound under conditions permissive for the activity of the polypeptide, b) assessing the activity of the polypeptide in the presence of the test compound, and c) comparing the activity of the polypeptide in the presence of the test compound with the activity of the polypeptide in the absence of the test compound, wherein a change in the activity of the polypeptide in the presence of
30 the test compound is indicative of a compound that modulates the activity of the polypeptide.

The invention further provides a method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a polynucleotide sequence selected from the group consisting of SEQ ID NO:19-36, the method comprising a) exposing a sample comprising the target polynucleotide to a compound, b) detecting
35 altered expression of the target polynucleotide, and c) comparing the expression of the target

polynucleotide in the presence of varying amounts of the compound and in the absence of the compound.

The invention further provides a method for assessing toxicity of a test compound, said method comprising a) treating a biological sample containing nucleic acids with the test compound; 5 b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide selected from the group consisting of i) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:19-36, ii) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:19-36, iii) a 10 polynucleotide having a sequence complementary to i), iv) a polynucleotide complementary to the polynucleotide of ii), and v) an RNA equivalent of i)-iv). Hybridization occurs under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide selected from the group consisting of i) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID 15 NO:19-36, ii) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:19-36, iii) a polynucleotide complementary to the polynucleotide of i), iv) a polynucleotide complementary to the polynucleotide of ii), and v) an RNA equivalent of i)-iv). Alternatively, the target polynucleotide comprises a fragment of a polynucleotide sequence selected from the group 20 consisting of i)-v) above; c) quantifying the amount of hybridization complex; and d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.

25 BRIEF DESCRIPTION OF THE TABLES

Table 1 summarizes the nomenclature for the full length polynucleotide and polypeptide sequences of the present invention.

Table 2 shows the GenBank identification number and annotation of the nearest GenBank homolog for polypeptides of the invention. The probability scores for the matches between each 30 polypeptide and its homolog(s) are also shown.

Table 3 shows structural features of polypeptide sequences of the invention, including predicted motifs and domains, along with the methods, algorithms, and searchable databases used for analysis of the polypeptides.

Table 4 lists the cDNA and/or genomic DNA fragments which were used to assemble polynucleotide sequences of the invention, along with selected fragments of the polynucleotide sequences.

Table 5 shows the representative cDNA library for polynucleotides of the invention.

5 Table 6 provides an appendix which describes the tissues and vectors used for construction of the cDNA libraries shown in Table 5.

Table 7 shows the tools, programs, and algorithms used to analyze the polynucleotides and polypeptides of the invention, along with applicable descriptions, references, and threshold parameters.

10

DESCRIPTION OF THE INVENTION

Before the present proteins, nucleotide sequences, and methods are described, it is understood that this invention is not limited to the particular machines, materials and methods described, as these may vary. It is also to be understood that the terminology used herein is for the
15 purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention which will be limited only by the appended claims.

It must be noted that as used herein and in the appended claims, the singular forms “a,” “an,” and “the” include plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to “a host cell” includes a plurality of such host cells, and a reference to “an
20 antibody” is a reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so forth.

Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any machines, materials, and methods similar or equivalent to those described
25 herein can be used to practice or test the present invention, the preferred machines, materials and methods are now described. All publications mentioned herein are cited for the purpose of describing and disclosing the cell lines, protocols, reagents and vectors which are reported in the publications and which might be used in connection with the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of
30 prior invention.

DEFINITIONS

“PMMM” refers to the amino acid sequences of substantially purified PMMM obtained from any species, particularly a mammalian species, including bovine, ovine, porcine, murine, equine, and human, and from any source, whether natural, synthetic, semi-synthetic, or recombinant.

The term “agonist” refers to a molecule which intensifies or mimics the biological activity of PMMM. Agonists may include proteins, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of PMMM either by directly interacting with PMMM or by acting on components of the biological pathway in which PMMM
5 participates.

An “allelic variant” is an alternative form of the gene encoding PMMM. Allelic variants may result from at least one mutation in the nucleic acid sequence and may result in altered mRNAs or in polypeptides whose structure or function may or may not be altered. A gene may have none, one, or many allelic variants of its naturally occurring form. Common mutational changes which
10 give rise to allelic variants are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

“Altered” nucleic acid sequences encoding PMMM include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polypeptide the same as PMMM
15 or a polypeptide with at least one functional characteristic of PMMM. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding PMMM, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding PMMM. The encoded protein may also be “altered,” and may
20 contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent PMMM. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues, as long as the biological or immunological activity of PMMM is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid,
25 and positively charged amino acids may include lysine and arginine. Amino acids with uncharged polar side chains having similar hydrophilicity values may include: asparagine and glutamine; and serine and threonine. Amino acids with uncharged side chains having similar hydrophilicity values may include: leucine, isoleucine, and valine; glycine and alanine; and phenylalanine and tyrosine.

The terms “amino acid” and “amino acid sequence” refer to an oligopeptide, peptide,
30 polypeptide, or protein sequence, or a fragment of any of these, and to naturally occurring or synthetic molecules. Where “amino acid sequence” is recited to refer to a sequence of a naturally occurring protein molecule, “amino acid sequence” and like terms are not meant to limit the amino acid sequence to the complete native amino acid sequence associated with the recited protein molecule.

“Amplification” relates to the production of additional copies of a nucleic acid sequence. Amplification is generally carried out using polymerase chain reaction (PCR) technologies well known in the art.

The term “antagonist” refers to a molecule which inhibits or attenuates the biological activity of PMMM. Antagonists may include proteins such as antibodies, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of PMMM either by directly interacting with PMMM or by acting on components of the biological pathway in which PMMM participates.

The term “antibody” refers to intact immunoglobulin molecules as well as to fragments thereof, such as Fab, F(ab')₂, and Fv fragments, which are capable of binding an epitopic determinant. Antibodies that bind PMMM polypeptides can be prepared using intact polypeptides or using fragments containing small peptides of interest as the immunizing antigen. The polypeptide or oligopeptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the translation of RNA, or synthesized chemically, and can be conjugated to a carrier protein if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin (KLH). The coupled peptide is then used to immunize the animal.

The term “antigenic determinant” refers to that region of a molecule (i.e., an epitope) that makes contact with a particular antibody. When a protein or a fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the production of antibodies which bind specifically to antigenic determinants (particular regions or three-dimensional structures on the protein). An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

The term “aptamer” refers to a nucleic acid or oligonucleotide molecule that binds to a specific molecular target. Aptamers are derived from an in vitro evolutionary process (e.g., SELEX (Systematic Evolution of Ligands by EXponential Enrichment), described in U.S. Patent No. 5,270,163), which selects for target-specific aptamer sequences from large combinatorial libraries. Aptamer compositions may be double-stranded or single-stranded, and may include deoxyribonucleotides, ribonucleotides, nucleotide derivatives, or other nucleotide-like molecules. The nucleotide components of an aptamer may have modified sugar groups (e.g., the 2'-OH group of a ribonucleotide may be replaced by 2'-F or 2'-NH₂), which may improve a desired property, e.g., resistance to nucleases or longer lifetime in blood. Aptamers may be conjugated to other molecules, e.g., a high molecular weight carrier to slow clearance of the aptamer from the circulatory system. Aptamers may be specifically cross-linked to their cognate ligands, e.g., by photo-activation of a cross-linker. (See, e.g., Brody, E.N. and L. Gold (2000) J. Biotechnol. 74:5-13.)

The term “intramer” refers to an aptamer which is expressed in vivo. For example, a vaccinia virus-based RNA expression system has been used to express specific RNA aptamers at high levels in the cytoplasm of leukocytes (Blind, M. et al. (1999) Proc. Natl Acad. Sci. USA 96:3606-3610).

- 5 The term “spiegelmer” refers to an aptamer which includes L-DNA, L-RNA, or other left-handed nucleotide derivatives or nucleotide-like molecules. Aptamers containing left-handed nucleotides are resistant to degradation by naturally occurring enzymes, which normally act on substrates containing right-handed nucleotides.

- 10 The term “antisense” refers to any composition capable of base-pairing with the “sense” (coding) strand of a specific nucleic acid sequence. Antisense compositions may include DNA; RNA; peptide nucleic acid (PNA); oligonucleotides having modified backbone linkages such as phosphorothioates, methylphosphonates, or benzylphosphonates; oligonucleotides having modified sugar groups such as 2'-methoxyethyl sugars or 2'-methoxyethoxy sugars; or oligonucleotides having modified bases such as 5-methyl cytosine, 2'-deoxyuracil, or 7-deaza-2'-deoxyguanosine. Antisense
15 molecules may be produced by any method including chemical synthesis or transcription. Once introduced into a cell, the complementary antisense molecule base-pairs with a naturally occurring nucleic acid sequence produced by the cell to form duplexes which block either transcription or translation. The designation “negative” or “minus” can refer to the antisense strand, and the designation “positive” or “plus” can refer to the sense strand of a reference DNA molecule.

- 20 The term “biologically active” refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, “immunologically active” or “immunogenic” refers to the capability of the natural, recombinant, or synthetic PMMM, or of any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

- 25 “Complementary” describes the relationship between two single-stranded nucleic acid sequences that anneal by base-pairing. For example, 5'-AGT-3' pairs with its complement, 3'-TCA-5'.

- A “composition comprising a given polynucleotide sequence” and a “composition comprising a given amino acid sequence” refer broadly to any composition containing the given
30 polynucleotide or amino acid sequence. The composition may comprise a dry formulation or an aqueous solution. Compositions comprising polynucleotide sequences encoding PMMM or fragments of PMMM may be employed as hybridization probes. The probes may be stored in freeze-dried form and may be associated with a stabilizing agent such as a carbohydrate. In hybridizations, the probe may be deployed in an aqueous solution containing salts (e.g., NaCl),

detergents (e.g., sodium dodecyl sulfate; SDS), and other components (e.g., Denhardt's solution, dry milk, salmon sperm DNA, etc.).

"Consensus sequence" refers to a nucleic acid sequence which has been subjected to repeated DNA sequence analysis to resolve uncalled bases, extended using the XL-PCR kit (Applied Biosystems, Foster City CA) in the 5' and/or the 3' direction, and resequenced, or which has been assembled from one or more overlapping cDNA, EST, or genomic DNA fragments using a computer program for fragment assembly, such as the GELVIEW fragment assembly system (GCG, Madison WI) or Phrap (University of Washington, Seattle WA). Some sequences have been both extended and assembled to produce the consensus sequence.

"Conservative amino acid substitutions" are those substitutions that are predicted to least interfere with the properties of the original protein, i.e., the structure and especially the function of the protein is conserved and not significantly changed by such substitutions. The table below shows amino acids which may be substituted for an original amino acid in a protein and which are regarded as conservative amino acid substitutions.

	Original Residue	Conservative Substitution
15	Ala	Gly, Ser
	Arg	His, Lys
	Asn	Asp, Gln, His
	Asp	Asn, Glu
20	Cys	Ala, Ser
	Gln	Asn, Glu, His
	Glu	Asp, Gln, His
	Gly	Ala
	His	Asn, Arg, Gln, Glu
25	Ile	Leu, Val
	Leu	Ile, Val
	Lys	Arg, Gln, Glu
	Met	Leu, Ile
	Phe	His, Met, Leu, Trp, Tyr
30	Ser	Cys, Thr
	Thr	Ser, Val
	Trp	Phe, Tyr
	Tyr	His, Phe, Trp
	Val	Ile, Leu, Thr
35		

Conservative amino acid substitutions generally maintain (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a beta sheet or alpha helical conformation, (b) the charge or hydrophobicity of the molecule at the site of the substitution, and/or (c) the bulk of the side chain.

A "deletion" refers to a change in the amino acid or nucleotide sequence that results in the absence of one or more amino acid residues or nucleotides.

The term “derivative” refers to a chemically modified polynucleotide or polypeptide. Chemical modifications of a polynucleotide can include, for example, replacement of hydrogen by an alkyl, acyl, hydroxyl, or amino group. A derivative polynucleotide encodes a polypeptide which retains at least one biological or immunological function of the natural molecule. A derivative
5 polypeptide is one modified by glycosylation, pegylation, or any similar process that retains at least one biological or immunological function of the polypeptide from which it was derived.

A “detectable label” refers to a reporter molecule or enzyme that is capable of generating a measurable signal and is covalently or noncovalently joined to a polynucleotide or polypeptide.

“Differential expression” refers to increased or upregulated; or decreased, downregulated,
10 or absent gene or protein expression, determined by comparing at least two different samples. Such comparisons may be carried out between, for example, a treated and an untreated sample, or a diseased and a normal sample.

“Exon shuffling” refers to the recombination of different coding regions (exons). Since an exon may represent a structural or functional domain of the encoded protein, new proteins may be
15 assembled through the novel reassortment of stable substructures, thus allowing acceleration of the evolution of new protein functions.

A “fragment” is a unique portion of PMMM or the polynucleotide encoding PMMM which is identical in sequence to but shorter in length than the parent sequence. A fragment may comprise up to the entire length of the defined sequence, minus one nucleotide/amino acid residue. For
20 example, a fragment may comprise from 5 to 1000 contiguous nucleotides or amino acid residues. A fragment used as a probe, primer, antigen, therapeutic molecule, or for other purposes, may be at least 5, 10, 15, 16, 20, 25, 30, 40, 50, 60, 75, 100, 150, 250 or at least 500 contiguous nucleotides or amino acid residues in length. Fragments may be preferentially selected from certain regions of a molecule. For example, a polypeptide fragment may comprise a certain length of contiguous amino
25 acids selected from the first 250 or 500 amino acids (or first 25% or 50%) of a polypeptide as shown in a certain defined sequence. Clearly these lengths are exemplary, and any length that is supported by the specification, including the Sequence Listing, tables, and figures, may be encompassed by the present embodiments.

A fragment of SEQ ID NO:19-36 comprises a region of unique polynucleotide sequence
30 that specifically identifies SEQ ID NO:19-36, for example, as distinct from any other sequence in the genome from which the fragment was obtained. A fragment of SEQ ID NO:19-36 is useful, for example, in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:19-36 from related polynucleotide sequences. The precise length of a fragment of SEQ ID NO:19-36 and the region of SEQ ID NO:19-36 to which the fragment corresponds are routinely
35 determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A fragment of SEQ ID NO:1-18 is encoded by a fragment of SEQ ID NO:19-36. A fragment of SEQ ID NO:1-18 comprises a region of unique amino acid sequence that specifically identifies SEQ ID NO:1-18. For example, a fragment of SEQ ID NO:1-18 is useful as an immunogenic peptide for the development of antibodies that specifically recognize SEQ ID NO:1-18. The precise length of a fragment of SEQ ID NO:1-18 and the region of SEQ ID NO:1-18 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A “full length” polynucleotide sequence is one containing at least a translation initiation codon (e.g., methionine) followed by an open reading frame and a translation termination codon. A “full length” polynucleotide sequence encodes a “full length” polypeptide sequence.

“Homology” refers to sequence similarity or, interchangeably, sequence identity, between two or more polynucleotide sequences or two or more polypeptide sequences.

The terms “percent identity” and “% identity,” as applied to polynucleotide sequences, refer to the percentage of residue matches between at least two polynucleotide sequences aligned using a standardized algorithm. Such an algorithm may insert, in a standardized and reproducible way, gaps in the sequences being compared in order to optimize alignment between two sequences, and therefore achieve a more meaningful comparison of the two sequences.

Percent identity between polynucleotide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program. This program is part of the LASERGENE software package, a suite of molecular biological analysis programs (DNASTAR, Madison WI). CLUSTAL V is described in Higgins, D.G. and P.M. Sharp (1989) CABIOS 5:151-153 and in Higgins, D.G. et al. (1992) CABIOS 8:189-191. For pairwise alignments of polynucleotide sequences, the default parameters are set as follows: Ktuple=2, gap penalty=5, window=4, and “diagonals saved”=4. The “weighted” residue weight table is selected as the default. Percent identity is reported by CLUSTAL V as the “percent similarity” between aligned polynucleotide sequences.

Alternatively, a suite of commonly used and freely available sequence comparison algorithms is provided by the National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST) (Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410), which is available from several sources, including the NCBI, Bethesda, MD, and on the Internet at <http://www.ncbi.nlm.nih.gov/BLAST/>. The BLAST software suite includes various sequence analysis programs including “blastn,” that is used to align a known polynucleotide sequence with other polynucleotide sequences from a variety of databases. Also available is a tool called “BLAST 2 Sequences” that is used for direct pairwise comparison of two nucleotide sequences. “BLAST 2 Sequences” can be accessed and used interactively at <http://www.ncbi.nlm.nih.gov/gorf/bl2.html>.

The “BLAST 2 Sequences” tool can be used for both blastn and blastp (discussed below). BLAST programs are commonly used with gap and other parameters set to default settings. For example, to compare two nucleotide sequences, one may use blastn with the “BLAST 2 Sequences” tool Version 2.0.12 (April-21-2000) set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62

Reward for match: 1

Penalty for mismatch: -2

Open Gap: 5 and Extension Gap: 2 penalties

10 *Gap x drop-off: 50*

Expect: 10

Word Size: 11

Filter: on

Percent identity may be measured over the length of an entire defined sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined sequence, for instance, a fragment of at least 20, at least 30, at least 40, at least 50, at least 70, at least 100, or at least 200 contiguous nucleotides. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures, or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

Nucleic acid sequences that do not show a high degree of identity may nevertheless encode similar amino acid sequences due to the degeneracy of the genetic code. It is understood that changes in a nucleic acid sequence can be made using this degeneracy to produce multiple nucleic acid sequences that all encode substantially the same protein.

25 The phrases “percent identity” and “% identity,” as applied to polypeptide sequences, refer to the percentage of residue matches between at least two polypeptide sequences aligned using a standardized algorithm. Methods of polypeptide sequence alignment are well-known. Some alignment methods take into account conservative amino acid substitutions. Such conservative substitutions, explained in more detail above, generally preserve the charge and hydrophobicity at the site of substitution, thus preserving the structure (and therefore function) of the polypeptide.

Percent identity between polypeptide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program (described and referenced above). For pairwise alignments of polypeptide sequences using CLUSTAL V, the default parameters are set as follows: Ktuple=1, gap penalty=3, window=5, and “diagonals saved”=5. The PAM250 matrix is selected as the default

residue weight table. As with polynucleotide alignments, the percent identity is reported by CLUSTAL V as the “percent similarity” between aligned polypeptide sequence pairs.

Alternatively the NCBI BLAST software suite may be used. For example, for a pairwise comparison of two polypeptide sequences, one may use the “BLAST 2 Sequences” tool Version 2.0.12 (April-21-2000) with blastp set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62

Open Gap: 11 and Extension Gap: 1 penalties

Gap x drop-off: 50

10 *Expect: 10*

Word Size: 3

Filter: on

Percent identity may be measured over the length of an entire defined polypeptide sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined polypeptide sequence, for instance, a fragment of at least 15, at least 20, at least 30, at least 40, at least 50, at least 70 or at least 150 contiguous residues. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

20 “Human artificial chromosomes” (HACs) are linear microchromosomes which may contain DNA sequences of about 6 kb to 10 Mb in size and which contain all of the elements required for chromosome replication, segregation and maintenance.

The term “humanized antibody” refers to an antibody molecule in which the amino acid sequence in the non-antigen binding regions has been altered so that the antibody more closely resembles a human antibody, and still retains its original binding ability.

“Hybridization” refers to the process by which a polynucleotide strand anneals with a complementary strand through base pairing under defined hybridization conditions. Specific hybridization is an indication that two nucleic acid sequences share a high degree of complementarity. Specific hybridization complexes form under permissive annealing conditions and remain hybridized after the “washing” step(s). The washing step(s) is particularly important in determining the stringency of the hybridization process, with more stringent conditions allowing less non-specific binding, i.e., binding between pairs of nucleic acid strands that are not perfectly matched. Permissive conditions for annealing of nucleic acid sequences are routinely determinable by one of ordinary skill in the art and may be consistent among hybridization experiments, whereas wash conditions may be varied among experiments to achieve the desired stringency, and therefore

hybridization specificity. Permissive annealing conditions occur, for example, at 68°C in the presence of about 6 x SSC, about 1% (w/v) SDS, and about 100 µg/ml sheared, denatured salmon sperm DNA.

Generally, stringency of hybridization is expressed, in part, with reference to the
5 temperature under which the wash step is carried out. Such wash temperatures are typically selected to be about 5°C to 20°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. An equation for calculating T_m and conditions for nucleic acid hybridization are well known and can be found in
10 Sambrook, J. et al. (1989) Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; specifically see volume 2, chapter 9.

High stringency conditions for hybridization between polynucleotides of the present invention include wash conditions of 68°C in the presence of about 0.2 x SSC and about 0.1% SDS, for 1 hour. Alternatively, temperatures of about 65°C, 60°C, 55°C, or 42°C may be used. SSC
15 concentration may be varied from about 0.1 to 2 x SSC, with SDS being present at about 0.1%. Typically, blocking reagents are used to block non-specific hybridization. Such blocking reagents include, for instance, sheared and denatured salmon sperm DNA at about 100-200 µg/ml. Organic solvent, such as formamide at a concentration of about 35-50% v/v, may also be used under particular circumstances, such as for RNA:DNA hybridizations. Useful variations on these wash
20 conditions will be readily apparent to those of ordinary skill in the art. Hybridization, particularly under high stringency conditions, may be suggestive of evolutionary similarity between the nucleotides. Such similarity is strongly indicative of a similar role for the nucleotides and their encoded polypeptides.

The term "hybridization complex" refers to a complex formed between two nucleic acid
25 sequences by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g., C_0t or R_0t analysis) or formed between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been fixed).

30 The words "insertion" and "addition" refer to changes in an amino acid or nucleotide sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively.

"Immune response" can refer to conditions associated with inflammation, trauma, immune disorders, or infectious or genetic disease, etc. These conditions can be characterized by expression of various factors, e.g., cytokines, chemokines, and other signaling molecules, which may affect
35 cellular and systemic defense systems.

An "immunogenic fragment" is a polypeptide or oligopeptide fragment of PMMM which is capable of eliciting an immune response when introduced into a living organism, for example, a mammal. The term "immunogenic fragment" also includes any polypeptide or oligopeptide fragment of PMMM which is useful in any of the antibody production methods disclosed herein or
5 known in the art.

The term "microarray" refers to an arrangement of a plurality of polynucleotides, polypeptides, or other chemical compounds on a substrate.

The terms "element" and "array element" refer to a polynucleotide, polypeptide, or other chemical compound having a unique and defined position on a microarray.

10 The term "modulate" refers to a change in the activity of PMMM. For example, modulation may cause an increase or a decrease in protein activity, binding characteristics, or any other biological, functional, or immunological properties of PMMM.

The phrases "nucleic acid" and "nucleic acid sequence" refer to a nucleotide, oligonucleotide, polynucleotide, or any fragment thereof. These phrases also refer to DNA or RNA
15 of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material.

"Operably linked" refers to the situation in which a first nucleic acid sequence is placed in a functional relationship with a second nucleic acid sequence. For instance, a promoter is operably
20 linked to a coding sequence if the promoter affects the transcription or expression of the coding sequence. Operably linked DNA sequences may be in close proximity or contiguous and, where necessary to join two protein coding regions, in the same reading frame.

"Peptide nucleic acid" (PNA) refers to an antisense molecule or anti-gene agent which comprises an oligonucleotide of at least about 5 nucleotides in length linked to a peptide backbone
25 of amino acid residues ending in lysine. The terminal lysine confers solubility to the composition. PNAs preferentially bind complementary single stranded DNA or RNA and stop transcript elongation, and may be pegylated to extend their lifespan in the cell.

"Post-translational modification" of an PMMM may involve lipidation, glycosylation, phosphorylation, acetylation, racemization, proteolytic cleavage, and other modifications known in
30 the art. These processes may occur synthetically or biochemically. Biochemical modifications will vary by cell type depending on the enzymatic milieu of PMMM.

"Probe" refers to nucleic acid sequences encoding PMMM, their complements, or fragments thereof, which are used to detect identical, allelic or related nucleic acid sequences. Probes are isolated oligonucleotides or polynucleotides attached to a detectable label or reporter molecule.
35 Typical labels include radioactive isotopes, ligands, chemiluminescent agents, and enzymes.

“Primers” are short nucleic acids, usually DNA oligonucleotides, which may be annealed to a target polynucleotide by complementary base-pairing. The primer may then be extended along the target DNA strand by a DNA polymerase enzyme. Primer pairs can be used for amplification (and identification) of a nucleic acid sequence, e.g., by the polymerase chain reaction (PCR).

5 Probes and primers as used in the present invention typically comprise at least 15 contiguous nucleotides of a known sequence. In order to enhance specificity, longer probes and primers may also be employed, such as probes and primers that comprise at least 20, 25, 30, 40, 50, 60, 70, 80, 90, 100, or at least 150 consecutive nucleotides of the disclosed nucleic acid sequences. Probes and primers may be considerably longer than these examples, and it is understood that any
10 length supported by the specification, including the tables, figures, and Sequence Listing, may be used.

Methods for preparing and using probes and primers are described in the references, for example Sambrook, J. et al. (1989) Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; Ausubel, F.M. et al. (1987) Current Protocols in Molecular
15 Biology, Greene Publ. Assoc. & Wiley-Intersciences, New York NY; Innis, M. et al. (1990) PCR Protocols, A Guide to Methods and Applications, Academic Press, San Diego CA. PCR primer pairs can be derived from a known sequence, for example, by using computer programs intended for that purpose such as Primer (Version 0.5, 1991, Whitehead Institute for Biomedical Research, Cambridge MA).

20 Oligonucleotides for use as primers are selected using software known in the art for such purpose. For example, OLIGO 4.06 software is useful for the selection of PCR primer pairs of up to 100 nucleotides each, and for the analysis of oligonucleotides and larger polynucleotides of up to 5,000 nucleotides from an input polynucleotide sequence of up to 32 kilobases. Similar primer selection programs have incorporated additional features for expanded capabilities. For example,
25 the PrimOU primer selection program (available to the public from the Genome Center at University of Texas South West Medical Center, Dallas TX) is capable of choosing specific primers from megabase sequences and is thus useful for designing primers on a genome-wide scope. The Primer3 primer selection program (available to the public from the Whitehead Institute/MIT Center for Genome Research, Cambridge MA) allows the user to input a “mispriming library,” in which
30 sequences to avoid as primer binding sites are user-specified. Primer3 is useful, in particular, for the selection of oligonucleotides for microarrays. (The source code for the latter two primer selection programs may also be obtained from their respective sources and modified to meet the user’s specific needs.) The PrimeGen program (available to the public from the UK Human Genome Mapping Project Resource Centre, Cambridge UK) designs primers based on multiple
35 sequence alignments, thereby allowing selection of primers that hybridize to either the most

conserved or least conserved regions of aligned nucleic acid sequences. Hence, this program is useful for identification of both unique and conserved oligonucleotides and polynucleotide fragments. The oligonucleotides and polynucleotide fragments identified by any of the above selection methods are useful in hybridization technologies, for example, as PCR or sequencing
5 primers, microarray elements, or specific probes to identify fully or partially complementary polynucleotides in a sample of nucleic acids. Methods of oligonucleotide selection are not limited to those described above.

A "recombinant nucleic acid" is a sequence that is not naturally occurring or has a sequence that is made by an artificial combination of two or more otherwise separated segments of sequence.
10 This artificial combination is often accomplished by chemical synthesis or, more commonly, by the artificial manipulation of isolated segments of nucleic acids, e.g., by genetic engineering techniques such as those described in Sambrook, *supra*. The term recombinant includes nucleic acids that have been altered solely by addition, substitution, or deletion of a portion of the nucleic acid. Frequently, a recombinant nucleic acid may include a nucleic acid sequence operably linked to a promoter
15 sequence. Such a recombinant nucleic acid may be part of a vector that is used, for example, to transform a cell.

Alternatively, such recombinant nucleic acids may be part of a viral vector, e.g., based on a vaccinia virus, that could be used to vaccinate a mammal wherein the recombinant nucleic acid is expressed, inducing a protective immunological response in the mammal.

20 A "regulatory element" refers to a nucleic acid sequence usually derived from untranslated regions of a gene and includes enhancers, promoters, introns, and 5' and 3' untranslated regions (UTRs). Regulatory elements interact with host or viral proteins which control transcription, translation, or RNA stability.

"Reporter molecules" are chemical or biochemical moieties used for labeling a nucleic acid,
25 amino acid, or antibody. Reporter molecules include radionuclides; enzymes; fluorescent, chemiluminescent, or chromogenic agents; substrates; cofactors; inhibitors; magnetic particles; and other moieties known in the art.

An "RNA equivalent," in reference to a DNA sequence, is composed of the same linear sequence of nucleotides as the reference DNA sequence with the exception that all occurrences of
30 the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

The term "sample" is used in its broadest sense. A sample suspected of containing PMMM, nucleic acids encoding PMMM, or fragments thereof may comprise a bodily fluid; an extract from a cell, chromosome, organelle, or membrane isolated from a cell; a cell; genomic DNA, RNA, or
35 cDNA, in solution or bound to a substrate; a tissue; a tissue print; etc.

The terms "specific binding" and "specifically binding" refer to that interaction between a protein or peptide and an agonist, an antibody, an antagonist, a small molecule, or any natural or synthetic binding composition. The interaction is dependent upon the presence of a particular structure of the protein, e.g., the antigenic determinant or epitope, recognized by the binding
5 molecule. For example, if an antibody is specific for epitope "A," the presence of a polypeptide comprising the epitope A, or the presence of free unlabeled A, in a reaction containing free labeled A and the antibody will reduce the amount of labeled A that binds to the antibody.

The term "substantially purified" refers to nucleic acid or amino acid sequences that are removed from their natural environment and are isolated or separated, and are at least 60% free,
10 preferably at least 75% free, and most preferably at least 90% free from other components with which they are naturally associated.

A "substitution" refers to the replacement of one or more amino acid residues or nucleotides by different amino acid residues or nucleotides, respectively.

"Substrate" refers to any suitable rigid or semi-rigid support including membranes, filters,
15 chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, tubing, plates, polymers, microparticles and capillaries. The substrate can have a variety of surface forms, such as wells, trenches, pins, channels and pores, to which polynucleotides or polypeptides are bound.

A "transcript image" or "expression profile" refers to the collective pattern of gene expression by a particular cell type or tissue under given conditions at a given time.

20 "Transformation" describes a process by which exogenous DNA is introduced into a recipient cell. Transformation may occur under natural or artificial conditions according to various methods well known in the art, and may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method for transformation is selected based on the type of host cell being transformed and may include, but is not limited to,
25 bacteriophage or viral infection, electroporation, heat shock, lipofection, and particle bombardment. The term "transformed cells" includes stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome, as well as transiently transformed cells which express the inserted DNA or RNA for limited periods of time.

30 A "transgenic organism," as used herein, is any organism, including but not limited to animals and plants, in which one or more of the cells of the organism contains heterologous nucleic acid introduced by way of human intervention, such as by transgenic techniques well known in the art. The nucleic acid is introduced into the cell, directly or indirectly by introduction into a precursor of the cell, by way of deliberate genetic manipulation, such as by microinjection or by
35 infection with a recombinant virus. The term genetic manipulation does not include classical

cross-breeding, or in vitro fertilization, but rather is directed to the introduction of a recombinant DNA molecule. The transgenic organisms contemplated in accordance with the present invention include bacteria, cyanobacteria, fungi, plants and animals. The isolated DNA of the present invention can be introduced into the host by methods known in the art, for example infection, 5 transfection, transformation or transconjugation. Techniques for transferring the DNA of the present invention into such organisms are widely known and provided in references such as Sambrook et al. (1989), supra.

A “variant” of a particular nucleic acid sequence is defined as a nucleic acid sequence having at least 40% sequence identity to the particular nucleic acid sequence over a certain length of 10 one of the nucleic acid sequences using blastn with the “BLAST 2 Sequences” tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of nucleic acids may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% or greater sequence identity over a certain defined length. A variant may be described as, for 15 example, an “allelic” (as defined above), “splice,” “species,” or “polymorphic” variant. A splice variant may have significant identity to a reference molecule, but will generally have a greater or lesser number of polynucleotides due to alternate splicing of exons during mRNA processing. The corresponding polypeptide may possess additional functional domains or lack domains that are present in the reference molecule. Species variants are polynucleotide sequences that vary from one 20 species to another. The resulting polypeptides will generally have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. Polymorphic variants also may encompass “single nucleotide polymorphisms” (SNPs) in which the polynucleotide sequence varies by one nucleotide base. The presence of SNPs may be indicative of, for example, a certain population, a 25 disease state, or a propensity for a disease state.

A “variant” of a particular polypeptide sequence is defined as a polypeptide sequence having at least 40% sequence identity to the particular polypeptide sequence over a certain length of one of the polypeptide sequences using blastp with the “BLAST 2 Sequences” tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of polypeptides may show, for example, at 30 least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% or greater sequence identity over a certain defined length of one of the polypeptides.

THE INVENTION

The invention is based on the discovery of new human protein modification and maintenance molecules (PMMM), the polynucleotides encoding PMMM, and the use of these compositions for the diagnosis, treatment, or prevention of gastrointestinal, cardiovascular, autoimmune/inflammatory, cell proliferative, developmental, epithelial, neurological, and
 5 reproductive disorders.

Table 1 summarizes the nomenclature for the full length polynucleotide and polypeptide sequences of the invention. Each polynucleotide and its corresponding polypeptide are correlated to a single Incyte project identification number (Incyte Project ID). Each polypeptide sequence is denoted by both a polypeptide sequence identification number (Polypeptide SEQ ID NO:) and an
 10 Incyte polypeptide sequence number (Incyte Polypeptide ID) as shown. Each polynucleotide sequence is denoted by both a polynucleotide sequence identification number (Polynucleotide SEQ ID NO:) and an Incyte polynucleotide consensus sequence number (Incyte Polynucleotide ID) as shown.

Table 2 shows sequences with homology to the polypeptides of the invention as identified
 15 by BLAST analysis against the GenBank protein (genpept) database. Columns 1 and 2 show the polypeptide sequence identification number (Polypeptide SEQ ID NO:) and the corresponding Incyte polypeptide sequence number (Incyte Polypeptide ID) for polypeptides of the invention. Column 3 shows the GenBank identification number (GenBank ID NO:) of the nearest GenBank homolog. Column 4 shows the probability scores for the matches between each polypeptide and its
 20 homolog(s). Column 5 shows the annotation of the GenBank homolog(s) along with relevant citations where applicable, all of which are expressly incorporated by reference herein.

Table 3 shows various structural features of the polypeptides of the invention. Columns 1 and 2 show the polypeptide sequence identification number (SEQ ID NO:) and the corresponding Incyte polypeptide sequence number (Incyte Polypeptide ID) for each polypeptide of the invention.
 25 Column 3 shows the number of amino acid residues in each polypeptide. Column 4 shows potential phosphorylation sites, and column 5 shows potential glycosylation sites, as determined by the MOTIFS program of the GCG sequence analysis software package (Genetics Computer Group, Madison WI). Column 6 shows amino acid residues comprising signature sequences, domains, and motifs. Column 7 shows analytical methods for protein structure/function analysis and in some
 30 cases, searchable databases to which the analytical methods were applied.

Together, Tables 2 and 3 summarize the properties of polypeptides of the invention, and these properties establish that the claimed polypeptides are protein modification and maintenance molecules. For example, SEQ ID NO:1 is 43% identical to human ubiquitin hydrolase (GenBank ID g1666075) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.)
 35 The BLAST probability score is 1.8e-233, which indicates the probability of obtaining the observed

polypeptide sequence alignment by chance. SEQ ID NO:1 also contains a ubiquitin carboxyl-terminal hydrolase, family 2 active site domain as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.) Data from BLIMPS and MOTIFS analyses provide further
5 corroborative evidence that SEQ ID NO:1 is a ubiquitin carboxyl-terminal hydrolase.

In another example, SEQ ID NO:7 is 91% identical to canine signal peptidase 21 kDa subunit (GenBank ID g164084) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is $4.8e-84$, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:7 also contains
10 signal peptidase I domain as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.) Data from BLIMPS, MOTIFS, and PROFILESCAN analyses provide further corroborative evidence that SEQ ID NO:7 is a signal peptidase.

In another example, SEQ ID NO:12 is 60% identical over 856 amino acid residues to human
15 dipeptidyl peptidase 8 (GenBank ID g11095188) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is $2.3e-299$, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:12 also contains a prolyl oligopeptidase family domain as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of
20 conserved protein family domains. (See Table 3.) Data from BLIMPS, MOTIFS, and BLAST analyses provide further corroborative evidence that SEQ ID NO:12 is a serine protease.

In another example, SEQ ID NO:13 is 42% identical to rat prostatic (GenBank ID g11181573) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is $4.6e-56$, which indicates the probability of obtaining the observed
25 polypeptide sequence alignment by chance. SEQ ID NO:13 also contains a trypsin domain as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.) Data from BLIMPS, MOTIFS, and PROFILESCAN analyses provide further corroborative evidence that SEQ ID NO:13 is a serine protease.

30 As a further example, SEQ ID NO:18 is 40% identical to Xenopus laevis oviductin, an oviductal protease (GenBank ID g1754714) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is $1.1e-115$, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:18 also contains trypsin domains and a CUB domain as determined by searching for statistically
35 significant matches in the hidden Markov model (HMM)-based PFAM database of conserved

protein family domains. (See Table 3.) Data from BLIMPS, MOTIFS, and PROFILESCAN analyses provide further corroborative evidence that SEQ ID NO:18 is a trypsin family serine protease. SEQ ID NO:2-6, SEQ ID NO:8-11 and SEQ ID NO:14-17 were analyzed and annotated in a similar manner. The algorithms and parameters for the analysis of SEQ ID NO:1-18 are described
 5 in Table 7.

As shown in Table 4, the full length polynucleotide sequences of the present invention were assembled using cDNA sequences or coding (exon) sequences derived from genomic DNA, or any combination of these two types of sequences. Column 1 lists the polynucleotide sequence identification number (Polynucleotide SEQ ID NO:), the corresponding Incyte polynucleotide
 10 consensus sequence number (Incyte ID) for each polynucleotide of the invention, and the length of each polynucleotide sequence in basepairs. Column 2 shows the nucleotide start (5') and stop (3') positions of the cDNA and/or genomic sequences used to assemble the full length polynucleotide sequences of the invention, and of fragments of the polynucleotide sequences which are useful, for example, in hybridization or amplification technologies that identify SEQ ID NO:19-36 or that
 15 distinguish between SEQ ID NO:19-36 and related polynucleotide sequences.

The polynucleotide fragments described in Column 2 of Table 4 may refer specifically, for example, to Incyte cDNAs derived from tissue-specific cDNA libraries or from pooled cDNA libraries. Alternatively, the polynucleotide fragments described in column 2 may refer to GenBank cDNAs or ESTs which contributed to the assembly of the full length polynucleotide sequences. In
 20 addition, the polynucleotide fragments described in column 2 may identify sequences derived from the ENSEMBL (The Sanger Centre, Cambridge, UK) database (*i.e.*, those sequences including the designation "ENST"). Alternatively, the polynucleotide fragments described in column 2 may be derived from the NCBI RefSeq Nucleotide Sequence Records Database (*i.e.*, those sequences including the designation "NM" or "NT") or the NCBI RefSeq Protein Sequence Records (*i.e.*,
 25 those sequences including the designation "NP"). Alternatively, the polynucleotide fragments described in column 2 may refer to assemblages of both cDNA and Genscan-predicted exons brought together by an "exon stitching" algorithm. For example, a polynucleotide sequence identified as FL_XXXXXX_N₁_N₂_YYYYY_N₃_N₄ represents a "stitched" sequence in which XXXXXX is the identification number of the cluster of sequences to which the algorithm was
 30 applied, and YYYYY is the number of the prediction generated by the algorithm, and N_{1,2,3...}, if present, represent specific exons that may have been manually edited during analysis (See Example V). Alternatively, the polynucleotide fragments in column 2 may refer to assemblages of exons brought together by an "exon-stretching" algorithm. For example, a polynucleotide sequence identified as FLXXXXXX_gAAAAA_gBBBBB_1_N is a "stretched" sequence, with XXXXXX being
 35 the Incyte project identification number, gAAAAA being the GenBank identification number of the

human genomic sequence to which the “exon-stretching” algorithm was applied, *gBBBBB* being the GenBank identification number or NCBI RefSeq identification number of the nearest GenBank protein homolog, and *N* referring to specific exons (See Example V). In instances where a RefSeq sequence was used as a protein homolog for the “exon-stretching” algorithm, a RefSeq identifier
 5 (denoted by “NM,” “NP,” or “NT”) may be used in place of the GenBank identifier (*i.e.*, *gBBBBB*).

Alternatively, a prefix identifies component sequences that were hand-edited, predicted from genomic DNA sequences, or derived from a combination of sequence analysis methods. The following Table lists examples of component sequence prefixes and corresponding sequence analysis methods associated with the prefixes (see Example IV and Example V).

10	Prefix	Type of analysis and/or examples of programs
	GNN, GFG, ENST	Exon prediction from genomic sequences using, for example, GENSCAN (Stanford University, CA, USA) or FGENES (Computer Genomics Group, The Sanger Centre, Cambridge, UK).
	GBI	Hand-edited analysis of genomic sequences.
	FL	Stitched or stretched genomic sequences (see Example V).
15	INCY	Full length transcript and exon prediction from mapping of EST sequences to the genome. Genomic location and EST composition data are combined to predict the exons and resulting transcript.

In some cases, Incyte cDNA coverage redundant with the sequence coverage shown in Table 4 was obtained to confirm the final consensus polynucleotide sequence, but the relevant Incyte cDNA identification numbers are not shown.

20 Table 5 shows the representative cDNA libraries for those full length polynucleotide sequences which were assembled using Incyte cDNA sequences. The representative cDNA library is the Incyte cDNA library which is most frequently represented by the Incyte cDNA sequences which were used to assemble and confirm the above polynucleotide sequences. The tissues and vectors which were used to construct the cDNA libraries shown in Table 5 are described in Table 6.

25 The invention also encompasses PMMM variants. A preferred PMMM variant is one which has at least about 80%, or alternatively at least about 90%, or even at least about 95% amino acid sequence identity to the PMMM amino acid sequence, and which contains at least one functional or structural characteristic of PMMM.

The invention also encompasses polynucleotides which encode PMMM. In a particular
 30 embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:19-36, which encodes PMMM. The polynucleotide sequences of SEQ ID NO:19-36, as presented in the Sequence Listing, embrace the equivalent RNA

sequences, wherein occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

The invention also encompasses a variant of a polynucleotide sequence encoding PMMM. In particular, such a variant polynucleotide sequence will have at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to the polynucleotide sequence encoding PMMM. A particular aspect of the invention encompasses a variant of a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:19-36 which has at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NO:19-36. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of PMMM.

In addition, or in the alternative, a polynucleotide variant of the invention is a splice variant of a polynucleotide sequence encoding PMMM. A splice variant may have portions which have significant sequence identity to the polynucleotide sequence encoding PMMM, but will generally have a greater or lesser number of polynucleotides due to additions or deletions of blocks of sequence arising from alternate splicing of exons during mRNA processing. A splice variant may have less than about 70%, or alternatively less than about 60%, or alternatively less than about 50% polynucleotide sequence identity to the polynucleotide sequence encoding PMMM over its entire length; however, portions of the splice variant will have at least about 70%, or alternatively at least about 85%, or alternatively at least about 95%, or alternatively 100% polynucleotide sequence identity to portions of the polynucleotide sequence encoding PMMM. Any one of the splice variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of PMMM.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of polynucleotide sequences encoding PMMM, some bearing minimal similarity to the polynucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of polynucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the polynucleotide sequence of naturally occurring PMMM, and all such variations are to be considered as being specifically disclosed.

Although nucleotide sequences which encode PMMM and its variants are generally capable of hybridizing to the nucleotide sequence of the naturally occurring PMMM under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding PMMM or its derivatives possessing a substantially different codon usage, e.g., inclusion of non-

naturally occurring codons. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding PMMM and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of DNA sequences which encode PMMM and PMMM derivatives, or fragments thereof, entirely by synthetic chemistry. After production, the synthetic sequence may be inserted into any of the many available expression vectors and cell systems using reagents well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding PMMM or any fragment thereof.

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed polynucleotide sequences, and, in particular, to those shown in SEQ ID NO:19-36 and fragments thereof under various conditions of stringency. (See, e.g., Wahl, G.M. and S.L. Berger (1987) *Methods Enzymol.* 152:399-407; Kimmel, A.R. (1987) *Methods Enzymol.* 152:507-511.) Hybridization conditions, including annealing and wash conditions, are described in "Definitions."

Methods for DNA sequencing are well known in the art and may be used to practice any of the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE (US Biochemical, Cleveland OH), Taq polymerase (Applied Biosystems), thermostable T7 polymerase (Amersham Pharmacia Biotech, Piscataway NJ), or combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system (Life Technologies, Gaithersburg MD). Preferably, sequence preparation is automated with machines such as the MICROLAB 2200 liquid transfer system (Hamilton, Reno NV), PTC200 thermal cycler (MJ Research, Watertown MA) and ABI CATALYST 800 thermal cycler (Applied Biosystems). Sequencing is then carried out using either the ABI 373 or 377 DNA sequencing system (Applied Biosystems), the MEGABACE 1000 DNA sequencing system (Molecular Dynamics, Sunnyvale CA), or other systems known in the art. The resulting sequences are analyzed using a variety of algorithms which are well known in the art. (See, e.g., Ausubel, F.M. (1997) Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7; Meyers, R.A. (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856-853.)

The nucleic acid sequences encoding PMMM may be extended utilizing a partial nucleotide sequence and employing various PCR-based methods known in the art to detect upstream sequences, such as promoters and regulatory elements. For example, one method which may be

employed, restriction-site PCR, uses universal and nested primers to amplify unknown sequence from genomic DNA within a cloning vector. (See, e.g., Sarkar, G. (1993) PCR Methods Applic. 2:318-322.) Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown sequence from a circularized template. The template is derived from restriction fragments comprising a known genomic locus and surrounding sequences. (See, e.g., Triglia, T. et al. (1988) Nucleic Acids Res. 16:8186.) A third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA. (See, e.g., Lagerstrom, M. et al. (1991) PCR Methods Applic. 1:111-119.) In this method, multiple restriction enzyme digestions and ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art. (See, e.g., Parker, J.D. et al. (1991) Nucleic Acids Res. 19:3055-3060). Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries (Clontech, Palo Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO 4.06 primer analysis software (National Biosciences, Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of about 68°C to 72°C.

When screening for full length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. In addition, random-primed libraries, which often include sequences containing the 5' regions of genes, are preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence into 5' non-transcribed regulatory regions.

Capillary electrophoresis systems which are commercially available may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different nucleotide-specific, laser-stimulated fluorescent dyes, and a charge coupled device camera for detection of the emitted wavelengths. Output/light intensity may be converted to electrical signal using appropriate software (e.g., GENOTYPER and SEQUENCE NAVIGATOR, Applied Biosystems), and the entire process from loading of samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for sequencing small DNA fragments which may be present in limited amounts in a particular sample.

In another embodiment of the invention, polynucleotide sequences or fragments thereof which encode PMMM may be cloned in recombinant DNA molecules that direct expression of PMMM, or fragments or functional equivalents thereof, in appropriate host cells. Due to the

inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be produced and used to express PMMM.

The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter PMMM-encoding sequences for a variety of purposes including, but not limited to, modification of the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, oligonucleotide-mediated site-directed mutagenesis may be used to introduce mutations that create new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

The nucleotides of the present invention may be subjected to DNA shuffling techniques such as MOLECULARBREEDING (Maxygen Inc., Santa Clara CA; described in U.S. Patent No. 5,837,458; Chang, C.-C. et al. (1999) Nat. Biotechnol. 17:793-797; Christians, F.C. et al. (1999) Nat. Biotechnol. 17:259-264; and Cramer, A. et al. (1996) Nat. Biotechnol. 14:315-319) to alter or improve the biological properties of PMMM, such as its biological or enzymatic activity or its ability to bind to other molecules or compounds. DNA shuffling is a process by which a library of gene variants is produced using PCR-mediated recombination of gene fragments. The library is then subjected to selection or screening procedures that identify those gene variants with the desired properties. These preferred variants may then be pooled and further subjected to recursive rounds of DNA shuffling and selection/screening. Thus, genetic diversity is created through "artificial" breeding and rapid molecular evolution. For example, fragments of a single gene containing random point mutations may be recombined, screened, and then reshuffled until the desired properties are optimized. Alternatively, fragments of a given gene may be recombined with fragments of homologous genes in the same gene family, either from the same or different species, thereby maximizing the genetic diversity of multiple naturally occurring genes in a directed and controllable manner.

In another embodiment, sequences encoding PMMM may be synthesized, in whole or in part, using chemical methods well known in the art. (See, e.g., Caruthers, M.H. et al. (1980) Nucleic Acids Symp. Ser. 7:215-223; and Horn, T. et al. (1980) Nucleic Acids Symp. Ser. 7:225-232.) Alternatively, PMMM itself or a fragment thereof may be synthesized using chemical methods. For example, peptide synthesis can be performed using various solution-phase or solid-phase techniques. (See, e.g., Creighton, T. (1984) Proteins, Structures and Molecular Properties, WH Freeman, New York NY, pp. 55-60; and Roberge, J.Y. et al. (1995) Science 269:202-204.) Automated synthesis may be achieved using the ABI 431A peptide synthesizer (Applied Biosystems). Additionally, the amino acid sequence of PMMM, or any part thereof, may

be altered during direct synthesis and/or combined with sequences from other proteins, or any part thereof, to produce a variant polypeptide or a polypeptide having a sequence of a naturally occurring polypeptide.

The peptide may be substantially purified by preparative high performance liquid
5 chromatography. (See, e.g., Chiez, R.M. and F.Z. Regnier (1990) *Methods Enzymol.* 182:392-421.)
The composition of the synthetic peptides may be confirmed by amino acid analysis or by
sequencing. (See, e.g., Creighton, *supra*, pp. 28-53.)

In order to express a biologically active PMMM, the nucleotide sequences encoding
PMMM or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector
10 which contains the necessary elements for transcriptional and translational control of the inserted
coding sequence in a suitable host. These elements include regulatory sequences, such as
enhancers, constitutive and inducible promoters, and 5' and 3' untranslated regions in the vector and
in polynucleotide sequences encoding PMMM. Such elements may vary in their strength and
specificity. Specific initiation signals may also be used to achieve more efficient translation of
15 sequences encoding PMMM. Such signals include the ATG initiation codon and adjacent
sequences, e.g. the Kozak sequence. In cases where sequences encoding PMMM and its initiation
codon and upstream regulatory sequences are inserted into the appropriate expression vector, no
additional transcriptional or translational control signals may be needed. However, in cases where
only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals
20 including an in-frame ATG initiation codon should be provided by the vector. Exogenous
translational elements and initiation codons may be of various origins, both natural and synthetic.
The efficiency of expression may be enhanced by the inclusion of enhancers appropriate for the
particular host cell system used. (See, e.g., Scharf, D. et al. (1994) *Results Probl. Cell Differ.*
20:125-162.)

25 Methods which are well known to those skilled in the art may be used to construct
expression vectors containing sequences encoding PMMM and appropriate transcriptional and
translational control elements. These methods include *in vitro* recombinant DNA techniques,
synthetic techniques, and *in vivo* genetic recombination. (See, e.g., Sambrook, J. et al. (1989)
Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview NY, ch. 4, 8, and
30 16-17; Ausubel, F.M. et al. (1995) *Current Protocols in Molecular Biology*, John Wiley & Sons,
New York NY, ch. 9, 13, and 16.)

A variety of expression vector/host systems may be utilized to contain and express
sequences encoding PMMM. These include, but are not limited to, microorganisms such as bacteria
transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast
35 transformed with yeast expression vectors; insect cell systems infected with viral expression vectors

(e.g., baculovirus); plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems. (See, e.g., Sambrook, supra; Ausubel, supra; Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509; Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945; Takamatsu, N. (1987) EMBO J. 6:307-311; The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196; Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659; and Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355.) Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted organ, tissue, or cell population. (See, e.g., Di Nicola, M. et al. (1998) Cancer Gen. Ther. 5(6):350-356; Yu, M. et al. (1993) Proc. Natl. Acad. Sci. USA 90(13):6340-6344; Buller, R.M. et al. (1985) Nature 317(6040):813-815; McGregor, D.P. et al. (1994) Mol. Immunol. 31(3):219-226; and Verma, I.M. and N. Somia (1997) Nature 389:239-242.) The invention is not limited by the host cell employed.

15 In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotide sequences encoding PMMM. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding PMMM can be achieved using a multifunctional E. coli vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or PSPORT1 plasmid (Life Technologies). Ligation of sequences encoding PMMM into the vector's multiple cloning site disrupts the *lacZ* gene, allowing a colorimetric screening procedure for identification of transformed bacteria containing recombinant molecules. In addition, these vectors may be useful for in vitro transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence. (See, e.g., Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509.) When large quantities of PMMM are needed, e.g. for the production of antibodies, vectors which direct high level expression of PMMM may be used. For example, vectors containing the strong, inducible SP6 or T7 bacteriophage promoter may be used.

Yeast expression systems may be used for production of PMMM. A number of vectors containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH promoters, may be used in the yeast Saccharomyces cerevisiae or Pichia pastoris. In addition, such vectors direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign sequences into the host genome for stable propagation. (See, e.g., Ausubel, 1995, supra; Bitter, G.A. et al. (1987) Methods Enzymol. 153:516-544; and Scorer, C.A. et al. (1994) Bio/Technology 12:181-184.)

Plant systems may also be used for expression of PMMM. Transcription of sequences encoding PMMM may be driven by viral promoters, e.g., the 35S and 19S promoters of CaMV used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) EMBO J. 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used. (See, e.g., Coruzzi, G. et al. (1984) EMBO J. 3:1671-1680; Broglie, R. et al. (1984) Science 224:838-843; and Winter, J. et al. (1991) Results Probl. Cell Differ. 17:85-105.) These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. (See, e.g., The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196.)

10 In mammalian cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding PMMM may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain infective virus which expresses PMMM in host cells. (See, e.g., Logan, J. and T. Shenk
15 (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659.) In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-based vectors may also be used for high-level protein expression.

Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of DNA than can be contained in and expressed from a plasmid. HACs of about 6 kb to 10 Mb are
20 constructed and delivered via conventional delivery methods (liposomes, polycationic amino polymers, or vesicles) for therapeutic purposes. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355.)

For long term production of recombinant proteins in mammalian systems, stable expression of PMMM in cell lines is preferred. For example, sequences encoding PMMM can be transformed
25 into cell lines using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for about 1 to 2 days in enriched media before being switched to selective media. The purpose of the selectable marker is to confer resistance to a selective agent, and its presence allows growth and recovery of cells which
30 successfully express the introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk⁻* and *apr⁻* cells, respectively. (See, e.g., Wigler, M. et
35 al. (1977) Cell 11:223-232; Lowy, I. et al. (1980) Cell 22:817-823.) Also, antimetabolite, antibiotic,

or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate; *neo* confers resistance to the aminoglycosides neomycin and G-418; and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively. (See, e.g., Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. USA 77:3567-3570; Colbere-Garapin, F. et al. (1981) J. Mol. Biol. 150:1-14.) Additional selectable genes have been described, e.g., *trpB* and *hisD*, which alter cellular requirements for metabolites. (See, e.g., Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. USA 85:8047-8051.) Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP; Clontech), β glucuronidase and its substrate β -glucuronide, or luciferase and its substrate luciferin may be used. These markers can be used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system. (See, e.g., Rhodes, C.A. (1995) Methods Mol. Biol. 55:121-131.)

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, the presence and expression of the gene may need to be confirmed. For example, if the sequence encoding PMMM is inserted within a marker gene sequence, transformed cells containing sequences encoding PMMM can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding PMMM under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

In general, host cells that contain the nucleic acid sequence encoding PMMM and that express PMMM may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations, PCR amplification, and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein sequences.

Immunological methods for detecting and measuring the expression of PMMM using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on PMMM is preferred, but a competitive binding assay may be employed. These and other assays are well known in the art. (See, e.g., Hampton, R. et al. (1990) Serological Methods, a Laboratory Manual, APS Press, St. Paul MN, Sect. IV; Coligan, J.E. et al. (1997) Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York NY; and Pound, J.D. (1998) Immunochemical Protocols, Humana Press, Totowa NJ.)

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding PMMM include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled
5 nucleotide. Alternatively, the sequences encoding PMMM, or any fragments thereof, may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes *in vitro* by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits, such as those provided by
10 Amersham Pharmacia Biotech, Promega (Madison WI), and US Biochemical. Suitable reporter molecules or labels which may be used for ease of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with nucleotide sequences encoding PMMM may be cultured under
15 conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a transformed cell may be secreted or retained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode PMMM may be designed to contain signal sequences which direct secretion of PMMM through a prokaryotic or eukaryotic cell membrane.

20 In addition, a host cell strain may be chosen for its ability to modulate expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" or "pro" form of the protein may also be used to specify protein targeting, folding, and/or activity.
25 Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38) are available from the American Type Culture Collection (ATCC, Manassas VA) and may be chosen to ensure the correct modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid
30 sequences encoding PMMM may be ligated to a heterologous sequence resulting in translation of a fusion protein in any of the aforementioned host systems. For example, a chimeric PMMM protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of PMMM activity. Heterologous protein and peptide moieties may also facilitate purification of fusion proteins using commercially available
35 affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST),

maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable
5 immunoaffinity purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the PMMM encoding sequence and the heterologous protein sequence, so that PMMM may be cleaved away from the heterologous moiety following purification. Methods for fusion protein expression and purification are discussed
10 in Ausubel (1995, supra, ch. 10). A variety of commercially available kits may also be used to facilitate expression and purification of fusion proteins.

In a further embodiment of the invention, synthesis of radiolabeled PMMM may be achieved in vitro using the TNT rabbit reticulocyte lysate or wheat germ extract system (Promega). These systems couple transcription and translation of protein-coding sequences operably associated
15 with the T7, T3, or SP6 promoters. Translation takes place in the presence of a radiolabeled amino acid precursor, for example, ³⁵S-methionine.

PMMM of the present invention or fragments thereof may be used to screen for compounds that specifically bind to PMMM. At least one and up to a plurality of test compounds may be screened for specific binding to PMMM. Examples of test compounds include antibodies,
20 oligonucleotides, proteins (e.g., receptors), or small molecules.

In one embodiment, the compound thus identified is closely related to the natural ligand of PMMM, e.g., a ligand or fragment thereof, a natural substrate, a structural or functional mimetic, or a natural binding partner. (See, e.g., Coligan, J.E. et al. (1991) Current Protocols in Immunology 1(2):_Chapter 5.) Similarly, the compound can be closely related to the natural receptor to which
25 PMMM binds, or to at least a fragment of the receptor, e.g., the ligand binding site. In either case, the compound can be rationally designed using known techniques. In one embodiment, screening for these compounds involves producing appropriate cells which express PMMM, either as a secreted protein or on the cell membrane. Preferred cells include cells from mammals, yeast, Drosophila, or E. coli. Cells expressing PMMM or cell membrane fractions which contain PMMM
30 are then contacted with a test compound and binding, stimulation, or inhibition of activity of either PMMM or the compound is analyzed.

An assay may simply test binding of a test compound to the polypeptide, wherein binding is detected by a fluorophore, radioisotope, enzyme conjugate, or other detectable label. For example, the assay may comprise the steps of combining at least one test compound with PMMM, either in
35 solution or affixed to a solid support, and detecting the binding of PMMM to the compound.

Alternatively, the assay may detect or measure binding of a test compound in the presence of a labeled competitor. Additionally, the assay may be carried out using cell-free preparations, chemical libraries, or natural product mixtures, and the test compound(s) may be free in solution or affixed to a solid support.

5 PMMM of the present invention or fragments thereof may be used to screen for compounds that modulate the activity of PMMM. Such compounds may include agonists, antagonists, or partial or inverse agonists. In one embodiment, an assay is performed under conditions permissive for PMMM activity, wherein PMMM is combined with at least one test compound, and the activity of PMMM in the presence of a test compound is compared with the activity of PMMM in the
10 absence of the test compound. A change in the activity of PMMM in the presence of the test compound is indicative of a compound that modulates the activity of PMMM. Alternatively, a test compound is combined with an in vitro or cell-free system comprising PMMM under conditions suitable for PMMM activity, and the assay is performed. In either of these assays, a test compound which modulates the activity of PMMM may do so indirectly and need not come in direct contact
15 with the test compound. At least one and up to a plurality of test compounds may be screened.

In another embodiment, polynucleotides encoding PMMM or their mammalian homologs may be “knocked out” in an animal model system using homologous recombination in embryonic stem (ES) cells. Such techniques are well known in the art and are useful for the generation of animal models of human disease. (See, e.g., U.S. Patent No. 5,175,383 and U.S. Patent No.
20 5,767,337.) For example, mouse ES cells, such as the mouse 129/SvJ cell line, are derived from the early mouse embryo and grown in culture. The ES cells are transformed with a vector containing the gene of interest disrupted by a marker gene, e.g., the neomycin phosphotransferase gene (neo; Capecchi, M.R. (1989) Science 244:1288-1292). The vector integrates into the corresponding region of the host genome by homologous recombination. Alternatively, homologous
25 recombination takes place using the Cre-loxP system to knockout a gene of interest in a tissue- or developmental stage-specific manner (Marth, J.D. (1996) Clin. Invest. 97:1999-2002; Wagner, K.U. et al. (1997) Nucleic Acids Res. 25:4323-4330). Transformed ES cells are identified and microinjected into mouse cell blastocysts such as those from the C57BL/6 mouse strain. The blastocysts are surgically transferred to pseudopregnant dams, and the resulting chimeric progeny
30 are genotyped and bred to produce heterozygous or homozygous strains. Transgenic animals thus generated may be tested with potential therapeutic or toxic agents.

Polynucleotides encoding PMMM may also be manipulated in vitro in ES cells derived from human blastocysts. Human ES cells have the potential to differentiate into at least eight separate cell lineages including endoderm, mesoderm, and ectodermal cell types. These cell

lineages differentiate into, for example, neural cells, hematopoietic lineages, and cardiomyocytes (Thomson, J.A. et al. (1998) Science 282:1145-1147).

Polynucleotides encoding PMMM can also be used to create “knockin” humanized animals (pigs) or transgenic animals (mice or rats) to model human disease. With knockin technology, a
 5 region of a polynucleotide encoding PMMM is injected into animal ES cells, and the injected sequence integrates into the animal cell genome. Transformed cells are injected into blastulae, and the blastulae are implanted as described above. Transgenic progeny or inbred lines are studied and treated with potential pharmaceutical agents to obtain information on treatment of a human disease. Alternatively, a mammal inbred to overexpress PMMM, e.g., by secreting PMMM in its milk, may
 10 also serve as a convenient source of that protein (Janne, J. et al. (1998) Biotechnol. Annu. Rev. 4:55-74).

THERAPEUTICS

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between regions of PMMM and protein modification and maintenance molecules. In addition,
 15 examples of tissues expressing PMMM are brain, lung, digestive, urogenital, small intestine, kidney, tumorous tissues, such as endocrine, esophageal and prostate tumors and tissue affected by Huntington’s disease. Examples can also be found in Table 6. Therefore, PMMM appears to play a role in gastrointestinal, cardiovascular, autoimmune/inflammatory, cell proliferative, developmental, epithelial, neurological, and reproductive disorders. In the treatment of disorders associated with
 20 increased PMMM expression or activity, it is desirable to decrease the expression or activity of PMMM. In the treatment of disorders associated with decreased PMMM expression or activity, it is desirable to increase the expression or activity of PMMM.

Therefore, in one embodiment, PMMM or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or
 25 activity of PMMM. Examples of such disorders include, but are not limited to, a gastrointestinal disorder, such as dysphagia, peptic esophagitis, esophageal spasm, esophageal stricture, esophageal carcinoma, dyspepsia, indigestion, gastritis, gastric carcinoma, anorexia, nausea, emesis, gastroparesis, antral or pyloric edema, abdominal angina, pyrosis, gastroenteritis, intestinal obstruction, infections of the intestinal tract, peptic ulcer, cholelithiasis, cholecystitis, cholestasis,
 30 pancreatitis, pancreatic carcinoma, biliary tract disease, hepatitis, hyperbilirubinemia, cirrhosis, passive congestion of the liver, hepatoma, infectious colitis, ulcerative colitis, ulcerative proctitis, Crohn’s disease, Whipple’s disease, Mallory-Weiss syndrome, colonic carcinoma, colonic obstruction, irritable bowel syndrome, short bowel syndrome, diarrhea, constipation, gastrointestinal hemorrhage, acquired immunodeficiency syndrome (AIDS) enteropathy, jaundice, hepatic
 35 encephalopathy, hepatorenal syndrome, hepatic steatosis, hemochromatosis, Wilson’s disease,

alpha₁-antitrypsin deficiency, Reye's syndrome, primary sclerosing cholangitis, liver infarction, portal vein obstruction and thrombosis, centrilobular necrosis, peliosis hepatis, hepatic vein thrombosis, veno-occlusive disease, preeclampsia, eclampsia, acute fatty liver of pregnancy, intrahepatic cholestasis of pregnancy, and hepatic tumors including nodular hyperplasias, adenomas,

5 and carcinomas; a cardiovascular disorder, such as arteriovenous fistula, atherosclerosis, hypertension, vasculitis, Raynaud's disease, aneurysms, arterial dissections, varicose veins, thrombophlebitis and phlebothrombosis, vascular tumors, and complications of thrombolysis, balloon angioplasty, vascular replacement, and coronary artery bypass graft surgery, congestive heart failure, ischemic heart disease, angina pectoris, myocardial infarction, hypertensive heart

10 disease, degenerative valvular heart disease, calcific aortic valve stenosis, congenitally bicuspid aortic valve, mitral annular calcification, mitral valve prolapse, rheumatic fever and rheumatic heart disease, infective endocarditis, nonbacterial thrombotic endocarditis, endocarditis of systemic lupus erythematosus, carcinoid heart disease, cardiomyopathy, myocarditis, pericarditis, neoplastic heart disease, congenital heart disease, and complications of cardiac transplantation; an

15 autoimmune/inflammatory disease, such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, atherosclerotic plaque rupture, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis,

20 dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, degradation of articular cartilage, osteoporosis, pancreatitis, polymyositis, psoriasis,

25 Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis,

30 hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas,

35 parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; a

developmental disorder, such as renal tubular acidosis, anemia, Cushing's syndrome, achondroplastic dwarfism, Duchenne and Becker muscular dystrophy, bone resorption, epilepsy, gonadal dysgenesis, WAGR syndrome (Wilms' tumor, aniridia, genitourinary abnormalities, and mental retardation), Smith-Magenis syndrome, myelodysplastic syndrome, hereditary

5 mucoepithelial dysplasia, hereditary keratodermas, hereditary neuropathies such as Charcot-Marie-Tooth disease and neurofibromatosis, hypothyroidism, hydrocephalus, seizure disorders such as Sydenham's chorea and cerebral palsy, spina bifida, anencephaly, craniorachischisis, congenital glaucoma, cataract, age-related macular degeneration, and sensorineural hearing loss; an epithelial disorder, such as dyshidrotic eczema, allergic contact dermatitis, keratosis pilaris, melasma,

10 vitiligo, actinic keratosis, basal cell carcinoma, squamous cell carcinoma, seborrheic keratosis, folliculitis, herpes simplex, herpes zoster, varicella, candidiasis, dermatophytosis, scabies, insect bites, cherry angioma, keloid, dermatofibroma, acrochordons, urticaria, transient acantholytic dermatosis, xerosis, eczema, atopic dermatitis, contact dermatitis, hand eczema, nummular eczema, lichen simplex chronicus, asteatotic eczema, stasis dermatitis and stasis ulceration, seborrheic

15 dermatitis, psoriasis, lichen planus, pityriasis rosea, impetigo, ecthyma, dermatophytosis, tinea versicolor, warts, acne vulgaris, acne rosacea, pemphigus vulgaris, pemphigus foliaceus, paraneoplastic pemphigus, bullous pemphigoid, herpes gestationis, dermatitis herpetiformis, linear IgA disease, epidermolysis bullosa acquisita, dermatomyositis, lupus erythematosus, scleroderma and morphea, erythroderma, alopecia, figurate skin lesions, telangiectasias, hypopigmentation,

20 hyperpigmentation, vesicles/bullae, exanthems, cutaneous drug reactions, papulonodular skin lesions, chronic non-healing wounds, photosensitivity diseases, epidermolysis bullosa simplex, epidermolytic hyperkeratosis, epidermolytic and nonepidermolytic palmoplantar keratoderma, ichthyosis bullosa of Siemens, ichthyosis exfoliativa, keratosis palmaris et plantaris, keratosis palmoplantaris, palmoplantar keratoderma, keratosis punctata, Meesmann's corneal dystrophy,

25 pachyonychia congenita, white sponge nevus, steatocystoma multiplex, epidermal nevi/epidermolytic hyperkeratosis type, monilethrix, trichothiodystrophy, chronic hepatitis/cryptogenic cirrhosis, and colorectal hyperplasia; a neurological disorder, such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders,

30 amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-

35 Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the

nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system including Down syndrome, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and
5 other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathies, myasthenia gravis, periodic paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, seasonal affective disorder (SAD), akathisia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, Tourette's disorder, progressive supranuclear palsy,
10 corticobasal degeneration, and familial frontotemporal dementia; and a reproductive disorder, such as infertility, including tubal disease, ovulatory defects, and endometriosis, a disorder of prolactin production, a disruption of the estrous cycle, a disruption of the menstrual cycle, polycystic ovary syndrome, ovarian hyperstimulation syndrome, an endometrial or ovarian tumor, a uterine fibroid, autoimmune disorders, an ectopic pregnancy, and teratogenesis; cancer of the breast, fibrocystic
15 breast disease, and galactorrhea; a disruption of spermatogenesis, abnormal sperm physiology, cancer of the testis, cancer of the prostate, benign prostatic hyperplasia, prostatitis, Peyronie's disease, impotence, carcinoma of the male breast, and gynecomastia.

In another embodiment, a vector capable of expressing PMMM or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased
20 expression or activity of PMMM including, but not limited to, those described above.

In a further embodiment, a composition comprising a substantially purified PMMM in conjunction with a suitable pharmaceutical carrier may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of PMMM including, but not limited to, those provided above.

25 In still another embodiment, an agonist which modulates the activity of PMMM may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of PMMM including, but not limited to, those listed above.

In a further embodiment, an antagonist of PMMM may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of PMMM. Examples of such
30 disorders include, but are not limited to, those gastrointestinal, cardiovascular, autoimmune/inflammatory, cell proliferative, developmental, epithelial, neurological, and reproductive disorders described above. In one aspect, an antibody which specifically binds PMMM may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissues which express PMMM.

In an additional embodiment, a vector expressing the complement of the polynucleotide encoding PMMM may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of PMMM including, but not limited to, those described above.

In other embodiments, any of the proteins, antagonists, antibodies, agonists, complementary
5 sequences, or vectors of the invention may be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic
10 efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

An antagonist of PMMM may be produced using methods which are generally known in the art. In particular, purified PMMM may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind PMMM. Antibodies to PMMM may also be generated using methods that are well known in the art. Such antibodies may include,
15 but are not limited to, polyclonal, monoclonal, chimeric, and single chain antibodies, Fab fragments, and fragments produced by a Fab expression library. Neutralizing antibodies (i.e., those which inhibit dimer formation) are generally preferred for therapeutic use.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others may be immunized by injection with PMMM or with any fragment or oligopeptide
20 thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are
25 especially preferable.

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to PMMM have an amino acid sequence consisting of at least about 5 amino acids, and generally will consist of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or fragments are identical to a portion of the amino acid sequence of the natural protein. Short
30 stretches of PMMM amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

Monoclonal antibodies to PMMM may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-
35 hybridoma technique. (See, e.g., Kohler, G. et al. (1975) Nature 256:495-497; Kozbor, D. et al.

(1985) J. Immunol. Methods 81:31-42; Cote, R.J. et al. (1983) Proc. Natl. Acad. Sci. USA 80:2026-2030; and Cole, S.P. et al. (1984) Mol. Cell Biol. 62:109-120.)

In addition, techniques developed for the production of "chimeric antibodies," such as the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate
5 antigen specificity and biological activity, can be used. (See, e.g., Morrison, S.L. et al. (1984) Proc. Natl. Acad. Sci. USA 81:6851-6855; Neuberger, M.S. et al. (1984) Nature 312:604-608; and Takeda, S. et al. (1985) Nature 314:452-454.) Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce PMMM-specific single chain antibodies. Antibodies with related specificity, but of distinct
10 idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries. (See, e.g., Burton, D.R. (1991) Proc. Natl. Acad. Sci. USA 88:10134-10137.)

Antibodies may also be produced by inducing in vivo production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as
15 disclosed in the literature. (See, e.g., Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. USA 86:3833-3837; Winter, G. et al. (1991) Nature 349:293-299.)

Antibody fragments which contain specific binding sites for PMMM may also be generated. For example, such fragments include, but are not limited to, F(ab')₂ fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by reducing the disulfide bridges of
20 the F(ab')₂ fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity. (See, e.g., Huse, W.D. et al. (1989) Science 246:1275-1281.)

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either
25 polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between PMMM and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering PMMM epitopes is generally used, but a competitive binding assay may also be employed (Pound, supra).

30 Various methods such as Scatchard analysis in conjunction with radioimmunoassay techniques may be used to assess the affinity of antibodies for PMMM. Affinity is expressed as an association constant, K_a , which is defined as the molar concentration of PMMM-antibody complex divided by the molar concentrations of free antigen and free antibody under equilibrium conditions. The K_a determined for a preparation of polyclonal antibodies, which are heterogeneous in their
35 affinities for multiple PMMM epitopes, represents the average affinity, or avidity, of the antibodies

for PMMM. The K_a determined for a preparation of monoclonal antibodies, which are monospecific for a particular PMMM epitope, represents a true measure of affinity. High-affinity antibody preparations with K_a ranging from about 10^9 to 10^{12} L/mole are preferred for use in immunoassays in which the PMMM-antibody complex must withstand rigorous manipulations.

5 Low-affinity antibody preparations with K_a ranging from about 10^6 to 10^7 L/mole are preferred for use in immunopurification and similar procedures which ultimately require dissociation of PMMM, preferably in active form, from the antibody (Catty, D. (1988) Antibodies, Volume I: A Practical Approach, IRL Press, Washington DC; Liddell, J.E. and A. Cryer (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY).

10 The titer and avidity of polyclonal antibody preparations may be further evaluated to determine the quality and suitability of such preparations for certain downstream applications. For example, a polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml, preferably 5-10 mg specific antibody/ml, is generally employed in procedures requiring precipitation of PMMM-antibody complexes. Procedures for evaluating antibody specificity, titer,
15 and avidity, and guidelines for antibody quality and usage in various applications, are generally available. (See, e.g., Catty, supra, and Coligan et al. supra.)

In another embodiment of the invention, the polynucleotides encoding PMMM, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, modifications of gene expression can be achieved by designing complementary sequences or
20 antisense molecules (DNA, RNA, PNA, or modified oligonucleotides) to the coding or regulatory regions of the gene encoding PMMM. Such technology is well known in the art, and antisense oligonucleotides or larger fragments can be designed from various locations along the coding or control regions of sequences encoding PMMM. (See, e.g., Agrawal, S., ed. (1996) Antisense Therapeutics, Humana Press Inc., Totawa NJ.)

25 In therapeutic use, any gene delivery system suitable for introduction of the antisense sequences into appropriate target cells can be used. Antisense sequences can be delivered intracellularly in the form of an expression plasmid which, upon transcription, produces a sequence complementary to at least a portion of the cellular sequence encoding the target protein. (See, e.g., Slater, J.E. et al. (1998) *J. Allergy Clin. Immunol.* 102(3):469-475; and Scanlon, K.J. et al. (1995)
30 9(13):1288-1296.) Antisense sequences can also be introduced intracellularly through the use of viral vectors, such as retrovirus and adeno-associated virus vectors. (See, e.g., Miller, A.D. (1990) *Blood* 76:271; Ausubel, supra; Uckert, W. and W. Walther (1994) *Pharmacol. Ther.* 63(3):323-347.) Other gene delivery mechanisms include liposome-derived systems, artificial viral envelopes, and other systems known in the art. (See, e.g., Rossi, J.J. (1995) *Br. Med. Bull.* 51(1):217-225;

Boado, R.J. et al. (1998) J. Pharm. Sci. 87(11):1308-1315; and Morris, M.C. et al. (1997) Nucleic Acids Res. 25(14):2730-2736.)

In another embodiment of the invention, polynucleotides encoding PMMM may be used for somatic or germline gene therapy. Gene therapy may be performed to (i) correct a genetic
 5 deficiency (e.g., in the cases of severe combined immunodeficiency (SCID)-X1 disease characterized by X-linked inheritance (Cavazzana-Calvo, M. et al. (2000) Science 288:669-672), severe combined immunodeficiency syndrome associated with an inherited adenosine deaminase (ADA) deficiency (Blaese, R.M. et al. (1995) Science 270:475-480; Bordignon, C. et al. (1995) Science 270:470-475), cystic fibrosis (Zabner, J. et al. (1993) Cell 75:207-216; Crystal, R.G. et al.
 10 (1995) Hum. Gene Therapy 6:643-666; Crystal, R.G. et al. (1995) Hum. Gene Therapy 6:667-703), thalassemias, familial hypercholesterolemia, and hemophilia resulting from Factor VIII or Factor IX deficiencies (Crystal, R.G. (1995) Science 270:404-410; Verma, I.M. and N. Somia (1997) Nature 389:239-242)), (ii) express a conditionally lethal gene product (e.g., in the case of cancers which result from unregulated cell proliferation), or (iii) express a protein which affords protection against
 15 intracellular parasites (e.g., against human retroviruses, such as human immunodeficiency virus (HIV) (Baltimore, D. (1988) Nature 335:395-396; Poeschla, E. et al. (1996) Proc. Natl. Acad. Sci. USA 93:11395-11399), hepatitis B or C virus (HBV, HCV); fungal parasites, such as Candida albicans and Paracoccidioides brasiliensis; and protozoan parasites such as Plasmodium falciparum and Trypanosoma cruzi). In the case where a genetic deficiency in PMMM expression or regulation
 20 causes disease, the expression of PMMM from an appropriate population of transduced cells may alleviate the clinical manifestations caused by the genetic deficiency.

In a further embodiment of the invention, diseases or disorders caused by deficiencies in PMMM are treated by constructing mammalian expression vectors encoding PMMM and introducing these vectors by mechanical means into PMMM-deficient cells. Mechanical transfer
 25 technologies for use with cells in vivo or ex vitro include (i) direct DNA microinjection into individual cells, (ii) ballistic gold particle delivery, (iii) liposome-mediated transfection, (iv) receptor-mediated gene transfer, and (v) the use of DNA transposons (Morgan, R.A. and W.F. Anderson (1993) Annu. Rev. Biochem. 62:191-217; Ivics, Z. (1997) Cell 91:501-510; Boulay, J-L. and H. Récipon (1998) Curr. Opin. Biotechnol. 9:445-450).

30 Expression vectors that may be effective for the expression of PMMM include, but are not limited to, the PCDNA 3.1, EPITAG, PRCCMV2, PREP, PVAX, PCR2-TOPOTA vectors (Invitrogen, Carlsbad CA), PCMV-SCRIPT, PCMV-TAG, PEGSH/PERV (Stratagene, La Jolla CA), and PTET-OFF, PTET-ON, PTRE2, PTRE2-LUC, PTK-HYG (Clontech, Palo Alto CA). PMMM may be expressed using (i) a constitutively active promoter, (e.g., from cytomegalovirus
 35 (CMV), Rous sarcoma virus (RSV), SV40 virus, thymidine kinase (TK), or β -actin genes), (ii) an

inducible promoter (e.g., the tetracycline-regulated promoter (Gossen, M. and H. Bujard (1992) Proc. Natl. Acad. Sci. USA 89:5547-5551; Gossen, M. et al. (1995) Science 268:1766-1769; Rossi, F.M.V. and H.M. Blau (1998) Curr. Opin. Biotechnol. 9:451-456), commercially available in the T-REX plasmid (Invitrogen)); the ecdysone-inducible promoter (available in the plasmids PVGRXR
 5 and PIND; Invitrogen); the FK506/rapamycin inducible promoter; or the RU486/mifepristone inducible promoter (Rossi, F.M.V. and H.M. Blau, *supra*), or (iii) a tissue-specific promoter or the native promoter of the endogenous gene encoding PMMM from a normal individual.

Commercially available liposome transformation kits (e.g., the PERFECT LIPID TRANSFECTION KIT, available from Invitrogen) allow one with ordinary skill in the art to deliver
 10 polynucleotides to target cells in culture and require minimal effort to optimize experimental parameters. In the alternative, transformation is performed using the calcium phosphate method (Graham, F.L. and A.J. Eb (1973) Virology 52:456-467), or by electroporation (Neumann, E. et al. (1982) EMBO J. 1:841-845). The introduction of DNA to primary cells requires modification of these standardized mammalian transfection protocols.

15 In another embodiment of the invention, diseases or disorders caused by genetic defects with respect to PMMM expression are treated by constructing a retrovirus vector consisting of (i) the polynucleotide encoding PMMM under the control of an independent promoter or the retrovirus long terminal repeat (LTR) promoter, (ii) appropriate RNA packaging signals, and (iii) a Rev-responsive element (RRE) along with additional retrovirus *cis*-acting RNA sequences and coding
 20 sequences required for efficient vector propagation. Retrovirus vectors (e.g., PFB and PFBNEO) are commercially available (Stratagene) and are based on published data (Riviere, I. et al. (1995) Proc. Natl. Acad. Sci. USA 92:6733-6737), incorporated by reference herein. The vector is propagated in an appropriate vector producing cell line (VPCL) that expresses an envelope gene with a tropism for receptors on the target cells or a promiscuous envelope protein such as VSVg
 25 (Armentano, D. et al. (1987) J. Virol. 61:1647-1650; Bender, M.A. et al. (1987) J. Virol. 61:1639-1646; Adam, M.A. and A.D. Miller (1988) J. Virol. 62:3802-3806; Dull, T. et al. (1998) J. Virol. 72:8463-8471; Zufferey, R. et al. (1998) J. Virol. 72:9873-9880). U.S. Patent No. 5,910,434 to Rigg ("Method for obtaining retrovirus packaging cell lines producing high transducing efficiency retroviral supernatant") discloses a method for obtaining retrovirus packaging cell lines and is
 30 hereby incorporated by reference. Propagation of retrovirus vectors, transduction of a population of cells (e.g., CD4⁺ T-cells), and the return of transduced cells to a patient are procedures well known to persons skilled in the art of gene therapy and have been well documented (Ranga, U. et al. (1997) J. Virol. 71:7020-7029; Bauer, G. et al. (1997) Blood 89:2259-2267; Bonyhadi, M.L. (1997) J. Virol. 71:4707-4716; Ranga, U. et al. (1998) Proc. Natl. Acad. Sci. USA 95:1201-1206; Su, L.
 35 (1997) Blood 89:2283-2290).

In the alternative, an adenovirus-based gene therapy delivery system is used to deliver polynucleotides encoding PMMM to cells which have one or more genetic abnormalities with respect to the expression of PMMM. The construction and packaging of adenovirus-based vectors are well known to those with ordinary skill in the art. Replication defective adenovirus vectors have
5 proven to be versatile for importing genes encoding immunoregulatory proteins into intact islets in the pancreas (Csete, M.E. et al. (1995) Transplantation 27:263-268). Potentially useful adenoviral vectors are described in U.S. Patent No. 5,707,618 to Armentano ("Adenovirus vectors for gene therapy"), hereby incorporated by reference. For adenoviral vectors, see also Antinozzi, P.A. et al. (1999) Annu. Rev. Nutr. 19:511-544 and Verma, I.M. and N. Somia (1997) Nature 389:239-242,
10 both incorporated by reference herein.

In another alternative, a herpes-based, gene therapy delivery system is used to deliver polynucleotides encoding PMMM to target cells which have one or more genetic abnormalities with respect to the expression of PMMM. The use of herpes simplex virus (HSV)-based vectors may be especially valuable for introducing PMMM to cells of the central nervous system, for which HSV
15 has a tropism. The construction and packaging of herpes-based vectors are well known to those with ordinary skill in the art. A replication-competent herpes simplex virus (HSV) type 1-based vector has been used to deliver a reporter gene to the eyes of primates (Liu, X. et al. (1999) Exp. Eye Res. 69:385-395). The construction of a HSV-1 virus vector has also been disclosed in detail in U.S. Patent No. 5,804,413 to DeLuca ("Herpes simplex virus strains for gene transfer"), which is
20 hereby incorporated by reference. U.S. Patent No. 5,804,413 teaches the use of recombinant HSV d92 which consists of a genome containing at least one exogenous gene to be transferred to a cell under the control of the appropriate promoter for purposes including human gene therapy. Also taught by this patent are the construction and use of recombinant HSV strains deleted for ICP4, ICP27 and ICP22. For HSV vectors, see also Goins, W.F. et al. (1999) J. Virol. 73:519-532 and Xu,
25 H. et al. (1994) Dev. Biol. 163:152-161, hereby incorporated by reference. The manipulation of cloned herpesvirus sequences, the generation of recombinant virus following the transfection of multiple plasmids containing different segments of the large herpesvirus genomes, the growth and propagation of herpesvirus, and the infection of cells with herpesvirus are techniques well known to those of ordinary skill in the art.

30 In another alternative, an alphavirus (positive, single-stranded RNA virus) vector is used to deliver polynucleotides encoding PMMM to target cells. The biology of the prototypic alphavirus, Semliki Forest Virus (SFV), has been studied extensively and gene transfer vectors have been based on the SFV genome (Garoff, H. and K.-J. Li (1998) Curr. Opin. Biotechnol. 9:464-469). During alphavirus RNA replication, a subgenomic RNA is generated that normally encodes the viral capsid
35 proteins. This subgenomic RNA replicates to higher levels than the full length genomic RNA,

resulting in the overproduction of capsid proteins relative to the viral proteins with enzymatic activity (e.g., protease and polymerase). Similarly, inserting the coding sequence for PMMM into the alphavirus genome in place of the capsid-coding region results in the production of a large number of PMMM-coding RNAs and the synthesis of high levels of PMMM in vector transduced
5 cells. While alphavirus infection is typically associated with cell lysis within a few days, the ability to establish a persistent infection in hamster normal kidney cells (BHK-21) with a variant of Sindbis virus (SIN) indicates that the lytic replication of alphaviruses can be altered to suit the needs of the gene therapy application (Dryga, S.A. et al. (1997) Virology 228:74-83). The wide host range of alphaviruses will allow the introduction of PMMM into a variety of cell types. The specific
10 transduction of a subset of cells in a population may require the sorting of cells prior to transduction. The methods of manipulating infectious cDNA clones of alphaviruses, performing alphavirus cDNA and RNA transfections, and performing alphavirus infections, are well known to those with ordinary skill in the art.

Oligonucleotides derived from the transcription initiation site, e.g., between about positions
15 -10 and +10 from the start site, may also be employed to inhibit gene expression. Similarly, inhibition can be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature. (See, e.g., Gee, J.E. et al. (1994) in Huber,
20 B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177.) A complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the
25 ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. For example, engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding PMMM.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites, including the following sequences: GUA,
30 GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides, corresponding to the region of the target gene containing the cleavage site, may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Complementary ribonucleic acid molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences encoding PMMM. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines, cells, or tissues.

RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule, or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

An additional embodiment of the invention encompasses a method for screening for a compound which is effective in altering expression of a polynucleotide encoding PMMM. Compounds which may be effective in altering expression of a specific polynucleotide may include, but are not limited to, oligonucleotides, antisense oligonucleotides, triple helix-forming oligonucleotides, transcription factors and other polypeptide transcriptional regulators, and non-macromolecular chemical entities which are capable of interacting with specific polynucleotide sequences. Effective compounds may alter polynucleotide expression by acting as either inhibitors or promoters of polynucleotide expression. Thus, in the treatment of disorders associated with increased PMMM expression or activity, a compound which specifically inhibits expression of the polynucleotide encoding PMMM may be therapeutically useful, and in the treatment of disorders associated with decreased PMMM expression or activity, a compound which specifically promotes expression of the polynucleotide encoding PMMM may be therapeutically useful.

At least one, and up to a plurality, of test compounds may be screened for effectiveness in altering expression of a specific polynucleotide. A test compound may be obtained by any method commonly known in the art, including chemical modification of a compound known to be effective in altering polynucleotide expression; selection from an existing, commercially-available or proprietary library of naturally-occurring or non-natural chemical compounds; rational design of a compound based on chemical and/or structural properties of the target polynucleotide; and selection from a library of chemical compounds created combinatorially or randomly. A sample comprising

a polynucleotide encoding PMMM is exposed to at least one test compound thus obtained. The sample may comprise, for example, an intact or permeabilized cell, or an in vitro cell-free or reconstituted biochemical system. Alterations in the expression of a polynucleotide encoding PMMM are assayed by any method commonly known in the art. Typically, the expression of a specific nucleotide is detected by hybridization with a probe having a nucleotide sequence complementary to the sequence of the polynucleotide encoding PMMM. The amount of hybridization may be quantified, thus forming the basis for a comparison of the expression of the polynucleotide both with and without exposure to one or more test compounds. Detection of a change in the expression of a polynucleotide exposed to a test compound indicates that the test compound is effective in altering the expression of the polynucleotide. A screen for a compound effective in altering expression of a specific polynucleotide can be carried out, for example, using a Schizosaccharomyces pombe gene expression system (Atkins, D. et al. (1999) U.S. Patent No. 5,932,435; Arndt, G.M. et al. (2000) Nucleic Acids Res. 28:E15) or a human cell line such as HeLa cell (Clarke, M.L. et al. (2000) Biochem. Biophys. Res. Commun. 268:8-13). A particular embodiment of the present invention involves screening a combinatorial library of oligonucleotides (such as deoxyribonucleotides, ribonucleotides, peptide nucleic acids, and modified oligonucleotides) for antisense activity against a specific polynucleotide sequence (Bruice, T.W. et al. (1997) U.S. Patent No. 5,686,242; Bruice, T.W. et al. (2000) U.S. Patent No. 6,022,691).

Many methods for introducing vectors into cells or tissues are available and equally suitable for use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection, by liposome injections, or by polycationic amino polymers may be achieved using methods which are well known in the art. (See, e.g., Goldman, C.K. et al. (1997) Nat. Biotechnol. 15:462-466.)

Any of the therapeutic methods described above may be applied to any subject in need of such therapy, including, for example, mammals such as humans, dogs, cats, cows, horses, rabbits, and monkeys.

An additional embodiment of the invention relates to the administration of a composition which generally comprises an active ingredient formulated with a pharmaceutically acceptable excipient. Excipients may include, for example, sugars, starches, celluloses, gums, and proteins. Various formulations are commonly known and are thoroughly discussed in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing, Easton PA). Such compositions may consist of PMMM, antibodies to PMMM, and mimetics, agonists, antagonists, or inhibitors of PMMM.

The compositions utilized in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, pulmonary, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

5 Compositions for pulmonary administration may be prepared in liquid or dry powder form. These compositions are generally aerosolized immediately prior to inhalation by the patient. In the case of small molecules (e.g. traditional low molecular weight organic drugs), aerosol delivery of fast-acting formulations is well-known in the art. In the case of macromolecules (e.g. larger peptides and proteins), recent developments in the field of pulmonary delivery via the alveolar
10 region of the lung have enabled the practical delivery of drugs such as insulin to blood circulation (see, e.g., Patton, J.S. et al., U.S. Patent No. 5,997,848). Pulmonary delivery has the advantage of administration without needle injection, and obviates the need for potentially toxic penetration enhancers.

 Compositions suitable for use in the invention include compositions wherein the active
15 ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

 Specialized forms of compositions may be prepared for direct intracellular delivery of macromolecules comprising PMMM or fragments thereof. For example, liposome preparations containing a cell-impermeable macromolecule may promote cell fusion and intracellular delivery of
20 the macromolecule. Alternatively, PMMM or a fragment thereof may be joined to a short cationic N-terminal portion from the HIV Tat-1 protein. Fusion proteins thus generated have been found to transduce into the cells of all tissues, including the brain, in a mouse model system (Schwarze, S.R. et al. (1999) Science 285:1569-1572).

 For any compound, the therapeutically effective dose can be estimated initially either in cell
25 culture assays, e.g., of neoplastic cells, or in animal models such as mice, rats, rabbits, dogs, monkeys, or pigs. An animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

 A therapeutically effective dose refers to that amount of active ingredient, for example
30 PMMM or fragments thereof, antibodies of PMMM, and agonists, antagonists or inhibitors of PMMM, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the ED₅₀ (the dose therapeutically effective in 50% of the population) or LD₅₀ (the dose lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is
35 the therapeutic index, which can be expressed as the LD₅₀/ED₅₀ ratio. Compositions which exhibit

large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the ED₅₀ with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, 5 the sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, the general health of the subject, the age, weight, and gender of the 10 subject, time and frequency of administration, drug combination(s), reaction sensitivities, and response to therapy. Long-acting compositions may be administered every 3 to 4 days, every week, or biweekly depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from about 0.1 μ g to 100,000 μ g, up to a total dose of about 1 gram, depending upon the route of administration. Guidance as to particular dosages and 15 methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

DIAGNOSTICS

20 In another embodiment, antibodies which specifically bind PMMM may be used for the diagnosis of disorders characterized by expression of PMMM, or in assays to monitor patients being treated with PMMM or agonists, antagonists, or inhibitors of PMMM. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for PMMM include methods which utilize the antibody and a label to detect 25 PMMM in human body fluids or in extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by covalent or non-covalent attachment of a reporter molecule. A wide variety of reporter molecules, several of which are described above, are known in the art and may be used.

A variety of protocols for measuring PMMM, including ELISAs, RIAs, and FACS, are 30 known in the art and provide a basis for diagnosing altered or abnormal levels of PMMM expression. Normal or standard values for PMMM expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, for example, human subjects, with antibodies to PMMM under conditions suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, such as photometric means. Quantities 35 of PMMM expressed in subject, control, and disease samples from biopsied tissues are compared

with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, the polynucleotides encoding PMMM may be used for diagnostic purposes. The polynucleotides which may be used include oligonucleotide
5 sequences, complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantify gene expression in biopsied tissues in which expression of PMMM may be correlated with disease. The diagnostic assay may be used to determine absence, presence, and excess expression of PMMM, and to monitor regulation of PMMM levels during therapeutic intervention.

10 In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding PMMM or closely related molecules may be used to identify nucleic acid sequences which encode PMMM. The specificity of the probe, whether it is made from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a conserved motif, and the stringency of the hybridization or amplification will
15 determine whether the probe identifies only naturally occurring sequences encoding PMMM, allelic variants, or related sequences.

Probes may also be used for the detection of related sequences, and may have at least 50% sequence identity to any of the PMMM encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:19-36
20 or from genomic sequences including promoters, enhancers, and introns of the PMMM gene.

Means for producing specific hybridization probes for DNAs encoding PMMM include the cloning of polynucleotide sequences encoding PMMM or PMMM derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA
25 polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, by radionuclides such as ^{32}P or ^{35}S , or by enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotide sequences encoding PMMM may be used for the diagnosis of disorders associated with expression of PMMM. Examples of such disorders include, but are not limited to, a
30 gastrointestinal disorder, such as dysphagia, peptic esophagitis, esophageal spasm, esophageal stricture, esophageal carcinoma, dyspepsia, indigestion, gastritis, gastric carcinoma, anorexia, nausea, emesis, gastroparesis, antral or pyloric edema, abdominal angina, pyrosis, gastroenteritis, intestinal obstruction, infections of the intestinal tract, peptic ulcer, cholelithiasis, cholecystitis, cholestasis, pancreatitis, pancreatic carcinoma, biliary tract disease, hepatitis, hyperbilirubinemia,
35 cirrhosis, passive congestion of the liver, hepatoma, infectious colitis, ulcerative colitis, ulcerative

proctitis, Crohn's disease, Whipple's disease, Mallory-Weiss syndrome, colonic carcinoma, colonic obstruction, irritable bowel syndrome, short bowel syndrome, diarrhea, constipation, gastrointestinal hemorrhage, acquired immunodeficiency syndrome (AIDS) enteropathy, jaundice, hepatic encephalopathy, hepatorenal syndrome, hepatic steatosis, hemochromatosis, Wilson's disease,

5 α_1 -antitrypsin deficiency, Reye's syndrome, primary sclerosing cholangitis, liver infarction, portal vein obstruction and thrombosis, centrilobular necrosis, peliosis hepatis, hepatic vein thrombosis, veno-occlusive disease, preeclampsia, eclampsia, acute fatty liver of pregnancy, intrahepatic cholestasis of pregnancy, and hepatic tumors including nodular hyperplasias, adenomas, and carcinomas; a cardiovascular disorder, such as arteriovenous fistula, atherosclerosis,

10 hypertension, vasculitis, Raynaud's disease, aneurysms, arterial dissections, varicose veins, thrombophlebitis and phlebothrombosis, vascular tumors, and complications of thrombolysis, balloon angioplasty, vascular replacement, and coronary artery bypass graft surgery, congestive heart failure, ischemic heart disease, angina pectoris, myocardial infarction, hypertensive heart disease, degenerative valvular heart disease, calcific aortic valve stenosis, congenitally bicuspid

15 aortic valve, mitral annular calcification, mitral valve prolapse, rheumatic fever and rheumatic heart disease, infective endocarditis, nonbacterial thrombotic endocarditis, endocarditis of systemic lupus erythematosus, carcinoid heart disease, cardiomyopathy, myocarditis, pericarditis, neoplastic heart disease, congenital heart disease, and complications of cardiac transplantation; an autoimmune/inflammatory disease, such as acquired immunodeficiency syndrome (AIDS),

20 Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, atherosclerotic plaque rupture, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins,

25 erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, degradation of articular cartilage, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis,

30 systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal

35 hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including

adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; a

5 developmental disorder, such as renal tubular acidosis, anemia, Cushing's syndrome, achondroplastic dwarfism, Duchenne and Becker muscular dystrophy, bone resorption, epilepsy, gonadal dysgenesis, WAGR syndrome (Wilms' tumor, aniridia, genitourinary abnormalities, and mental retardation), Smith-Magenis syndrome, myelodysplastic syndrome, hereditary mucoepithelial dysplasia, hereditary keratodermas, hereditary neuropathies such as Charcot-Marie-

10 Tooth disease and neurofibromatosis, hypothyroidism, hydrocephalus, seizure disorders such as Sydenham's chorea and cerebral palsy, spina bifida, anencephaly, craniorachischisis, congenital glaucoma, cataract, age-related macular degeneration, and sensorineural hearing loss; an epithelial disorder, such as dyshidrotic eczema, allergic contact dermatitis, keratosis pilaris, melasma, vitiligo, actinic keratosis, basal cell carcinoma, squamous cell carcinoma, seborrheic keratosis,

15 folliculitis, herpes simplex, herpes zoster, varicella, candidiasis, dermatophytosis, scabies, insect bites, cherry angioma, keloid, dermatofibroma, acrochordons, urticaria, transient acantholytic dermatosis, xerosis, eczema, atopic dermatitis, contact dermatitis, hand eczema, nummular eczema, lichen simplex chronicus, asteatotic eczema, stasis dermatitis and stasis ulceration, seborrheic dermatitis, psoriasis, lichen planus, pityriasis rosea, impetigo, ecthyma, dermatophytosis, tinea

20 versicolor, warts, acne vulgaris, acne rosacea, pemphigus vulgaris, pemphigus foliaceus, paraneoplastic pemphigus, bullous pemphigoid, herpes gestationis, dermatitis herpetiformis, linear IgA disease, epidermolysis bullosa acquisita, dermatomyositis, lupus erythematosus, scleroderma and morphea, erythroderma, alopecia, figurate skin lesions, telangiectasias, hypopigmentation, hyperpigmentation, vesicles/bullae, exanthems, cutaneous drug reactions, papulonodular skin

25 lesions, chronic non-healing wounds, photosensitivity diseases, epidermolysis bullosa simplex, epidermolytic hyperkeratosis, epidermolytic and nonepidermolytic palmoplantar keratoderma, ichthyosis bullosa of Siemens, ichthyosis exfoliativa, keratosis palmaris et plantaris, keratosis palmoplantaris, palmoplantar keratoderma, keratosis punctata, Meesmann's corneal dystrophy, pachyonychia congenita, white sponge nevus, steatocystoma multiplex, epidermal

30 nevi/epidermolytic hyperkeratosis type, monilethrix, trichothiodystrophy, chronic hepatitis/cryptogenic cirrhosis, and colorectal hyperplasia; a neurological disorder, such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular

35 atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating

diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system including Down syndrome, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathies, myasthenia gravis, periodic paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, seasonal affective disorder (SAD), akathisia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, Tourette's disorder, progressive supranuclear palsy, corticobasal degeneration, and familial frontotemporal dementia; and a reproductive disorder, such as infertility, including tubal disease, ovulatory defects, and endometriosis, a disorder of prolactin production, a disruption of the estrous cycle, a disruption of the menstrual cycle, polycystic ovary syndrome, ovarian hyperstimulation syndrome, an endometrial or ovarian tumor, a uterine fibroid, autoimmune disorders, an ectopic pregnancy, and teratogenesis; cancer of the breast, fibrocystic breast disease, and galactorrhea; a disruption of spermatogenesis, abnormal sperm physiology, cancer of the testis, cancer of the prostate, benign prostatic hyperplasia, prostatitis, Peyronie's disease, impotence, carcinoma of the male breast, and gynecomastia. The polynucleotide sequences encoding PMMM may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; in dipstick, pin, and multiformat ELISA-like assays; and in microarrays utilizing fluids or tissues from patients to detect altered PMMM expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding PMMM may be useful in assays that detect the presence of associated disorders, particularly those mentioned above. The nucleotide sequences encoding PMMM may be labeled by standard methods and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantified and compared with a standard value. If the amount of signal in the patient sample is significantly altered in comparison to a control sample then the presence of altered levels of nucleotide sequences encoding PMMM in the sample indicates the presence of the associated disorder. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the treatment of an individual patient.

In order to provide a basis for the diagnosis of a disorder associated with expression of PMMM, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, encoding PMMM, under conditions suitable for hybridization or
5 amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. Standard values obtained in this manner may be compared with values obtained from samples from patients who are symptomatic for a disorder. Deviation from standard values is used to establish the presence of a disorder.

10 Once the presence of a disorder is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to determine if the level of expression in the patient begins to approximate that which is observed in the normal subject. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

15 With respect to cancer, the presence of an abnormal amount of transcript (either under- or overexpressed) in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing
20 the development or further progression of the cancer.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding PMMM may involve the use of PCR. These oligomers may be chemically synthesized, generated enzymatically, or produced in vitro. Oligomers will preferably contain a fragment of a polynucleotide encoding PMMM, or a fragment of a polynucleotide complementary to the
25 polynucleotide encoding PMMM, and will be employed under optimized conditions for identification of a specific gene or condition. Oligomers may also be employed under less stringent conditions for detection or quantification of closely related DNA or RNA sequences.

In a particular aspect, oligonucleotide primers derived from the polynucleotide sequences encoding PMMM may be used to detect single nucleotide polymorphisms (SNPs). SNPs are
30 substitutions, insertions and deletions that are a frequent cause of inherited or acquired genetic disease in humans. Methods of SNP detection include, but are not limited to, single-stranded conformation polymorphism (SSCP) and fluorescent SSCP (fSSCP) methods. In SSCP, oligonucleotide primers derived from the polynucleotide sequences encoding PMMM are used to amplify DNA using the polymerase chain reaction (PCR). The DNA may be derived, for example,
35 from diseased or normal tissue, biopsy samples, bodily fluids, and the like. SNPs in the DNA cause

differences in the secondary and tertiary structures of PCR products in single-stranded form, and these differences are detectable using gel electrophoresis in non-denaturing gels. In fSCCP, the oligonucleotide primers are fluorescently labeled, which allows detection of the amplimers in high-throughput equipment such as DNA sequencing machines. Additionally, sequence database
5 analysis methods, termed in silico SNP (isSNP), are capable of identifying polymorphisms by comparing the sequence of individual overlapping DNA fragments which assemble into a common consensus sequence. These computer-based methods filter out sequence variations due to laboratory preparation of DNA and sequencing errors using statistical models and automated analyses of DNA sequence chromatograms. In the alternative, SNPs may be detected and
10 characterized by mass spectrometry using, for example, the high throughput MASSARRAY system (Sequenom, Inc., San Diego CA).

Methods which may also be used to quantify the expression of PMMM include radiolabeling or biotinylating nucleotides, coamplification of a control nucleic acid, and interpolating results from standard curves. (See, e.g., Melby, P.C. et al. (1993) J. Immunol.
15 Methods 159:235-244; Duplaa, C. et al. (1993) Anal. Biochem. 212:229-236.) The speed of quantitation of multiple samples may be accelerated by running the assay in a high-throughput format where the oligomer or polynucleotide of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

In further embodiments, oligonucleotides or longer fragments derived from any of the
20 polynucleotide sequences described herein may be used as elements on a microarray. The microarray can be used in transcript imaging techniques which monitor the relative expression levels of large numbers of genes simultaneously as described below. The microarray may also be used to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disorder, to diagnose a disorder, to
25 monitor progression/regression of disease as a function of gene expression, and to develop and monitor the activities of therapeutic agents in the treatment of disease. In particular, this information may be used to develop a pharmacogenomic profile of a patient in order to select the most appropriate and effective treatment regimen for that patient. For example, therapeutic agents which are highly effective and display the fewest side effects may be selected for a patient based on
30 his/her pharmacogenomic profile.

In another embodiment, PMMM, fragments of PMMM, or antibodies specific for PMMM may be used as elements on a microarray. The microarray may be used to monitor or measure protein-protein interactions, drug-target interactions, and gene expression profiles, as described above.

A particular embodiment relates to the use of the polynucleotides of the present invention to generate a transcript image of a tissue or cell type. A transcript image represents the global pattern of gene expression by a particular tissue or cell type. Global gene expression patterns are analyzed by quantifying the number of expressed genes and their relative abundance under given conditions and at a given time. (See Seilhamer et al., "Comparative Gene Transcript Analysis," U.S. Patent No. 5,840,484, expressly incorporated by reference herein.) Thus a transcript image may be generated by hybridizing the polynucleotides of the present invention or their complements to the totality of transcripts or reverse transcripts of a particular tissue or cell type. In one embodiment, the hybridization takes place in high-throughput format, wherein the polynucleotides of the present invention or their complements comprise a subset of a plurality of elements on a microarray. The resultant transcript image would provide a profile of gene activity.

Transcript images may be generated using transcripts isolated from tissues, cell lines, biopsies, or other biological samples. The transcript image may thus reflect gene expression in vivo, as in the case of a tissue or biopsy sample, or in vitro, as in the case of a cell line.

Transcript images which profile the expression of the polynucleotides of the present invention may also be used in conjunction with in vitro model systems and preclinical evaluation of pharmaceuticals, as well as toxicological testing of industrial and naturally-occurring environmental compounds. All compounds induce characteristic gene expression patterns, frequently termed molecular fingerprints or toxicant signatures, which are indicative of mechanisms of action and toxicity (Nuwaysir, E.F. et al. (1999) Mol. Carcinog. 24:153-159; Steiner, S. and N.L. Anderson (2000) Toxicol. Lett. 112-113:467-471, expressly incorporated by reference herein). If a test compound has a signature similar to that of a compound with known toxicity, it is likely to share those toxic properties. These fingerprints or signatures are most useful and refined when they contain expression information from a large number of genes and gene families. Ideally, a genome-wide measurement of expression provides the highest quality signature. Even genes whose expression is not altered by any tested compounds are important as well, as the levels of expression of these genes are used to normalize the rest of the expression data. The normalization procedure is useful for comparison of expression data after treatment with different compounds. While the assignment of gene function to elements of a toxicant signature aids in interpretation of toxicity mechanisms, knowledge of gene function is not necessary for the statistical matching of signatures which leads to prediction of toxicity. (See, for example, Press Release 00-02 from the National Institute of Environmental Health Sciences, released February 29, 2000, available at <http://www.niehs.nih.gov/oc/news/toxchip.htm>.) Therefore, it is important and desirable in toxicological screening using toxicant signatures to include all expressed gene sequences.

In one embodiment, the toxicity of a test compound is assessed by treating a biological sample containing nucleic acids with the test compound. Nucleic acids that are expressed in the treated biological sample are hybridized with one or more probes specific to the polynucleotides of the present invention, so that transcript levels corresponding to the polynucleotides of the present invention may be quantified. The transcript levels in the treated biological sample are compared with levels in an untreated biological sample. Differences in the transcript levels between the two samples are indicative of a toxic response caused by the test compound in the treated sample.

Another particular embodiment relates to the use of the polypeptide sequences of the present invention to analyze the proteome of a tissue or cell type. The term proteome refers to the global pattern of protein expression in a particular tissue or cell type. Each protein component of a proteome can be subjected individually to further analysis. Proteome expression patterns, or profiles, are analyzed by quantifying the number of expressed proteins and their relative abundance under given conditions and at a given time. A profile of a cell's proteome may thus be generated by separating and analyzing the polypeptides of a particular tissue or cell type. In one embodiment, the separation is achieved using two-dimensional gel electrophoresis, in which proteins from a sample are separated by isoelectric focusing in the first dimension, and then according to molecular weight by sodium dodecyl sulfate slab gel electrophoresis in the second dimension (Steiner and Anderson, *supra*). The proteins are visualized in the gel as discrete and uniquely positioned spots, typically by staining the gel with an agent such as Coomassie Blue or silver or fluorescent stains. The optical density of each protein spot is generally proportional to the level of the protein in the sample. The optical densities of equivalently positioned protein spots from different samples, for example, from biological samples either treated or untreated with a test compound or therapeutic agent, are compared to identify any changes in protein spot density related to the treatment. The proteins in the spots are partially sequenced using, for example, standard methods employing chemical or enzymatic cleavage followed by mass spectrometry. The identity of the protein in a spot may be determined by comparing its partial sequence, preferably of at least 5 contiguous amino acid residues, to the polypeptide sequences of the present invention. In some cases, further sequence data may be obtained for definitive protein identification.

A proteomic profile may also be generated using antibodies specific for PMMM to quantify the levels of PMMM expression. In one embodiment, the antibodies are used as elements on a microarray, and protein expression levels are quantified by exposing the microarray to the sample and detecting the levels of protein bound to each array element (Lueking, A. et al. (1999) *Anal. Biochem.* 270:103-111; Mendoze, L.G. et al. (1999) *Biotechniques* 27:778-788). Detection may be performed by a variety of methods known in the art, for example, by reacting the proteins in the

sample with a thiol- or amino-reactive fluorescent compound and detecting the amount of fluorescence bound at each array element.

Toxicant signatures at the proteome level are also useful for toxicological screening, and should be analyzed in parallel with toxicant signatures at the transcript level. There is a poor correlation between transcript and protein abundances for some proteins in some tissues (Anderson, N.L. and J. Seilhamer (1997) *Electrophoresis* 18:533-537), so proteome toxicant signatures may be useful in the analysis of compounds which do not significantly affect the transcript image, but which alter the proteomic profile. In addition, the analysis of transcripts in body fluids is difficult, due to rapid degradation of mRNA, so proteomic profiling may be more reliable and informative in such cases.

In another embodiment, the toxicity of a test compound is assessed by treating a biological sample containing proteins with the test compound. Proteins that are expressed in the treated biological sample are separated so that the amount of each protein can be quantified. The amount of each protein is compared to the amount of the corresponding protein in an untreated biological sample. A difference in the amount of protein between the two samples is indicative of a toxic response to the test compound in the treated sample. Individual proteins are identified by sequencing the amino acid residues of the individual proteins and comparing these partial sequences to the polypeptides of the present invention.

In another embodiment, the toxicity of a test compound is assessed by treating a biological sample containing proteins with the test compound. Proteins from the biological sample are incubated with antibodies specific to the polypeptides of the present invention. The amount of protein recognized by the antibodies is quantified. The amount of protein in the treated biological sample is compared with the amount in an untreated biological sample. A difference in the amount of protein between the two samples is indicative of a toxic response to the test compound in the treated sample.

Microarrays may be prepared, used, and analyzed using methods known in the art. (See, e.g., Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) *Proc. Natl. Acad. Sci. USA* 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) *Proc. Natl. Acad. Sci. USA* 94:2150-2155; and Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662.) Various types of microarrays are well known and thoroughly described in DNA Microarrays: A Practical Approach, M. Schena, ed. (1999) Oxford University Press, London, hereby expressly incorporated by reference.

In another embodiment of the invention, nucleic acid sequences encoding PMMM may be used to generate hybridization probes useful in mapping the naturally occurring genomic sequence.

Either coding or noncoding sequences may be used, and in some instances, noncoding sequences may be preferable over coding sequences. For example, conservation of a coding sequence among members of a multi-gene family may potentially cause undesired cross hybridization during chromosomal mapping. The sequences may be mapped to a particular chromosome, to a specific
5 region of a chromosome, or to artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355; Price, C.M. (1993) Blood Rev. 7:127-134; and Trask, B.J. (1991) Trends Genet. 7:149-154.) Once mapped, the nucleic acid sequences of the invention may
10 be used to develop genetic linkage maps, for example, which correlate the inheritance of a disease state with the inheritance of a particular chromosome region or restriction fragment length polymorphism (RFLP). (See, for example, Lander, E.S. and D. Botstein (1986) Proc. Natl. Acad. Sci. USA 83:7353-7357.)

Fluorescent in situ hybridization (FISH) may be correlated with other physical and genetic
15 map data. (See, e.g., Heinz-Ulrich, et al. (1995) in Meyers, supra, pp. 965-968.) Examples of genetic map data can be found in various scientific journals or at the Online Mendelian Inheritance in Man (OMIM) World Wide Web site. Correlation between the location of the gene encoding PMMM on a physical map and a specific disorder, or a predisposition to a specific disorder, may help define the region of DNA associated with that disorder and thus may further positional cloning
20 efforts.

In situ hybridization of chromosomal preparations and physical mapping techniques, such as linkage analysis using established chromosomal markers, may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the exact chromosomal locus is not known. This information
25 is valuable to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the gene or genes responsible for a disease or syndrome have been crudely localized by genetic linkage to a particular genomic region, e.g., ataxia-telangiectasia to 11q22-23, any sequences mapping to that area may represent associated or regulatory genes for further investigation. (See, e.g., Gatti, R.A. et al. (1988) Nature 336:577-580.) The nucleotide
30 sequence of the instant invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier, or affected individuals.

In another embodiment of the invention, PMMM, its catalytic or immunogenic fragments, or oligopeptides thereof can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to

a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes between PMMM and the agent being tested may be measured.

Another technique for drug screening provides for high throughput screening of compounds having suitable binding affinity to the protein of interest. (See, e.g., Geysen, et al. (1984) PCT application WO84/03564.) In this method, large numbers of different small test compounds are synthesized on a solid substrate. The test compounds are reacted with PMMM, or fragments thereof, and washed. Bound PMMM is then detected by methods well known in the art. Purified PMMM can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding PMMM specifically compete with a test compound for binding PMMM. In this manner, antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with PMMM.

In additional embodiments, the nucleotide sequences which encode PMMM may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following preferred specific embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

The disclosures of all patents, applications, and publications mentioned above and below, including U.S. Ser. No. 60/254,399, U.S. Ser. No. 60/257,803, U.S. Ser. No. 60/260,110, U.S. Ser. No. 60/262,851 and U.S. Ser. No. 60/264,623 are hereby expressly incorporated by reference.

EXAMPLES

I. Construction of cDNA Libraries

Incyte cDNAs were derived from cDNA libraries described in the LIFESEQ GOLD database (Incyte Genomics, Palo Alto CA). Some tissues were homogenized and lysed in guanidinium isothiocyanate, while others were homogenized and lysed in phenol or in a suitable mixture of denaturants, such as TRIZOL (Life Technologies), a monophasic solution of phenol and guanidine isothiocyanate. The resulting lysates were centrifuged over CsCl cushions or extracted with chloroform. RNA was precipitated from the lysates with either isopropanol or sodium acetate and ethanol, or by other routine methods.

Phenol extraction and precipitation of RNA were repeated as necessary to increase RNA purity. In some cases, RNA was treated with DNase. For most libraries, poly(A)+ RNA was isolated using oligo d(T)-coupled paramagnetic particles (Promega), OLIGOTEX latex particles (QIAGEN, Chatsworth CA), or an OLIGOTEX mRNA purification kit (QIAGEN). Alternatively, 5 RNA was isolated directly from tissue lysates using other RNA isolation kits, e.g., the POLY(A)PURE mRNA purification kit (Ambion, Austin TX).

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP vector system (Stratagene) or SUPERScript plasmid system (Life Technologies), using 10 the recommended procedures or similar methods known in the art. (See, e.g., Ausubel, 1997, *supra*, units 5.1-6.6.) Reverse transcription was initiated using oligo d(T) or random primers. Synthetic oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-1000 bp) using SEPHACRYL S1000, SEPHAROSE CL2B, or SEPHAROSE CL4B column 15 chromatography (Amersham Pharmacia Biotech) or preparative agarose gel electrophoresis. cDNAs were ligated into compatible restriction enzyme sites of the polylinker of a suitable plasmid, e.g., PBLUESCRIPT plasmid (Stratagene), PSORT1 plasmid (Life Technologies), PCDNA2.1 plasmid (Invitrogen, Carlsbad CA), PBK-CMV plasmid (Stratagene), PCR2-TOPOTA plasmid (Invitrogen), PCMV-ICIS plasmid (Stratagene), pIGEN (Incyte Genomics, Palo Alto CA), pRARE 20 (Incyte Genomics), or pINCY (Incyte Genomics), or derivatives thereof. Recombinant plasmids were transformed into competent *E. coli* cells including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene or DH5 α , DH10B, or ElectroMAX DH10B from Life Technologies.

II. Isolation of cDNA Clones

Plasmids obtained as described in Example I were recovered from host cells by *in vivo* 25 excision using the UNIZAP vector system (Stratagene) or by cell lysis. Plasmids were purified using at least one of the following: a Magic or WIZARD Minipreps DNA purification system (Promega); an AGTC Miniprep purification kit (Edge Biosystems, Gaithersburg MD); and QIAWELL 8 Plasmid, QIAWELL 8 Plus Plasmid, QIAWELL 8 Ultra Plasmid purification systems or the R.E.A.L. PREP 96 plasmid purification kit from QIAGEN. Following precipitation, plasmids 30 were resuspended in 0.1 ml of distilled water and stored, with or without lyophilization, at 4°C.

Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR in a high-throughput format (Rao, V.B. (1994) Anal. Biochem. 216:1-14). Host cell lysis and thermal cycling steps were carried out in a single reaction mixture. Samples were processed and stored in 384-well plates, and the concentration of amplified plasmid DNA was quantified fluorometrically

using PICOGREEN dye (Molecular Probes, Eugene OR) and a FLUOROSKAN II fluorescence scanner (Labsystems Oy, Helsinki, Finland).

III. Sequencing and Analysis

Incyte cDNA recovered in plasmids as described in Example II were sequenced as follows.

5 Sequencing reactions were processed using standard methods or high-throughput instrumentation such as the ABI CATALYST 800 (Applied Biosystems) thermal cycler or the PTC-200 thermal cycler (MJ Research) in conjunction with the HYDRA microdispenser (Robbins Scientific) or the MICROLAB 2200 (Hamilton) liquid transfer system. cDNA sequencing reactions were prepared using reagents provided by Amersham Pharmacia Biotech or supplied in ABI sequencing kits such
10 as the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Applied Biosystems). Electrophoretic separation of cDNA sequencing reactions and detection of labeled polynucleotides were carried out using the MEGABACE 1000 DNA sequencing system (Molecular Dynamics); the ABI PRISM 373 or 377 sequencing system (Applied Biosystems) in conjunction with standard ABI protocols and base calling software; or other sequence analysis systems known in the art. Reading
15 frames within the cDNA sequences were identified using standard methods (reviewed in Ausubel, 1997, *supra*, unit 7.7). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example VIII.

The polynucleotide sequences derived from Incyte cDNAs were validated by removing vector, linker, and poly(A) sequences and by masking ambiguous bases, using algorithms and
20 programs based on BLAST, dynamic programming, and dinucleotide nearest neighbor analysis. The Incyte cDNA sequences or translations thereof were then queried against a selection of public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS, PRINTS, DOMO, PRODOM; PROTEOME databases with sequences from Homo sapiens, Rattus norvegicus, Mus musculus, Caenorhabditis elegans, Saccharomyces cerevisiae,
25 Schizosaccharomyces pombe, and Candida albicans (Incyte Genomics, Palo Alto CA); and hidden Markov model (HMM)-based protein family databases such as PFAM. (HMM is a probabilistic approach which analyzes consensus primary structures of gene families. See, for example, Eddy, S.R. (1996) Curr. Opin. Struct. Biol. 6:361-365.) The queries were performed using programs based on BLAST, FASTA, BLIMPS, and HMMER. The Incyte cDNA sequences were assembled
30 to produce full length polynucleotide sequences. Alternatively, GenBank cDNAs, GenBank ESTs, stitched sequences, stretched sequences, or Genscan-predicted coding sequences (see Examples IV and V) were used to extend Incyte cDNA assemblages to full length. Assembly was performed using programs based on Phred, Phrap, and Consed, and cDNA assemblages were screened for open reading frames using programs based on GeneMark, BLAST, and FASTA. The full length
35 polynucleotide sequences were translated to derive the corresponding full length polypeptide

sequences. Alternatively, a polypeptide of the invention may begin at any of the methionine residues of the full length translated polypeptide. Full length polypeptide sequences were subsequently analyzed by querying against databases such as the GenBank protein databases (genpept), SwissProt, the PROTEOME databases, BLOCKS, PRINTS, DOMO, PRODOM, Prosite, 5 and hidden Markov model (HMM)-based protein family databases such as PFAM. Full length polynucleotide sequences are also analyzed using MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA) and LASERGENE software (DNASTAR). Polynucleotide and polypeptide sequence alignments are generated using default parameters specified by the CLUSTAL algorithm as incorporated into the MEGALIGN multisequence alignment program 10 (DNASTAR), which also calculates the percent identity between aligned sequences.

Table 7 summarizes the tools, programs, and algorithms used for the analysis and assembly of Incyte cDNA and full length sequences and provides applicable descriptions, references, and threshold parameters. The first column of Table 7 shows the tools, programs, and algorithms used, the second column provides brief descriptions thereof, the third column presents appropriate 15 references, all of which are incorporated by reference herein in their entirety, and the fourth column presents, where applicable, the scores, probability values, and other parameters used to evaluate the strength of a match between two sequences (the higher the score or the lower the probability value, the greater the identity between two sequences).

The programs described above for the assembly and analysis of full length polynucleotide 20 and polypeptide sequences were also used to identify polynucleotide sequence fragments from SEQ ID NO:19-36. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies are described in Table 4, column 2.

IV. Identification and Editing of Coding Sequences from Genomic DNA

Putative protein modification and maintenance molecules were initially identified by 25 running the Genscan gene identification program against public genomic sequence databases (e.g., gbpri and gbhtg). Genscan is a general-purpose gene identification program which analyzes genomic DNA sequences from a variety of organisms (See Burge, C. and S. Karlin (1997) J. Mol. Biol. 268:78-94, and Burge, C. and S. Karlin (1998) Curr. Opin. Struct. Biol. 8:346-354). The program concatenates predicted exons to form an assembled cDNA sequence extending from a 30 methionine to a stop codon. The output of Genscan is a FASTA database of polynucleotide and polypeptide sequences. The maximum range of sequence for Genscan to analyze at once was set to 30 kb. To determine which of these Genscan predicted cDNA sequences encode protein modification and maintenance molecules, the encoded polypeptides were analyzed by querying against PFAM models for protein modification and maintenance molecules. Potential protein 35 modification and maintenance molecules were also identified by homology to Incyte cDNA

sequences that had been annotated as protein modification and maintenance molecules. These selected Genscan-predicted sequences were then compared by BLAST analysis to the genpept and gbpr public databases. Where necessary, the Genscan-predicted sequences were then edited by comparison to the top BLAST hit from genpept to correct errors in the sequence predicted by Genscan, such as extra or omitted exons. BLAST analysis was also used to find any Incyte cDNA or public cDNA coverage of the Genscan-predicted sequences, thus providing evidence for transcription. When Incyte cDNA coverage was available, this information was used to correct or confirm the Genscan predicted sequence. Full length polynucleotide sequences were obtained by assembling Genscan-predicted coding sequences with Incyte cDNA sequences and/or public cDNA sequences using the assembly process described in Example III. Alternatively, full length polynucleotide sequences were derived entirely from edited or unedited Genscan-predicted coding sequences.

V. Assembly of Genomic Sequence Data with cDNA Sequence Data

"Stitched" Sequences

Partial cDNA sequences were extended with exons predicted by the Genscan gene identification program described in Example IV. Partial cDNAs assembled as described in Example III were mapped to genomic DNA and parsed into clusters containing related cDNAs and Genscan exon predictions from one or more genomic sequences. Each cluster was analyzed using an algorithm based on graph theory and dynamic programming to integrate cDNA and genomic information, generating possible splice variants that were subsequently confirmed, edited, or extended to create a full length sequence. Sequence intervals in which the entire length of the interval was present on more than one sequence in the cluster were identified, and intervals thus identified were considered to be equivalent by transitivity. For example, if an interval was present on a cDNA and two genomic sequences, then all three intervals were considered to be equivalent. This process allows unrelated but consecutive genomic sequences to be brought together, bridged by cDNA sequence. Intervals thus identified were then "stitched" together by the stitching algorithm in the order that they appear along their parent sequences to generate the longest possible sequence, as well as sequence variants. Linkages between intervals which proceed along one type of parent sequence (cDNA to cDNA or genomic sequence to genomic sequence) were given preference over linkages which change parent type (cDNA to genomic sequence). The resultant stitched sequences were translated and compared by BLAST analysis to the genpept and gbpr public databases. Incorrect exons predicted by Genscan were corrected by comparison to the top BLAST hit from genpept. Sequences were further extended with additional cDNA sequences, or by inspection of genomic DNA, when necessary.

"Stretched" Sequences

Partial DNA sequences were extended to full length with an algorithm based on BLAST analysis. First, partial cDNAs assembled as described in Example III were queried against public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases using the BLAST program. The nearest GenBank protein homolog was then compared by BLAST
5 analysis to either Incyte cDNA sequences or GenScan exon predicted sequences described in Example IV. A chimeric protein was generated by using the resultant high-scoring segment pairs (HSPs) to map the translated sequences onto the GenBank protein homolog. Insertions or deletions may occur in the chimeric protein with respect to the original GenBank protein homolog. The GenBank protein homolog, the chimeric protein, or both were used as probes to search for
10 homologous genomic sequences from the public human genome databases. Partial DNA sequences were therefore "stretched" or extended by the addition of homologous genomic sequences. The resultant stretched sequences were examined to determine whether it contained a complete gene.

VI. Chromosomal Mapping of PMMM Encoding Polynucleotides

The sequences which were used to assemble SEQ ID NO:19-36 were compared with
15 sequences from the Incyte LIFESEQ database and public domain databases using BLAST and other implementations of the Smith-Waterman algorithm. Sequences from these databases that matched SEQ ID NO:19-36 were assembled into clusters of contiguous and overlapping sequences using assembly algorithms such as Phrap (Table 7). Radiation hybrid and genetic mapping data available from public resources such as the Stanford Human Genome Center (SHGC), Whitehead Institute for
20 Genome Research (WIGR), and Généthon were used to determine if any of the clustered sequences had been previously mapped. Inclusion of a mapped sequence in a cluster resulted in the assignment of all sequences of that cluster, including its particular SEQ ID NO:, to that map location.

Map locations are represented by ranges, or intervals, of human chromosomes. The map
25 position of an interval, in centiMorgans, is measured relative to the terminus of the chromosome's p-arm. (The centiMorgan (cM) is a unit of measurement based on recombination frequencies between chromosomal markers. On average, 1 cM is roughly equivalent to 1 megabase (Mb) of DNA in humans, although this can vary widely due to hot and cold spots of recombination.) The cM distances are based on genetic markers mapped by Généthon which provide boundaries for
30 radiation hybrid markers whose sequences were included in each of the clusters. Human genome maps and other resources available to the public, such as the NCBI "GeneMap'99" World Wide Web site (<http://www.ncbi.nlm.nih.gov/genemap/>), can be employed to determine if previously identified disease genes map within or in proximity to the intervals indicated above.

In this manner, SEQ ID NO:19 was mapped to chromosome 1 within the interval from 75.3 to 81.6 centiMorgans. In this manner, SEQ ID NO:27 was mapped to chromosome 1 within the interval from 153.30 to 156.10 centiMorgans.

VII. Analysis of Polynucleotide Expression

5 Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound. (See, e.g., Sambrook, supra, ch. 7; Ausubel (1995) supra, ch. 4 and 16.)

Analogous computer techniques applying BLAST were used to search for identical or
10 related molecules in cDNA databases such as GenBank or LIFESEQ (Incyte Genomics). This analysis is much faster than multiple membrane-based hybridizations. In addition, the sensitivity of the computer search can be modified to determine whether any particular match is categorized as exact or similar. The basis of the search is the product score, which is defined as:

$$15 \quad \frac{\text{BLAST Score} \times \text{Percent Identity}}{5 \times \text{minimum} \{ \text{length}(\text{Seq. 1}), \text{length}(\text{Seq. 2}) \}}$$

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. The product score is a normalized value between 0 and 100, and is
20 calculated as follows: the BLAST score is multiplied by the percent nucleotide identity and the product is divided by (5 times the length of the shorter of the two sequences). The BLAST score is calculated by assigning a score of +5 for every base that matches in a high-scoring segment pair (HSP), and -4 for every mismatch. Two sequences may share more than one HSP (separated by gaps). If there is more than one HSP, then the pair with the highest BLAST score is used to
25 calculate the product score. The product score represents a balance between fractional overlap and quality in a BLAST alignment. For example, a product score of 100 is produced only for 100% identity over the entire length of the shorter of the two sequences being compared. A product score of 70 is produced either by 100% identity and 70% overlap at one end, or by 88% identity and 100% overlap at the other. A product score of 50 is produced either by 100% identity and 50% overlap at
30 one end, or 79% identity and 100% overlap.

Alternatively, polynucleotide sequences encoding PMMM are analyzed with respect to the tissue sources from which they were derived. For example, some full length sequences are assembled, at least in part, with overlapping Incyte cDNA sequences (see Example III). Each cDNA sequence is derived from a cDNA library constructed from a human tissue. Each human tissue is
35 classified into one of the following organ/tissue categories: cardiovascular system; connective

tissue; digestive system; embryonic structures; endocrine system; exocrine glands; genitalia, female; genitalia, male; germ cells; hemic and immune system; liver; musculoskeletal system; nervous system; pancreas; respiratory system; sense organs; skin; stomatognathic system; unclassified/mixed; or urinary tract. The number of libraries in each category is counted and
 5 divided by the total number of libraries across all categories. Similarly, each human tissue is classified into one of the following disease/condition categories: cancer, cell line, developmental, inflammation, neurological, trauma, cardiovascular, pooled, and other, and the number of libraries in each category is counted and divided by the total number of libraries across all categories. The resulting percentages reflect the tissue- and disease-specific expression of cDNA encoding PMMM.
 10 cDNA sequences and cDNA library/tissue information are found in the LIFESEQ GOLD database (Incyte Genomics, Palo Alto CA).

VIII. Extension of PMMM Encoding Polynucleotides

Full length polynucleotide sequences were also produced by extension of an appropriate fragment of the full length molecule using oligonucleotide primers designed from this fragment.
 15 One primer was synthesized to initiate 5' extension of the known fragment, and the other primer was synthesized to initiate 3' extension of the known fragment. The initial primers were designed using OLIGO 4.06 software (National Biosciences), or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would
 20 result in hairpin structures and primer-primer dimerizations was avoided.

Selected human cDNA libraries were used to extend the sequence. If more than one extension was necessary or desired, additional or nested sets of primers were designed.

High fidelity amplification was obtained by PCR using methods well known in the art. PCR was performed in 96-well plates using the PTC-200 thermal cycler (MJ Research, Inc.). The
 25 reaction mix contained DNA template, 200 nmol of each primer, reaction buffer containing Mg^{2+} , $(NH_4)_2SO_4$, and 2-mercaptoethanol, Taq DNA polymerase (Amersham Pharmacia Biotech), ELONGASE enzyme (Life Technologies), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair PCI A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C,
 30 5 min; Step 7: storage at 4°C. In the alternative, the parameters for primer pair T7 and SK+ were as follows: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C.

The concentration of DNA in each well was determined by dispensing 100 μ l PICOGREEN quantitation reagent (0.25% (v/v) PICOGREEN; Molecular Probes, Eugene OR) dissolved in 1X TE
 35 and 0.5 μ l of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Costar,

Acton MA), allowing the DNA to bind to the reagent. The plate was scanned in a Fluoroskan II (Labsystems Oy, Helsinki, Finland) to measure the fluorescence of the sample and to quantify the concentration of DNA. A 5 μ l to 10 μ l aliquot of the reaction mixture was analyzed by electrophoresis on a 1 % agarose gel to determine which reactions were successful in extending the
5 sequence.

The extended nucleotides were desalted and concentrated, transferred to 384-well plates, digested with CviII cholera virus endonuclease (Molecular Biology Research, Madison WI), and sonicated or sheared prior to religation into pUC 18 vector (Amersham Pharmacia Biotech). For
10 shotgun sequencing, the digested nucleotides were separated on low concentration (0.6 to 0.8%) agarose gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended clones were religated using T4 ligase (New England Biolabs, Beverly MA) into pUC 18 vector (Amersham Pharmacia Biotech), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transfected into competent *E. coli* cells. Transformed cells were selected on antibiotic-containing media, and individual colonies were picked and cultured overnight at 37°C in
15 384-well plates in LB/2x carb liquid media.

The cells were lysed, and DNA was amplified by PCR using Taq DNA polymerase (Amersham Pharmacia Biotech) and Pfu DNA polymerase (Stratagene) with the following parameters: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 72°C, 2 min; Step 5: steps 2, 3, and 4 repeated 29 times; Step 6: 72°C, 5 min; Step 7: storage at 4°C. DNA was
20 quantified by PICOGREEN reagent (Molecular Probes) as described above. Samples with low DNA recoveries were reamplified using the same conditions as described above. Samples were diluted with 20% dimethylsulfoxide (1:2, v/v), and sequenced using DYENAMIC energy transfer sequencing primers and the DYENAMIC DIRECT kit (Amersham Pharmacia Biotech) or the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Applied Biosystems).

25 In like manner, full length polynucleotide sequences are verified using the above procedure or are used to obtain 5' regulatory sequences using the above procedure along with oligonucleotides designed for such extension, and an appropriate genomic library.

IX. Labeling and Use of Individual Hybridization Probes

Hybridization probes derived from SEQ ID NO:19-36 are employed to screen cDNAs,
30 genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250 μ Ci of [γ -³²P] adenosine triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase
35 (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a

SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech). An aliquot containing 10^7 counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba I, or Pvu II (DuPont NEN).

5 The DNA from each digest is fractionated on a 0.7% agarose gel and transferred to nylon membranes (Nytran Plus, Schleicher & Schuell, Durham NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under conditions of up to, for example, 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate. Hybridization patterns are visualized using autoradiography or an alternative imaging means and
10 compared.

X. Microarrays

 The linkage or synthesis of array elements upon a microarray can be achieved utilizing photolithography, piezoelectric printing (ink-jet printing, See, e.g., Baldeschweiler, supra), mechanical microspotting technologies, and derivatives thereof. The substrate in each of the
15 aforementioned technologies should be uniform and solid with a non-porous surface (Skena (1999), supra). Suggested substrates include silicon, silica, glass slides, glass chips, and silicon wafers. Alternatively, a procedure analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced using available methods and machines well known to
20 those of ordinary skill in the art and may contain any appropriate number of elements. (See, e.g., Skena, M. et al. (1995) Science 270:467-470; Shalon, D. et al. (1996) Genome Res. 6:639-645; Marshall, A. and J. Hodgson (1998) Nat. Biotechnol. 16:27-31.)

 Full length cDNAs, Expressed Sequence Tags (ESTs), or fragments or oligomers thereof may comprise the elements of the microarray. Fragments or oligomers suitable for hybridization
25 can be selected using software well known in the art such as LASERGENE software (DNASTAR). The array elements are hybridized with polynucleotides in a biological sample. The polynucleotides in the biological sample are conjugated to a fluorescent label or other molecular tag for ease of detection. After hybridization, nonhybridized nucleotides from the biological sample are removed, and a fluorescence scanner is used to detect hybridization at each array element. Alternatively, laser
30 desorption and mass spectrometry may be used for detection of hybridization. The degree of complementarity and the relative abundance of each polynucleotide which hybridizes to an element on the microarray may be assessed. In one embodiment, microarray preparation and usage is described in detail below.

Tissue or Cell Sample Preparation

Total RNA is isolated from tissue samples using the guanidinium thiocyanate method and poly(A)⁺ RNA is purified using the oligo-(dT) cellulose method. Each poly(A)⁺ RNA sample is reverse transcribed using MMLV reverse-transcriptase, 0.05 pg/ μ l oligo-(dT) primer (21mer), 1X first strand buffer, 0.03 units/ μ l RNase inhibitor, 500 μ M dATP, 500 μ M dGTP, 500 μ M dTTP, 40 μ M dCTP, 40 μ M dCTP-Cy3 (BDS) or dCTP-Cy5 (Amersham Pharmacia Biotech). The reverse transcription reaction is performed in a 25 ml volume containing 200 ng poly(A)⁺ RNA with GEMBRIGHT kits (Incyte). Specific control poly(A)⁺ RNAs are synthesized by in vitro transcription from non-coding yeast genomic DNA. After incubation at 37° C for 2 hr, each reaction sample (one with Cy3 and another with Cy5 labeling) is treated with 2.5 ml of 0.5M sodium hydroxide and incubated for 20 minutes at 85° C to stop the reaction and degrade the RNA. Samples are purified using two successive CHROMA SPIN 30 gel filtration spin columns (CLONTECH Laboratories, Inc. (CLONTECH), Palo Alto CA) and after combining, both reaction samples are ethanol precipitated using 1 ml of glycogen (1 mg/ml), 60 ml sodium acetate, and 300 ml of 100% ethanol. The sample is then dried to completion using a SpeedVAC (Savant Instruments Inc., Holbrook NY) and resuspended in 14 μ l 5X SSC/0.2% SDS.

Microarray Preparation

Sequences of the present invention are used to generate array elements. Each array element is amplified from bacterial cells containing vectors with cloned cDNA inserts. PCR amplification uses primers complementary to the vector sequences flanking the cDNA insert. Array elements are amplified in thirty cycles of PCR from an initial quantity of 1-2 ng to a final quantity greater than 5 μ g. Amplified array elements are then purified using SEPHACRYL-400 (Amersham Pharmacia Biotech).

Purified array elements are immobilized on polymer-coated glass slides. Glass microscope slides (Corning) are cleaned by ultrasound in 0.1% SDS and acetone, with extensive distilled water washes between and after treatments. Glass slides are etched in 4% hydrofluoric acid (VWR Scientific Products Corporation (VWR), West Chester PA), washed extensively in distilled water, and coated with 0.05% aminopropyl silane (Sigma) in 95% ethanol. Coated slides are cured in a 110° C oven.

Array elements are applied to the coated glass substrate using a procedure described in U.S. Patent No. 5,807,522, incorporated herein by reference. 1 μ l of the array element DNA, at an average concentration of 100 ng/ μ l, is loaded into the open capillary printing element by a high-speed robotic apparatus. The apparatus then deposits about 5 nl of array element sample per slide.

Microarrays are UV-crosslinked using a STRATALINKER UV-crosslinker (Stratagene). Microarrays are washed at room temperature once in 0.2% SDS and three times in distilled water. Non-specific binding sites are blocked by incubation of microarrays in 0.2% casein in phosphate

buffered saline (PBS) (Tropix, Inc., Bedford MA) for 30 minutes at 60° C followed by washes in 0.2% SDS and distilled water as before.

Hybridization

Hybridization reactions contain 9 μ l of sample mixture consisting of 0.2 μ g each of Cy3 and
5 Cy5 labeled cDNA synthesis products in 5X SSC, 0.2% SDS hybridization buffer. The sample mixture is heated to 65° C for 5 minutes and is aliquoted onto the microarray surface and covered with an 1.8 cm² coverslip. The arrays are transferred to a waterproof chamber having a cavity just slightly larger than a microscope slide. The chamber is kept at 100% humidity internally by the addition of 140 μ l of 5X SSC in a corner of the chamber. The chamber containing the arrays is
10 incubated for about 6.5 hours at 60° C. The arrays are washed for 10 min at 45° C in a first wash buffer (1X SSC, 0.1% SDS), three times for 10 minutes each at 45° C in a second wash buffer (0.1X SSC), and dried.

Detection

Reporter-labeled hybridization complexes are detected with a microscope equipped with an
15 Innova 70 mixed gas 10 W laser (Coherent, Inc., Santa Clara CA) capable of generating spectral lines at 488 nm for excitation of Cy3 and at 632 nm for excitation of Cy5. The excitation laser light is focused on the array using a 20X microscope objective (Nikon, Inc., Melville NY). The slide containing the array is placed on a computer-controlled X-Y stage on the microscope and raster-scanned past the objective. The 1.8 cm x 1.8 cm array used in the present example is scanned with
20 a resolution of 20 micrometers.

In two separate scans, a mixed gas multiline laser excites the two fluorophores sequentially. Emitted light is split, based on wavelength, into two photomultiplier tube detectors (PMT R1477, Hamamatsu Photonics Systems, Bridgewater NJ) corresponding to the two fluorophores. Appropriate filters positioned between the array and the photomultiplier tubes are used to filter the
25 signals. The emission maxima of the fluorophores used are 565 nm for Cy3 and 650 nm for Cy5. Each array is typically scanned twice, one scan per fluorophore using the appropriate filters at the laser source, although the apparatus is capable of recording the spectra from both fluorophores simultaneously.

The sensitivity of the scans is typically calibrated using the signal intensity generated by a
30 cDNA control species added to the sample mixture at a known concentration. A specific location on the array contains a complementary DNA sequence, allowing the intensity of the signal at that location to be correlated with a weight ratio of hybridizing species of 1:100,000. When two samples from different sources (e.g., representing test and control cells), each labeled with a different fluorophore, are hybridized to a single array for the purpose of identifying genes that are

differentially expressed, the calibration is done by labeling samples of the calibrating cDNA with the two fluorophores and adding identical amounts of each to the hybridization mixture.

The output of the photomultiplier tube is digitized using a 12-bit RTI-835H analog-to-digital (A/D) conversion board (Analog Devices, Inc., Norwood MA) installed in an IBM-compatible PC computer. The digitized data are displayed as an image where the signal intensity is mapped using a linear 20-color transformation to a pseudocolor scale ranging from blue (low signal) to red (high signal). The data is also analyzed quantitatively. Where two different fluorophores are excited and measured simultaneously, the data are first corrected for optical crosstalk (due to overlapping emission spectra) between the fluorophores using each fluorophore's emission spectrum.

A grid is superimposed over the fluorescence signal image such that the signal from each spot is centered in each element of the grid. The fluorescence signal within each element is then integrated to obtain a numerical value corresponding to the average intensity of the signal. The software used for signal analysis is the GEMTOOLS gene expression analysis program (Incyte).

15 **XI. Complementary Polynucleotides**

Sequences complementary to the PMMM-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring PMMM. Although use of oligonucleotides comprising from about 15 to 30 base pairs is described, essentially the same procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are designed using OLIGO 4.06 software (National Biosciences) and the coding sequence of PMMM. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding to the PMMM-encoding transcript.

25 **XII. Expression of PMMM**

Expression and purification of PMMM is achieved using bacterial or virus-based expression systems. For expression of PMMM in bacteria, cDNA is subcloned into an appropriate vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the *trp-lac (tac)* hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express PMMM upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of PMMM in eukaryotic cells is achieved by infecting insect or mammalian cell lines with recombinant Autographica californica nuclear polyhedrosis virus (AcMNPV), commonly known as baculovirus. The nonessential polyhedrin gene of

baculovirus is replaced with cDNA encoding PMMM by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription.

Recombinant baculovirus is used to infect Spodoptera frugiperda (Sf9) insect cells in most cases, or
 5 human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus. (See Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945.)

In most expression systems, PMMM is synthesized as a fusion protein with, e.g., glutathione S-transferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid,
 10 single-step, affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-kilodalton enzyme from Schistosoma japonicum, enables the purification of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Amersham Pharmacia Biotech). Following purification, the GST moiety can be proteolytically cleaved from PMMM at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity
 15 purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak). 6-His, a stretch of six consecutive histidine residues, enables purification on metal-chelate resins (QIAGEN). Methods for protein expression and purification are discussed in Ausubel (1995, supra, ch. 10 and 16). Purified PMMM obtained by these methods can be used directly in the assays shown in Examples XVI, XVII, XVIII and XIX, where applicable.

20 **XIII. Functional Assays**

PMMM function is assessed by expressing the sequences encoding PMMM at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include PCMV SPORT (Life Technologies) and PCR3.1 (Invitrogen,
 25 Carlsbad CA), both of which contain the cytomegalovirus promoter. 5-10 μ g of recombinant vector are transiently transfected into a human cell line, for example, an endothelial or hematopoietic cell line, using either liposome formulations or electroporation. 1-2 μ g of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor
 30 of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP; Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP and to evaluate the apoptotic state of the cells and other cellular properties. FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or
 35 coincident with cell death. These events include changes in nuclear DNA content as measured by

staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane
 5 composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M.G. (1994) Flow Cytometry, Oxford, New York NY.

The influence of PMMM on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding PMMM and either CD64 or CD64-GFP.
 10 CD64 and CD64-GFP are expressed on the surface of transfected cells and bind to conserved regions of human immunoglobulin G (IgG). Transfected cells are efficiently separated from nontransfected cells using magnetic beads coated with either human IgG or antibody against CD64 (DYNAL, Lake Success NY). mRNA can be purified from the cells using methods well known by those of skill in the art. Expression of mRNA encoding PMMM and other genes of interest can be
 15 analyzed by northern analysis or microarray techniques.

XIV. Production of PMMM Specific Antibodies

PMMM substantially purified using polyacrylamide gel electrophoresis (PAGE; see, e.g., Harrington, M.G. (1990) *Methods Enzymol.* 182:488-495), or other purification techniques, is used to immunize rabbits and to produce antibodies using standard protocols.

20 Alternatively, the PMMM amino acid sequence is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity, and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are well described in the art. (See, e.g., Ausubel, 1995, supra, ch. 11.)

25 Typically, oligopeptides of about 15 residues in length are synthesized using an ABI 431A peptide synthesizer (Applied Biosystems) using FMOC chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity. (See, e.g., Ausubel, 1995, supra.) Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for
 30 anti-peptide and anti-PMMM activity by, for example, binding the peptide or PMMM to a substrate, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

XV. Purification of Naturally Occurring PMMM Using Specific Antibodies

Naturally occurring or recombinant PMMM is substantially purified by immunoaffinity
 35 chromatography using antibodies specific for PMMM. An immunoaffinity column is constructed

by covalently coupling anti-PMMM antibody to an activated chromatographic resin, such as CNBr-activated SEPHAROSE (Amersham Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing PMMM are passed over the immunoaffinity column, and the column is
5 washed under conditions that allow the preferential absorbance of PMMM (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/PMMM binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and PMMM is collected.

XVI. Identification of Molecules Which Interact with PMMM

10 PMMM, or biologically active fragments thereof, are labeled with ¹²⁵I Bolton-Hunter reagent. (See, e.g., Bolton, A.E. and W.M. Hunter (1973) *Biochem. J.* 133:529-539.) Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled PMMM, washed, and any wells with labeled PMMM complex are assayed. Data obtained using different concentrations of PMMM are used to calculate values for the number, affinity, and
15 association of PMMM with the candidate molecules.

Alternatively, molecules interacting with PMMM are analyzed using the yeast two-hybrid system as described in Fields, S. and O. Song (1989) *Nature* 340:245-246, or using commercially available kits based on the two-hybrid system, such as the MATCHMAKER system (Clontech).

PMMM may also be used in the PATHCALLING process (CuraGen Corp., New Haven
20 CT) which employs the yeast two-hybrid system in a high-throughput manner to determine all interactions between the proteins encoded by two large libraries of genes (Nandabalan, K. et al. (2000) U.S. Patent No. 6,057,101).

XVII. Demonstration of PMMM Activity

PMMM activity can be demonstrated using a generic immunoblotting strategy or through a
25 variety of specific activity assays, some of which are outlined below. As a general approach, cell lines or tissues transformed with a vector containing PMMM coding sequences can be assayed for PMMM activity by immunoblotting. Transformed cells are denatured in SDS in the presence of β -mercaptoethanol, nucleic acids are removed by ethanol precipitation, and proteins are purified by acetone precipitation. Pellets are resuspended in 20 mM Tris buffer at pH 7.5 and incubated with
30 Protein G-Sepharose pre-coated with an antibody specific for PMMM. After washing, the Sepharose beads are boiled in electrophoresis sample buffer, and the eluted proteins subjected to SDS-PAGE. The SDS-PAGE is transferred to a membrane for immunoblotting, and the PMMM activity is assessed by visualizing and quantifying bands on the blot using the antibody specific for PMMM as the primary antibody and ¹²⁵I-labeled IgG specific for the primary antibody as the
35 secondary antibody.

PMMM kinase activity is measured by quantifying the phosphorylation of a protein substrate by PMMM in the presence of gamma-labeled ^{32}P -ATP. PMMM is incubated with the protein substrate, ^{32}P -ATP, and an appropriate kinase buffer. The ^{32}P incorporated into the substrate is separated from free ^{32}P -ATP by electrophoresis and the incorporated ^{32}P is counted using a
5 radioisotope counter. The amount of incorporated ^{32}P is proportional to the activity of PMMM. A determination of the specific amino acid residue phosphorylated is made by phosphoamino acid analysis of the hydrolyzed protein.

PMMM phosphatase activity is measured by the hydrolysis of p-nitrophenyl phosphate (PNPP). PMMM is incubated together with PNPP in HEPES buffer, pH 7.5, in the presence of
10 0.1% β -mercaptoethanol at 37 °C for 60 min. The reaction is stopped by the addition of 6 ml of 10 N NaOH and the increase in light absorbance at 410 nm resulting from the hydrolysis of PNPP is measured using a spectrophotometer. The increase in light absorbance is proportional to the activity of PMMM in the assay (Diamond, R.H. et al. (1994) Mol. Cell. Biol. 14:3752-62).

In the alternative, PMMM phosphatase activity is determined by measuring the amount of
15 phosphate removed from a phosphorylated protein substrate. Reactions are performed with 2 or 4 nM enzyme in a final volume of 30 μl containing 60 mM Tris, pH 7.6, 1 mM EDTA, 1 mM EGTA, 0.1% 2-mercaptoethanol and 10 μM substrate, ^{32}P -labeled on serine/threonine or tyrosine, as appropriate. Reactions are initiated with substrate and incubated at 30° C for 10-15 min. Reactions are quenched with 450 μl of 4% (w/v) activated charcoal in 0.6 M HCl, 90 mM $\text{Na}_4\text{P}_2\text{O}_7$, and 2 mM
20 NaH_2PO_4 , then centrifuged at 12,000 $\times$ g for 5 min. Acid-soluble ^{32}Pi is quantified by liquid scintillation counting (Sinclair, C. et al. (1999) J. Biol. Chem. 274:23666-23672).

PMMM protease activity is measured by the hydrolysis of appropriate synthetic peptide substrates conjugated with various chromogenic molecules in which the degree of hydrolysis is quantified by spectrophotometric (or fluorometric) absorption of the released chromophore
25 (Beynon, R.J. and J.S. Bond (1994) Proteolytic Enzymes: A Practical Approach, Oxford University Press, New York, NY, pp.25-55). Peptide substrates are designed according to the category of protease activity as endopeptidase (serine, cysteine, aspartic proteases, or metalloproteases), aminopeptidase (leucine aminopeptidase), or carboxypeptidase (carboxypeptidases A and B, procollagen C-proteinase). Commonly used chromogens are 2-naphthylamine, 4-nitroaniline, and
30 furylacrylic acid. Assays are performed at ambient temperature and contain an aliquot of the enzyme and the appropriate substrate in a suitable buffer. Reactions are carried out in an optical cuvette, and the increase/decrease in absorbance of the chromogen released during hydrolysis of the peptide substrate is measured. The change in absorbance is proportional to the enzyme activity in the assay.

In the alternative, an assay for PMMM protease activity takes advantage of fluorescence resonance energy transfer (FRET) that occurs when one donor and one acceptor fluorophore with an appropriate spectral overlap are in close proximity. A flexible peptide linker containing a cleavage site specific for PMMM is fused between a red-shifted variant (RSGFP4) and a blue variant (BFP5) of Green Fluorescent Protein. This fusion protein has spectral properties that suggest energy transfer is occurring from BFP5 to RSGFP4. When the fusion protein is incubated with PMMM, the substrate is cleaved, and the two fluorescent proteins dissociate. This is accompanied by a marked decrease in energy transfer which is quantified by comparing the emission spectra before and after the addition of PMMM (Mitra, R.D. et al (1996) Gene 173:13-17). This assay can also be performed in living cells. In this case the fluorescent substrate protein is expressed constitutively in cells and PMMM is introduced on an inducible vector so that FRET can be monitored in the presence and absence of PMMM (Sagot, I. et al (1999) FEBS Letters 447:53-57).

An assay for ubiquitin hydrolase activity measures the hydrolysis of a ubiquitin precursor. The assay is performed at ambient temperature and contains an aliquot of PMMM and the appropriate substrate in a suitable buffer. Chemically synthesized human ubiquitin-valine may be used as substrate. Cleavage of the C-terminal valine residue from the substrate is monitored by capillary electrophoresis (Franklin, K. et al. (1997) Anal. Biochem. 247:305-309).

PMMM protease inhibitor activity for alpha 2-HS-glycoprotein (AHSG) can be measured as a decrease in osteogenic activity in dexamethasone-treated rat bone marrow cell cultures (dex-RBMC). Assays are carried out in 96-well culture plates containing minimal essential medium supplemented with 15% fetal bovine serum, ascorbic acid (50 µg/ml), antibiotics (100 µg/ml penicillin G, 50 µg/ml gentamicin, 0.3 µg/ml fungizone), 10 mM B-glycerophosphate, dexamethasone (10^{-8} M) and various concentrations of PMMM for 12-14 days. Mineralized tissue formation in the cultures is quantified by measuring the absorbance at 525 nm using a 96-well plate reader (Binkert, C. et al. supra).

PMMM protease inhibitor activity for inter-alpha-trypsin inhibitor (ITI) can be measured by a continuous spectrophotometric rate determination of trypsin activity. The assay is performed at ambient temperature in a quartz cuvette in pH 7.6 assay buffer containing 63 mM sodium phosphate, 0.23 mM N α -benzoyle-L-arginine ethyl ester, 0.06 mM hydrochloric acid, 100 units trypsin, and various concentrations of PMMM. Immediately after mixing by inversion, the increase in $A_{253\text{ nm}}$ is recorded for approximately 5 minutes and the enzyme activity is calculated (Bergmeyer, H.U. et al. (1974) Meth. Enzym. Anal. 1:515-516).

PMMM isomerase activity such as peptidyl prolyl *cis/trans* isomerase activity can be assayed by an enzyme assay described by Rahfeld, J.U., et al. (1994) (FEBS Lett. 352: 180-184). The assay is performed at 10 °C in 35 mM HEPES buffer, pH 7.8, containing chymotrypsin (0.5

mg/ml) and PMMM at a variety of concentrations. Under these assay conditions, the substrate, Suc-Ala-Xaa-Pro-Phe-4-NA, is in equilibrium with respect to the prolyl bond, with 80-95% in *trans* and 5-20% in *cis* conformation. An aliquot (2 μ l) of the substrate dissolved in dimethyl sulfoxide (10 mg/ml) is added to the reaction mixture described above. Only the *cis* isomer of the substrate is
 5 a substrate for cleavage by chymotrypsin. Thus, as the substrate is isomerized by PMMM, the product is cleaved by chymotrypsin to produce 4-nitroanilide, which is detected by its absorbance at 390 nm. 4-nitroanilide appears in a time-dependent and a PMMM concentration-dependent manner.

PMMM galactosyltransferase activity can be determined by measuring the transfer of
 10 radiolabeled galactose from UDP-galactose to a GlcNAc-terminated oligosaccharide chain (Kolbinger, F. et al. (1998) J. Biol. Chem. 273:58-65). The sample is incubated with 14 μ l of assay stock solution (180 mM sodium cacodylate, pH 6.5, 1 mg/ml bovine serum albumin, 0.26 mM UDP-galactose, 2 μ l of UDP-[³H]galactose), 1 μ l of MnCl₂ (500 mM), and 2.5 μ l of GlcNAc β O-(CH₂)₈-CO₂Me (37 mg/ml in dimethyl sulfoxide) for 60 minutes at 37 °C. The reaction is quenched
 15 by the addition of 1 ml of water and loaded on a C18 Sep-Pak cartridge (Waters), and the column is washed twice with 5 ml of water to remove unreacted UDP-[³H]galactose. The [³H]galactosylated GlcNAc β O-(CH₂)₈-CO₂Me remains bound to the column during the water washes and is eluted with 5 ml of methanol. Radioactivity in the eluted material is measured by liquid scintillation counting and is proportional to galactosyltransferase activity in the starting sample.

PMMM induction by heat or toxins may be demonstrated using primary cultures of human fibroblasts or human cell lines such as CCL-13, HEK293, or HEP G2 (ATCC). To heat induce PMMM expression, aliquots of cells are incubated at 42°C for 15, 30, or 60 minutes. Control aliquots are incubated at 37°C for the same time periods. To induce PMMM expression by toxins, aliquots of cells are treated with 100 μ M arsenite or 20 mM azetidine-2-carboxylic acid for 0, 3, 6,
 25 or 12 hours. After exposure to heat, arsenite, or the amino acid analogue, samples of the treated cells are harvested and cell lysates prepared for analysis by western blot. Cells are lysed in lysis buffer containing 1% Nonidet P-40, 0.15 M NaCl, 50 mM Tris-HCl, 5 mM EDTA, 2 mM N-ethylmaleimide, 2 mM phenylmethylsulfonyl fluoride, 1 mg/ml leupeptin, and 1 mg/ml pepstatin. Twenty micrograms of the cell lysate is separated on an 8% SDS-PAGE gel and transferred to a
 30 membrane. After blocking with 5% nonfat dry milk/phosphate-buffered saline for 1 h, the membrane is incubated overnight at 4°C or at room temperature for 2-4 hours with an appropriate dilution of anti-PMMM serum in 2% nonfat dry milk/phosphate-buffered saline. The membrane is then washed and incubated with a 1:1000 dilution of horseradish peroxidase-conjugated goat anti-rabbit IgG in 2% dry milk/phosphate-buffered saline. After washing with 0.1% Tween 20 in

phosphate-buffered saline, the PMMM protein is detected and compared to controls using chemiluminescence.

PMMM lysyl hydroxylase activity is determined by measuring the production of hydroxy[^{14}C]lysine from [^{14}C]lysine. Radiolabeled procollagen is incubated with PMMM in
5 buffer containing ascorbic acid, iron sulfate, dithiothreitol, bovine serum albumin, and catalase. Following a 30 minute incubation, the reaction is stopped by the addition of acetone, and centrifuged. The sedimented material is dried, and the hydroxy[^{14}C]lysine is converted to [^{14}C]formaldehyde by oxidation with periodate, and then extracted into toluene. The amount of ^{14}C extracted into toluene is quantified by scintillation counting, and is proportional to the activity of
10 PMMM in the sample (Kivirikko, K., and Myllyla, R. (1982) *Methods Enzymol.* 82:245-304).

XVIII. Identification of PMMM Substrates

Phage display libraries can be used to identify optimal substrate sequences for PMMM. A random hexamer followed by a linker and a known antibody epitope is cloned as an N-terminal extension of gene III in a filamentous phage library. Gene III codes for a coat protein, and the
15 epitope will be displayed on the surface of each phage particle. The library is incubated with PMMM under proteolytic conditions so that the epitope will be removed if the hexamer codes for a PMMM cleavage site. An antibody that recognizes the epitope is added along with immobilized protein A. Uncleaved phage, which still bear the epitope, are removed by centrifugation. Phage in the supernatant are then amplified and undergo several more rounds of screening. Individual phage
20 clones are then isolated and sequenced. Reaction kinetics for these peptide substrates can be studied using an assay in Example XVII, and an optimal cleavage sequence can be derived (Ke, S.H. et al. (1997) *J. Biol. Chem.* 272:16603-16609).

To screen for in vivo PMMM substrates, this method can be expanded to screen a cDNA expression library displayed on the surface of phage particles (T7SELECT™10-3 Phage display
25 vector, Novagen, Madison, WI) or yeast cells (pYD1 yeast display vector kit, Invitrogen, Carlsbad, CA). In this case, entire cDNAs are fused between Gene III and the appropriate epitope.

XIX. Identification of PMMM Inhibitors

Compounds to be tested are arrayed in the wells of a multi-well plate in varying concentrations along with an appropriate buffer and substrate, as described in the assays in Example
30 XVII. PMMM activity is measured for each well and the ability of each compound to inhibit PMMM activity can be determined, as well as the dose-response kinetics. This assay could also be used to identify molecules which enhance PMMM activity.

In the alternative, phage display libraries can be used to screen for peptide PMMM inhibitors. Candidates are found among peptides which bind tightly to a protease. In this case,
35 multi-well plate wells are coated with PMMM and incubated with a random peptide phage display

library or a cyclic peptide library (Koivunen, E. et al. (1999) Nature Biotech 17:768-774). Unbound phage are washed away and selected phage amplified and rescreened for several more rounds.

Candidates are tested for PMMM inhibitory activity using an assay described in Example XVII.

Various modifications and variations of the described methods and systems of the invention
5 will be apparent to those skilled in the art without departing from the scope and spirit of the
invention. Although the invention has been described in connection with certain embodiments, it
should be understood that the invention as claimed should not be unduly limited to such specific
embodiments. Indeed, various modifications of the described modes for carrying out the invention
which are obvious to those skilled in molecular biology or related fields are intended to be within
10 the scope of the following claims.

Table 1

Incyte Project ID	Polypeptide SEQ ID NO:	Incyte Polypeptide ID	Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID
3230318	1	3230318CD1	19	3230318CB1
5928830	2	5928830CD1	20	5928830CB1
7473607	3	7473607CD1	21	7473607CB1
7481673	4	7481673CD1	22	7481673CB1
7484316	5	7484316CD1	23	7484316CB1
7485008	6	7485008CD1	24	7485008CB1
4820375	7	4820375CD1	25	4820375CB1
7483698	8	7483698CD1	26	7483698CB1
7485421	9	7485421CD1	27	7485421CB1
7485720	10	7485720CD1	28	7485720CB1
7485896	11	7485896CD1	29	7485896CB1
7972712	12	7972712CD1	30	7972712CB1
2751509	13	2751509CD1	31	2751509CB1
7480192	14	7480192CD1	32	7480192CB1
55047465	15	55047465CD1	33	55047465CB1
55063036	16	55063036CD1	34	55063036CB1
6178623	17	6178623CD1	35	6178623CB1
7484157	18	7484157CD1	36	7484157CB1

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO: or PROTEOME ID NO:	Probability Score	Annotation
1	3230318CD1	g1666075	1.8E-233	ubiquitin hydrolase [Homo sapiens] (Jones, M.H. et al. (1996) Hum. Mol. Genet. 5:1695-1701; Jones, M.H. et al. (1997) Hum. Mol. Genet. 6:334-5)
2	5928830CD1	g306708	1.5E-231	dipeptidyl aminopeptidase like protein [Homo sapiens]
3	7473607CD1	g1865716	2.7E-93	procollagen I N-proteinase [Bos taurus]
4	7481673CD1	g2330791	5.2E-46	carboxypeptidase s precursor [Schizosaccharomyces pombe]
6	7485008CD1	g5923788	4.3E-120	zinc metalloprotease ADAMTS7 [Homo sapiens]
7	4820375CD1	g164084	4.8E-84	[Canis familiaris] signal peptidase 21 kDa subunit Greenburg G. et al., (1989) J. Biol. Chem. 264:15762-15765
8	7483698CD1	g9581879	0	[Homo sapiens] disintegrin metalloproteinase with thrombospondin repeats Clark, M.E. et al., (2000) Genomics 67 :343-350
9	7485421CD1	g7673618	3.7E-219	[Mus musculus] ubiquitin specific protease Kim, J. et al. (2000) Genome Res. 10:1138-47
10	7485720CD1	g971270	1.2E-133	[Homo sapiens] proteasome subunit p40 / Mov34 protein Tsurumi, C. et al. (1995) Biochem. Biophys. Res. Commun. 210:600-608
11	7485896CD1	g11077995	7.4E-68	[Homo sapiens] (AF013990) ubiquitin C-terminal hydrolase
12	7972712CD1	g11095188	2.3E-299	[Homo sapiens] (AF221634) dipeptidyl peptidase 8 Abbott, C.A. et al. (2000) Eur. J. Biochem. 267:6140-6150
13	2751509CD1	g11181573	4.6E-56	Prostasin [Rattus norvegicus]
14	7480192CD1	g190263	1.0E-12	[Homo sapiens] plasma prekallikrein Chung, D.W., et al. (1986) Human plasma prekallikrein, a zymogen to a serine protease that contains four tandem repeats. Biochemistry 25, 2410-2417
15	55047465CD1	g1429371	8.0E-91	[Drosophila melanogaster] ubiquitin-specific protease Henchoz, S., et al. (1996) The dose of a putative ubiquitin-specific protease affects position-effect variegation in Drosophila melanogaster. Mol. Cell. Biol. 16, 5717-5725

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO: or PROTEOME ID NO:	Probability Score	Annotation
16	55063036CD1	g818209	8.0E-43	[Thermus sp.] serine proteinase Munro, G.K., et al. (1995) A gene encoding a thermophilic alkaline serine proteinase from Thermus sp. strain Rt41A and its expression in Escherichia coli. Microbiology 141, 1731-1738
17	6178623CD1	g7671653	1.3E-07	[Homo sapiens] dJ298J18.3 (novel protein similar to putative HIV-1 induced protein HIN-1 and Drosophila ovarian tumor locus protein OTU) Makarova, K.S., et al. (2000) A novel superfamily of predicted cysteine proteases from eukaryotes, viruses and Chlamydia pneumoniae. Trends Biochem. Sci. 25, 50-52.
18	7484157CD1	g1754714	1.1E-115	[Xenopus laevis] oviductin Lindsay, L.L., et al. (1999) Oviductin, the Xenopus laevis oviductal protease that processes egg envelope glycoprotein gp43, increases sperm binding to envelopes, and is translated as part of an unusual mosaic protein composed of two protease and several CUB domains. Biol. Reprod. 60, 989-995

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
1	3230318CD1	2642	S44 S102 S187 S283 S296 S337 S342 S358 S407 S515 S542 S561 S593 S855 S884 S940 S990 S1007 S1018 S1065 S1160 S1167 S1178 S1204 S1222 S1247 S1291 S1316 S1321 S1323 S1327 S1401 S1403 S1422 S1484 S1547 S1581 S1642 S1741 S1802 S1916 S1919 S1947 S1965 S2094 S2103 S2352 S2377 S2515 S2558 S2583 S2635 T277 T317 T365 T424 T718 T725 T762 T816	N51 N262 N870 N985 N1002 N1028 N1297 N1398 N1573 N2012 N2171 N2567 N2570 N2605	Ubiquitin carboxyl-terminal hydrolases family 2 signatures: V1711-V1742, L1972-R2063	HMMER_PFAM

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
1 (cont.)			T897 T959 T1053 T1120 T1151 T1193 T1210 T1276 T1302 T1545 T1558 T1565 T1890 T2014 T2202 T2262 T2574 Y1481		Ubiquitin carboxyl-terminal hydrolases family 2 signatures BL00972: G1712-L1729, R1791-F1800, F1820-C1834, N1975-D1999, K2005-L2026	BLIMPS_BLOCKS
					PROBABLE UBIQUITIN CARBOXYL TERMINAL HYDROLASE THIOLESTERASE UBIQUITIN-SPECIFIC PROCESSING PROTEASE DE-UBIQUITINATING ENZYME PD014578: W653-L1483, K424-F1710, L2285-L2309, M2440-L2482, P161-S187	BLAST_PRODOR
					UBIQUITIN CARBOXYL-TERMINAL HYDROLASES FAMILY 2 DM00659 Q09879 217-457: N1715-G1961, S1965-G1986 DM00659 P50101 209-458: N1715-E1950, Y1976-G1986 DM00659 P50102 141-420: W1778-E1953, S85-L124 DM00659 P39967 359-610: P1774-M1934, E1974-G1986	BLAST_DOMO
					Ubiquitin carboxyl-terminal hydrolases family 2 signature 2: Y1976-Y1993	MOTIFS

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
					Transmembrane domains: L534-S561, D564-H582, F688-G710, L1093-Q1110, S1449-V1475, E1886-R1909, M2424-S2452 N terminus is cytosolic	TMAP
2	5928830CD1	796	S70 S210 S216 S313 S357 S408 S446 S642 S646 S688 S696 S762 T18 T57 T124 T291 T347 T414 T435 T476 T577 Y561	N2 N63 N90 N111 N119 N257 N342 N748 N760	Dipeptidyl peptidase IV (DPP IV) N-term: L69-D578	HMIMER_PFAM
					Prolyl oligopeptidase family: F580-I656	HMIMER_PFAM
					Dipeptidyl peptidase IV PF00930: H130-Y151, R212-P262, Y274-Y300, E318-E350, L418-I428, E473-N518, P552-I589, D590-K632, F668-P695, N718-L738	BLIMPS_PFAM
					Prolyl endopeptidase family BL00708B: D626-I656	BLIMPS_BLOCKS
					DIPEPTIDYL IV HYDROLASE PROTEASE SERINE PEPTIDASE DIPEPTIDASE TRANSMEMBRANE GLYCOPROTEIN PROTEIN PD003086: L69-P546, S328-T577 PD003048: I656-C780 PD025216: L22-R68	BLAST_PRODOR
					PROLYL ENDOPEPTIDASE FAMILY SERINE DM02461 P42659 335-862: P275-E796 DM02461 P27487 192-765: E220-C780 DM02461 I38593 190-759: I222-C780 DM02461 P33894 340-930: I222-V747	BLAST_DOMO

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
					signal_cleavage: M1-S47	SPSCAN
					Transmembrane domain: Q29-L55 N terminus is non-cytosolic	TMAP
3	7473607CD1	1445	S46 S269 S309 S376 S480 S504 S529 S559 S564 S622 S712 S851 S878 S889 S944 S1090 S1255 S1281 S1383 S1402 S1418 S1438 T35 T103 T138 T182 T217 T409 T444 T672 T1035 T1076 T1083 T1101 T1124 T1244 T1267	N145 N149 N555 N582 N617 N670 N710 N831 N1253 N1372	Neutral zinc metalloprotease BL00142: T224-G234	BLIMPS_BLOCKS
					Neutral zinc metalloproteases, zinc-binding region signature T224-F233	MOTIFS
					Matrixin signature PR00138: T224-G249	BLIMPS_PRINTS
					Astacin signature PR00480: F219-H237	BLIMPS_PRINTS
					NEPRILYSIN METALLOPROTEASE PR00786C: T217-F233	BLIMPS_PRINTS
					Thrombospondin type 1 domain: S391-C441, W689-C745, W1079-P1122, G752-C807, A956-C1015	HMMER_PFAM

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
					PROTEIN PROCOLLAGEN SERINE PROTEASE INHIBITOR ALTERNATIVE PD007018: R957-C1070, Y748-P869	BLAST_PRODOME
3 (cont.)					THROMBOSPONDIN TYPE 1 REPEAT DM00275 P35440 485-548: G372-C436 DM00275 P35442 479-542: V386-C436	BLAST_DOMO
					Signal Peptide: M1-G33	HMMER
					Transmembrane domain: R77-R105 N terminus is non-cytosolic	TMAP
4	7481673CD1	473	S44 S68 S142 S240 S275 S363 S371 T155 T340 T462	N72 N252 N395 N413	Peptidase family M20/M25/M40: L55-F286	HMMER_PFAM
					CARBOXYPEPTIDASE S PRECURSOR YSCS GLYX HYDROLASE GLYCOPROTEIN SIGNAL PUTATIVE PD040868: I174-F286, D366-I457	BLAST_PRODOME
					PROTEIN HYDROLASE DESUCCINYLASE ZINC SUCCINYLDIAMINOPIMELATE DEACETYLASE CARBOXYPEPTIDASE ACETYLORNITHINE BIOSYNTHESIS COBAL T PD001449: E53-L172	BLAST_PRODOME
					Transmembrane domain: R4-G27	TMAP
					Signal Peptide: M1-P20, M1-S23, M1-S25, M1-G27, M1-G27	HMMER
					signal_cleavage: M1-S23	SPSCAN

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
5	7484316CD1	1627	S2 S60 S90 S358 S395 S399 S406 S407 S424 S512 S525 S552 S603 S608 S633 S643 S654 S674 S679 S680 S682 S739 S745 S797 S803 S854 S871 S945 S988 S1044 S1067 S1074 S1086 S1101 S1106 S1107 S1131 S1141 S1193 S1229 S1260 S1305 S1315 S1359 S1368 S1479 S1529 T66 T98 T106 T140 T159 T218 T374 T385 T394 T428 T445 T503 T518 T900 T924 T1050 T1060 T1238 T1255 T1328 T1443 T1466 T1514 T1599 Y1061 Y1176 Y1350	N215 N437 N611 N949 N1002 N1208	Ubiquitin carboxyl-terminal hydrolases family 2: K31-L64	HMMER_PFAM

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
5					Ubiquitin carboxyl-terminal family 2 BL00972: S34-L51, I312-V333	BLIMPS_BLOCKS
6	7485008CD1	1035	S71 S79 S109 S348 S385 S412 S517 S541 S588 S626 S675 S757 S796 S810 S838 S936 S960 S1016 T270 T314 T323 T451 T490 T499 T513 T809 T826 T849 T925 T947 T1006 Y972	N684 N794 N836 N896	Neutral zinc metalloprotease BL00142: T366-G376	BLIMPS_BLOCKS
					Neutral zinc metalloproteases, zinc-binding region signature: T366-M375	MOTIFS
					ASTACIN FAMILY SIGNATURE PR00480B: L361-H379	BLIMPS_PRINTS
					METALLOPROTEASE PRECURSOR HYDROLASE SIGNAL ZINC VENOM CELL PROTEIN TRANSMEMBRANE ADHESION PD000791: W290-L429 PD013511: K444-C512	BLAST_PRODROM
					PROTEIN PROCOLLAGEN THROMBOSPONDIN MOTIFS N-PROTEINASE A DISINTEGRIN METALLOPROTEASE WITH ADAMTS1 PD011654: V603-C676	BLAST_PRODROM

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
6 (cont.)					ZINC; METALLOPEPTIDASE; NEUTRAL; ATROLYSIN; DM00368 P28891 1-202; N308-L429 DM00368 P22796 1-199; D336-L429 DM00368 S48160 193-396; T337-L428 DM00368 Q05910 189-395; W290-L428	BLAST_DOMO
					Thrombospondin type 1 domain: T870-C923, G981-A1030, T807-C863, G524-C572, T925-C975	HMMER_PFAM
					Transmembrane domain: M1-N20	TMAP
					Signal Peptide: M1-G16, M1-L18, M1-G21, M1-S24	HMMER
7	4820375CD1	185	S185 T48 T118 Y25		Signal peptidase I: V17-G176	HMMER_PFAM
					Transmembrane domains: V17-K42, V144-Y169 N-terminus is non-cytosolic	TMAP
					Signal peptidases I serine proteases BL00501: L38-E63, P64-E79, V88-I99, N111-L130	BLIMPS_BLOCKS
					Signal peptidases I signatures: M32-I87	PROFILESKAN
					Eukaryotic signal peptidase (S27) family signature PR00728: F116-D127, V57-L72	BLIMPS_PRINTS
					SIGNAL PEPTIDASES I DM03793 P13679 44-182; L38-A177 DM03793 P21378 33-171; L38-G176 DM03793 P42667 33-171; L38-G176 DM03793 P15367 21-166; M39-E184	BLAST_DOMO
					Cell attachment sequence R68-D70	MOTIFS
					Signal peptidases I serine active site S59-F66	MOTIFS
					Signal peptidases I signature 3 F116-G129	MOTIFS

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
8	7483698CD1	962	S64 S129 S176 S276 S291 S374 S408 S632 S647 S694 S711 S811 S863 T81 T124 T163 T256 T266 T285 T306 T440 T442 T508 T673 T710 T722 T739 T830 Y399	N19 N118 N372 N630 N645 N656 N737 N798	Neutral zinc metalloproteinases, zinc-binding region proteins BL00142: T327-G337	BLIMPS_BLOCKS
					Reprolysin family propeptide: N19-F142	HMMER_PFAM
					Reprolysin (M12B) family zinc metalloprotease: R186-P394	HMMER_PFAM
					Thrombospondin type 1 domain: G488-C538, W898- C950	HMMER_PFAM
					PROTEIN PROCOLLAGEN THROMBOSPONDIN MOTIFS N-PROTEINASE A DISINTEGRIN METALLOPROTEASE WITH ADAMTS1 PD011654: K579-C648	BLAST_PRODROM
					METALLOPROTEASE PRECURSOR HYDROLASE SIGNAL ZINC VENOM CELL PROTEIN TRANSMEMBRANE ADHESION PD000791: I182-P394	BLAST_PRODROM
					PROTEIN PROCOLLAGEN THROMBOSPONDIN MOTIFS N-PROTEINASE A DISINTEGRIN METALLOPROTEASE WITH ADAMTS1 PD014161: K649-I767	BLAST_PRODROM

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
8 (cont.)					SIMILAR TO THROMBOSPONDIN PD036756: V483-G661, G459-V474	BLAST_PRODROM
					ZINC; METALLOPEPTIDASE; NEUTRAL; ATROLYSIN; DM00368 P15167 190-392: R186-P394 DM00368 P22796 1-199: R186-P394 DM00368 JC4342 190-392: Y184-P394 DM00368 P34179 1-202: R186-P394	BLAST_DOMO
					Neutral zinc metalloproteinases, zinc-binding region signature: T327-L336	MOTIFS
9	7485421CD1	508	S129 S189 S361 S371 S375 S398 S423 S486 T29 T69 T134 T151 T302 T344 T353 T380 T406 Y496	N153	signal_cleavage: M1-S22	SPSCAN
					Signal Peptide: M1-S22	HMMER
					Ubiquitin carboxyl-terminal hydrolases family 2: R159-D190	HMMER_PFAM
					Ubiquitin carboxyl-terminal hydrolase family 2: N442-K502	HMMER_PFAM
					Zn-finger in ubiquitin-hydrolases and other proteins: C46-Q115	HMMER_PFAM
					Ubiquitin carboxyl-terminal hydrolases family 2 proteins BL00972: K445-Q469, K471-D492, G160-P177, Y241-L250, I280-C294	BLIMPS_BLOCKS

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
9 (cont.)					PROTEASE UBIQUITIN HYDROLASE UBIQUITIN SPECIFIC ENZYME DEUBIQUITINATING CARBOXYL TERMINAL THIOLESTERASE PROCESSING CONJUGATION PD017412: V333-E426, D305-S318	BLAST_PRODOM
					UBIQUITIN CARBOXYL-TERMINAL HYDROLASES FAMILY 2 DM00659 P40818 782- 1103: T344-L498, L164-H275, I279-S318	BLAST_DOMO
					UBIQUITIN CARBOXYL-TERMINAL HYDROLASES FAMILY 2 DM00659 P50102 141- 420: N163-I314, L345-G456	BLAST_DOMO
					UBIQUITIN CARBOXYL-TERMINAL HYDROLASES FAMILY 2 DM00659 Q09738 149- 388: N163-I314, T344-M417, S439-G456	BLAST_DOMO
					UBIQUITIN CARBOXYL-TERMINAL HYDROLASES FAMILY 2 DM00659 P32571 566- 873: T344-G456, N163-H258, I279-P315	BLAST_DOMO
					Ubiquitin carboxyl-terminal hydrolases family 2 signature 1: G160-Q175	MOTIFS
					Ubiquitin carboxyl-terminal hydrolases family 2 signature 2: Y446-Y463	MOTIFS
10	7485720CD1	321	S208 S238 S293 T77 T160 T215	N103	Signal Cleavage: M254-A268	HMMER
					Mov34/MPN/PAD-1 family: L5-K276	HMMER_PFAM
					Transmembrane Domain: F246-N274 N-terminus is Non-cytosolic	TMAP

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
10 (cont.)					PROTEIN PROTEASOME SUBUNIT 26S MOV34 REGULATORY A S12 PAD1 HOMOLOG PD005425: L132-K276, Q8-P98	BLAST_PRODUM
					26S; S12; REGULATORY; PROTEASOME; DM07904 P51665 1-323: D4-K321 DM07904 P26270 1-337: D4-K320	BLAST_DOMO
11	7485896CD1	1123	S23 S74 S120 S182 S204 S210 S290 S337 S377 S458 S472 S508 S516 S635 S683 S687 S712 S721 S729 S744 S796 S838 S865 S981 S986 S1057 S1069 S1077 S1088 T346 T541 T550 T580 T624 T802 T856 T911 T985 T1117 Y682		Ubiquitin carboxyl-terminal hydrolases family: Q101- G132, L648-K709	HMMER_PFAM
					Transmembrane Domain: R292-R310, V450-A467 N-terminus is cytosolic	TMAP
					Ubiquitin carboxyl-terminal hydrolases family 2 proteins BL00972: G102-L119, N181-L190, V238- C252, L651-N675, D678-T699	BLIMPS_BLOCKS

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
11 (cont.)					PROTEASE UBIQUITIN HYDROLASE ENZYME UBIQUITIN SPECIFIC CARBOXYL-TERMINAL DE-UBIQUITINATING THIOLESTERASE PROCESSING CONJUGATION PD000590: P80-T261 PD017412: Q531-Q621	BLAST_PRODOM
					UBIQUITIN CARBOXYL-TERMINAL HYDROLASES FAMILY 2 DM00659 P40818 782-1103: G107-I272, C540- V619, Y652-L705 DM00659 P35123 139-432: G107-T366 DM00659 P51784 41-331: G107-V358 DM00659 P35125 220-508: G107-P378	BLAST_DOMO
					Ubiquitin carboxyl-terminal hydrolases family 2 signature 1: G102-Q117	MOTIFS
					Ubiquitin carboxyl-terminal hydrolases family 2 signature 2: Y652-Y669	MOTIFS
12	7972712CD1	892	S16 S70 S74 S98 S133 S151 S312 S331 S361 S428 S462 S502 S607 S623 S684 S870 T169 T245 T252 T301 T419 T519 T550 T644 T796 Y334	N234 N682	Prolyl oligopeptidase family: R692-L764	HMMER_PFAM
					Transmembrane Domain: W758-T786 N-terminus is cytosolic	TMAP

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
12					Prolyl aminopeptidase (S33) family signature PR00793: V669-Q677, V703-R714, I755-L769	BLIMPS_PRINTS
(cont.)					Prolyl oligopeptidase serine protease (S9A) signature PR00862: P667-F685, G757-K777	BLIMPS_PRINTS
					Dipeptidyl peptidase IV PF00930: E326-Q352, L518-R528, T572-Q617, A697-Q739, F776-P803, R828-L848	BLIMPS_PFAM
					DIPEPTIDYL IV HYDROLASE PROTEASE SERINE PEPTIDASE DIPEPTIDASE TRANSMEMBRANE GLYCOPROTEIN PROTEIN PD003048: L764-E890 PD003086: P511-V681, G269-Q402	BLAST_PRODOM
					HYDROLASE PROTEIN PROTEASE DIPEPTIDYL SERINE IV PEPTIDASE ENZYME AMINOPEPTIDASE DIPEPTIDASE PD001060: L693-L764	BLAST_PRODOM
					PROLYL ENDOPEPTIDASE FAMILY SERINE DM02461 P27487 192-765: F508-E890, G269-W404 DM02461 P18962 229-817: L518-R872, G269-L346 DM02461 P42659 335-862: A517-E890, Y332-Q424 DM02461 P33894 340-930: K513-R872, R333-Q352	BLAST_DOMO
					Cell attachment sequence R40-D42	MOTIFS
13	2751509CD1	818	S44 S132 S164 S630 S649 S650 T29 T238 T306 T696 T769 Y704	N92 N130 N217 N317 N369 N402 N407 N421	Signal peptide: M1-Q22	SPScan

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
13 (cont.)					Signal peptide: M1-Q22	HMMER
					Trypsin: I47-I286, P323-P454, W564-W679	HMMER-PFAM
					Transmembrane domains: A2-Q22, H70-H87	TMAP
					Serine proteases, trypsin family, histidine active site: L83-C88	MOTIFS
					Serine proteases, trypsin family, serine active site: 237-V248	MOTIFS
					Kringle domain proteins BL00021: C72-F89, V158-G179, G245-I286	BLIMPS-BLOCKS
					Apple domain proteins BL00495: P18-W58, Q110-P148, A149-E183, A229-F263, G264-G292	BLIMPS-BLOCKS
					Type I fibronectin domain BL01253: C72-A85, S147-E183, R236-C249, W255-Q289	BLIMPS-BLOCKS
					Serine proteases, trypsin family, active sites: W60-G116, V570-G621, I222-R269	ProfileScan
					Chymotrypsin serine protease family (S1) signature PR00722: G73-C88, E135-A149, R236-V248	BLIMPS-PRINTS
					Protease, serine, hydrolase, zymogen PD0000046: I47-P273, A121-I286, G326-F438	BLAST-PRODOM
					Trypsin: DM00018 A57014 45-284: I47-V290 DM00018 P15944 31-270: I47-V290 DM00018 P21845 31-271: I47-V290 DM00018 P15157 31-270: I47-V290	BLAST-DOMO

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
14	7480192CD1	296	S167 S203 S229 S244 T110 T156 T166 T198 T201	N24 N194 N228	Trypsin domain: H8-R86, K115-I165	HMMER_PFAM
					Serine proteases, trypsin signature BL00134: T116-R139, L152-I165	BLIMPS_BLOCKS
					TRYPSIN DM00018 P03952 392-624: V13-K168 DM00018 P26262 391-624: V13-K168 DM00018 P14272 391-624: V13-K168 DM00018 P10323 42-288: P3-K168	BLAST_DOMO
15	55047465CD1	420	S88 S91 S123 S128 S145 S204 S226 S322 T108 T225 T303 T390 T399 Y296 Y377	N65 N161	Ubiquitin carboxyl-terminal hydrolases family domain: V168-W199	HMMER_PFAM
					Ubiquitin carboxyl-terminal hydrolases signature BL00972: G169-L186, W248-C257, I277-C291	BLIMPS_BLOCKS
					PROTEASE UBIQUITIN HYDROLASE UBIQUITIN-SPECIFIC ENZYME DE- UBIQUITINATING CARBOXYL TERMINAL THIOLESTERASE PROCESSING CONJUGATION PD017412: R300-E406	BLAST_PRODOM

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
15 (cont.)					UBIQUITIN CARBOXYL-TERMINAL HYDROLASES FAMILY DM00659 P50101 209-458: N172-D402 DM00659 Q09879 217-457: N172-H420 DM00659 P38187 434-611: T271-D402 DM00659 P39967 359-610: D275-F400	BLAST_DOMO
					Ubiquitin carboxyl-terminal hydrolases family 2 signature 1: G169-Q184	MOTIFS
16	55063036CD1	629	S5 S6 S48 S115 S158 S172 S183 S186 S266 S356 S429 S501 S603 T140 T344 T510 T604	N470	signal_cleavage: M1-A31	SPSCAN
					Signal Peptide: M1-A29, M1-A31	HMMER
					Subtilase family domain: S84-N362	HMMER_PFAM
					Transmembrane domains: S6-A29 L213-Y230 N-terminus is cytosolic	TMAP
					Serine proteases, subtilase family signature BL00136: L121-I133, D161-G173, G321-G331	BLIMPS_BLOCKS
					Serine proteases, subtilase family, active sites: A300- Q350	PROFILES SCAN
					Subtilisin serine protease family (S8) signature PR00723: G114-I133, K159-S172, S320-M336	BLIMPS_PRINTS

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
16 (cont.)					SERINE PROTEASES, SUBTILASE FAMILY, HISTIDINE DM00108 P80146 150-377: G114-T344 DM00108 P08594 146-372: G114-A346 DM00108 P16588 160-386: D106-L337 DM00108 P06873 123-352: V119-S338	BLAST_DOMO
17	6178623CD1	398	S2 S15 S169 S185 S211 S224 S287 S326 S372 S379 T165 T342 Y102	N241 N377	OTU-like cysteine protease: V70-D184	HMMER_PFAM
18	7484157CD1	863	S49 S63 S91 S93 S319 S326 S355 S362 S394 S533 S647 S798 S824 T103 T126 T297 T422 T513 T683 T766 T812 T851 T856 Y618 Y674	N52 N99 N324 N392 N468 N605 N645	signal_cleavage: M1-G22	SPSCAN
					Signal Peptide: M1-P18, M1-G22	HMMER
					CUB domain: C380-F489	HMMER_PFAM
					Trypsin domain: I536-N605, I47-I291, D650-I748	HMMER_PFAM
					Transmembrane domains: L4-L21 C243-N271 L625-L649 P732-I748 N-terminus is non-cytosolic	TMAP
					Serine proteases, trypsin signature BL00134: C561-C577, D698-G721, P735-I748	BLIMPS_BLOCKS
					Serine proteases, trypsin family, active sites: L64-Q115, L216-G264, L553-T602, I682-Q731	PROFILES SCAN

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
18 (cont.)					PROTEASE SERINE PRECURSOR SIGNAL HYDROLASE ZYMOGEN GLYCOPROTEIN FAMILY MULTIGENE FACTOR PD000046: P59-I291, L657-I748, I536-I660	BLAST_PRODROM
					GLYCOPROTEIN DOMAIN EGF-LIKE PROTEIN PRECURSOR SIGNAL RECEPTOR INTRINSIC FACTOR B12 REPEAT PD000165: G390-I491	BLAST_PRODROM
					TRYPSIN DM00018 P03952 392-624: N52-Q293, C561-K751 DM00018 P26262 391-624: N52-Q293, W547-K751 DM00018 P40313 33-260: R46-Q293, R535-I752 DM00018 P20918 576-808: R535-I752, P59-L295	BLAST_DOMO
					Serine proteases, trypsin family, histidine active site: V83-C88, I572-C577	MOTIFS
					Serine proteases, trypsin family, serine active site: D231-V242	MOTIFS

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
19/3230318CB1/8285	1-366, 6-132, 132-468, 238-953, 299-366, 396-920, 426-918, 511-2646, 1162-1362, 1181-1362, 1861-2310, 1861-2334, 1943-2514, 2099-2756, 2180-2514, 2287-2742, 2287-2810, 2287-3065, 2380-2756, 2701-2756, 2701-2945, 2701-3265, 2734-3559, 2847-3423, 2863-3559, 2872-3512, 2908-3460, 2920-3508, 2930-3468, 2934-3595, 3002-3595, 3069-3460, 3169-3566, 3177-3701, 3200-3701, 3309-3701, 3314-3795, 3378-3701, 3397-3699, 3397-3701, 3541-3701, 3580-3701, 3702-4413, 3815-4303, 3815-4337, 3815-4342, 4004-4636, 4115-4856, 4168-4342, 4342-4413, 4430-4687, 4430-4929, 4829-4946, 4829-6269, 4893-5431, 5156-5371, 5156-5526, 5259-5660, 5385-5655, 5410-5657, 5410-5909, 5588-6036, 5638-6152, 5675-6098, 5675-6183, 5731-6340, 5795-6065, 5795-6134, 6121-6294, 6245-6361, 6245-6802, 6337-6569, 6473-6970, 6527-6734, 6534-6734, 6542-6632, 6542-6868, 6542-6926, 6543-6629, 6543-6630, 6549-6773, 6549-6901, 6621-7380, 6721-6980, 6744-7220, 6752-6889, 6814-7129, 6814-7141, 6831-7065, 6834-7250, 6867-7096, 6964-7220, 6964-7399, 6964-7618, 7031-7443, 7049-7673, 7055-7452, 7055-7501, 7055-7645, 7056-7460, 7123-7552, 7123-7737, 7124-7779, 7126-7686, 7178-7276, 7184-7550, 7185-7660, 7319-8102, 7438-7923, 7486-8044, 7496-8031, 7502-7806, 7507-8059, 7512-7737, 7513-7756, 7546-8054, 7546-8127, 7547-7800, 7555-7866, 7566-8114, 7569-7817, 7573-8144, 7593-7885, 7614-8082, 7614-8156, 7617-8156, 7623-7846, 7660-7875, 7660-8285, 7693-8138, 7701-8242, 7704-7756, 7707-8230, 7715-8176, 7718-8285, 7721-8002, 7728-7854
20/5928830CB1/2767	1-603, 1-639, 68-751, 75-750, 255-616, 407-666, 407-687, 407-749, 407-859, 478-751, 768-1417, 774-1422, 820-1067, 860-1569, 872-1529, 887-1390, 894-1424, 898-1688, 924-1013, 932-1482, 939-1045, 969-1649, 971-1232, 971-1471, 971-1474, 971-1486, 971-1500, 971-1524, 981-1602, 982-1826, 990-1636, 994-1618, 1000-1610, 1003-1688, 1040-1623, 1059-1487, 1060-1688, 1062-1684, 1063-1684, 1065-1740, 1071-1523, 1080-1714, 1153-1623, 1158-1776, 1188-1780, 1189-1688, 1199-1832, 1203-1529, 1207-1889, 1220-1842, 1222-2014, 1223-1899, 1225-1798, 1227-1731, 1233-1775, 1233-1776, 1234-1776, 1239-1530, 1239-1725, 1239-1831, 1239-1886, 1239-1894, 1252-1906, 1252-1954, 1279-1621, 1292-1941, 1292-1952, 1295-1740, 1297-1496, 1297-1625, 1297-1687, 1297-1688, 1297-1694, 1305-2017, 1308-1974, 1310-1808, 1316-1688, 1324-2036, 1343-1519, 1343-2021, 1353-1915, 1355-1989, 1356-1903, 1362-1688, 1367-1886, 1384-1989, 1386-2101, 1388-1888, 1391-1967, 1392-1673, 1395-2054, 1400-1776, 1417-1640, 1418-2017, 1418-2109, 1423-2086, 1425-1898, 1431-2122, 1433-2065, 1436-2162, 1443-2069, 1444-2047, 1447-2110, 1448-1684, 1448-1688, 1450-

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
20 (cont.)	2247, 1457-1688, 1462-1684, 1462-1832, 1470-2113, 1471-2283, 1525-2109, 1533-2204, 1533-2229, 1534-2315, 1556-2241, 1558-2236, 1564-2333, 1572-2279, 1574-2270, 1578-2162, 1586-2255, 1594-2392, 1597-2236, 1597-2247, 1599-2223, 1604-2306, 1607-2236, 1607-2277, 1610-2331, 1612-2396, 1631-2332, 1650-2277, 1650-2285, 1656-2218, 1657-2236, 1660-2220, 1662-2286, 1665-2256, 1674-2207, 1674-2357, 1676-2250, 1676-2291, 1680-2339, 1684-2390, 1698-2344, 1699-2321, 1702-2432, 1707-2396, 1712-2255, 1721-2288, 1722-2281, 1722-2369, 1730-2379, 1735-2291, 1735-2325, 1743-2315, 1745-2431, 1746-2421, 1748-2195, 1749-2220, 1751-2396, 1751-2412, 1753-2585, 1760-2418, 1770-2343, 1770-2545, 1771-2265, 1772-2290, 1774-2394, 1779-2299, 1779-2378, 1779-2479, 1783-2354, 1784-2308, 1787-2361, 1788-2374, 1792-2360, 1792-2650, 1796-2442, 1803-2498, 1806-2340, 1819-2389, 1821-2483, 1823-2651, 1827-2548, 1829-2604, 1839-2545, 1842-2516, 1846-2505, 1848-2590, 1851-2533, 1854-2527, 1861-2492, 1871-2402, 1872-2420, 1874-1978, 1877-2501, 1879-2451, 1879-2536, 1880-2458, 1882-2579, 1887-2505, 1891-2361, 1905-2566, 1908-2346, 1909-2572, 1915-2624, 1917-2512, 1918-2589, 1920-2596, 1923-2577, 1931-2257, 1940-2679, 1940-2683, 1955-2624, 1956-2644, 1960-2574, 1962-2738, 1974-2566, 1975-2590, 1980-2670, 1983-2672, 1984-2516, 1987-2651, 1992-2566, 1994-2758, 1997-2635, 1997-2664, 1997-2742, 2001-2670, 2009-2664, 2010-2269, 2010-2372, 2010-2379, 2010-2503, 2010-2602, 2010-2614, 2010-2618, 2010-2621, 2010-2644, 2015-2681, 2019-2616, 2020-2656, 2024-2624, 2035-2577, 2035-2763, 2038-2696, 2047-2681, 2049-2350, 2049-2648, 2050-2671, 2052-2648, 2055-2735, 2062-2564, 2068-2700, 2075-2735, 2078-2746, 2080-2131, 2085-2663, 2087-2684, 2087-2744, 2091-2726, 2093-2643, 2115-2763, 2116-2767
21/7473607CB1/5266	1-548, 1-634, 1-641, 1-642, 1-646, 1-790, 1-1047, 709-3761, 890-1131, 949-1285, 1065-1528, 1077-1534, 1085-1541, 1098-1541, 1111-1517, 1131-1541, 1186-1541, 1254-1541, 1267-1528, 1810-5046, 1892-2561, 1981-2480, 2224-2836, 2224-2889, 2321-2422, 2528-3150, 3135-3246, 3135-3777, 3138-3327, 3389-3914, 3433-4058, 3442-3704, 3442-4240, 3446-3590, 3504-3788, 3767-4331, 3967-5046, 4331-4422, 4337-4961, 4350-4602, 4368-4655, 4396-4646, 4423-4754, 4502-4806, 4600-5183, 4652-5260, 4653-4839, 4673-5181, 4673-5266, 4693-4881, 4693-5236, 4693-5247, 4747-5021, 4747-5025, 4839-5250, 4840-5032, 4851-5259, 4919-5250, 4947-5251

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
22/7481673CB1/1779	1-280, 33-458, 42-649, 65-392, 109-1530, 363-539, 363-598, 364-598, 387-597, 387-598, 389-598, 400-598, 411-598, 502-598, 598-641, 598-645, 598-653, 598-699, 598-717, 598-723, 598-725, 598-740, 598-747, 598-781, 598-833, 598-842, 598-933, 598-936, 598-968, 598-1025, 598-1030, 598-1051, 598-1057, 598-1058, 598-1069, 598-1077, 598-1080, 598-1084, 598-1088, 598-1214, 598-1217, 598-1298, 598-1303, 602-1169, 609-904, 615-1084, 619-1085, 624-1309, 625-1291, 625-1295, 697-1294, 698-1402, 699-1402, 702-1294, 732-1452, 733-1084, 745-1356, 745-1407, 750-1412, 799-1282, 808-1282, 842-1428, 842-1453, 860-1529, 869-1307, 934-1485, 938-1529, 955-1531, 967-1531, 1005-1085, 1010-1084, 1017-1084, 1017-1085, 1028-1084, 1039-1085, 1084-1109, 1084-1122, 1084-1130, 1084-1133, 1084-1137, 1084-1153, 1084-1395, 1084-1451, 1084-1517, 1084-1521, 1084-1531, 1084-1726, 1085-1777, 1087-1117, 1087-1166, 1089-1531, 1093-1531, 1095-1531, 1095-1716, 1098-1710, 1103-1779, 1123-1532, 1125-1643
23/7484316CB1/5187	1-609, 1-812, 40-422, 40-510, 40-610, 46-610, 72-311, 75-610, 83-638, 83-791, 102-696, 582-1216, 643-1242, 643-1287, 699-1327, 959-1385, 1216-1277, 1216-1458, 1269-1458, 1343-2023, 1364-1945, 1391-2023, 1393-1625, 1393-1886, 1405-2027, 1468-2027, 1506-2023, 1525-1589, 1536-2182, 1585-2185, 1615-2128, 1701-2164, 1718-2261, 1959-2094, 2010-2687, 2020-2687, 2068-2687, 2078-2596, 2095-3952, 2194-2555, 2212-2555, 2576-3241, 2578-2719, 2578-2727, 2579-2934, 2580-2775, 2606-3415, 2607-3117, 2690-3363, 2700-3390, 2788-3251, 2789-3182, 2883-3400, 2905-3457, 2913-3537, 2917-3762, 2934-3577, 2952-3541, 2969-3580, 2979-3798, 3073-3589, 3111-3547, 3115-3608, 3120-3412, 3126-3596, 3285-3587, 3285-3735, 3287-3467, 3299-3951, 3309-3761, 3310-3609, 3316-3942, 3327-3951, 3335-3929, 3368-3951, 3390-3953, 3392-3951, 3441-3951, 3442-3927, 3452-4024, 3452-4071, 3462-3951, 3543-4246, 3585-3764, 3601-4071, 3644-4056, 3650-4070, 3663-4328, 3697-3951, 3804-4046, 3840-4070, 3850-4493, 3862-4070, 3993-4071, 4026-4552, 4077-4126, 4112-4390, 4184-4482, 4212-4420, 4212-4578, 4212-4596, 4212-4739, 4216-4672, 4238-4740, 4253-4813, 4285-4545, 4302-4569, 4320-4607, 4328-4739, 4329-4548, 4364-4607, 4392-4537, 4392-4917, 4393-4946, 4415-5015, 4428-4748, 4448-5069, 4456-5175, 4507-4729, 4507-5109, 4512-5008, 4554-4829, 4681-5021, 4708-5187, 4712-5187, 4779-5187, 4828-5129
24/7485008CB1/3165	1-971, 355-615, 730-904, 730-1132, 730-1411, 780-1414, 924-1415, 1103-1636, 1106-1646, 1123-1819, 1249-1646, 1724-2113, 1820-2068, 1820-2461, 1851-2389, 1876-2506, 1906-2423, 1912-2439, 1928-2320, 1984-2632, 2010-2513, 2088-2943, 2116-2371, 2116-2376, 2116-2735, 2134-2904, 2160-2753, 2177-2990, 2192-2789, 2201-2744, 2272-2958, 2276-2868, 2280-2846, 2283-2846, 2297-3070, 2335-2759, 2362-2919, 2420-2922, 2420-3164, 2446-3079, 2462-2654, 2493-3165, 2498-2842, 2527-3029, 2528-2722, 2550-2869, 2550-3078, 2555-3030, 2562-2972, 2565-3063, 2578-2860, 2596-2698, 2606-3102, 2609-2785, 2648-2821

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
25/4820375CB1/1567	1-347, 1-489, 25-276, 25-596, 248-546, 351-575, 351-577, 351-844, 491-757, 491-1057, 494-1044, 765-1134, 786-1046, 864-1496, 920-1108, 920-1113, 920-1121, 920-1127, 920-1154, 920-1163, 920-1168, 920-1206, 920-1211, 920-1214, 920-1217, 920-1256, 920-1301, 920-1467, 922-1096, 922-1119, 923-1095, 924-1258, 935-1122, 935-1273, 941-1179, 944-1557, 945-1508, 947-1455, 951-1177, 951-1222, 956-1523, 958-1185, 963-1557, 964-1307, 964-1447, 964-1469, 971-1243, 973-1262, 975-1161, 977-1235, 986-1252, 991-1260, 994-1037, 995-1462, 1000-1462, 1018-1462, 1030-1556, 1030-1567, 1031-1285, 1033-1567, 1037-1307, 1037-1326, 1040-1276, 1044-1321, 1044-1521, 1045-1554, 1051-1306, 1052-1519, 1053-1300, 1055-1556, 1058-1251, 1065-1567, 1068-1545, 1069-1255, 1074-1560, 1075-1528, 1078-1457, 1085-1561, 1087-1460, 1089-1363, 1095-1297, 1095-1371, 1095-1528, 1096-1373, 1096-1413, 1101-1323, 1101-1350, 1101-1480, 1101-1557, 1108-1556, 1108-1567, 1109-1564, 1110-1559, 1114-1410, 1117-1363, 1122-1382, 1122-1556, 1128-1557, 1131-1345, 1131-1460, 1135-1322, 1138-1387, 1139-1381, 1139-1567, 1140-1562, 1145-1386, 1149-1499, 1159-1557, 1160-1411, 1160-1416, 1165-1557, 1165-1567, 1167-1528, 1168-1407, 1168-1557, 1170-1301, 1170-1560, 1171-1398, 1171-1433, 1171-1557, 1173-1436, 1182-1386, 1182-1470, 1189-1384, 1195-1439, 1201-1557, 1216-1446, 1223-1465, 1223-1511, 1227-1557, 1228-1482, 1228-1525, 1240-1387, 1240-1508, 1240-1561, 1241-1517, 1249-1451, 1256-1551, 1256-1557, 1256-1565, 1263-1557, 1265-1401, 1269-1567, 1359-1467, 1359-1509, 1359-1564
26/7483698CB1/3308	1-902, 111-174, 112-902, 117-902, 126-902, 163-899, 184-902, 369-902, 413-902, 435-902, 657-764, 812-1151, 812-1421, 1317-1567, 1317-1958, 1424-1709, 1568-1709, 1568-1817, 1710-1958, 1768-2033, 1768-2062, 1768-2063, 1768-2067, 1777-1932, 1818-1958, 1818-2132, 1964-2145, 1964-2282, 1964-3308, 1965-2145, 1965-2617, 2146-2400, 2283-2493, 2322-2645, 2322-2654, 2323-2654, 2325-2654, 2401-2493, 2401-2666, 2494-2796, 2667-2923, 2797-3134, 2924-3134, 2924-3308, 3135-3308

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
27/7485421CB1/2207	1-677, 10-681, 112-600, 439-715, 467-592, 467-752, 467-760, 467-839, 467-843, 467-867, 467-891, 467-892, 467-900, 467-921, 467-1000, 468-722, 469-529, 469-664, 474-912, 474-1102, 485-741, 485-1155, 486-760, 493-1157, 540-716, 544-619, 544-620, 598-1226, 624-1035, 632-1165, 633-716, 643-1214, 666-866, 666-1245, 668-1339, 673-877, 673-1320, 686-1368, 690-982, 692-1030, 692-1124, 692-1167, 693-1161, 693-1167, 701-1161, 703-1165, 704-1135, 704-1336, 714-1265, 731-1241, 731-1303, 731-1304, 741-1304, 750-1369, 763-1317, 763-1395, 763-1418, 783-1167, 806-1333, 817-1832, 826-1417, 826-1464, 861-1508, 865-1153, 882-1169, 888-1508, 911-1372, 911-1373, 945-1598, 945-1613, 945-1622, 945-1631, 949-1606, 956-1368, 957-1613, 984-1237, 995-1663, 1025-1545, 1029-1545, 1040-1450, 1075-1617, 1089-1619, 1091-1619, 1099-1663, 1102-1441, 1136-1400, 1147-1400, 1156-1788, 1158-1400, 1222-1400, 1237-1869, 1242-1492, 1243-1882, 1243-1899, 1245-1922, 1289-1705, 1302-1952, 1311-1936, 1312-1400, 1316-1950, 1319-1952, 1328-1399, 1351-1936, 1368-1944, 1429-2029, 1450-2048, 1480-1873, 1497-2057, 1511-1791, 1523-1835, 1551-1854, 1617-1922, 1619-1949, 1619-1961, 1619-2030, 1619-2112, 1619-2116, 1621-1873, 1635-1851, 1640-1864, 1668-2116, 1671-1947, 1675-2184, 1681-1991, 1688-2201, 1764-2201, 1778-2199, 1778-2207, 1787-1946
28/7485720CB1/986	1-726, 21-726, 29-986
29/7485896CB1/3492	1-504, 286-504, 304-825, 371-539, 371-740, 505-636, 524-951, 525-950, 538-951, 741-1105, 762-1132, 876-1290, 1106-1480, 1310-1586, 1337-1793, 1337-1919, 1337-1923, 1337-1927, 1344-1963, 1523-2054, 1628-2335, 1658-2233, 1659-1886, 1660-1956, 1660-2140, 1662-2080, 1733-2391, 1749-2428, 1753-2073, 1768-1886, 1769-2132, 1822-2359, 1830-2239, 1852-2420, 1875-2533, 1904-2617, 1918-2470, 1938-2554, 2000-2706, 2007-2284, 2007-2428, 2007-2477, 2007-2486, 2007-2543, 2007-2552, 2007-2584, 2007-2630, 2007-2636, 2007-2660, 2007-2678, 2007-2690, 2009-2204, 2009-2640, 2009-2683, 2010-2196, 2010-2197, 2011-2197, 2013-2197, 2051-2785, 2081-2746, 2089-2667, 2105-2650, 2143-2854, 2162-2853, 2165-2541, 2172-2423, 2180-2336, 2206-2890, 2225-2736, 2237-2867, 2244-2974, 2259-2815, 2281-2838, 2283-2838, 2296-2812, 2299-2964, 2309-2861, 2316-2945, 2347-2878, 2347-2886, 2347-2920, 2349-2890, 2367-2940, 2371-2885, 2372-2918, 2382-2905, 2388-2524, 2388-2885, 2390-2951, 2398-2951, 2405-3012, 2425-3019, 2428-3097, 2435-3053, 2438-2915, 2450-3109, 2457-3138, 2464-3184, 2473-2869, 2477-3016, 2490-2887, 2501-3080, 2502-3030, 2509-3184,

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
29 (cont.)	2517-3155, 2518-3111, 2519-3030, 2527-2951, 2528-3058, 2530-3199, 2532-3154, 2547-2843, 2547-3019, 2547-3090, 2547-3098, 2547-3124, 2547-3138, 2548-2874, 2552-3041, 2561-2800, 2565-3067, 2580-3191, 2597-2949, 2603-3020, 2603-3244, 2607-3314, 2610-3113, 2610-3277, 2612-3101, 2612-3192, 2620-3243, 2621-3305, 2637-3168, 2640-2882, 2645-3343, 2653-3160, 2657-3285, 2659-3051, 2659-3185, 2681-3341, 2684-3308, 2689-3275, 2694-3319, 2697-3195, 2699-3390, 2703-3419, 2706-3272, 2707-3327, 2708-3044, 2717-3201, 2720-3309, 2728-3401, 2730-3360, 2735-3338, 2737-3169, 2745-3280, 2748-2989, 2754-3272, 2755-2926, 2755-3278, 2755-3365, 2757-3033, 2758-3297, 2759-3046, 2766-3311, 2771-3296, 2771-3403, 2776-3290, 2776-3375, 2790-2996, 2799-3435, 2801-3345, 2801-3397, 2803-3160, 2808-3491, 2810-3463, 2810-3492, 2811-3472, 2813-2989, 2829-3488, 2833-3492, 3319-3433
30/7972712CB1/3716	1-243, 1-298, 1-504, 1-507, 13-313, 18-495, 27-574, 33-206, 46-675, 50-298, 55-564, 64-617, 74-469, 74-524, 77-739, 78-210, 78-371, 78-687, 78-716, 78-734, 78-761, 79-620, 80-754, 83-671, 89-382, 90-630, 91-610, 91-686, 165-791, 296-422, 341-844, 359-849, 474-725, 474-906, 484-952, 503-849, 518-791, 518-792, 545-1085, 620-877, 620-1007, 620-1060, 620-1132, 620-1138, 620-1155, 620-1191, 620-1289, 650-1312, 655-1351, 675-1363, 704-1128, 733-1562, 770-1654, 804-1273, 807-1065, 808-1434, 821-1397, 835-1417, 842-1241, 845-1434, 860-1305, 889-1213, 896-1506, 896-1512, 896-1528, 896-1540, 921-1483, 1022-1552, 1029-1593, 1041-1652, 1046-1689, 1076-1566, 1095-1586, 1099-1621, 1145-1665, 1174-1824, 1184-1417, 1186-1719, 1218-1967, 1223-1878, 1223-1904, 1223-1916, 1223-1927, 1223-1949, 1223-2000, 1241-1891, 1287-1943, 1311-1417, 1348-1954, 1379-1967, 1379-1997, 1379-2003, 1432-2135, 1441-2078, 1441-2106, 1448-2077, 1472-2315, 1474-2315, 1481-2315, 1486-2128, 1486-2315, 1500-2137, 1509-2213, 1517-2365, 1519-2315, 1528-2315, 1536-2315, 1542-2014, 1547-2152, 1550-2132, 1550-2135, 1550-2136, 1550-2146, 1561-2123,

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
30 (cont.)	1580-2176, 1583-2063, 1594-1963, 1596-2047, 1605-2186, 1614-2315, 1625-2105, 1635-1911, 1637-2151, 1665-1868, 1678-2315, 1685-2276, 1686-2315, 1691-2315, 1693-2315, 1699-2315, 1701-2240, 1704-2178, 1704-2328, 1704-2348, 1704-2412, 1707-2298, 1711-2315, 1715-2120, 1751-2256, 1754-1979, 1754-1985, 1767-2370, 1790-2478, 1798-1873, 1809-2057, 1810-2031, 1810-2037, 1811-2086, 1823-2197, 1833-2296, 1837-1991, 1863-2546, 1933-2562, 1938-2634, 1956-2478, 1960-2421, 1961-2657, 1978-2433, 1982-2657, 1998-2579, 2013-2432, 2024-2573, 2034-2657, 2039-2634, 2045-2712, 2079-2634, 2133-2632, 2166-2657, 2180-2638, 2183-2923, 2240-2657, 2241-2441, 2262-2299, 2267-2421, 2267-2535, 2267-2640, 2270-2551, 2298-2420, 2298-2424, 2298-2657, 2305-2903, 2314-3010, 2324-2944, 2336-3000, 2347-2944, 2349-2571, 2357-2593, 2377-2833, 2382-2829, 2410-3085, 2412-3054, 2416-2931, 2419-2903, 2426-2573, 2433-3128, 2434-3063, 2446-3010, 2468-2936, 2521-2766, 2521-3056, 2521-3089, 2521-3124, 2521-3287, 2544-2789, 2546-2568, 2546-2573, 2546-2772, 2546-3288, 2546-3310, 2546-3311, 2546-3340, 2546-3341, 2546-3357, 2546-3362, 2548-3307, 2552-3378, 2562-2969, 2572-2666, 2572-2770, 2572-2862, 2572-2884, 2572-3052, 2572-3111, 2572-3112, 2572-3168, 2572-3212, 2586-3079, 2627-3069, 2653-3264, 2661-3068, 2661-3360, 2663-3494, 2667-3104, 2672-3343, 2678-3276, 2678-3289, 2679-3327, 2681-3264, 2690-3329, 2690-3374, 2696-2948, 2696-2967, 2696-3099, 2696-3148, 2696-3191, 2696-3224, 2696-3236, 2696-3238, 2696-3246, 2696-3272, 2696-3326, 2696-3342, 2696-3367, 2696-3374, 2699-3334, 2705-3289, 2708-3515, 2710-3343, 2713-3504, 2726-3356, 2727-3327, 2728-3339, 2728-3510, 2740-3476, 2782-3499, 2784-3039, 2795-3327, 2808-3149, 2811-3407, 2821-3467, 2823-3675, 2854-3546, 2870-3122, 2870-3528, 2880-3543, 2881-3599, 2891-3541, 2909-3350, 2916-3492, 2924-3286, 2925-3392, 2933-3698, 2942-3675, 2944-3591, 2953-3327, 2965-3343, 2966-3573, 2968-3668, 2986-3394, 2991-3716, 2996-3267, 2998-3665, 3005-3273, 3008-3716, 3019-3296, 3019-3516, 3037-3289, 3049-3297, 3049-3313, 3049-3385, 3049-3510, 3049-3601, 3049-3672, 3049-3687, 3049-3689, 3049-3696, 3049-3698, 3064-3686, 3065-3554, 3071-3611, 3084-3716, 3088-3716, 3097-3632, 3102-3271, 3102-3374, 3104-3336, 3107-3256, 3110-3709, 3113-3408, 3133-3716, 3145-3427, 3145-3716, 3148-3708, 3150-3348, 3151-3630, 3155-3386, 3163-3716, 3165-3716, 3167-3716, 3172-3642, 3172-3713, 3179-3403, 3193-3389, 3195-3450, 3197-3710, 3204-3426, 3204-3489, 3204-3710, 3205-3716, 3213-3716, 3216-3447, 3216-3716, 3223-3716, 3228-3709, 3235-3518, 3239-3637, 3415-3625
31/2751509CB1/2681	1-2183, 38-83, 38-98, 38-112, 38-114, 71-273, 92-549, 101-750, 273-2178, 293-744, 798-1416, 1215-1491, 2057-2344, 2078-2503, 2078-2506, 2122-2389, 2177-2681, 2178-2466, 2178-2620, 2240-2486, 2240-2651, 2316-2676, 2371-2681, 2459-2681

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
32/7480192CB1/1293	1-459, 12-459, 24-468, 33-468, 118-376, 118-381, 170-1039, 170-1042, 280-459, 508-532, 508-560, 508-811, 508-884, 508-985, 509-967, 509-1087, 792-1272, 811-1155, 814-1272, 824-1272, 848-1293, 897-1263, 921-1263, 946-1268, 1044-1262, 1078-1271
33/55047465CB1/1579	1-190, 1-195, 1-212, 1-249, 1-297, 1-306, 1-324, 1-350, 1-363, 1-368, 1-436, 1-441, 1-447, 1-470, 1-486, 1-491, 1-492, 1-518, 1-521, 1-533, 1-553, 1-554, 1-557, 1-558, 1-564, 1-576, 1-587, 1-594, 1-611, 1-618, 1-623, 1-624, 1-627, 1-634, 1-639, 1-642, 1-643, 1-653, 1-686, 1-745, 1-767, 1-772, 1-781, 2-294, 2-701, 4-811, 5-687, 5-781, 6-679, 7-637, 8-786, 9-687, 12-855, 30-247, 30-577, 37-636, 53-636, 59-644, 59-646, 71-601, 71-612, 71-618, 73-678, 74-687, 74-706, 87-173, 87-231, 87-233, 87-402, 87-550, 87-679, 91-298, 97-234, 99-663, 117-704, 117-803, 132-808, 139-759, 164-515, 173-1021, 201-934, 214-620, 214-624, 250-804, 251-579, 270-556, 293-1094, 297-1132, 321-983, 341-453, 362-600, 364-899, 365-955, 365-991, 381-903, 393-1230, 395-1221, 416-1230, 424-1230, 425-1230, 428-1092, 438-1230, 461-1228, 462-1230, 481-1230, 484-945, 484-961, 502-1371, 508-722, 508-1065, 515-1071, 516-1160, 521-1071, 530-670, 541-1073, 541-1077, 541-1083, 541-1091, 541-1100, 541-1102, 541-1103, 548-1230, 554-1230, 561-1230, 571-1325, 575-1230, 602-1230, 606-1230, 622-1230, 628-1230, 632-1230, 636-1191, 651-1230, 651-1516, 666-1230, 668-1230, 676-1230, 679-1047, 685-1230, 686-1230, 690-1517, 702-1230, 713-1138, 735-1499, 749-1001, 749-1230, 752-1230, 754-1230, 755-1230, 758-1230, 762-1574, 765-1047, 765-1551, 779-1143, 783-1230, 784-1230, 795-1230, 796-1143, 798-1230, 801-1230, 809-1143, 809-1230, 819-1023, 830-1143, 835-1143, 837-1576, 848-1230, 862-1579, 866-1143, 898-1468, 899-1579, 912-1576, 917-1143, 921-1576, 925-1224, 937-1576, 939-1576, 963-1391, 963-1576, 1000-1579, 1038-1143, 1047-1214, 1047-1252, 1047-1391, 1050-1549, 1140-1230, 1140-1507, 1157-1391, 1179-1579, 1323-1391
34/55063036CB1/2591	1-557, 2-461, 2-727, 2-747, 6-267, 18-406, 149-748, 160-714, 170-393, 170-773, 195-753, 240-494, 246-2173, 349-494, 383-1110, 401-1110, 420-1110, 437-1111, 468-1110, 474-768, 493-617, 493-752, 493-958, 506-1111, 507-1086, 507-1113, 508-1113, 511-1113, 540-1111, 588-1113, 589-1111, 676-1080, 682-1110, 756-1113, 885-1103, 885-1111, 885-1112, 885-1113, 894-2173, 915-1104, 915-1109, 915-1111, 915-1113, 994-1110, 1050-1113, 1087-1295, 1087-1370, 1087-1432, 1087-1440, 1087-1447, 1087-1464, 1087-1548, 1087-1658, 1087-1670, 1087-1678, 1087-1746, 1087-1760, 1087-1813, 1087-1864, 1087-1871, 1087-1875, 1090-1336, 1104-1779, 1172-1582, 1222-2023, 1225-1475, 1242-2026, 1262-1895, 1279-1836, 1284-1925, 1361-1996, 1374-1823, 1396-1882, 1432-1805, 1455-1772, 1455-1915, 1455-1955, 1455-1994, 1455-2049, 1455-2085, 1463-1855, 1463-2065, 1475-1876, 1485-2206, 1490-2203, 1503-2073, 1508-1995, 1518-2097, 1529-2248, 1537-2156, 1545-2321, 1549-2202, 1571-2034, 1592-1997, 1600-2011, 1608-2248, 1618-2176, 1622-1750, 1624-2153, 1644-2218, 1654-1966, 1658-2221, 1666-2253, 1668-2009, 1679-2347, 1681-2110, 1693-2208, 1695-2392, 1731-

Table 4

Polynucleotide SEQ ID NO:/ Incyte ID/ Sequence Length	Sequence Fragments
34 (cont.)	2339, 1738-2107, 1738-2129, 1738-2284, 1743-2296, 1743-2420, 1750-2389, 1751-2418, 1760-2365, 1776-2201, 1780-2458, 1785-2146, 1798-2134, 1798-2435, 1804-2287, 1804-2490, 1827-2522, 1849-2504, 1856-2494, 1875-2492, 1899-2139, 1899-2162, 1913-2479, 1922-2591, 1928-2486, 1930-2159, 1931-2494, 1934-2569, 1937-2587, 1943-2523, 1944-2512, 1946-2492, 1951-2528, 1954-2535, 1957-2502, 1979-2475, 1980-2550, 2237-2326
35/6178623CB1/1197	1-606, 145-221, 248-370, 307-896, 369-412, 422-835, 565-635, 613-1197, 661-962, 676-835
36/7484157CB1/2627	1-2627, 1028-1163, 1038-1144, 1097-1163, 1104-1160, 1104-1163, 1515-1847, 1515-1848, 1517-1847, 1521-1847, 1567-1847, 1982-2131, 1982-2152, 1982-2166, 1982-2200, 1985-2196, 2387-2524

Table 5

Polynucleotide SEQ ID NO:	Incyte Project ID:	Representative Library
19	3230318CB1	BRAWTDR02
20	5928830CB1	STOMTDE01
21	7473607CB1	BSTMNON02
22	7481673CB1	OVARDIT04
23	7484316CB1	BRAUNOR01
24	7485008CB1	KIDEUNC10
25	4820375CB1	PROSTUT17
27	7485421CB1	ADRETUR01
29	7485896CB1	SINTNOR01
30	7972712CB1	KIDPTDE01
31	2751509CB1	ADRETUE02
32	7480192CB1	LUNGDIN02
33	55047465CB1	BRABDIR01
34	55063036CB1	ESOGTUE01
35	6178623CB1	BRAITDR03
36	7484157CB1	BRABDIR03

Table 6

Library	Vector	Library Description
ADRETUE02	PCDNA2.1	This 5' biased random primed library was constructed using RNA isolated from right adrenal tumor tissue removed from a 49-year-old Caucasian male during unilateral adrenalectomy. Pathology indicated adrenal cortical carcinoma comprising nearly the entire specimen. The tumor was attached to the adrenal gland which showed mild cortical atrophy. The tumor was encapsulated, being surrounded by a thin (1-3 mm) rim of connective tissue. The patient presented with adrenal cancer, abdominal pain, pyrexia of unknown origin, and deficiency anemia. Patient history included benign hypertension. Previous surgeries included adenotonsillectomy. Patient medications included aspirin, calcium, and iron. Family history included atherosclerotic coronary artery disease in the mother; cerebrovascular accident and atherosclerotic coronary artery disease in the father; and benign hypertension in the grandparent(s).
ADRETUR01	PCDNA2.1	This random primed library was constructed using RNA isolated from left upper pole, adrenal gland tumor tissue removed from a 52-year-old Caucasian male during nephroureterectomy and local destruction of renal lesion. Pathology indicated grade 3 adrenal cortical carcinoma forming a mass that infiltrated almost the whole adrenal parenchyma and extended to adjacent adipose tissue. A metastatic tumor nodule was identified in the hilar region. The renal vein was infiltrated by tumor and the neoplastic process was present at the resection margin of the renal vein. Fragments of adrenal cortical carcinoma and thrombus were found in the inferior vena cava. Patient history included abnormal weight loss. Family history included skin cancer, type I diabetes, and neurotic depression.
BRABDIR01	pINCY	Library was constructed using RNA isolated from diseased cerebellum tissue removed from the brain of a 57-year-old Caucasian male, who died from a cerebrovascular accident. Patient history included Huntington's disease, emphysema, and tobacco abuse.
BRABDIR03	pINCY	This random primed library was constructed using RNA isolated from diseased cerebellum tissue removed from the brain of a 57-year-old Caucasian male who died from a cerebrovascular accident. Serologies were negative. Patient history included Huntington's disease, emphysema, and tobacco abuse (3-4 packs per day for 40 years).

Table 6

Library	Vector	Library Description
BRAITDR03	PCDNA2.1	This random primed library was constructed using RNA isolated from allocortex, cingulate posterior tissue removed from a 55-year-old Caucasian female who died from cholangiocarcinoma. Pathology indicated mild meningeal fibrosis predominately over the convexities, scattered axonal spheroids in the white matter of the cingulate cortex and the thalamus, and a few scattered neurofibrillary tangles in the entorhinal cortex and the periaqueductal gray region. Pathology for the associated tumor tissue indicated well-differentiated cholangiocarcinoma of the liver with residual or relapsed tumor. Patient history included cholangiocarcinoma, post-operative Budd-Chiari syndrome, biliary ascites, hydrothorax, dehydration, malnutrition, oliguria and acute renal failure. Previous surgeries included cholecystectomy and resection of 85% of the liver.
BRAUNOR01	pINCY	This random primed library was constructed using RNA isolated from striatum, globus pallidus and posterior putamen tissue removed from an 81-year-old Caucasian female who died from a hemorrhage and ruptured thoracic aorta due to atherosclerosis. Pathology indicated moderate atherosclerosis involving the internal carotids, bilaterally; microscopic infarcts of the frontal cortex and hippocampus; and scattered diffuse amyloid plaques and neurofibrillary tangles, consistent with age. Grossly, the leptomeninges showed only mild thickening and hyalinization along the superior sagittal sinus. The remainder of the leptomeninges was thin and contained some congested blood vessels. Mild atrophy was found mostly in the frontal poles and lobes, and temporal lobes, bilaterally. Microscopically, there were pairs of Alzheimer type II astrocytes within the deep layers of the neocortex. There was increased satellitosis around neurons in the deep gray matter in the middle frontal cortex. The amygdala contained rare diffuse plaques and neurofibrillary tangles. The posterior hippocampus contained a microscopic area of cystic cavitation with hemosiderin-laden macrophages
		surrounded by reactive gliosis. Patient history included sepsis, cholangitis, post-operative atelectasis, pneumonia CAD, cardiomegaly due to left ventricular hypertrophy, splenomegaly, arteriolonephrosclerosis, nodular colloidal goiter, emphysema, CHF, hypothyroidism, and peripheral vascular disease.
BRAWTDR02	PCDNA2.1	This random primed library was constructed using RNA isolated from dentate nucleus tissue removed from a 55-year-old Caucasian female who died from cholangiocarcinoma. Pathology indicated mild meningeal fibrosis predominately over the convexities, scattered axonal spheroids in the white matter of the cingulate cortex and the thalamus, and a few scattered neurofibrillary tangles in the entorhinal cortex and the periaqueductal gray region. Pathology for the associated tumor tissue indicated well-differentiated cholangiocarcinoma of the liver with residual or relapsed tumor. Patient history included cholangiocarcinoma, post-operative Budd-Chiari syndrome, biliary ascites, hydrothorax, dehydration, malnutrition, oliguria and acute renal failure. Previous surgeries included cholecystectomy and resection of 85% of the liver.

Table 6

Library	Vector	Library Description
BSTMNON02	PSPORT1	This normalized brain stem library was constructed from 2.84 million independent clones from a brain stem library. Starting RNA was made from the brain stem tissue of a 72-year-old Caucasian male who died from myocardial infarction. Patient history included coronary artery disease, insulin-dependent diabetes mellitus, and arthritis. Normalization and hybridization conditions were adapted from Soares et al. (PNAS (1994) 91:9228).
ESOGTUE01	pINCY	This 5' biased random primed library was constructed using RNA isolated from esophageal tumor tissue removed from a 61 year-old Caucasian male during a partial esophagectomy, proximal gastrectomy, pyloromyotomy, and regional lymph node excision. Pathology indicated an invasive grade 3 adenocarcinoma in the esophagus, extending distally to involve the gastroesophageal junction. The tumor extended through the muscularis to involve periesophageal and perigastric soft tissues. One perigastric and two periesophageal lymph nodes were positive for tumor. There were multiple perigastric and periesophageal tumor implants. The patient presented with deficiency anemia and myelodysplasia. Patient history included hyperlipidemia, and tobacco and alcohol abuse in remission. Previous surgeries included adenotonsillectomy, rhinoplasty, vasectomy, and hemorrhoidectomy. A previous bone marrow aspiration found the marrow to be hypercellular for age and had a cellularity-to-fat ratio of 95:5. The marrow was focally densely fibrotic. Granulocytic precursors were slightly increased with normal maturation. The estimate of blast cells was greater than 5%. Megakaryocytes were increased and appeared atypical in clusters. Storage cells and granulomata were absent. Patient medications included Epoetin, Danocrine, Berocca Plus tablets, Selenium, vitamin B6 phosphate, vitamins E & C, and beta carotene. Family history included alcohol abuse, atherosclerotic coronary artery disease, type II diabetes, chronic liver disease, and primary cardiomyopathy in the father; and benign hypertension and cerebrovascular disease in the mother.
KIDEUNC10	pINCY	The library was constructed using RNA isolated from an untreated transformed embryonal cell line (293-EBNA) derived from kidney epithelial tissue. The cells were transformed with adenovirus 5 DNA. cDNA synthesis was initiated using a NotI-oligo(dT) primer.

Table 6

Library	Vector	Library Description
KIDPTDE01	PCDNA2.1	This 5' biased random primed library was constructed using RNA isolated from diseased left kidney tissue removed from a 63-year-old Caucasian male during nephroureterectomy, radical prostatectomy, urinary diversion to bowel, and regional lymph node excision. Pathology indicated minimal chronic interstitial nephritis. Pathology for the associated tumor tissue indicated grade 4 transitional cell carcinoma with squamous differentiation involving the lateral bladder wall and extending superficially into the left trigone which shows deep muscle invasion. Tumor encases the proximal left ureter. The patient presented with unspecified abdominal and pelvic symptoms, hematuria, and hyperplasia of the prostate. Patient history included atrial fibrillation and tobacco abuse. The patient was not taking any medications. Family history included cerebrovascular accident and atherosclerotic coronary artery disease in the mother and acute myocardial infarction and atherosclerotic coronary artery disease in the father.
LUNGDI02	pINCY	This normalized lung tissue library was constructed from 7.6x10e5 independent clones from a diseased lung tissue library. Starting RNA was made from RNA isolated from diseased lung tissue. Pathology indicated idiopathic pulmonary disease. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91:9228-9232 and Bonaldo et al., Genome Research 6 (1996):791, except that a significantly longer (48 hours/round) reannealing hybridization was used.
OVARDIT04	pINCY	The library was constructed using RNA isolated from diseased left ovary tissue removed from a 22-year-old Caucasian female during a left ovarian cystectomy. Pathology indicated a mature cystic teratoma (dermoid cyst) of the left ovary which consisted of aggregate pieces of fibrous tissue and hair. Patient history included capsular disruption of spleen, mononucleosis, and chlamydia (treated at age 21).
PROSTUT17	pINCY	The library was constructed using RNA isolated from prostate tumor tissue removed from a 66-year-old Caucasian male during radical prostatectomy, regional lymph node excision, and prostate needle biopsy. Pathology indicated adenocarcinoma Gleason grade 3+4, forming a predominant mass involving the right lobe and the left side centrally. The patient presented with elevated prostate specific antigen (PSA) and induration. Family history included acute myocardial infarction, atherosclerotic coronary artery disease, type II diabetes, hyperlipidemia, and Jakob-Creutzfeldt disease.
SINTNOR01	PCDNA2.1	This random primed library was constructed using RNA isolated from small intestine tissue removed from a 31-year-old Caucasian female during Roux-en-Y gastric bypass. Patient history included clinical obesity.

Table 6

Library	Vector	Library Description
STOMTDE01	PCDNA2.1	This 5' biased random primed library was constructed using RNA isolated from stomach tissue removed from a 61-year-old Caucasian male during a partial esophagectomy, proximal gastrectomy, pyloromyotomy, and regional lymph node excision. Pathology for the associated tumor indicated an invasive grade 3 adenocarcinoma in the esophagus, extending distally to involve the gastroesophageal junction. The tumor extended through the muscularis to involve periesophageal and perigastric soft tissues. One perigastric and two periesophageal lymph nodes were positive for tumor. There were multiple perigastric and periesophageal tumor implants. The patient presented with deficiency anemia and myelodysplasia. Patient history included hyperlipidemia, and tobacco and alcohol abuse in remission. Previous surgeries included adenotonsillectomy, rhinoplasty, vasectomy, and hemorrhoidectomy. A previous bone marrow aspiration found the marrow to be hypercellular for age and had a cellularity-to-fat ratio of 95:5. The marrow was focally densely fibrotic. Granulocytic precursors were slightly increased with normal maturation. The estimate of blast cells was greater than 5%. Megakaryocytes

Table 7

Program	Description	Reference	Parameter Threshold
ABIFACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25:3389-3402.	<i>ESTs</i> : Probability value= 1.0E-8 or less <i>Full Length sequences</i> : Probability value= 1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises at least five functions: fasta, tfasta, fastx, tfastx, and ssearch.	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad Sci. USA 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183:63-98; and Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489.	<i>ESTs</i> : fasta E value=1.06E-6 <i>Assembled ESTs</i> : fasta Identity= 95% or greater and Match length=200 bases or greater; fastx E value=1.0E-8 or less <i>Full Length sequences</i> : fastx score=100 or greater
BLIMPS	A BLocks IMProved Searcher that matches a sequence against those in BLOCKS, PRINTS, DOMO, PRODOM, and PFAM databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S. and J.G. Henikoff (1991) Nucleic Acids Res. 19:6565-6572; Henikoff, J.G. and S. Henikoff (1996) Methods Enzymol. 266:88-105; and Attwood, T.K. et al. (1997) J. Chem. Inf. Comput. Sci. 37:417-424.	Probability value= 1.0E-3 or less
HMMER	An algorithm for searching a query sequence against hidden Markov model (HMM)-based databases of protein family consensus sequences, such as PFAM.	Krogh, A. et al. (1994) J. Mol. Biol. 235:1501-1531; Sonnhammer, E.L.L. et al. (1988) Nucleic Acids Res. 26:320-322; Durbin, R. et al. (1998) Our World View, in a Nutshell, Cambridge Univ. Press, pp. 1-350.	<i>PFAM hits</i> : Probability value= 1.0E-3 or less <i>Signal peptide hits</i> : Score= 0 or greater

Table 7 (cont.)

Program	Description	Reference	Parameter Threshold
ProfileScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Gribskov, M. et al. (1988) CABIOS 4:61-66; Gribskov, M. et al. (1989) Methods Enzymol. 183:146-159; Bairoch, A. et al. (1997) Nucleic Acids Res. 25:217-221.	Normalized quality score≥GCG-specified “HIGH” value for that particular Prosite motif. Generally, score=1.4-2.1.
Phred	A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Phils Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M.S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	Score= 120 or greater; Match length= 56 or greater
Consed	A graphical tool for viewing and editing Phrap assemblies.	Gordon, D. et al. (1998) Genome Res. 8:195-202.	
SPScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielson, H. et al. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12:431-439.	Score=3.5 or greater
TMAP	A program that uses weight matrices to delineate transmembrane segments on protein sequences and determine orientation.	Persson, B. and P. Argos (1994) J. Mol. Biol. 237:182-192; Persson, B. and P. Argos (1996) Protein Sci. 5:363-371.	
TMHMMER	A program that uses a hidden Markov model (HMM) to delineate transmembrane segments on protein sequences and determine orientation.	Sonnhammer, E.L. et al. (1998) Proc. Sixth Intl. Conf. on Intelligent Systems for Mol. Biol., Glasgow et al., eds., The Am. Assoc. for Artificial Intelligence Press, Menlo Park, CA, pp. 175-182.	
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Bairoch, A. et al. (1997) Nucleic Acids Res. 25:217-221; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

<110> INCYTE GENOMICS, INC.
 YUE, Henry
 AZIMZAI, Yalda
 KALLICK, Deborah A.
 BAUGHN, Mariah R.
 GRIFFIN, Jennifer A.
 SWARNAKAR, Anita
 LAL, Preeti
 WALIA, Narinder K.
 HAFALIA, April J.A.
 GANDHI, Ameena R.
 AU-YOUNG, Janice
 ELLIOTT, Vicki S.
 RAMKUMAR, Jayalaxmi
 THANGAVELU, Kavitha
 LU, Yan
 WARREN, Bridget A.
 LU, Dyung Aina M.
 LEE, Ernestine A.
 TRIBOULEY, Catherine M.
 ARVIZU, Chandra
 DELEGEANE, Angele M.
 YAO, Monique G.
 KHAN, Farrah A.
 SANJANWALA, Madhusudan

<120> PROTEIN MODIFICATION AND MAINTENANCE MOLECULES

<130> PI-0310 PCT

<140> To Be Assigned

<141> Herewith

<150> 60/254,399; 60/257,803; 60/260,110; 60/262,851; 60/264,623

<151> 2000-12-08; 2000-12-21; 2001-01-05; 2001-01-19; 2001-01-25

<160> 36

<170> PERL Program

<210> 1

<211> 2642

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 3230318CD1

<400> 1

Met	Thr	Leu	Thr	Val	Ala	Ile	Leu	Glu	Asn	Arg	Asp	Ser	Gly	Ile
1				5				10						15
Gln	Ile	Gly	Val	Leu	Ser	Gly	Met	Ser	Gln	Trp	Cys	Gly	Asp	Glu
				20				25						30
Asp	Gly	Lys	Tyr	Arg	Tyr	Leu	Phe	Glu	Glu	Phe	Ile	Pro	Ser	Lys
				35				40						45
Asn	Asp	Glu	Asn	Gly	Asn	Cys	Ser	Gly	Glu	Gly	Ile	Glu	Phe	Pro
				50				55						60
Thr	Thr	Asn	Leu	Tyr	Glu	Leu	Glu	Ser	Arg	Val	Leu	Thr	Asp	His
				65				70						75
Trp	Ser	Ile	Pro	Tyr	Lys	Arg	Glu	Glu	Ser	Leu	Gly	Lys	Cys	Leu
				80				85						90
Leu	Ala	Ser	Thr	Tyr	Leu	Ala	Arg	Leu	Gly	Leu	Ser	Glu	Ser	Asp

				95					100					105
Glu	Asn	Cys	Arg	Arg	Phe	Met	Asp	Arg	Cys	Met	Pro	Glu	Ala	Phe
				110					115					120
Lys	Lys	Leu	Leu	Thr	Ser	Ser	Ala	Val	His	Lys	Trp	Gly	Thr	Glu
				125					130					135
Ile	His	Glu	Gly	Ile	Tyr	Asn	Met	Leu	Met	Leu	Leu	Ile	Glu	Leu
				140					145					150
Val	Ala	Glu	Arg	Ile	Lys	Gln	Asp	Pro	Ile	Pro	Ile	Gly	Leu	Leu
				155					160					165
Gly	Val	Leu	Thr	Met	Ala	Phe	Asn	Pro	Asp	Asn	Glu	Tyr	His	Phe
				170					175					180
Lys	Asn	Arg	Met	Lys	Val	Ser	Gln	Arg	Asn	Trp	Ala	Glu	Val	Phe
				185					190					195
Gly	Glu	Gly	Asn	Met	Phe	Ala	Val	Ser	Pro	Val	Ser	Thr	Phe	Gln
				200					205					210
Lys	Glu	Pro	His	Gly	Trp	Val	Val	Asp	Leu	Val	Asn	Lys	Phe	Gly
				215					220					225
Glu	Leu	Gly	Gly	Phe	Ala	Ala	Ile	Gln	Ala	Lys	Leu	His	Ser	Glu
				230					235					240
Asp	Ile	Glu	Leu	Gly	Ala	Val	Ser	Ala	Leu	Ile	Gln	Pro	Leu	Gly
				245					250					255
Val	Cys	Ala	Glu	Tyr	Leu	Asn	Ser	Ser	Val	Val	Gln	Pro	Met	Leu
				260					265					270
Asp	Pro	Val	Ile	Leu	Thr	Thr	Ile	Gln	Asp	Val	Arg	Ser	Val	Glu
				275					280					285
Glu	Lys	Asp	Leu	Lys	Asp	Lys	Arg	Leu	Val	Ser	Ile	Pro	Glu	Leu
				290					295					300
Leu	Ser	Ala	Val	Lys	Leu	Leu	Cys	Met	Arg	Phe	Gln	Pro	Asp	Leu
				305					310					315
Val	Thr	Ile	Val	Asp	Asp	Leu	Arg	Leu	Asp	Ile	Leu	Leu	Arg	Met
				320					325					330
Leu	Lys	Ser	Pro	His	Phe	Ser	Ala	Lys	Met	Asn	Ser	Leu	Lys	Glu
				335					340					345
Val	Thr	Lys	Leu	Ile	Glu	Asp	Ser	Thr	Leu	Ser	Lys	Ser	Val	Lys
				350					355					360
Asn	Ala	Ile	Asp	Thr	Asp	Arg	Leu	Leu	Asp	Trp	Leu	Val	Glu	Asn
				365					370					375
Ser	Val	Leu	Ser	Ile	Ala	Leu	Glu	Gly	Asn	Ile	Asp	Gln	Ala	Gln
				380					385					390
Tyr	Cys	Asp	Arg	Ile	Lys	Gly	Ile	Ile	Glu	Leu	Leu	Gly	Ser	Lys
				395					400					405
Leu	Ser	Leu	Asp	Glu	Leu	Thr	Lys	Ile	Trp	Lys	Ile	Gln	Ser	Gly
				410					415					420
Gln	Ser	Ser	Thr	Val	Ile	Glu	Asn	Ile	His	Thr	Ile	Ile	Ala	Ala
				425					430					435
Ala	Ala	Val	Lys	Phe	Asn	Ser	Asp	Gln	Leu	Asn	His	Leu	Phe	Val
				440					445					450
Leu	Ile	Gln	Lys	Val	Leu	Asp	Val	Leu	Trp	Glu	Leu	Ala	His	Leu
				455					460					465
Pro	Thr	Leu	Pro	Ser	Ser	Leu	Ile	Gln	Gln	Ala	Leu	Glu	Glu	His
				470					475					480
Leu	Thr	Ile	Leu	Ser	Asp	Ala	Tyr	Ala	Val	Lys	Glu	Ala	Ile	Lys
				485					490					495
Arg	Ser	Tyr	Ile	Ile	Lys	Cys	Ile	Glu	Asp	Ile	Lys	Arg	Val	Val
				500					505					510
Val	Ser	Arg	Leu	Ser	Gly	Asn	Asp	Cys	Ser	Ser	Pro	Val	Val	Pro
				515					520					525
Val	Leu	Lys	Pro	Gln	Ala	Ser	Pro	Leu	Arg	Gly	Leu	Ile	Thr	Ala
				530					535					540
Ala	Ser	Ser	Val	Asp	Cys	Ala	Ser	Val	Val	Ala	Ala	Ala	Leu	Ile
				545					550					555
Gly	Ala	Ala	Leu	Ser	Ser	His	Leu	Asp	Pro	Gln	Ala	Leu	Phe	Ser
				560					565					570

Leu	Leu	Ser	Ala	Phe	Met	Asp	Phe	Tyr	Lys	Val	His	Ile	Ala	Glu
				575					580					585
Gly	Gly	Gln	Trp	Glu	Asp	Gln	Ser	Pro	Leu	Asp	Met	Ala	Pro	Gly
				590					595					600
Arg	Gly	Val	Asn	Tyr	Leu	Leu	Pro	Leu	Lys	Val	Phe	Phe	Tyr	Ala
				605					610					615
Met	Pro	Phe	Pro	Ala	Arg	Gln	Gln	Gly	Gly	Leu	Thr	Gly	Asp	Tyr
				620					625					630
Val	Ser	Leu	Pro	Gly	Tyr	Thr	Glu	Thr	Lys	Gln	Arg	Ser	Ser	Gln
				635					640					645
Leu	Asn	Asn	Pro	Gln	Phe	Val	Trp	Val	Val	Pro	Ala	Leu	Arg	Gln
				650					655					660
Leu	His	Glu	Ile	Thr	Arg	Ser	Phe	Ile	Lys	Gln	Thr	Tyr	Gln	Lys
				665					670					675
Gln	Asp	Lys	Ser	Ile	Ile	Gln	Asp	Leu	Lys	Lys	Asn	Phe	Glu	Ile
				680					685					690
Val	Lys	Leu	Val	Thr	Gly	Ser	Leu	Ile	Ala	Cys	His	Arg	Leu	Ala
				695					700					705
Ala	Ala	Val	Ala	Gly	Pro	Gly	Gly	Leu	Ser	Gly	Ser	Thr	Leu	Val
				710					715					720
Asp	Gly	Arg	Tyr	Thr	Tyr	Arg	Glu	Tyr	Leu	Glu	Ala	His	Leu	Lys
				725					730					735
Phe	Leu	Ala	Phe	Phe	Leu	Gln	Glu	Ala	Thr	Leu	Tyr	Leu	Gly	Trp
				740					745					750
Asn	Arg	Ala	Lys	Glu	Ile	Trp	Glu	Cys	Leu	Val	Thr	Gly	Gln	Asp
				755					760					765
Val	Cys	Glu	Leu	Asp	Arg	Glu	Met	Cys	Phe	Glu	Trp	Phe	Thr	Lys
				770					775					780
Gly	Gln	His	Asp	Leu	Glu	Ser	Asp	Val	Gln	Gln	Gln	Leu	Phe	Lys
				785					790					795
Glu	Lys	Ile	Leu	Lys	Leu	Glu	Ser	Tyr	Glu	Ile	Thr	Met	Asn	Gly
				800					805					810
Phe	Asn	Leu	Phe	Lys	Thr	Phe	Phe	Glu	Asn	Val	Asn	Leu	Cys	Asp
				815					820					825
His	Arg	Leu	Lys	Arg	Gln	Gly	Ala	Gln	Leu	Tyr	Val	Glu	Lys	Leu
				830					835					840
Glu	Leu	Ile	Gly	Met	Asp	Phe	Ile	Trp	Lys	Ile	Ala	Met	Glu	Ser
				845					850					855
Pro	Asp	Glu	Glu	Ile	Ala	Asn	Glu	Ala	Ile	Gln	Leu	Ile	Ile	Asn
				860					865					870
Tyr	Ser	Tyr	Ile	Asn	Leu	Asn	Pro	Arg	Leu	Lys	Lys	Asp	Ser	Val
				875					880					885
Ser	Leu	His	Lys	Lys	Phe	Ile	Ala	Asp	Cys	Tyr	Thr	Arg	Leu	Glu
				890					895					900
Ala	Ala	Ser	Ser	Ala	Leu	Gly	Gly	Pro	Thr	Leu	Thr	His	Ala	Val
				905					910					915
Thr	Arg	Ala	Thr	Lys	Met	Leu	Thr	Ala	Thr	Ala	Met	Pro	Thr	Val
				920					925					930
Ala	Thr	Ser	Val	Gln	Ser	Pro	Tyr	Arg	Ser	Thr	Lys	Leu	Val	Ile
				935					940					945
Ile	Glu	Arg	Leu	Leu	Leu	Leu	Ala	Glu	Arg	Tyr	Val	Ile	Thr	Ile
				950					955					960
Glu	Asp	Phe	Tyr	Ser	Val	Pro	Arg	Thr	Ile	Leu	Pro	His	Gly	Ala
				965					970					975
Ser	Phe	His	Gly	His	Leu	Leu	Thr	Leu	Asn	Val	Thr	Tyr	Glu	Ser
				980					985					990
Thr	Lys	Asp	Thr	Phe	Thr	Val	Glu	Ala	His	Ser	Asn	Glu	Thr	Ile
				995					1000					1005
Gly	Ser	Val	Arg	Trp	Lys	Ile	Ala	Lys	Gln	Leu	Cys	Ser	Pro	Val
				1010					1015					1020
Asp	Asn	Ile	Gln	Ile	Phe	Thr	Asn	Asp	Ser	Leu	Leu	Thr	Val	Asn
				1025					1030					1035
Lys	Asp	Gln	Lys	Leu	Leu	His	Gln	Leu	Gly	Phe	Ser	Asp	Glu	Gln

1040	1045	1050
Ile Leu Thr Val Lys Thr Ser Gly Ser Gly Thr Pro Ser Gly Ser		
1055	1060	1065
Ser Ala Asp Ser Ser Thr Ser Ser Ser Ser Ser Ser Ser Gly Val		
1070	1075	1080
Phe Ser Ser Ser Tyr Ala Met Glu Gln Glu Lys Ser Leu Pro Gly		
1085	1090	1095
Val Val Met Ala Leu Val Cys Asn Val Phe Asp Met Leu Tyr Gln		
1100	1105	1110
Leu Ala Asn Leu Glu Glu Pro Arg Ile Thr Leu Arg Val Arg Lys		
1115	1120	1125
Leu Leu Leu Leu Ile Pro Thr Asp Pro Ala Ile Gln Glu Ala Leu		
1130	1135	1140
Asp Gln Leu Asp Ser Leu Gly Arg Lys Lys Thr Leu Leu Ser Glu		
1145	1150	1155
Ser Ser Ser Gln Ser Ser Lys Ser Pro Ser Leu Ser Ser Lys Gln		
1160	1165	1170
Gln His Gln Pro Ser Ala Ser Ser Ile Leu Glu Ser Leu Phe Arg		
1175	1180	1185
Ser Phe Ala Pro Gly Met Ser Thr Phe Arg Val Leu Tyr Asn Leu		
1190	1195	1200
Glu Val Leu Ser Ser Lys Leu Met Pro Thr Ala Asp Asp Asp Met		
1205	1210	1215
Ala Arg Ser Cys Ala Lys Ser Phe Cys Glu Asn Phe Leu Lys Ala		
1220	1225	1230
Gly Gly Leu Ser Leu Val Val Asn Val Met Gln Arg Asp Ser Ile		
1235	1240	1245
Pro Ser Glu Val Asp Tyr Glu Thr Arg Gln Gly Val Tyr Ser Ile		
1250	1255	1260
Cys Leu Gln Leu Ala Arg Phe Leu Leu Val Gly Gln Thr Met Pro		
1265	1270	1275
Thr Leu Leu Asp Glu Asp Leu Thr Lys Asp Gly Ile Glu Ala Leu		
1280	1285	1290
Ser Ser Arg Pro Phe Arg Asn Val Ser Arg Gln Thr Ser Arg Gln		
1295	1300	1305
Met Ser Leu Cys Gly Thr Pro Glu Lys Ser Ser Tyr Arg Gln Leu		
1310	1315	1320
Ser Val Ser Asp Arg Ser Ser Ile Arg Val Glu Glu Ile Ile Pro		
1325	1330	1335
Ala Ala Arg Val Ala Ile Gln Thr Met Glu Val Ser Asp Phe Thr		
1340	1345	1350
Ser Thr Val Ala Cys Phe Met Arg Leu Ser Trp Ala Ala Ala Ala		
1355	1360	1365
Gly Arg Leu Asp Leu Val Gly Ser Ser Gln Pro Ile Lys Glu Ser		
1370	1375	1380
Asn Ser Leu Cys Pro Ala Gly Ile Arg Asn Arg Leu Ser Ser Ser		
1385	1390	1395
Gly Ser Asn Cys Ser Ser Gly Ser Glu Gly Glu Pro Val Ala Leu		
1400	1405	1410
His Ala Gly Ile Cys Val Arg Gln Gln Ser Val Ser Thr Lys Asp		
1415	1420	1425
Ser Leu Ile Ala Gly Glu Ala Leu Ser Leu Leu Val Thr Cys Leu		
1430	1435	1440
Gln Leu Arg Ser Gln Gln Leu Ala Ser Phe Tyr Asn Leu Pro Cys		
1445	1450	1455
Val Ala Asp Phe Ile Ile Asp Ile Leu Leu Gly Ser Pro Ser Ala		
1460	1465	1470
Glu Ile Arg Arg Val Ala Cys Asp Gln Leu Tyr Thr Leu Ser Gln		
1475	1480	1485
Thr Asp Thr Ser Ala His Pro Asp Val Gln Lys Pro Asn Gln Phe		
1490	1495	1500
Leu Leu Gly Val Ile Leu Thr Ala Gln Leu Pro Leu Trp Ser Pro		
1505	1510	1515

Thr	Ser	Ile	Met	Arg	Gly	Val	Asn	Gln	Arg	Leu	Leu	Ser	Gln	Cys	1520	1525	1530
Met	Glu	Tyr	Phe	Asp	Leu	Arg	Cys	Gln	Leu	Leu	Asp	Asp	Leu	Thr	1535	1540	1545
Thr	Ser	Glu	Met	Glu	Gln	Leu	Arg	Ile	Ser	Pro	Ala	Thr	Met	Leu	1550	1555	1560
Glu	Asp	Glu	Ile	Thr	Trp	Leu	Asp	Asn	Phe	Glu	Pro	Asn	Arg	Thr	1565	1570	1575
Ala	Glu	Cys	Glu	Thr	Ser	Glu	Ala	Asp	Asn	Ile	Leu	Leu	Ala	Gly	1580	1585	1590
His	Leu	Arg	Leu	Ile	Lys	Thr	Leu	Leu	Ser	Leu	Cys	Gly	Ala	Glu	1595	1600	1605
Lys	Glu	Met	Leu	Gly	Ser	Ser	Leu	Ile	Lys	Pro	Leu	Leu	Asp	Asp	1610	1615	1620
Phe	Leu	Phe	Arg	Ala	Ser	Arg	Ile	Ile	Leu	Asn	Ser	His	Ser	Pro	1625	1630	1635
Ala	Gly	Ser	Ala	Ala	Ile	Ser	Gln	Gln	Asp	Phe	His	Pro	Lys	Cys	1640	1645	1650
Ser	Thr	Ala	Asn	Ser	Arg	Leu	Ala	Ala	Tyr	Glu	Val	Leu	Val	Met	1655	1660	1665
Leu	Ala	Asp	Ser	Ser	Pro	Ser	Asn	Leu	Gln	Ile	Ile	Ile	Lys	Glu	1670	1675	1680
Leu	Leu	Ser	Met	His	His	Gln	Pro	Asp	Pro	Ala	Leu	Thr	Lys	Glu	1685	1690	1695
Phe	Asp	Tyr	Leu	Pro	Pro	Val	Asp	Ser	Arg	Ser	Ser	Ser	Gly	Phe	1700	1705	1710
Val	Gly	Leu	Arg	Asn	Gly	Gly	Ala	Thr	Cys	Tyr	Met	Asn	Ala	Val	1715	1720	1725
Phe	Gln	Gln	Leu	Tyr	Met	Gln	Pro	Gly	Leu	Pro	Glu	Ser	Leu	Leu	1730	1735	1740
Ser	Val	Asp	Asp	Asp	Thr	Asp	Asn	Pro	Asp	Asp	Ser	Val	Phe	Tyr	1745	1750	1755
Gln	Val	Gln	Ser	Leu	Phe	Gly	His	Leu	Met	Glu	Ser	Lys	Leu	Gln	1760	1765	1770
Tyr	Tyr	Val	Pro	Glu	Asn	Phe	Trp	Lys	Ile	Phe	Lys	Met	Trp	Asn	1775	1780	1785
Lys	Glu	Leu	Tyr	Val	Arg	Glu	Gln	Gln	Asp	Ala	Tyr	Glu	Phe	Phe	1790	1795	1800
Thr	Ser	Leu	Ile	Asp	Gln	Met	Asp	Glu	Tyr	Leu	Lys	Lys	Met	Gly	1805	1810	1815
Arg	Asp	Gln	Ile	Phe	Lys	Asn	Thr	Phe	Gln	Gly	Ile	Tyr	Ser	Asp	1820	1825	1830
Gln	Lys	Ile	Cys	Lys	Asp	Cys	Pro	His	Arg	Tyr	Glu	Arg	Glu	Glu	1835	1840	1845
Ala	Phe	Met	Ala	Leu	Asn	Leu	Gly	Val	Thr	Ser	Cys	Gln	Ser	Leu	1850	1855	1860
Glu	Ile	Ser	Leu	Asp	Gln	Phe	Val	Arg	Gly	Glu	Val	Leu	Glu	Gly	1865	1870	1875
Ser	Asn	Ala	Tyr	Tyr	Cys	Glu	Lys	Cys	Lys	Glu	Lys	Arg	Ile	Thr	1880	1885	1890
Val	Lys	Arg	Thr	Cys	Ile	Lys	Ser	Leu	Pro	Ser	Val	Leu	Val	Ile	1895	1900	1905
His	Leu	Met	Arg	Phe	Gly	Phe	Asp	Trp	Glu	Ser	Gly	Arg	Ser	Ile	1910	1915	1920
Lys	Tyr	Asp	Glu	Gln	Ile	Arg	Phe	Pro	Trp	Met	Leu	Asn	Met	Glu	1925	1930	1935
Pro	Tyr	Thr	Val	Ser	Gly	Met	Ala	Arg	Gln	Asp	Ser	Ser	Ser	Glu	1940	1945	1950
Val	Gly	Glu	Asn	Gly	Arg	Ser	Val	Asp	Gln	Gly	Gly	Gly	Gly	Ser	1955	1960	1965
Pro	Arg	Lys	Lys	Val	Ala	Leu	Thr	Glu	Asn	Tyr	Glu	Leu	Val	Gly	1970	1975	1980
Val	Ile	Val	His	Ser	Gly	Gln	Ala	His	Ala	Gly	His	Tyr	Tyr	Ser			

1985	1990	1995
Phe Ile Lys Asp Arg	Arg Gly Cys Gly Lys	Gly Lys Trp Tyr Lys
2000	2005	2010
Phe Asn Asp Thr Val	Ile Glu Glu Phe Asp	Leu Asn Asp Glu Thr
2015	2020	2025
Leu Glu Tyr Glu Cys	Phe Gly Gly Glu Tyr	Arg Pro Lys Val Tyr
2030	2035	2040
Asp Gln Thr Asn Pro	Tyr Thr Asp Val Arg	Arg Arg Tyr Trp Asn
2045	2050	2055
Ala Tyr Met Leu Phe	Tyr Gln Arg Val Ser	Asp Gln Asn Ser Pro
2060	2065	2070
Val Leu Pro Lys Lys	Ser Arg Val Ser Val	Val Arg Gln Glu Ala
2075	2080	2085
Glu Asp Leu Ser Leu	Ser Ala Pro Ser Ser	Pro Glu Ile Ser Pro
2090	2095	2100
Gln Ser Ser Pro Arg	Pro His Arg Pro Asn	Asn Asp Arg Leu Ser
2105	2110	2115
Ile Leu Thr Lys Leu	Val Lys Lys Gly Glu	Lys Lys Gly Leu Phe
2120	2125	2130
Val Glu Lys Met Pro	Ala Arg Ile Tyr Gln	Met Val Arg Asp Glu
2135	2140	2145
Asn Leu Lys Phe Met	Lys Asn Arg Asp Val	Tyr Ser Ser Asp Tyr
2150	2155	2160
Phe Ser Phe Val Leu	Ser Leu Ala Ser Leu	Asn Ala Thr Lys Leu
2165	2170	2175
Lys His Pro Tyr Tyr	Pro Cys Met Ala Lys	Val Ser Leu Gln Leu
2180	2185	2190
Ala Ile Gln Phe Leu	Phe Gln Thr Tyr Leu	Arg Thr Lys Lys Lys
2195	2200	2205
Leu Arg Val Asp Thr	Glu Glu Trp Ile Ala	Thr Ile Glu Ala Leu
2210	2215	2220
Leu Ser Lys Ser Phe	Asp Ala Cys Gln Trp	Leu Val Glu Tyr Phe
2225	2230	2235
Ile Ser Ser Glu Gly	Arg Glu Leu Ile Lys	Ile Phe Leu Leu Glu
2240	2245	2250
Cys Asn Val Arg Glu	Val Arg Val Ala Val	Ala Thr Ile Leu Glu
2255	2260	2265
Lys Thr Leu Asp Ser	Ala Leu Phe Tyr Gln	Asp Lys Leu Lys Ser
2270	2275	2280
Leu His Gln Leu Leu	Glu Val Leu Leu Ala	Leu Leu Asp Lys Asp
2285	2290	2295
Val Pro Glu Asn Cys	Lys Asn Cys Ala Gln	Tyr Phe Phe Leu Phe
2300	2305	2310
Asn Thr Phe Val Gln	Lys Gln Gly Ile Arg	Ala Gly Asp Leu Leu
2315	2320	2325
Leu Arg His Ser Ala	Leu Arg His Met Ile	Ser Phe Leu Leu Gly
2330	2335	2340
Ala Ser Arg Gln Asn	Asn Gln Ile Arg Arg	Trp Ser Ser Ala Gln
2345	2350	2355
Ala Arg Glu Phe Gly	Asn Leu His Asn Thr	Val Ala Leu Leu Val
2360	2365	2370
Leu His Ser Asp Val	Ser Ser Gln Arg Asn	Val Ala Pro Gly Ile
2375	2380	2385
Phe Lys Gln Arg Pro	Pro Ile Ser Ile Ala	Pro Ser Ser Pro Leu
2390	2395	2400
Leu Pro Leu His Glu	Glu Val Glu Ala Leu	Leu Phe Met Ser Glu
2405	2410	2415
Gly Lys Pro Tyr Leu	Leu Glu Val Met Phe	Ala Leu Arg Glu Leu
2420	2425	2430
Thr Gly Ser Leu Leu	Ala Leu Ile Glu Met	Val Val Tyr Cys Cys
2435	2440	2445
Phe Cys Asn Glu His	Phe Ser Phe Thr Met	Leu His Phe Ile Lys
2450	2455	2460

```

Asn Gln Leu Glu Thr Ala Pro Pro His Glu Leu Lys Asn Thr Phe
      2465                      2470                      2475
Gln Leu Leu His Glu Ile Leu Val Ile Glu Asp Pro Ile Gln Ala
      2480                      2485                      2490
Glu Arg Val Lys Phe Val Phe Glu Thr Glu Asn Gly Leu Leu Ala
      2495                      2500                      2505
Leu Met His His Ser Asn His Val Asp Ser Ser Arg Cys Tyr Gln
      2510                      2515                      2520
Cys Val Lys Phe Leu Val Thr Leu Ala Gln Lys Cys Pro Ala Ala
      2525                      2530                      2535
Lys Glu Tyr Phe Lys Glu Asn Ser His His Trp Ser Trp Ala Val
      2540                      2545                      2550
Gln Trp Leu Gln Lys Lys Met Ser Glu His Tyr Trp Thr Pro Gln
      2555                      2560                      2565
Ser Asn Val Ser Asn Glu Thr Ser Thr Gly Lys Thr Phe Gln Arg
      2570                      2575                      2580
Thr Ile Ser Ala Gln Asp Thr Leu Ala Tyr Ala Thr Ala Leu Leu
      2585                      2590                      2595
Asn Glu Lys Glu Gln Ser Gly Ser Ser Asn Gly Ser Glu Ser Ser
      2600                      2605                      2610
Pro Ala Asn Glu Asn Gly Asp Arg His Leu Gln Gln Gly Ser Glu
      2615                      2620                      2625
Ser Pro Met Met Ile Gly Glu Leu Arg Ser Asp Leu Asp Asp Val
      2630                      2635                      2640
Asp Pro

```

<210> 2

<211> 796

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 5928830CD1

<400> 2

```

Met Asn Gln Thr Ala Ser Val Ser His His Ile Lys Cys Gln Pro
  1                      5                      10                      15
Ser Lys Thr Ile Lys Glu Leu Gly Ser Asn Ser Pro Pro Gln Arg
      20                      25                      30
Asn Trp Lys Gly Ile Ala Ile Ala Leu Leu Val Ile Leu Val Val
      35                      40                      45
Cys Ser Leu Ile Thr Met Ser Val Ile Leu Leu Thr Pro Asp Glu
      50                      55                      60
Leu Thr Asn Ser Ser Glu Thr Arg Leu Ser Leu Glu Asp Leu Phe
      65                      70                      75
Arg Lys Asp Phe Val Leu His Asp Pro Glu Ala Arg Trp Ile Asn
      80                      85                      90
Asp Thr Asp Val Val Tyr Lys Ser Glu Asn Gly His Val Ile Lys
      95                      100                     105
Leu Asn Ile Glu Thr Asn Ala Thr Thr Leu Leu Leu Glu Asn Thr
      110                     115                     120
Thr Phe Val Thr Phe Lys Ala Ser Arg His Ser Val Ser Pro Asp
      125                     130                     135
Leu Lys Tyr Val Leu Leu Ala Tyr Asp Val Lys Gln Ile Phe His
      140                     145                     150
Tyr Ser Tyr Thr Ala Ser Tyr Val Ile Tyr Asn Ile His Thr Arg
      155                     160                     165
Glu Val Trp Glu Leu Asn Pro Pro Glu Val Glu Asp Ser Val Leu
      170                     175                     180
Gln Tyr Ala Ala Trp Gly Val Gln Gly Gln Gln Leu Ile Tyr Ile
      185                     190                     195

```

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

Pro	Ile	Thr	Asp	Leu	Lys	Leu	Tyr	Ala	Ser	Ala	Phe	Ser	Glu	Arg	665	670	675
Tyr	Leu	Gly	Met	Pro	Ser	Lys	Glu	Glu	Ser	Thr	Tyr	Gln	Ala	Ala	680	685	690
Ser	Val	Leu	His	Asn	Val	His	Gly	Leu	Lys	Glu	Glu	Asn	Ile	Leu	695	700	705
Ile	Ile	His	Gly	Thr	Ala	Asp	Thr	Lys	Val	His	Phe	Gln	His	Ser	710	715	720
Ala	Glu	Leu	Ile	Lys	His	Leu	Ile	Lys	Ala	Gly	Val	Asn	Tyr	Thr	725	730	735
Met	Gln	Val	Tyr	Pro	Asp	Glu	Gly	His	Asn	Val	Ser	Glu	Lys	Ser	740	745	750
Lys	Tyr	His	Leu	Tyr	Ser	Thr	Ile	Leu	Lys	Phe	Phe	Ser	Asp	Cys	755	760	765
Leu	Lys	Glu	Glu	Ile	Ser	Val	Leu	Pro	Gln	Glu	Pro	Glu	Glu	Asp	770	775	780
Glu															785	790	795

<210> 3

<211> 1445

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7473607CD1

<400> 3

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

				245					250					255
Gly	Ala	Ala	Pro	Arg	Ala	Gly	Leu	Ala	Trp	Ser	Pro	Cys	Ser	Arg
				260					265					270
Arg	Gln	Leu	Leu	Ser	Leu	Leu	Ser	Ala	Gly	Arg	Ala	Arg	Cys	Val
				275					280					285
Trp	Asp	Pro	Pro	Arg	Pro	Gln	Pro	Gly	Ser	Ala	Gly	His	Pro	Pro
				290					295					300
Asp	Ala	Gln	Pro	Gly	Leu	Tyr	Tyr	Ser	Ala	Asn	Glu	Gln	Cys	Arg
				305					310					315
Val	Ala	Phe	Gly	Pro	Lys	Ala	Val	Ala	Cys	Thr	Phe	Ala	Arg	Glu
				320					325					330
His	Leu	Asp	Met	Cys	Gln	Ala	Leu	Ser	Cys	His	Thr	Asp	Pro	Leu
				335					340					345
Asp	Gln	Ser	Ser	Cys	Ser	Arg	Leu	Leu	Val	Pro	Leu	Leu	Asp	Gly
				350					355					360
Thr	Glu	Cys	Gly	Val	Glu	Lys	Trp	Cys	Ser	Lys	Gly	Arg	Cys	Arg
				365					370					375
Ser	Leu	Val	Glu	Leu	Thr	Pro	Ile	Ala	Ala	Val	His	Gly	Arg	Trp
				380					385					390
Ser	Ser	Trp	Gly	Pro	Arg	Ser	Pro	Cys	Ser	Arg	Ser	Cys	Gly	Gly
				395					400					405
Gly	Val	Val	Thr	Arg	Arg	Arg	Gln	Cys	Asn	Asn	Pro	Arg	Pro	Ala
				410					415					420
Phe	Gly	Gly	Arg	Ala	Cys	Val	Gly	Ala	Asp	Leu	Gln	Ala	Glu	Met
				425					430					435
Cys	Asn	Thr	Gln	Ala	Cys	Glu	Lys	Thr	Gln	Leu	Glu	Phe	Met	Ser
				440					445					450
Gln	Gln	Cys	Ala	Arg	Thr	Asp	Gly	Gln	Pro	Leu	Arg	Ser	Ser	Pro
				455					460					465
Gly	Gly	Ala	Ser	Phe	Tyr	His	Trp	Gly	Ala	Ala	Val	Pro	His	Ser
				470					475					480
Gln	Gly	Asp	Ala	Leu	Cys	Arg	His	Met	Cys	Arg	Ala	Ile	Gly	Glu
				485					490					495
Ser	Phe	Ile	Met	Lys	Arg	Gly	Asp	Ser	Phe	Leu	Asp	Gly	Thr	Arg
				500					505					510
Cys	Met	Pro	Ser	Gly	Pro	Arg	Glu	Asp	Gly	Thr	Leu	Ser	Leu	Cys
				515					520					525
Val	Ser	Gly	Ser	Cys	Arg	Thr	Phe	Gly	Cys	Asp	Gly	Arg	Met	Asp
				530					535					540
Ser	Gln	Gln	Val	Trp	Asp	Arg	Cys	Gln	Val	Cys	Gly	Gly	Asp	Asn
				545					550					555
Ser	Thr	Cys	Ser	Pro	Arg	Lys	Gly	Ser	Phe	Thr	Ala	Gly	Arg	Ala
				560					565					570
Arg	Glu	Tyr	Val	Thr	Phe	Leu	Thr	Val	Thr	Pro	Asn	Leu	Thr	Ser
				575					580					585
Val	Tyr	Ile	Ala	Asn	His	Arg	Pro	Leu	Phe	Thr	His	Leu	Ala	Val
				590					595					600
Arg	Ile	Gly	Gly	Arg	Tyr	Val	Val	Ala	Gly	Lys	Met	Ser	Ile	Ser
				605					610					615
Pro	Asn	Thr	Thr	Tyr	Ala	Ser	Leu	Leu	Glu	Asp	Gly	Arg	Val	Glu
				620					625					630
Tyr	Arg	Val	Ala	Leu	Thr	Glu	Asp	Arg	Leu	Pro	Arg	Leu	Glu	Glu
				635					640					645
Ile	Arg	Ile	Trp	Gly	Pro	Leu	Gln	Glu	Asp	Ala	Asp	Ile	Gln	Val
				650					655					660
Tyr	Arg	Arg	Tyr	Gly	Glu	Glu	Tyr	Gly	Asn	Leu	Thr	Arg	Pro	Asp
				665					670					675
Ile	Thr	Phe	Thr	Tyr	Phe	Gln	Pro	Lys	Pro	Arg	Gln	Ala	Trp	Val
				680					685					690
Trp	Ala	Ala	Val	Arg	Gly	Pro	Cys	Ser	Val	Ser	Cys	Gly	Ala	Gly
				695					700					705
Leu	Arg	Trp	Val	Asn	Tyr	Ser	Cys	Leu	Asp	Gln	Ala	Arg	Lys	Glu
				710					715					720

Leu	Val	Glu	Thr	Val	Gln	Cys	Gln	Gly	Ser	Gln	Gln	Pro	Pro	Ala
				725					730					735
Trp	Pro	Glu	Ala	Cys	Val	Leu	Glu	Pro	Cys	Pro	Pro	Tyr	Trp	Ala
				740					745					750
Val	Gly	Asp	Phe	Gly	Pro	Cys	Ser	Ala	Ser	Cys	Gly	Gly	Gly	Leu
				755					760					765
Arg	Glu	Arg	Pro	Val	Arg	Cys	Val	Glu	Ala	Gln	Gly	Ser	Leu	Leu
				770					775					780
Lys	Thr	Leu	Pro	Pro	Ala	Arg	Cys	Arg	Ala	Gly	Ala	Gln	Gln	Pro
				785					790					795
Ala	Val	Ala	Leu	Glu	Thr	Cys	Asn	Pro	Gln	Pro	Cys	Pro	Ala	Arg
				800					805					810
Trp	Glu	Val	Ser	Glu	Pro	Ser	Ser	Cys	Thr	Ser	Ala	Gly	Gly	Ala
				815					820					825
Gly	Leu	Ala	Leu	Glu	Asn	Glu	Thr	Cys	Val	Pro	Gly	Ala	Asp	Gly
				830					835					840
Leu	Glu	Ala	Pro	Val	Thr	Glu	Gly	Pro	Gly	Ser	Val	Asp	Glu	Lys
				845					850					855
Leu	Pro	Ala	Pro	Glu	Pro	Cys	Val	Gly	Met	Ser	Cys	Pro	Pro	Gly
				860					865					870
Trp	Gly	His	Leu	Asp	Ala	Thr	Ser	Ala	Gly	Glu	Lys	Ala	Pro	Ser
				875					880					885
Pro	Trp	Gly	Ser	Ile	Arg	Thr	Gly	Ala	Gln	Ala	Ala	His	Val	Trp
				890					895					900
Thr	Pro	Ala	Ala	Gly	Ser	Cys	Ser	Val	Ser	Cys	Gly	Arg	Gly	Leu
				905					910					915
Met	Glu	Leu	Arg	Phe	Leu	Cys	Met	Asp	Ser	Ala	Leu	Arg	Val	Pro
				920					925					930
Val	Gln	Glu	Glu	Leu	Cys	Gly	Leu	Ala	Ser	Lys	Pro	Gly	Ser	Arg
				935					940					945
Arg	Glu	Val	Cys	Gln	Ala	Val	Pro	Cys	Pro	Ala	Arg	Trp	Gln	Tyr
				950					955					960
Lys	Leu	Ala	Ala	Cys	Ser	Val	Ser	Cys	Gly	Arg	Gly	Val	Val	Arg
				965					970					975
Arg	Ile	Leu	Tyr	Cys	Ala	Arg	Ala	His	Gly	Glu	Asp	Asp	Gly	Glu
				980					985					990
Glu	Ile	Leu	Leu	Asp	Thr	Gln	Cys	Gln	Gly	Leu	Pro	Arg	Pro	Glu
				995					1000					1005
Pro	Gln	Glu	Ala	Cys	Ser	Leu	Glu	Pro	Cys	Pro	Pro	Arg	Trp	Lys
				1010					1015					1020
Val	Met	Ser	Leu	Gly	Pro	Cys	Ser	Ala	Ser	Cys	Gly	Leu	Gly	Thr
				1025					1030					1035
Ala	Arg	Arg	Ser	Val	Ala	Cys	Val	Gln	Leu	Asp	Gln	Gly	Gln	Asp
				1040					1045					1050
Val	Glu	Val	Asp	Glu	Ala	Ala	Cys	Ala	Ala	Leu	Val	Arg	Pro	Glu
				1055					1060					1065
Ala	Ser	Val	Pro	Cys	Leu	Ile	Ala	Asp	Cys	Thr	Tyr	Arg	Trp	His
				1070					1075					1080
Val	Gly	Thr	Trp	Met	Glu	Cys	Ser	Val	Ser	Cys	Gly	Asp	Gly	Ile
				1085					1090					1095
Gln	Arg	Arg	Arg	Asp	Thr	Cys	Leu	Gly	Pro	Gln	Ala	Gln	Ala	Pro
				1100					1105					1110
Val	Pro	Ala	Asp	Phe	Cys	Gln	His	Leu	Pro	Lys	Pro	Val	Thr	Val
				1115					1120					1125
Arg	Gly	Cys	Trp	Ala	Gly	Pro	Cys	Val	Gly	Gln	Gly	Thr	Pro	Ser
				1130					1135					1140
Leu	Val	Pro	His	Glu	Glu	Ala	Ala	Ala	Pro	Gly	Arg	Thr	Thr	Ala
				1145					1150					1155
Thr	Pro	Ala	Gly	Ala	Ser	Leu	Glu	Trp	Ser	Gln	Ala	Arg	Gly	Leu
				1160					1165					1170
Leu	Phe	Ser	Pro	Ala	Pro	Gln	Pro	Arg	Arg	Leu	Leu	Pro	Gly	Pro
				1175					1180					1185
Gln	Glu	Asn	Ser	Val	Gln	Ser	Ser	Tyr	Val	Leu	Ser	Ser	Phe	Leu

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

<210> 4

<211> 473

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7481673CD1

<400> 4

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

WO 02/46383

PCT/US01/46964

Trp	Glu	Val	Pro	125	Pro	Phe	Ser	Gly	Leu	130	Glu	Arg	Asp	Gly	Val	135	Ile
				140						145						150	
Tyr	Gly	Arg	Gly	Thr	Leu	Asp	Asp	Lys	Asn	Ser	Val	Met	Ala	Leu			
				155						160						165	
Leu	Gln	Ala	Leu	Glu	Leu	Leu	Leu	Ile	Arg	Lys	Tyr	Ile	Pro	Arg			
				170						175						180	
Arg	Ser	Phe	Phe	Ile	Ser	Leu	Gly	His	Asp	Glu	Glu	Ser	Ser	Gly			
				185						190						195	
Thr	Gly	Ala	Gln	Arg	Ile	Ser	Ala	Leu	Leu	Gln	Ser	Arg	Gly	Val			
				200						205						210	
Gln	Leu	Ala	Phe	Ile	Val	Asp	Glu	Gly	Gly	Phe	Ile	Leu	Asp	Asp			
				215						220						225	
Phe	Ile	Pro	Asn	Phe	Lys	Lys	Pro	Ile	Ala	Leu	Ile	Ala	Val	Ser			
				230						235						240	
Glu	Lys	Gly	Ser	Met	Asn	Leu	Met	Leu	Gln	Val	Asn	Met	Thr	Ser			
				245						250						255	
Gly	His	Ser	Ser	Ala	Pro	Pro	Lys	Glu	Thr	Ser	Ile	Gly	Ile	Leu			
				260						265						270	
Ala	Ala	Ala	Val	Ser	Arg	Leu	Glu	Gln	Thr	Pro	Met	Pro	Ile	Ile			
				275						280						285	
Phe	Gly	Ser	Gly	Thr	Val	Val	Thr	Val	Leu	Gln	Gln	Leu	Ala	Asn			
				290						295						300	
Glu	Val	Tyr	Gly	Glu	Lys	Ser	Leu	Asn	Gln	Cys	Asn	Asn	Gln	Asp			
				305						310						315	
His	His	Gly	Thr	His	His	Ile	Gln	Ser	Arg	Val	Ala	Gln	Ala	Thr			
				320						325						330	
Val	Asn	Phe	Arg	Ile	His	Pro	Gly	Gln	Thr	Val	Gln	Glu	Val	Leu			
				335						340						345	
Glu	Leu	Thr	Lys	Asn	Ile	Val	Ala	Asp	Asn	Arg	Val	Gln	Phe	His			
				350						355						360	
Val	Leu	Ser	Ala	Phe	Asp	Pro	Leu	Pro	Val	Ser	Pro	Ser	Asp	Asp			
				365						370						375	
Lys	Ala	Leu	Gly	Tyr	Gln	Leu	Leu	Arg	Gln	Thr	Val	Gln	Ser	Val			
				380						385						390	
Phe	Pro	Glu	Val	Asn	Ile	Thr	Ala	Pro	Val	Thr	Ser	Ile	Gly	Asn			
				395						400						405	
Thr	Asp	Ser	Arg	Phe	Phe	Thr	Asn	Leu	Thr	Thr	Gly	Ile	Tyr	Arg			
				410						415						420	
Phe	Tyr	Pro	Ile	Tyr	Ile	Gln	Pro	Glu	Asp	Phe	Lys	Arg	Ile	His			
				425						430						435	
Gly	Val	Asn	Glu	Lys	Ile	Ser	Val	Gln	Ala	Tyr	Glu	Thr	Gln	Val			
				440						445						450	
Lys	Phe	Ile	Phe	Glu	Leu	Ile	Gln	Asn	Ala	Asp	Thr	Asp	Gln	Glu			
				455						460						465	
Pro	Val	Ser	His	Leu	His	Lys	Leu										
				470													

<210> 5

<211> 1627

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7484316CD1

<400> 5

Met	Ser	Trp	Lys	Arg	Asn	Tyr	Phe	Ser	Gly	Gly	Arg	Gly	Ser	Val
1				5					10					15
Gln	Gly	Met	Phe	Ala	Pro	Arg	Ser	Ser	Thr	Ser	Ile	Ala	Pro	Ser
				20					25					30
Lys	Gly	Leu	Ser	Asn	Glu	Pro	Gly	Gln	Asn	Ser	Cys	Phe	Leu	Asn

	35		40		45									
Ser	Ala	Leu	Gln	Val	Leu	Trp	His	Leu	Asp	Ile	Phe	Arg	Arg	Ser
	50		55		60									
Phe	Arg	Gln	Leu	Thr	Thr	His	Lys	Cys	Met	Gly	Asp	Ser	Cys	Ile
	65		70		75									
Phe	Cys	Ala	Leu	Lys	Gly	Ile	Phe	Asn	Gln	Phe	Gln	Cys	Ser	Ser
	80		85		90									
Glu	Lys	Val	Leu	Pro	Ser	Asp	Thr	Leu	Arg	Ser	Ala	Leu	Ala	Lys
	95		100		105									
Thr	Phe	Gln	Asp	Glu	Gln	Arg	Phe	Gln	Leu	Gly	Ile	Met	Asp	Asp
	110		115		120									
Ala	Ala	Glu	Cys	Phe	Glu	Asn	Leu	Leu	Met	Arg	Ile	His	Phe	His
	125		130		135									
Ile	Ala	Asp	Glu	Thr	Lys	Glu	Asp	Ile	Cys	Thr	Ala	Gln	His	Cys
	140		145		150									
Ile	Ser	His	Gln	Lys	Phe	Ala	Met	Thr	Leu	Phe	Glu	Gln	Cys	Val
	155		160		165									
Cys	Thr	Ser	Cys	Gly	Ala	Thr	Ser	Asp	Pro	Leu	Pro	Phe	Ile	Gln
	170		175		180									
Met	Val	His	Tyr	Ile	Ser	Thr	Thr	Ser	Leu	Cys	Asn	Gln	Ala	Ile
	185		190		195									
Cys	Met	Leu	Glu	Arg	Arg	Glu	Lys	Pro	Ser	Pro	Ser	Met	Phe	Gly
	200		205		210									
Glu	Leu	Leu	Gln	Asn	Ala	Ser	Thr	Met	Gly	Asp	Leu	Arg	Asn	Cys
	215		220		225									
Pro	Ser	Asn	Cys	Gly	Glu	Arg	Ile	Arg	Ile	Arg	Arg	Val	Leu	Met
	230		235		240									
Asn	Ala	Pro	Gln	Ile	Ile	Thr	Ile	Gly	Leu	Val	Trp	Asp	Ser	Asp
	245		250		255									
His	Ser	Asp	Leu	Ala	Glu	Asp	Val	Ile	His	Ser	Leu	Gly	Thr	Cys
	260		265		270									
Leu	Lys	Leu	Gly	Asp	Leu	Phe	Phe	Arg	Val	Thr	Asp	Asp	Arg	Ala
	275		280		285									
Lys	Gln	Ser	Glu	Leu	Tyr	Leu	Val	Gly	Met	Ile	Cys	Tyr	Tyr	Gly
	290		295		300									
Lys	His	Tyr	Ser	Thr	Phe	Phe	Phe	Gln	Thr	Lys	Ile	Arg	Lys	Trp
	305		310		315									
Met	Tyr	Phe	Asp	Asp	Ala	His	Val	Lys	Glu	Ile	Gly	Pro	Lys	Trp
	320		325		330									
Lys	Asp	Val	Val	Thr	Lys	Cys	Ile	Lys	Gly	His	Tyr	Gln	Pro	Leu
	335		340		345									
Leu	Leu	Leu	Tyr	Ala	Asp	Pro	Gln	Gly	Thr	Pro	Val	Ser	Thr	Gln
	350		355		360									
Asp	Leu	Pro	Pro	Gln	Ala	Glu	Phe	Gln	Ser	Tyr	Ser	Arg	Thr	Cys
	365		370		375									
Tyr	Asp	Ser	Glu	Asp	Ser	Gly	His	Leu	Thr	Asp	Ser	Glu	Cys	Asn
	380		385		390									
Gln	Lys	His	Thr	Ser	Lys	Lys	Gly	Ser	Leu	Ile	Glu	Arg	Lys	Arg
	395		400		405									
Ser	Ser	Gly	Arg	Val	Arg	Arg	Lys	Gly	Asp	Glu	Pro	Gln	Ala	Ser
	410		415		420									
Gly	Tyr	His	Ser	Glu	Gly	Glu	Thr	Leu	Lys	Glu	Lys	Gln	Ala	Pro
	425		430		435									
Arg	Asn	Ala	Ser	Lys	Pro	Ser	Ser	Ser	Thr	Asn	Arg	Leu	Arg	Asp
	440		445		450									
Phe	Lys	Glu	Thr	Val	Ser	Asn	Met	Ile	His	Asn	Arg	Pro	Ser	Leu
	455		460		465									
Ala	Ser	Gln	Thr	Asn	Val	Gly	Ser	His	Cys	Arg	Gly	Arg	Gly	Gly
	470		475		480									
Asp	Gln	Pro	Asp	Lys	Lys	Pro	Pro	Arg	Thr	Leu	Pro	Leu	His	Ser
	485		490		495									
Arg	Asp	Trp	Glu	Ile	Glu	Ser	Thr	Ser	Ser	Glu	Ser	Lys	Ser	Ser
	500		505		510									

Ser	Ser	Ser	Lys	Tyr	Arg	Pro	Thr	Trp	Arg	Pro	Lys	Arg	Glu	Ser
				515					520					525
Leu	Asn	Ile	Asp	Ser	Ile	Phe	Ser	Lys	Asp	Lys	Arg	Lys	His	Cys
				530					535					540
Gly	Tyr	Thr	Gln	Leu	Ser	Pro	Phe	Ser	Glu	Asp	Ser	Ala	Lys	Glu
				545					550					555
Phe	Ile	Pro	Asp	Glu	Pro	Ser	Lys	Pro	Pro	Ser	Tyr	Asp	Ile	Lys
				560					565					570
Phe	Gly	Gly	Pro	Ser	Pro	Gln	Tyr	Lys	Arg	Trp	Gly	Pro	Ala	Arg
				575					580					585
Pro	Gly	Ser	His	Leu	Leu	Glu	Gln	His	Pro	Arg	Leu	Ile	Gln	Arg
				590					595					600
Met	Glu	Ser	Gly	Tyr	Glu	Ser	Ser	Glu	Arg	Asn	Ser	Ser	Ser	Pro
				605					610					615
Val	Ser	Leu	Asp	Ala	Ala	Leu	Pro	Glu	Ser	Ser	Asn	Val	Tyr	Arg
				620					625					630
Asp	Pro	Ser	Ala	Lys	Arg	Ser	Ala	Gly	Leu	Val	Pro	Ser	Trp	Arg
				635					640					645
His	Ile	Pro	Lys	Ser	His	Ser	Ser	Ser	Ile	Leu	Glu	Val	Asp	Ser
				650					655					660
Thr	Ala	Ser	Met	Gly	Gly	Trp	Thr	Lys	Ser	Gln	Pro	Phe	Ser	Gly
				665					670					675
Glu	Glu	Ile	Ser	Ser	Lys	Ser	Glu	Leu	Asp	Glu	Leu	Gln	Glu	Glu
				680					685					690
Val	Ala	Arg	Arg	Ala	Gln	Glu	Gln	Glu	Leu	Arg	Arg	Lys	Arg	Glu
				695					700					705
Lys	Glu	Leu	Glu	Ala	Ala	Lys	Gly	Phe	Asn	Pro	His	Pro	Ser	Arg
				710					715					720
Phe	Met	Asp	Leu	Asp	Glu	Leu	Gln	Asn	Gln	Gly	Arg	Ser	Asp	Gly
				725					730					735
Phe	Glu	Arg	Ser	Leu	Gln	Glu	Ala	Glu	Ser	Val	Phe	Glu	Glu	Ser
				740					745					750
Leu	His	Leu	Glu	Gln	Lys	Gly	Asp	Cys	Ala	Ala	Ala	Leu	Ala	Leu
				755					760					765
Cys	Asn	Glu	Ala	Ile	Ser	Lys	Leu	Arg	Leu	Ala	Leu	His	Gly	Ala
				770					775					780
Ser	Cys	Ser	Thr	His	Ser	Arg	Ala	Leu	Val	Asp	Lys	Lys	Leu	Gln
				785					790					795
Ile	Ser	Ile	Arg	Lys	Ala	Arg	Ser	Leu	Gln	Asp	Arg	Met	Gln	Gln
				800					805					810
Gln	Gln	Ser	Pro	Gln	Gln	Pro	Ser	Gln	Pro	Ser	Ala	Cys	Leu	Pro
				815					820					825
Thr	Gln	Ala	Gly	Thr	Leu	Ser	Gln	Pro	Thr	Ser	Glu	Gln	Pro	Ile
				830					835					840
Pro	Leu	Gln	Val	Leu	Leu	Ser	Gln	Glu	Ala	Gln	Leu	Glu	Ser	Gly
				845					850					855
Met	Asp	Thr	Glu	Phe	Gly	Ala	Ser	Ser	Phe	Phe	His	Ser	Pro	Ala
				860					865					870
Ser	Cys	His	Glu	Ser	His	Ser	Ser	Leu	Ser	Pro	Glu	Ser	Ser	Ala
				875					880					885
Pro	Gln	His	Ser	Ser	Pro	Ser	Arg	Ser	Ala	Leu	Lys	Leu	Leu	Thr
				890					895					900
Ser	Val	Glu	Val	Asp	Asn	Ile	Glu	Pro	Ser	Ala	Phe	His	Arg	Gln
				905					910					915
Gly	Leu	Pro	Lys	Ala	Pro	Gly	Trp	Thr	Glu	Lys	Asn	Ser	His	His
				920					925					930
Ser	Trp	Glu	Pro	Leu	Asp	Ala	Pro	Glu	Gly	Lys	Leu	Gln	Gly	Ser
				935					940					945
Arg	Cys	Asp	Asn	Ser	Ser	Cys	Ser	Lys	Leu	Pro	Pro	Gln	Glu	Gly
				950					955					960
Arg	Gly	Ile	Ala	Gln	Glu	Gln	Leu	Phe	Gln	Glu	Lys	Lys	Asp	Pro
				965					970					975
Ala	Asn	Pro	Ser	Pro	Val	Met	Pro	Gly	Ile	Ala	Thr	Ser	Glu	Arg

				980					985					990
Gly	Asp	Glu	His	Ser	Leu	Gly	Cys	Ser	Pro	Ser	Asn	Ser	Ser	Ala
				995					1000					1005
Gln	Pro	Ser	Leu	Pro	Leu	Tyr	Arg	Thr	Cys	His	Pro	Ile	Met	Pro
				1010					1015					1020
Val	Ala	Ser	Ser	Phe	Val	Leu	His	Cys	Pro	Asp	Pro	Val	Gln	Lys
				1025					1030					1035
Thr	Asn	Gln	Cys	Leu	Gln	Gly	Gln	Ser	Leu	Lys	Thr	Ser	Leu	Thr
				1040					1045					1050
Leu	Lys	Val	Asp	Arg	Gly	Ser	Glu	Glu	Thr	Tyr	Arg	Pro	Glu	Phe
				1055					1060					1065
Pro	Ser	Thr	Lys	Gly	Leu	Val	Arg	Ser	Leu	Ala	Glu	Gln	Phe	Gln
				1070					1075					1080
Arg	Met	Gln	Gly	Val	Ser	Met	Arg	Asp	Ser	Thr	Gly	Phe	Lys	Asp
				1085					1090					1095
Arg	Ser	Leu	Ser	Gly	Ser	Leu	Arg	Lys	Asn	Ser	Ser	Pro	Ser	Asp
				1100					1105					1110
Ser	Lys	Pro	Pro	Phe	Ser	Gln	Gly	Gln	Glu	Lys	Gly	His	Trp	Pro
				1115					1120					1125
Trp	Ala	Lys	Gln	Gln	Ser	Ser	Leu	Glu	Gly	Gly	Asp	Arg	Pro	Leu
				1130					1135					1140
Ser	Trp	Glu	Glu	Ser	Thr	Glu	His	Ser	Ser	Leu	Ala	Leu	Asn	Ser
				1145					1150					1155
Gly	Leu	Pro	Asn	Gly	Glu	Thr	Ser	Ser	Gly	Gly	Gln	Pro	Arg	Leu
				1160					1165					1170
Ala	Glu	Pro	Asp	Ile	Tyr	Gln	Glu	Lys	Leu	Ser	Gln	Val	Arg	Asp
				1175					1180					1185
Val	Arg	Ser	Lys	Asp	Leu	Gly	Ser	Ser	Thr	Asp	Leu	Gly	Thr	Ser
				1190					1195					1200
Leu	Pro	Leu	Asp	Ser	Trp	Val	Asn	Ile	Thr	Arg	Phe	Cys	Asp	Ser
				1205					1210					1215
Gln	Leu	Lys	His	Gly	Ala	Pro	Arg	Pro	Gly	Met	Lys	Ser	Ser	Pro
				1220					1225					1230
His	Asp	Ser	His	Thr	Cys	Val	Thr	Tyr	Pro	Glu	Arg	Asn	His	Ile
				1235					1240					1245
Leu	Leu	His	Pro	His	Trp	Asn	Gln	Asp	Thr	Glu	Gln	Glu	Thr	Ser
				1250					1255					1260
Glu	Leu	Glu	Ser	Leu	Tyr	Gln	Ala	Ser	Leu	Gln	Ala	Ser	Gln	Ala
				1265					1270					1275
Gly	Cys	Ser	Gly	Trp	Gly	Gln	Gln	Asp	Thr	Ala	Trp	His	Pro	Leu
				1280					1285					1290
Ser	Gln	Thr	Gly	Ser	Ala	Asp	Gly	Met	Gly	Arg	Arg	Leu	His	Ser
				1295					1300					1305
Ala	His	Asp	Pro	Gly	Leu	Ser	Lys	Thr	Ser	Thr	Ala	Glu	Met	Glu
				1310					1315					1320
His	Gly	Leu	His	Glu	Ala	Arg	Thr	Val	Arg	Thr	Ser	Gln	Ala	Thr
				1325					1330					1335
Pro	Cys	Arg	Gly	Leu	Ser	Arg	Glu	Cys	Gly	Glu	Asp	Glu	Gln	Tyr
				1340					1345					1350
Ser	Ala	Glu	Asn	Leu	Arg	Arg	Ile	Ser	Arg	Ser	Leu	Ser	Gly	Thr
				1355					1360					1365
Val	Val	Ser	Glu	Arg	Glu	Glu	Ala	Pro	Val	Ser	Ser	His	Ser	Phe
				1370					1375					1380
Asp	Ser	Ser	Asn	Val	Arg	Lys	Pro	Leu	Glu	Thr	Gly	His	Arg	Cys
				1385					1390					1395
Ser	Ser	Ser	Ser	Ser	Leu	Pro	Val	Ile	His	Asp	Pro	Ser	Val	Phe
				1400					1405					1410
Leu	Leu	Gly	Pro	Gln	Leu	Tyr	Leu	Pro	Gln	Pro	Gln	Phe	Leu	Ser
				1415					1420					1425
Pro	Asp	Val	Leu	Met	Pro	Thr	Met	Ala	Gly	Glu	Pro	Asn	Arg	Leu
				1430					1435					1440
Pro	Gly	Thr	Ser	Arg	Ser	Val	Gln	Gln	Phe	Leu	Ala	Met	Cys	Asp
				1445					1450					1455

WO 02/46383

PCT/US01/46964

Arg	Gly	Glu	Thr	Ser	Gln	Gly	Ala	Lys	Tyr	Thr	Gly	Arg	Thr	Leu
				1460					1465					1470
Asn	Tyr	Gln	Ser	Leu	Pro	His	Arg	Ser	Arg	Thr	Asp	Asn	Ser	Trp
				1475					1480					1485
Ala	Pro	Trp	Ser	Glu	Thr	Asn	Gln	His	Ile	Gly	Thr	Arg	Phe	Leu
				1490					1495					1500
Thr	Thr	Pro	Gly	Cys	Asn	Pro	Gln	Leu	Thr	Tyr	Thr	Ala	Thr	Leu
				1505					1510					1515
Pro	Glu	Arg	Ser	Lys	Gly	Leu	Gln	Val	Pro	His	Thr	Gln	Ser	Trp
				1520					1525					1530
Ser	Asp	Leu	Phe	His	Ser	Pro	Ser	His	Pro	Pro	Ile	Val	His	Pro
				1535					1540					1545
Val	Tyr	Pro	Pro	Ser	Ser	Ser	Leu	His	Val	Pro	Leu	Arg	Ser	Ala
				1550					1555					1560
Trp	Asn	Ser	Asp	Pro	Val	Pro	Gly	Ser	Arg	Thr	Pro	Gly	Pro	Arg
				1565					1570					1575
Arg	Val	Asp	Met	Pro	Pro	Asp	Asp	Asp	Trp	Arg	Gln	Ser	Ser	Tyr
				1580					1585					1590
Ala	Ser	His	Ser	Gly	His	Arg	Arg	Thr	Val	Gly	Glu	Gly	Phe	Leu
				1595					1600					1605
Phe	Val	Leu	Ser	Asp	Ala	Pro	Arg	Arg	Glu	Gln	Ile	Arg	Ala	Arg
				1610					1615					1620
Val	Leu	Gln	His	Ser	Gln	Trp								
				1625										

<210> 6

<211> 1035

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7485008CD1

<400> 6

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

Gln	Pro	Pro	Ala	Pro	Pro	Asp	Ala	Gly	Cys	Phe	Tyr	Thr	Gly	Ala
				215					220					225
Val	Leu	Arg	His	Pro	Gly	Ser	Leu	Ala	Ser	Phe	Ser	Thr	Cys	Gly
				230					235					240
Gly	Gly	Leu	Val	Phe	Asn	Leu	Phe	Gln	His	Lys	Ser	Leu	Gly	Val
				245					250					255
Gln	Val	Asn	Leu	Arg	Val	Ile	Lys	Leu	Ile	Leu	Leu	His	Glu	Thr
				260					265					270
Pro	Pro	Glu	Leu	Tyr	Ile	Gly	His	His	Gly	Glu	Lys	Met	Leu	Glu
				275					280					285
Ser	Phe	Cys	Lys	Trp	Gln	His	Glu	Glu	Phe	Gly	Lys	Lys	Asn	Asp
				290					295					300
Ile	His	Leu	Glu	Met	Ser	Thr	Asn	Trp	Gly	Glu	Asp	Met	Thr	Ser
				305					310					315
Val	Asp	Ala	Ala	Ile	Leu	Ile	Thr	Arg	Lys	Asp	Phe	Cys	Val	His
				320					325					330
Lys	Asp	Glu	Pro	Cys	Asp	Thr	Val	Gly	Ile	Ala	Tyr	Leu	Ser	Gly
				335					340					345
Met	Cys	Ser	Glu	Lys	Arg	Lys	Cys	Ile	Ile	Ala	Glu	Asp	Asn	Gly
				350					355					360
Leu	Asn	Leu	Ala	Phe	Thr	Ile	Ala	His	Glu	Met	Gly	His	Asn	Met
				365					370					375
Gly	Ile	Asn	His	Asp	Asn	Asp	His	Pro	Ser	Cys	Ala	Asp	Gly	Leu
				380					385					390
His	Ile	Met	Ser	Gly	Glu	Trp	Ile	Lys	Gly	Gln	Asn	Leu	Gly	Asp
				395					400					405
Val	Ser	Trp	Ser	Arg	Cys	Ser	Lys	Glu	Asp	Leu	Glu	Arg	Phe	Leu
				410					415					420
Arg	Ser	Lys	Ala	Ser	Asn	Cys	Leu	Leu	Gln	Thr	Asn	Pro	Gln	Ser
				425					430					435
Val	Asn	Ser	Val	Met	Val	Pro	Ser	Lys	Leu	Pro	Gly	Met	Thr	Tyr
				440					445					450
Thr	Ala	Asp	Glu	Gln	Cys	Gln	Ile	Leu	Phe	Gly	Pro	Leu	Ala	Ser
				455					460					465
Phe	Cys	Gln	Glu	Met	Gln	His	Val	Ile	Cys	Thr	Gly	Leu	Trp	Cys
				470					475					480
Lys	Val	Glu	Gly	Glu	Lys	Glu	Cys	Arg	Thr	Lys	Leu	Asp	Pro	Pro
				485					490					495
Met	Asp	Gly	Thr	Asp	Cys	Asp	Leu	Gly	Lys	Trp	Cys	Lys	Ala	Gly
				500					505					510
Glu	Cys	Thr	Ser	Arg	Thr	Ser	Ala	Pro	Glu	His	Leu	Ala	Gly	Glu
				515					520					525
Trp	Ser	Leu	Trp	Ser	Pro	Cys	Ser	Arg	Thr	Cys	Ser	Ala	Gly	Ile
				530					535					540
Ser	Ser	Arg	Glu	Arg	Lys	Cys	Pro	Gly	Leu	Asp	Ser	Glu	Ala	Arg
				545					550					555
Asp	Cys	Asn	Gly	Pro	Arg	Lys	Gln	Tyr	Arg	Ile	Cys	Glu	Asn	Pro
				560					565					570
Pro	Cys	Pro	Ala	Gly	Leu	Pro	Gly	Phe	Arg	Asp	Trp	Gln	Cys	Gln
				575					580					585
Ala	Tyr	Ser	Val	Arg	Thr	Ser	Pro	Pro	Lys	His	Ile	Leu	Gln	Trp
				590					595					600
Gln	Ala	Val	Leu	Asp	Glu	Glu	Lys	Pro	Cys	Ala	Leu	Phe	Cys	Ser
				605					610					615
Pro	Val	Gly	Lys	Glu	Gln	Pro	Ile	Leu	Leu	Ser	Glu	Lys	Val	Met
				620					625					630
Asp	Gly	Thr	Ser	Cys	Gly	Tyr	Gln	Gly	Leu	Asp	Ile	Cys	Ala	Asn
				635					640					645
Gly	Arg	Cys	Gln	Lys	Val	Gly	Cys	Asp	Gly	Leu	Leu	Gly	Ser	Leu
				650					655					660
Ala	Arg	Glu	Asp	His	Cys	Gly	Val	Cys	Asn	Gly	Asn	Gly	Lys	Ser
				665					670					675
Cys	Lys	Ile	Ile	Lys	Gly	Asp	Phe	Asn	His	Thr	Arg	Gly	Ala	Gly

Tyr Val Glu Val	680	Tyr Val Glu Val	685	Tyr Val Glu Val	690
Leu Val Ile Pro Ala	695	Gly Ala Arg Arg Ile	700	Lys	705
Val Val Glu Glu	710	Lys Pro Ala His Ser	715	Tyr Leu Ala Leu Arg	720
Ala Gly Lys Gln	725	Ser Ile Asn Ser Asp	730	Trp Lys Ile Glu His	735
Gly Ala Phe Asn	740	Leu Ala Gly Thr Thr	745	Val His Tyr Val Arg	750
Gly Leu Trp Glu	755	Lys Ile Ser Ala Lys	760	Gly Pro Thr Thr Ala	765
Leu His Leu Leu	770	Val Leu Leu Phe Gln	775	Asp Gln Asn Tyr Gly	780
His Tyr Glu Tyr	785	Thr Ile Pro Ser Asp	790	Pro Leu Pro Glu Asn	795
Ser Ser Lys Ala	800	Pro Glu Pro Leu Phe	805	Met Trp Thr His Thr	810
Trp Glu Asp Cys	815	Asp Ala Thr Cys Gly	820	Gly Gly Glu Arg Lys	825
Thr Val Ser Cys	830	Thr Lys Ile Met Ser	835	Lys Asn Ile Ser Ile	840
Asp Asn Glu Lys	845	Cys Lys Tyr Leu Thr	850	Lys Pro Glu Pro Gln	855
Arg Lys Cys Asn	860	Glu Gln Pro Cys Gln	865	Thr Arg Trp Met Met	870
Glu Trp Thr Pro	875	Cys Ser Arg Thr Cys	880	Gly Lys Gly Met Gln	885
Arg Gln Val Ala	890	Cys Thr Gln Gln Leu	895	Ser Asn Gly Thr Leu	900
Arg Ala Arg Glu	905	Arg Asp Cys Ile Gly	910	Pro Lys Pro Ala Ser	915
Gln Arg Cys Glu	920	Gly Gln Asp Cys Met	925	Thr Val Trp Glu Ala	930
Val Trp Ser Glu	935	Cys Ser Val Lys Cys	940	Gly Lys Gly Ile Arg	945
Arg Thr Val Arg	950	Cys Thr Asn Pro Arg	955	Lys Lys Cys Val Leu	960
Thr Arg Pro Arg	965	Glu Ala Glu Asp Cys	970	Glu Asp Tyr Ser Lys	975
Tyr Val Trp Arg	980	Met Gly Asp Trp Ser	985	Lys Cys Ser Ile Thr	990
Gly Lys Gly Met	995	Gln Ser Arg Val Ile	1000	Gln Cys Met His Lys	1005
Thr Gly Arg His	1010	Gly Asn Glu Cys Phe	1015	Ser Ser Glu Lys Pro	1020
Ala Tyr Arg Pro	1025	Cys His Leu Gln Pro	1030	Ala Met Arg Lys Leu	1035

<210> 7

<211> 185

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 4820375CD1

<400> 7

Met Asp His Pro Pro	Gly Ser Ala Asp	His Pro Asn Asn Cys Arg
1	5	10
Ile Val Lys Arg Lys	Ile Glu Leu Tyr Tyr	Gln Val Leu Asn Phe
20	25	30

WO 02/46383

PCT/US01/46964

Ala	Met	Ile	Val	Ser	Ser	Ala	Leu	Met	Ile	Trp	Lys	Gly	Leu	Ile			
				35					40					45			
Val	Leu	Thr	Gly	Ser	Glu	Ser	Pro	Ile	Val	Val	Val	Leu	Ser	Gly			
				50					55					60			
Ser	Met	Glu	Pro	Ala	Phe	His	Arg	Gly	Asp	Leu	Leu	Phe	Leu	Thr			
				65					70					75			
Asn	Phe	Arg	Glu	Asp	Pro	Ile	Arg	Ala	Gly	Glu	Ile	Val	Val	Phe			
				80					85					90			
Lys	Val	Glu	Gly	Arg	Asp	Ile	Pro	Ile	Val	His	Arg	Val	Ile	Lys			
				95					100					105			
Val	His	Glu	Lys	Asp	Asn	Gly	Asp	Ile	Lys	Phe	Leu	Thr	Lys	Gly			
				110					115					120			
Asp	Asn	Asn	Glu	Val	Asp	Asp	Arg	Gly	Leu	Tyr	Lys	Glu	Gly	Gln			
				125					130					135			
Asn	Trp	Leu	Glu	Lys	Lys	Asp	Val	Val	Gly	Arg	Ala	Arg	Gly	Phe			
				140					145					150			
Leu	Pro	Tyr	Val	Gly	Met	Val	Thr	Ile	Ile	Met	Asn	Asp	Tyr	Pro			
				155					160					165			
Lys	Phe	Lys	Tyr	Ala	Leu	Leu	Ala	Val	Met	Gly	Ala	Tyr	Val	Leu			
				170					175					180			
Leu	Lys	Arg	Glu	Ser													
				185													

<210> 8

<211> 962

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7483698CD1

<400> 8

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

Ile	Gly	Asn	Leu	Ile	His	Ile	Val	Val	Val	Lys	Leu	Val	Met	Ile
				230					235					240
His	Arg	Glu	Glu	Glu	Gly	Pro	Val	Ile	Asn	Phe	Asp	Gly	Ala	Thr
				245					250					255
Thr	Leu	Lys	Asn	Phe	Cys	Ser	Trp	Gln	Gln	Thr	Gln	Asn	Asp	Leu
				260					265					270
Asp	Asp	Val	His	Pro	Ser	His	His	Asp	Thr	Ala	Val	Leu	Ile	Thr
				275					280					285
Arg	Glu	Asp	Ile	Cys	Ser	Ser	Lys	Glu	Lys	Cys	Asn	Met	Leu	Gly
				290					295					300
Leu	Ser	Tyr	Leu	Gly	Thr	Ile	Cys	Asp	Pro	Leu	Gln	Ser	Cys	Phe
				305					310					315
Ile	Asn	Glu	Glu	Lys	Gly	Leu	Ile	Ser	Ala	Phe	Thr	Ile	Ala	His
				320					325					330
Glu	Leu	Gly	His	Thr	Leu	Gly	Val	Gln	His	Asp	Asp	Asn	Pro	Arg
				335					340					345
Cys	Lys	Glu	Met	Lys	Val	Thr	Lys	Tyr	His	Val	Met	Ala	Pro	Ala
				350					355					360
Leu	Ser	Phe	His	Met	Ser	Pro	Trp	Ser	Trp	Ser	Asn	Cys	Ser	Arg
				365					370					375
Lys	Tyr	Val	Thr	Glu	Phe	Leu	Asp	Thr	Gly	Tyr	Gly	Glu	Cys	Leu
				380					385					390
Leu	Asp	Lys	Pro	Asp	Glu	Glu	Ile	Tyr	Asn	Leu	Pro	Ser	Glu	Leu
				395					400					405
Pro	Gly	Ser	Arg	Tyr	Asp	Gly	Asn	Lys	Gln	Cys	Glu	Leu	Ala	Phe
				410					415					420
Gly	Pro	Gly	Ser	Gln	Met	Cys	Pro	His	Ile	Glu	Asn	Ile	Cys	Met
				425					430					435
His	Leu	Trp	Cys	Thr	Ser	Thr	Glu	Lys	Leu	His	Lys	Gly	Cys	Phe
				440					445					450
Thr	Gln	His	Val	Pro	Pro	Ala	Asp	Gly	Thr	Asp	Cys	Gly	Pro	Gly
				455					460					465
Met	His	Cys	Arg	His	Gly	Leu	Cys	Val	Asn	Lys	Glu	Thr	Glu	Thr
				470					475					480
Arg	Pro	Val	Asn	Gly	Glu	Trp	Gly	Pro	Trp	Glu	Pro	Tyr	Ser	Ser
				485					490					495
Cys	Ser	Arg	Thr	Cys	Gly	Gly	Gly	Ile	Glu	Ser	Ala	Thr	Arg	Arg
				500					505					510
Cys	Asn	Arg	Pro	Glu	Pro	Arg	Asn	Gly	Gly	Asn	Tyr	Cys	Val	Gly
				515					520					525
Arg	Arg	Met	Lys	Phe	Arg	Ser	Cys	Asn	Thr	Asp	Ser	Cys	Pro	Lys
				530					535					540
Gly	Thr	Gln	Asp	Phe	Arg	Glu	Lys	Gln	Cys	Ser	Asp	Phe	Asn	Gly
				545					550					555
Lys	His	Leu	Asp	Ile	Ser	Gly	Ile	Pro	Ser	Asn	Val	Arg	Trp	Leu
				560					565					570
Pro	Arg	Tyr	Ser	Gly	Ile	Gly	Thr	Lys	Asp	Arg	Cys	Lys	Leu	Tyr
				575					580					585
Cys	Gln	Val	Ala	Gly	Thr	Asn	Tyr	Phe	Tyr	Leu	Leu	Lys	Asp	Met
				590					595					600
Val	Glu	Asp	Gly	Thr	Pro	Cys	Gly	Thr	Glu	Thr	His	Asp	Ile	Cys
				605					610					615
Val	Gln	Gly	Gln	Cys	Met	Ala	Ala	Gly	Cys	Asp	His	Val	Leu	Asn
				620					625					630
Ser	Ser	Ala	Lys	Ile	Asp	Lys	Cys	Gly	Val	Cys	Gly	Gly	Asp	Asn
				635					640					645
Ser	Ser	Cys	Lys	Thr	Ile	Thr	Gly	Val	Phe	Asn	Ser	Ser	His	Tyr
				650					655					660
Gly	Tyr	Asn	Val	Val	Val	Lys	Ile	Pro	Ala	Gly	Ala	Thr	Asn	Val
				665					670					675
Asp	Ile	Arg	Gln	Tyr	Ser	Tyr	Ser	Gly	Gln	Pro	Asp	Asp	Ser	Tyr
				680					685					690
Leu	Ala	Leu	Ser	Asp	Ala	Glu	Gly	Asn	Phe	Leu	Phe	Asn	Gly	Asn

WO 02/46383

PCT/US01/46964

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

Val His

<210> 9
 <211> 508
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 7485421CD1

<400> 9

Met	Glu	Phe	Pro	Val	Leu	Ser	Ser	Ser	Ser	Cys	Leu	Gly	Gly	Met
1				5					10					15
Leu	Cys	Leu	Thr	Val	Ser	Ser	Glu	His	Pro	Cys	Leu	Ile	Thr	Gln
				20					25					30
Arg	Ser	Leu	Leu	Leu	Phe	Ser	Glu	Phe	Gln	Ala	Lys	Ser	Cys	Ile
				35					40					45
Cys	His	Val	Cys	Gly	Val	His	Leu	Asn	Arg	Leu	His	Ser	Cys	Leu
				50					55					60
Tyr	Cys	Val	Phe	Phe	Gly	Cys	Phe	Thr	Lys	Lys	His	Ile	His	Glu
				65					70					75
His	Ala	Lys	Ala	Lys	Arg	His	Asn	Leu	Ala	Ile	Asp	Leu	Met	Tyr
				80					85					90
Gly	Gly	Ile	Tyr	Cys	Phe	Leu	Cys	Gln	Asp	Tyr	Ile	Tyr	Asp	Lys
				95					100					105
Asp	Met	Glu	Ile	Ile	Ala	Lys	Glu	Glu	Gln	Arg	Lys	Ala	Trp	Lys

Met	Gln	Gly	Val	110	Gly	Glu	Lys	Phe	Ser	115	Thr	Trp	Glu	Pro	Thr	Lys	120
				125						130							135
Arg	Glu	Leu	Glu	140	Leu	Leu	Lys	His	Asn	145	Pro	Lys	Arg	Arg	Lys	Ile	150
Thr	Ser	Asn	Cys	155	Thr	Ile	Gly	Leu	Arg	160	Gly	Leu	Ile	Asn	Leu	Gly	165
Asn	Thr	Cys	Phe	170	Met	Asn	Cys	Ile	Val	175	Gln	Ala	Pro	Thr	His	Thr	180
Pro	Leu	Leu	Arg	185	Asp	Phe	Phe	Leu	Ser	190	Asp	Arg	His	Arg	Cys	Glu	195
Met	Gln	Ser	Pro	200	Ser	Ser	Cys	Leu	Val	205	Cys	Glu	Met	Ser	Ser	Leu	210
Phe	Gln	Glu	Phe	215	Tyr	Ser	Gly	His	Arg	220	Ser	Pro	His	Ile	Pro	Tyr	225
Lys	Leu	Leu	His	230	Leu	Val	Trp	Thr	His	235	Ala	Arg	His	Leu	Ala	Gly	240
Tyr	Glu	Gln	Gln	245	Asp	Ala	His	Glu	Phe	250	Leu	Ile	Ala	Ala	Leu	Asp	255
Val	Leu	His	Arg	260	His	Cys	Lys	Gly	Asp	265	Asp	Asn	Gly	Lys	Lys	Ala	270
Asn	Asn	Pro	Asn	275	His	Cys	Asn	Cys	Ile	280	Ile	Asp	Gln	Ile	Phe	Thr	285
Gly	Gly	Leu	Gln	290	Ser	Asp	Val	Thr	Cys	295	Gln	Val	Cys	His	Gly	Val	300
Ser	Thr	Thr	Ile	305	Asp	Pro	Phe	Trp	Asp	310	Ile	Ser	Leu	Asp	Ile	Pro	315
Gly	Ser	Ser	Thr	320	Pro	Phe	Trp	Pro	Leu	325	Ser	Pro	Gly	Ser	Glu	Gly	330
Asn	Val	Val	Asn	335	Gly	Glu	Ser	His	Val	340	Ser	Gly	Thr	Thr	Thr	Leu	345
Thr	Asp	Cys	Leu	350	Arg	Arg	Phe	Thr	Arg	355	Pro	Glu	His	Leu	Gly	Ser	360
Ser	Ala	Lys	Ile	365	Lys	Cys	Ser	Gly	Cys	370	His	Ser	Tyr	Gln	Glu	Ser	375
Thr	Lys	Gln	Leu	380	Thr	Met	Lys	Lys	Leu	385	Pro	Ile	Val	Ala	Cys	Phe	390
His	Leu	Lys	Arg	395	Phe	Glu	His	Ser	Ala	400	Lys	Leu	Arg	Arg	Lys	Ile	405
Thr	Thr	Tyr	Val	410	Ser	Phe	Pro	Leu	Glu	415	Leu	Asp	Met	Thr	Pro	Phe	420
Met	Ala	Ser	Ser	425	Lys	Glu	Ser	Arg	Met	430	Asn	Gly	Gln	Tyr	Gln	Gln	435
Pro	Thr	Asp	Ser	440	Leu	Asn	Asn	Asp	Asn	445	Lys	Tyr	Ser	Leu	Phe	Ala	450
Val	Val	Asn	His	455	Gln	Gly	Thr	Leu	Glu	460	Ser	Gly	His	Tyr	Thr	Ser	465
Phe	Ile	Arg	Gln	470	His	Lys	Asp	Gln	Trp	475	Phe	Lys	Cys	Asp	Asp	Ala	480
Ile	Ile	Thr	Lys	485	Ala	Ser	Ile	Lys	Asp	490	Val	Leu	Asp	Ser	Glu	Gly	495
Tyr	Leu	Leu	Phe	500	Tyr	His	Lys	Gln	Phe	505	Leu	Glu	Tyr	Glu			

<210> 10

<211> 321

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7485720CD1

WO 02/46383

PCT/US01/46964

<400> 10

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

<210> 11

<211> 1123

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7485896CD1

<400> 11

Met	Asp	Leu	Gly	Pro	Gly	Asp	Ala	Ala	Gly	Gly	Gly	Pro	Leu	Ala
1				5					10					15
Pro	Arg	Pro	Arg	Arg	Arg	Arg	Ser	Leu	Arg	Arg	Leu	Phe	Ser	Arg
				20					25					30
Phe	Leu	Leu	Ala	Leu	Gly	Ser	Arg	Ser	Arg	Pro	Gly	Asp	Ser	Pro
				35					40					45
Pro	Arg	Pro	Gln	Pro	Gly	His	Cys	Asp	Gly	Asp	Gly	Glu	Gly	Gly

				50					55					60
Phe	Ala	Cys	Ala	Pro	Gly	Pro	Val	Pro	Ala	Ala	Pro	Gly	Ser	Pro
				65					70					75
Gly	Glu	Glu	Arg	Pro	Pro	Gly	Pro	Gln	Pro	Gln	Leu	Gln	Leu	Pro
				80					85					90
Ala	Gly	Asp	Gly	Ala	Arg	Pro	Pro	Gly	Ala	Gln	Gly	Leu	Lys	Asn
				95					100					105
His	Gly	Asn	Thr	Cys	Phe	Met	Asn	Ala	Val	Val	Gln	Cys	Leu	Ser
				110					115					120
Asn	Thr	Asp	Leu	Leu	Ala	Glu	Phe	Leu	Ala	Leu	Gly	Arg	Tyr	Arg
				125					130					135
Ala	Ala	Pro	Gly	Arg	Ala	Glu	Val	Thr	Glu	Gln	Leu	Ala	Ala	Leu
				140					145					150
Val	Arg	Ala	Leu	Trp	Thr	Arg	Glu	Tyr	Thr	Pro	Gln	Leu	Ser	Ala
				155					160					165
Glu	Phe	Lys	Asn	Ala	Val	Ser	Lys	Tyr	Gly	Ser	Gln	Phe	Gln	Gly
				170					175					180
Asn	Ser	Gln	His	Asp	Ala	Leu	Glu	Phe	Leu	Leu	Trp	Leu	Leu	Asp
				185					190					195
Arg	Val	His	Glu	Asp	Leu	Glu	Gly	Ser	Ser	Arg	Gly	Pro	Val	Ser
				200					205					210
Glu	Lys	Leu	Pro	Pro	Glu	Ala	Thr	Lys	Thr	Ser	Glu	Asn	Cys	Leu
				215					220					225
Ser	Pro	Ser	Ala	Gln	Leu	Pro	Leu	Gly	Gln	Ser	Phe	Val	Gln	Ser
				230					235					240
His	Phe	Gln	Ala	Gln	Tyr	Arg	Ser	Ser	Leu	Thr	Cys	Pro	His	Cys
				245					250					255
Leu	Lys	Gln	Ser	Asn	Thr	Phe	Asp	Pro	Phe	Leu	Cys	Val	Ser	Leu
				260					265					270
Pro	Ile	Pro	Leu	Arg	Gln	Thr	Arg	Phe	Leu	Ser	Val	Thr	Leu	Val
				275					280					285
Phe	Pro	Ser	Lys	Ser	Gln	Arg	Phe	Leu	Arg	Val	Gly	Leu	Ala	Val
				290					295					300
Pro	Ile	Leu	Ser	Thr	Val	Ala	Ala	Leu	Arg	Lys	Met	Val	Ala	Glu
				305					310					315
Glu	Gly	Gly	Val	Pro	Ala	Asp	Glu	Val	Ile	Leu	Val	Glu	Leu	Tyr
				320					325					330
Pro	Ser	Gly	Phe	Gln	Arg	Ser	Phe	Phe	Asp	Glu	Glu	Asp	Leu	Asn
				335					340					345
Thr	Ile	Ala	Glu	Gly	Asp	Asn	Val	Tyr	Ala	Phe	Gln	Val	Pro	Pro
				350					355					360
Ser	Pro	Ser	Gln	Gly	Thr	Leu	Ser	Ala	His	Pro	Leu	Gly	Leu	Ser
				365					370					375
Ala	Ser	Pro	Arg	Leu	Ala	Ala	Arg	Glu	Gly	Gln	Arg	Phe	Ser	Leu
				380					385					390
Ser	Leu	His	Ser	Glu	Ser	Lys	Val	Leu	Ile	Leu	Phe	Cys	Asn	Leu
				395					400					405
Val	Gly	Ser	Gly	Gln	Gln	Ala	Ser	Arg	Phe	Gly	Pro	Pro	Phe	Leu
				410					415					420
Ile	Arg	Glu	Asp	Arg	Ala	Val	Ser	Trp	Ala	Gln	Leu	Gln	Gln	Ser
				425					430					435
Ile	Leu	Ser	Lys	Val	Arg	His	Leu	Met	Lys	Ser	Glu	Ala	Pro	Val
				440					445					450
Gln	Asn	Leu	Gly	Ser	Leu	Phe	Ser	Ile	Arg	Val	Val	Gly	Leu	Ser
				455					460					465
Val	Ala	Cys	Ser	Tyr	Leu	Ser	Pro	Lys	Asp	Ser	Arg	Pro	Leu	Cys
				470					475					480
His	Trp	Ala	Val	Asp	Arg	Val	Leu	His	Leu	Arg	Arg	Pro	Gly	Gly
				485					490					495
Pro	Pro	His	Val	Lys	Leu	Ala	Val	Glu	Trp	Asp	Ser	Ser	Val	Lys
				500					505					510
Glu	Arg	Leu	Phe	Gly	Ser	Leu	Gln	Glu	Glu	Arg	Ala	Gln	Asp	Ala
				515					520					525

Asp	Ser	Val	Trp	Gln	Gln	Gln	Gln	Ala	His	Gln	Gln	His	Ser	Cys
				530					535					540
Thr	Leu	Asp	Glu	Cys	Phe	Gln	Phe	Tyr	Thr	Lys	Glu	Glu	Gln	Leu
				545					550					555
Ala	Gln	Asp	Asp	Ala	Trp	Lys	Cys	Pro	His	Cys	Gln	Val	Leu	Gln
				560					565					570
Gln	Gly	Met	Val	Lys	Leu	Ser	Leu	Trp	Thr	Leu	Pro	Asp	Ile	Leu
				575					580					585
Ile	Ile	His	Leu	Lys	Arg	Phe	Cys	Gln	Val	Gly	Glu	Arg	Arg	Asn
				590					595					600
Lys	Leu	Ser	Thr	Leu	Val	Lys	Phe	Pro	Leu	Ser	Gly	Leu	Asn	Met
				605					610					615
Ala	Pro	His	Val	Ala	Gln	Arg	Ser	Thr	Ser	Pro	Glu	Ala	Gly	Leu
				620					625					630
Gly	Pro	Trp	Pro	Ser	Trp	Lys	Gln	Pro	Asp	Cys	Leu	Pro	Thr	Ser
				635					640					645
Tyr	Pro	Leu	Asp	Phe	Leu	Tyr	Asp	Leu	Tyr	Ala	Val	Cys	Asn	His
				650					655					660
His	Gly	Asn	Leu	Gln	Gly	Gly	His	Tyr	Thr	Ala	Tyr	Cys	Arg	Asn
				665					670					675
Ser	Leu	Asp	Gly	Gln	Trp	Tyr	Ser	Tyr	Asp	Asp	Ser	Thr	Val	Glu
				680					685					690
Pro	Leu	Arg	Glu	Asp	Glu	Val	Asn	Thr	Arg	Gly	Ala	Tyr	Ile	Leu
				695					700					705
Phe	Tyr	Gln	Lys	Arg	Asn	Ser	Ile	Pro	Pro	Trp	Ser	Ala	Ser	Ser
				710					715					720
Ser	Met	Arg	Gly	Ser	Thr	Ser	Ser	Ser	Leu	Ser	Asp	His	Trp	Leu
				725					730					735
Leu	Arg	Leu	Gly	Ser	His	Ala	Gly	Ser	Thr	Arg	Gly	Ser	Leu	Leu
				740					745					750
Ser	Trp	Ser	Ser	Ala	Pro	Cys	Pro	Ser	Leu	Pro	Gln	Val	Pro	Asp
				755					760					765
Ser	Pro	Ile	Phe	Thr	Asn	Ser	Leu	Cys	Asn	Gln	Glu	Lys	Gly	Gly
				770					775					780
Leu	Glu	Pro	Arg	Arg	Leu	Val	Arg	Gly	Val	Lys	Gly	Arg	Ser	Ile
				785					790					795
Ser	Met	Lys	Ala	Pro	Thr	Thr	Ser	Arg	Ala	Lys	Gln	Gly	Pro	Phe
				800					805					810
Lys	Thr	Met	Pro	Leu	Arg	Trp	Ser	Phe	Gly	Ser	Lys	Glu	Lys	Pro
				815					820					825
Pro	Gly	Ala	Ser	Val	Glu	Leu	Val	Glu	Tyr	Leu	Glu	Ser	Arg	Arg
				830					835					840
Arg	Pro	Arg	Ser	Thr	Ser	Gln	Ser	Ile	Val	Ser	Leu	Leu	Thr	Gly
				845					850					855
Thr	Ala	Gly	Glu	Asp	Glu	Lys	Ser	Ala	Ser	Pro	Arg	Ser	Asn	Val
				860					865					870
Ala	Leu	Pro	Ala	Asn	Ser	Glu	Asp	Gly	Gly	Arg	Ala	Ile	Glu	Arg
				875					880					885
Gly	Pro	Ala	Gly	Val	Pro	Cys	Pro	Ser	Ala	Gln	Pro	Asn	His	Cys
				890					895					900
Leu	Ala	Pro	Gly	Asn	Ser	Asp	Gly	Pro	Asn	Thr	Ala	Arg	Lys	Leu
				905					910					915
Lys	Glu	Asn	Ala	Gly	Gln	Asp	Ile	Lys	Leu	Pro	Arg	Lys	Phe	Asp
				920					925					930
Leu	Pro	Leu	Thr	Val	Met	Pro	Ser	Val	Glu	His	Glu	Lys	Pro	Ala
				935					940					945
Arg	Pro	Glu	Gly	Gln	Lys	Ala	Met	Asn	Trp	Lys	Glu	Ser	Phe	Gln
				950					955					960
Met	Gly	Ser	Lys	Ser	Ser	Pro	Pro	Ser	Pro	Tyr	Met	Gly	Phe	Ser
				965					970					975
Gly	Asn	Ser	Lys	Asp	Ser	Arg	Arg	Gly	Thr	Ser	Glu	Leu	Asp	Arg
				980					985					990
Pro	Leu	Gln	Gly	Thr	Leu	Thr	Leu	Leu	Arg	Ser	Val	Phe	Arg	Lys

	995		1000		1005									
Lys	Glu	Asn	Arg	Arg	Asn	Glu	Arg	Ala	Glu	Val	Ser	Pro	Gln	Val
	1010								1015					1020
Pro	Pro	Val	Ser	Leu	Val	Ser	Gly	Gly	Leu	Ser	Pro	Ala	Met	Asp
	1025								1030					1035
Gly	Gln	Ala	Pro	Gly	Ser	Pro	Pro	Ala	Leu	Arg	Ile	Pro	Glu	Gly
	1040								1045					1050
Leu	Ala	Arg	Gly	Leu	Gly	Ser	Arg	Leu	Glu	Arg	Asp	Val	Trp	Ser
	1055								1060					1065
Ala	Pro	Ser	Ser	Leu	Arg	Leu	Pro	Arg	Lys	Ala	Ser	Arg	Ala	Pro
	1070								1075					1080
Arg	Gly	Ser	Ala	Leu	Gly	Met	Ser	Gln	Arg	Thr	Val	Pro	Gly	Glu
	1085								1090					1095
Gln	Ala	Ser	Tyr	Gly	Thr	Phe	Gln	Arg	Val	Lys	Tyr	His	Thr	Leu
	1100								1105					1110
Ser	Leu	Gly	Arg	Lys	Lys	Thr	Leu	Pro	Glu	Ser	Ser	Phe		
	1115								1120					

<210> 12

<211> 892

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7972712CD1

<400> 12

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

Ala Thr Phe Val	260	Ala Thr Phe Val	265	Ala Thr Phe Val	270
Ile Gln Glu Glu Phe	275	Asp Arg Phe Thr Gly Tyr	280	Asp Arg Phe Thr Gly Tyr	285
Trp Trp Cys Pro Thr	290	Gly Ser Glu Gly Leu Lys	295	Gly Ser Glu Gly Leu Lys	300
Leu Tyr Glu Glu Val	305	Asp Glu Ser Glu Val Glu	310	Asp Glu Ser Glu Val Glu	315
Val Ile His Val Pro	320	Glu Glu Arg Lys Thr Asp	325	Glu Glu Arg Lys Thr Asp	330
Ser Tyr Arg Tyr Pro	335	Lys Asn Pro Lys Ile Ala	340	Lys Asn Pro Lys Ile Ala	345
Leu Lys Leu Ala Glu	350	Phe Gln Thr Asp Ser	355	Phe Gln Thr Asp Ser	360
Ser Thr Gln Glu Lys	365	Glu Leu Val Gln Pro	370	Glu Leu Val Gln Pro	375
Pro Lys Val Glu Tyr	380	Ile Ala Arg Ala Gly	385	Ile Ala Arg Ala Gly	390
Lys Tyr Ala Trp Ala	395	Met Phe Leu Asp Arg	400	Met Phe Leu Asp Arg	405
Gln Leu Val Leu Leu	410	Pro Pro Ala Leu Phe	415	Pro Pro Ala Leu Phe	420
Asn Glu Glu Gln Arg	425	Leu Ala Ser Ala Arg	430	Leu Ala Ser Ala Arg	435
Val Gln Pro Tyr Val	440	Val Tyr Glu Glu Val	445	Val Tyr Glu Glu Val	450
Asn Val His Asp Ile	455	Phe Tyr Pro Phe Pro	460	Phe Tyr Pro Phe Pro	465
Asp Glu Leu Cys Phe	470	Leu Arg Ala Asn Glu	475	Leu Arg Ala Asn Glu	480
Cys His Leu Tyr Lys	485	Val Thr Ala Val Leu	490	Val Thr Ala Val Leu	495
Asp Trp Ser Glu Pro	500	Phe Ser Pro Gly Glu	505	Phe Ser Pro Gly Glu	510
Pro Ile Lys Glu Glu	515	Ile Ala Leu Thr Ser	520	Ile Ala Leu Thr Ser	525
Leu Ala Arg His Gly	530	Ser Lys Ile Trp Val	535	Ser Lys Ile Trp Val	540
Leu Val Tyr Phe Gln	545	Gly Thr Lys Asp Thr	550	Gly Thr Lys Asp Thr	555
Leu Tyr Val Val Ser	560	Tyr Glu Ala Ala Gly	565	Tyr Glu Ala Ala Gly	570
Thr Thr Pro Gly Phe	575	Ser His Ser Cys Ser	580	Ser His Ser Cys Ser	585
Asp Met Phe Val Ser	590	His Tyr Ser Ser Val	595	His Tyr Ser Ser Val	600
Val His Val Tyr Lys	605	Leu Ser Gly Pro Asp	610	Leu Ser Gly Pro Asp	615
Lys Gln Pro Arg Phe	620	Trp Ala Ser Met Met	625	Trp Ala Ser Met Met	630
Pro Pro Asp Tyr Val	635	Pro Pro Glu Ile Phe	640	Pro Pro Glu Ile Phe	645
Ser Asp Val Arg Leu	650	Tyr Gly Met Ile Tyr	655	Tyr Gly Met Ile Tyr	660
Gln Pro Gly Lys Lys	665	His Pro Thr Val Leu	670	His Pro Thr Val Leu	675
Pro Gln Val Gln Leu	680	Val Asn Asn Ser Phe	685	Val Asn Asn Ser Phe	690
Leu Arg Leu Asn Thr	695	Leu Ala Ser Leu Gly	700	Leu Ala Ser Leu Gly	705
Ile Asp Gly Arg Gly	710	Ser Cys Gln Arg Gly	715	Ser Cys Gln Arg Gly	720
Ala Leu Lys Asn Gln	725	Met Gly Gln Val Glu	730	Met Gly Gln Val Glu	735

WO 02/46383

PCT/US01/46964

Glu	Gly	Leu	Gln	Phe	Val	Ala	Glu	Lys	Tyr	Gly	Phe	Ile	Asp	Leu	
				740					745					750	
Ser	Arg	Val	Ala	Ile	His	Gly	Trp	Ser	Tyr	Gly	Gly	Phe	Leu	Ser	
				755					760					765	
Leu	Met	Gly	Leu	Ile	His	Lys	Pro	Gln	Val	Phe	Lys	Val	Ala	Ile	
				770					775					780	
Ala	Gly	Ala	Pro	Val	Thr	Val	Trp	Met	Ala	Tyr	Asp	Thr	Gly	Tyr	
				785					790					795	
Thr	Glu	Arg	Tyr	Met	Asp	Val	Pro	Glu	Asn	Asn	Gln	His	Gly	Tyr	
				800					805					810	
Glu	Ala	Gly	Ser	Val	Ala	Leu	His	Val	Glu	Lys	Leu	Pro	Asn	Glu	
				815					820					825	
Pro	Asn	Arg	Leu	Leu	Ile	Leu	His	Gly	Phe	Leu	Asp	Glu	Asn	Val	
				830					835					840	
His	Phe	Phe	His	Thr	Asn	Phe	Leu	Val	Ser	Gln	Leu	Ile	Arg	Ala	
				845					850					855	
Gly	Lys	Pro	Tyr	Gln	Leu	Gln	Ile	Tyr	Pro	Asn	Glu	Arg	His	Ser	
				860					865					870	
Ile	Arg	Cys	Pro	Glu	Ser	Gly	Glu	His	Tyr	Glu	Val	Thr	Leu	Leu	
				875					880					885	
His	Phe	Leu	Gln	Glu	Tyr	Leu									
				890											

<210> 13

<211> 818

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2751509CD1

<400> 13

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

Met	Leu	Cys	Ala	Gly	Tyr	Pro	Glu	Gly	Arg	Arg	Asp	Thr	Cys	Gln
				230					235					240
Gly	Asp	Ser	Gly	Gly	Pro	Leu	Val	Cys	Glu	Glu	Gly	Gly	Arg	Trp
				245					250					255
Phe	Gln	Ala	Gly	Ile	Thr	Ser	Phe	Gly	Phe	Gly	Cys	Gly	Arg	Arg
				260					265					270
Asn	Arg	Pro	Gly	Val	Phe	Thr	Ala	Val	Ala	Thr	Tyr	Glu	Ala	Trp
				275					280					285
Ile	Arg	Glu	Gln	Val	Met	Gly	Ser	Glu	Pro	Gly	Pro	Ala	Phe	Pro
				290					295					300
Thr	Gln	Pro	Gln	Lys	Thr	Gln	Ser	Asp	Pro	Gln	Glu	Pro	Arg	Glu
				305					310					315
Glu	Asn	Cys	Thr	Ile	Ala	Leu	Pro	Glu	Cys	Gly	Lys	Ala	Pro	Arg
				320					325					330
Pro	Gly	Ala	Trp	Pro	Trp	Glu	Ala	Gln	Val	Met	Val	Pro	Gly	Ser
				335					340					345
Arg	Pro	Cys	His	Gly	Ala	Leu	Val	Ser	Glu	Ser	Trp	Val	Leu	Ala
				350					355					360
Pro	Ala	Ser	Cys	Phe	Leu	Asp	Pro	Asn	Ser	Ser	Asp	Ser	Pro	Pro
				365					370					375
Arg	Asp	Leu	Asp	Ala	Trp	Arg	Val	Leu	Leu	Pro	Ser	Arg	Pro	Arg
				380					385					390
Ala	Glu	Arg	Val	Ala	Arg	Leu	Val	Gln	His	Glu	Asn	Ala	Ser	Trp
				395					400					405
Asp	Asn	Ala	Ser	Asp	Leu	Ala	Leu	Leu	Gln	Leu	Arg	Thr	Pro	Val
				410					415					420
Asn	Leu	Ser	Ala	Ala	Ser	Arg	Pro	Val	Cys	Leu	Pro	His	Pro	Glu
				425					430					435
His	Tyr	Phe	Leu	Pro	Gly	Ser	Arg	Cys	Arg	Leu	Ala	Arg	Trp	Gly
				440					445					450
Arg	Gly	Glu	Pro	Ala	Leu	Gly	Pro	Gly	Ala	Leu	Leu	Glu	Ala	Glu
				455					460					465
Leu	Leu	Gly	Gly	Trp	Trp	Cys	His	Cys	Leu	Tyr	Gly	Arg	Gln	Gly
				470					475					480
Ala	Ala	Val	Pro	Leu	Pro	Gly	Asp	Pro	Pro	His	Ala	Leu	Cys	Pro
				485					490					495
Ala	Tyr	Gln	Glu	Lys	Glu	Glu	Val	Gly	Ser	Cys	Trp	Thr	His	Gly
				500					505					510
Pro	Trp	Ile	Ser	His	Val	Thr	Arg	Gly	Ala	Tyr	Leu	Glu	Asp	Gln
				515					520					525
Leu	Ala	Trp	Asp	Trp	Gly	Pro	Asp	Gly	Glu	Glu	Thr	Glu	Thr	Gln
				530					535					540
Thr	Cys	Pro	Pro	His	Thr	Glu	His	Gly	Ala	Cys	Gly	Leu	Arg	Leu
				545					550					555
Glu	Ala	Ala	Pro	Val	Gly	Val	Leu	Trp	Pro	Trp	Leu	Ala	Glu	Val
				560					565					570
His	Val	Ala	Gly	Asp	Arg	Val	Cys	Thr	Gly	Ile	Leu	Leu	Ala	Pro
				575					580					585
Gly	Trp	Val	Leu	Ala	Ala	Thr	His	Cys	Val	Leu	Arg	Pro	Gly	Ser
				590					595					600
Thr	Thr	Val	Pro	Tyr	Ile	Glu	Val	Tyr	Leu	Gly	Arg	Ala	Gly	Ala
				605					610					615
Ser	Ser	Leu	Pro	Gln	Gly	His	Gln	Val	Ser	Arg	Leu	Val	Ile	Ser
				620					625					630
Ile	Arg	Leu	Pro	Gln	His	Leu	Gly	Leu	Arg	Pro	Pro	Leu	Ala	Leu
				635					640					645
Leu	Glu	Leu	Ser	Ser	Arg	Val	Glu	Pro	Ser	Pro	Ser	Ala	Leu	Pro
				650					655					660
Ile	Cys	Leu	His	Pro	Ala	Gly	Ile	Pro	Pro	Gly	Ala	Ser	Cys	Trp
				665					670					675
Val	Leu	Gly	Trp	Lys	Glu	Pro	Gln	Asp	Arg	Val	Pro	Val	Ala	Ala
				680					685					690
Ala	Val	Ser	Ile	Leu	Thr	Gln	Arg	Ile	Cys	Asp	Cys	Leu	Tyr	Gln

WO 02/46383

PCT/US01/46964

				695					700					705
Gly	Ile	Leu	Pro	Pro	Gly	Thr	Leu	Cys	Val	Leu	Tyr	Ala	Glu	Gly
				710					715					720
Gln	Glu	Asn	Arg	Cys	Glu	Met	Thr	Ser	Ala	Pro	Pro	Leu	Leu	Cys
				725					730					735
Gln	Met	Thr	Glu	Gly	Ser	Trp	Ile	Leu	Val	Gly	Met	Ala	Val	Gln
				740					745					750
Gly	Ser	Arg	Glu	Leu	Phe	Ala	Ala	Ile	Gly	Pro	Glu	Glu	Ala	Trp
				755					760					765
Ile	Ser	Gln	Thr	Val	Gly	Glu	Ala	Asn	Phe	Leu	Pro	Pro	Ser	Gly
				770					775					780
Ser	Pro	His	Trp	Pro	Thr	Gly	Gly	Ser	Asn	Leu	Cys	Pro	Pro	Glu
				785					790					795
Leu	Ala	Lys	Ala	Ser	Gly	Ser	Pro	His	Ala	Val	Tyr	Phe	Leu	Leu
				800					805					810
Leu	Leu	Thr	Leu	Leu	Ile	Gln	Ser							
				815										

<210> 14

<211> 296

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7480192CD1

<400> 14

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

WO 02/46383

PCT/US01/46964

	260		265		270
Ala Gly Gln Asn Arg Leu Tyr Gln Pro Glu Glu Ile Ile Leu Val					
	275		280		285
Ser Phe Val Leu Val Phe Phe Cys Ser Ser Ile					
	290		295		

<210> 15
 <211> 420
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 55047465CD1

<400> 15

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

WO 02/46383

PCT/US01/46964

				350					355					360
His	Phe	Pro	Tyr	Leu	Leu	Thr	Leu	Gln	Leu	Lys	Arg	Phe	Asp	Phe
				365					370					375
Asp	Tyr	Thr	Thr	Met	His	Arg	Ile	Lys	Leu	Asn	Asp	Arg	Met	Thr
				380					385					390
Phe	Pro	Glu	Glu	Leu	Asp	Met	Ser	Thr	Phe	Ile	Asp	Val	Glu	Asp
				395					400					405
Glu	Val	Asn	Ile	Cys	Tyr	Phe	Lys	Val	Phe	Phe	Ile	Asn	Pro	His
				410					415					420

<210> 16

<211> 629

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 55063036CD1

<400> 16

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

Phe	Val	Ser	Gln	Ser	Gly	Thr	Ser	Gln	Ala	Ala	Ala	His	Val	Ala
				320					325					330
Gly	Ile	Ala	Ala	Met	Met	Leu	Ser	Ala	Glu	Pro	Glu	Leu	Thr	Leu
				335					340					345
Ala	Glu	Leu	Arg	Gln	Arg	Leu	Ile	His	Phe	Ser	Ala	Lys	Asp	Val
				350					355					360
Ile	Asn	Glu	Ala	Trp	Phe	Pro	Glu	Asp	Gln	Arg	Val	Leu	Thr	Pro
				365					370					375
Asn	Leu	Val	Ala	Ala	Leu	Pro	Pro	Ser	Thr	His	Gly	Ala	Gly	Trp
				380					385					390
Gln	Leu	Phe	Cys	Arg	Thr	Val	Trp	Ser	Ala	His	Ser	Gly	Pro	Thr
				395					400					405
Arg	Met	Ala	Thr	Ala	Ile	Ala	Arg	Cys	Ala	Pro	Asp	Glu	Glu	Leu
				410					415					420
Leu	Ser	Cys	Ser	Ser	Phe	Ser	Arg	Ser	Gly	Lys	Arg	Arg	Gly	Glu
				425					430					435
Arg	Met	Glu	Ala	Gln	Gly	Gly	Lys	Leu	Val	Cys	Arg	Ala	His	Asn
				440					445					450
Ala	Phe	Gly	Gly	Glu	Gly	Val	Tyr	Ala	Ile	Ala	Arg	Cys	Cys	Leu
				455					460					465
Leu	Pro	Gln	Ala	Asn	Cys	Ser	Val	His	Thr	Ala	Pro	Pro	Ala	Glu
				470					475					480
Ala	Ser	Met	Gly	Thr	Arg	Val	His	Cys	His	Gln	Gln	Gly	His	Val
				485					490					495
Leu	Thr	Gly	Cys	Ser	Ser	His	Trp	Glu	Val	Glu	Asp	Leu	Gly	Thr
				500					505					510
His	Lys	Pro	Pro	Val	Leu	Arg	Pro	Arg	Gly	Gln	Pro	Asn	Gln	Cys
				515					520					525
Val	Gly	His	Arg	Glu	Ala	Ser	Ile	His	Ala	Ser	Cys	Cys	His	Ala
				530					535					540
Pro	Gly	Leu	Glu	Cys	Lys	Val	Lys	Glu	His	Gly	Ile	Pro	Ala	Pro
				545					550					555
Gln	Glu	Gln	Val	Thr	Val	Ala	Cys	Glu	Glu	Gly	Trp	Thr	Leu	Thr
				560					565					570
Gly	Cys	Ser	Ala	Leu	Pro	Gly	Thr	Ser	His	Val	Leu	Gly	Ala	Tyr
				575					580					585
Ala	Val	Asp	Asn	Thr	Cys	Val	Val	Arg	Ser	Arg	Asp	Val	Ser	Thr
				590					595					600
Thr	Gly	Ser	Thr	Ser	Glu	Glu	Ala	Val	Thr	Ala	Val	Ala	Ile	Cys
				605					610					615
Cys	Arg	Ser	Arg	His	Leu	Ala	Gln	Ala	Ser	Gln	Glu	Leu	Gln	
				620					625					

<210> 17
 <211> 398
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 6178623CD1

<400> 17														
Met	Ser	Arg	Lys	Gln	Ala	Ala	Lys	Ser	Arg	Pro	Gly	Ser	Gly	Ser
1				5					10					15
Arg	Lys	Ala	Glu	Ala	Glu	Arg	Lys	Arg	Asp	Glu	Arg	Ala	Ala	Arg
				20					25					30
Arg	Ala	Leu	Ala	Lys	Glu	Arg	Arg	Asn	Arg	Pro	Glu	Ser	Gly	Gly
				35					40					45
Gly	Gly	Gly	Cys	Glu	Glu	Glu	Phe	Val	Ser	Phe	Ala	Asn	Gln	Leu
				50					55					60
Gln	Ala	Leu	Gly	Leu	Lys	Leu	Arg	Glu	Val	Pro	Gly	Asp	Gly	Asn
				65					70					75

Cys	Leu	Phe	Arg	Ala	Leu	Gly	Asp	Gln	Leu	Glu	Gly	His	Ser	Arg	
				80					85					90	
Asn	His	Leu	Lys	His	Arg	Gln	Glu	Thr	Val	Asp	Tyr	Met	Ile	Lys	
				95					100					105	
Gln	Arg	Glu	Asp	Phe	Glu	Pro	Phe	Val	Glu	Asp	Asp	Ile	Pro	Phe	
				110					115					120	
Glu	Lys	His	Val	Ala	Ser	Leu	Ala	Lys	Pro	Gly	Thr	Phe	Ala	Gly	
				125					130					135	
Asn	Asp	Ala	Ile	Val	Ala	Phe	Ala	Arg	Asn	His	Gln	Leu	Asn	Val	
				140					145					150	
Val	Ile	His	Gln	Leu	Asn	Ala	Pro	Leu	Trp	Gln	Ile	Arg	Gly	Thr	
				155					160					165	
Glu	Lys	Ser	Ser	Val	Arg	Glu	Leu	His	Ile	Ala	Tyr	Arg	Tyr	Gly	
				170					175					180	
Glu	His	Tyr	Asp	Ser	Val	Arg	Arg	Ile	Asn	Asp	Asn	Ser	Glu	Ala	
				185					190					195	
Pro	Ala	His	Leu	Gln	Thr	Asp	Phe	Gln	Met	Leu	His	Gln	Asp	Glu	
				200					205					210	
Ser	Asn	Lys	Arg	Glu	Lys	Ile	Lys	Thr	Lys	Gly	Met	Asp	Ser	Glu	
				215					220					225	
Asp	Asp	Leu	Arg	Asp	Glu	Val	Glu	Asp	Ala	Val	Gln	Lys	Val	Cys	
				230					235					240	
Asn	Ala	Thr	Gly	Cys	Ser	Asp	Phe	Asn	Leu	Ile	Val	Gln	Asn	Leu	
				245					250					255	
Glu	Ala	Glu	Asn	Tyr	Asn	Ile	Glu	Ser	Ala	Ile	Ile	Ala	Val	Leu	
				260					265					270	
Arg	Met	Asn	Gln	Gly	Lys	Arg	Asn	Asn	Ala	Glu	Glu	Asn	Leu	Glu	
				275					280					285	
Pro	Ser	Gly	Arg	Val	Leu	Lys	Gln	Cys	Gly	Pro	Leu	Trp	Glu	Glu	
				290					295					300	
Gly	Gly	Ser	Gly	Ala	Arg	Ile	Phe	Gly	Asn	Gln	Gly	Leu	Asn	Glu	
				305					310					315	
Gly	Arg	Thr	Glu	Asn	Asn	Lys	Ala	Gln	Ala	Ser	Pro	Ser	Glu	Glu	
				320					325					330	
Asn	Lys	Ala	Asn	Lys	Asn	Gln	Leu	Ala	Lys	Val	Thr	Asn	Lys	Gln	
				335					340					345	
Arg	Arg	Glu	Gln	Gln	Trp	Met	Glu	Lys	Lys	Lys	Arg	Gln	Glu	Glu	
				350					355					360	
Arg	His	Arg	His	Lys	Ala	Leu	Glu	Ser	Arg	Gly	Ser	His	Arg	Asp	
				365					370					375	
Asn	Asn	Arg	Ser	Glu	Ala	Glu	Ala	Asn	Thr	Gln	Val	Thr	Leu	Val	
				380					385					390	
Lys	Thr	Phe	Ala	Ala	Leu	Asn	Ile								
				395											

<210> 18
 <211> 863
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 7484157CD1

<400> 18
 Met Gly Leu Leu Ala Ser Ala Gly Leu Leu Leu Leu Val Ile
 1 5 10 15
 Gly His Pro Arg Ser Leu Gly Leu Lys Cys Gly Ile Arg Met Val
 20 25 30
 Asn Met Lys Ser Lys Glu Pro Ala Val Gly Ser Arg Phe Phe Ser
 35 40 45
 Arg Ile Ser Ser Trp Arg Asn Ser Thr Val Thr Gly His Pro Trp
 50 55 60

Gln	Val	Ser	Leu	Lys	Ser	Asp	Glu	His	His	Phe	Cys	Gly	Gly	Ser
				65					70					75
Leu	Ile	Gln	Glu	Asp	Arg	Val	Val	Thr	Ala	Ala	His	Cys	Leu	Asp
				80					85					90
Ser	Leu	Ser	Glu	Lys	Gln	Leu	Lys	Asn	Ile	Thr	Val	Thr	Ser	Gly
				95					100					105
Glu	Tyr	Ser	Leu	Phe	Gln	Lys	Asp	Lys	Gln	Glu	Gln	Asn	Ile	Pro
				110					115					120
Val	Ser	Lys	Ile	Ile	Thr	His	Pro	Glu	Tyr	Asn	Ser	Arg	Glu	Tyr
				125					130					135
Met	Ser	Pro	Asp	Ile	Ala	Leu	Leu	Tyr	Leu	Lys	His	Lys	Val	Lys
				140					145					150
Phe	Gly	Asn	Ala	Val	Gln	Pro	Ile	Cys	Leu	Pro	Asp	Ser	Asp	Asp
				155					160					165
Lys	Val	Glu	Pro	Gly	Ile	Leu	Cys	Leu	Ser	Ser	Gly	Trp	Gly	Lys
				170					175					180
Ile	Ser	Lys	Thr	Ser	Glu	Tyr	Ser	Asn	Val	Leu	Gln	Glu	Met	Glu
				185					190					195
Leu	Pro	Ile	Met	Asp	Asp	Arg	Ala	Cys	Asn	Thr	Val	Leu	Lys	Ser
				200					205					210
Met	Asn	Leu	Pro	Pro	Leu	Gly	Arg	Thr	Met	Leu	Cys	Ala	Gly	Phe
				215					220					225
Pro	Asp	Trp	Gly	Met	Asp	Ala	Cys	Gln	Gly	Asp	Ser	Gly	Gly	Pro
				230					235					240
Leu	Val	Cys	Arg	Arg	Gly	Gly	Gly	Ile	Trp	Ile	Leu	Ala	Gly	Ile
				245					250					255
Thr	Ser	Trp	Val	Ala	Gly	Cys	Ala	Gly	Gly	Ser	Val	Pro	Val	Arg
				260					265					270
Asn	Asn	His	Val	Lys	Ala	Ser	Leu	Gly	Ile	Phe	Ser	Lys	Val	Ser
				275					280					285
Glu	Leu	Met	Asp	Phe	Ile	Thr	Gln	Asn	Leu	Phe	Thr	Gly	Leu	Asp
				290					295					300
Arg	Gly	Gln	Pro	Leu	Ser	Lys	Val	Gly	Ser	Arg	Tyr	Ile	Thr	Lys
				305					310					315
Ala	Leu	Ser	Ser	Val	Gln	Glu	Val	Asn	Gly	Ser	Gln	Arg	Gly	Lys
				320					325					330
Gly	Lys	Val	Cys	Gly	Lys	Ile	Leu	Pro	Ser	Pro	Leu	Leu	Ala	Glu
				335					340					345
Thr	Ser	Glu	Ala	Met	Val	Pro	Phe	Val	Ser	Asp	Thr	Glu	Asp	Ser
				350					355					360
Gly	Ser	Gly	Phe	Glu	Leu	Thr	Val	Thr	Ala	Val	Gln	Lys	Ser	Glu
				365					370					375
Ala	Gly	Ser	Gly	Cys	Gly	Ser	Leu	Ala	Ile	Leu	Val	Glu	Glu	Gly
				380					385					390
Thr	Asn	His	Ser	Ala	Lys	Tyr	Pro	Asp	Leu	Tyr	Pro	Ser	Asn	Thr
				395					400					405
Arg	Cys	His	Trp	Phe	Ile	Cys	Ala	Pro	Glu	Lys	His	Ile	Ile	Lys
				410					415					420
Leu	Thr	Phe	Glu	Asp	Phe	Ala	Val	Lys	Phe	Ser	Pro	Asn	Cys	Ile
				425					430					435
Tyr	Asp	Ala	Val	Val	Ile	Tyr	Gly	Asp	Ser	Glu	Glu	Lys	His	Lys
				440					445					450
Leu	Ala	Lys	Leu	Cys	Gly	Met	Leu	Thr	Ile	Thr	Ser	Ile	Phe	Ser
				455					460					465
Ser	Ser	Asn	Met	Thr	Val	Ile	Tyr	Phe	Lys	Ser	Asp	Gly	Lys	Asn
				470					475					480
Arg	Leu	Gln	Gly	Phe	Lys	Ala	Arg	Phe	Thr	Ile	Leu	Pro	Ser	Glu
				485					490					495
Ser	Leu	Asn	Lys	Phe	Glu	Pro	Lys	Leu	Pro	Pro	Gln	Asn	Asn	Pro
				500					505					510
Val	Ser	Thr	Val	Lys	Ala	Ile	Leu	His	Asp	Val	Cys	Gly	Ile	Pro
				515					520					525
Pro	Phe	Ser	Pro	Gln	Trp	Leu	Ser	Arg	Arg	Ile	Ala	Gly	Gly	Glu

Glu Ala Cys Pro	530	Gln Val Gly Leu Arg	535	Phe	540
Leu Gly Asp Tyr	545	Ile Ile Asn Pro Val	550	Trp	555
Ile Leu Thr Ala	560	Leu Lys Asn Asn Pro	565	Leu	570
Ser Trp Thr Ile	575	Asp Arg Asn Leu Lys	580	Glu	585
Ser Thr Glu Gln	590	Gln Ala Lys Leu Asn	595	Asp	600
Phe Ser Tyr Val	605	Leu Asn Leu Asn Thr	610	Phe	615
Leu Thr Thr Leu	620	Ile Glu Leu Ser Leu	625	Asn	630
Val Ser Ser Leu	635	Ser Arg Leu Gln Gln	640	Ile	645
Gln Val His Val	650	Cys Glu His Thr Tyr	655	Tyr	660
Ser Ala His Pro	665	Lys Met Ile Cys Ala	670	Gly	675
Phe Ala Ala Ser	680	Cys Gln Gly Asp Ser	685	Gly	690
Gly Pro Leu Val	695	Gly Pro Phe Val Leu	700	Tyr	705
Gly Ile Val Ser	710	Val Gln Pro Trp Lys	715	Pro	720
Gly Val Phe Ala	725	Leu Asp Trp Ile Gln	730	Ser	735
Lys Ile Asn Gly	740	Thr Asn Asn Lys Cys	745	Lys	750
Thr Leu Lys Gln	755	Thr Pro Ser Pro Asp	760	Ser	765
Ala Ser Trp Pro	770	Lys Ile Thr Arg Leu	775	Ser	780
Gln Ser Ser Asn	785	Pro Cys Glu Asp Val	790	Leu	795
Leu Thr Lys Pro	800	Ile Pro Arg Asn Ser	805	His	810
Arg Thr Thr Met	815	Met Glu Ala Thr Ile	820	Gln	825
Gly Cys Pro Val	830	Val Thr Ser Val Glu	835	Ile	840
Thr Ser Leu Asp	845		850		855
	860				

<210> 19

<211> 8285

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 3230318CB1

<400> 19

```

atgaccttaa cagtagccat tttggagaat agggactctg gaatccagat tgggggtgttg 60
tcaggaatga gtcagtgggtg tggagatgaa gatggcaaat atagatacct ctttgaagaa 120
tttatcccct caaagaatga tgagaatgga aactgctcag gggaaggaat tgaattccct 180
acaacaaatt tatatgaact ggaaagccgt gttttgactg atcattgggc catcccttac 240
aagcgagaag aatcactagg caaatgcctg ttggcatcta cctacctagc aagacttggt 300
ctttccgagt ctgatgagaa ttgtagaagg tttatggaca ggtgtatgcc tgaagcattt 360
aaaaagctcc tgacatcaag tgctgttcac aagtggggta ctgaaattca tgaaggaatt 420

```

tacaacatgt	tgatgctatt	aatagaactg	gtcgcagaga	gaataaaaaca	agatccaatt	480
cccatgggtc	tcctgggtgt	gcttacaatg	gctttcaatc	ctgataatga	ataccatttt	540
aaaaacagaa	tgaaagtgtc	tcaaaggaat	tgggcagaag	tgtttggaga	gggaaatatg	600
tttgctgttt	cacctgtatc	gactttccaa	aaggagcctc	atggatgggt	tgtggatttg	660
gtaaataagt	ttggagaatt	aggtggattt	gcagcaatcc	aagccaagct	ccattcagaa	720
gatatagaac	ttggggctgt	ctcagcactg	attcagccct	taggagtgtg	tgcagagtac	780
ctcaattcct	ccgtggtaca	gcccatgcta	gacccagtca	ttcttactac	aatccaggat	840
gtacggagtg	tagaagagaa	agacctcaaa	gacaagagat	tggttagcat	ccctgagctc	900
ttgtctgccg	ttaagttact	ttgcatgcgc	ttccaaccgg	atctggtgac	aattgtggat	960
gaccttcgac	tagatattct	attgcgcatg	ctgaaatcac	cacatttcag	tgctaagatg	1020
aattctctca	aagaagtaac	caaactaata	gaagatagca	ctttatccaa	atctgtgaag	1080
aatgctatag	atacagacag	attattagat	tggctagtgt	aaaactcagt	tctgtcgatt	1140
gcactggaag	gcaacataga	ccaagcacaa	tactgtgacc	gtataaaggg	aattattgaa	1200
ctcttgggta	gtaaattgtc	gttagatgaa	ctcactaaaa	tttggagat	acagtcagga	1260
caatcatcta	ctgtgattga	gaacattcat	actattattg	ctgcagcggc	tgtgaaattt	1320
aattcagatc	agcttaatca	tttgtttgtt	ctcattcaga	aggtttttaga	cgtactctgg	1380
gaactggctc	accttccaac	cctgcccagt	agccttattc	agcaggcctt	ggaggagcac	1440
ctgacaatcc	ttagtgatgc	atatgcagtg	aaagaagcaa	tcaagaggag	ctacatcatc	1500
aagtgcatag	aagatattaa	gaggggtggg	gttagcagat	tgagcggtaa	tgattgcagc	1560
tctcctgttg	ttccagtcct	taagcctcaa	gcctctcctc	tgagaggggt	gatcacagca	1620
gccagctcag	tggactgtgc	ttctgttgtt	gcagcagccc	taattggagc	agcattgtcc	1680
tctcacctag	accctcaggc	acttttcagt	ttactaagtg	ctttcatgga	ctttttacaaa	1740
gtacacattg	ctgaaggggg	gcagtgggaa	gatcaaagcc	ccctggatat	ggctccaggc	1800
aggggggtga	attatttact	tccactaaag	gtgtttttct	atgctatgcc	ctttcctgcc	1860
aggcagcaag	gtgggctgac	tggggattat	gtctccctgc	caggctacac	agaaactaag	1920
caaaggctcat	ctcagcttaa	taatccccag	tttgtatggg	tggtaccagc	tttgcgtcag	1980
ctccatgaaa	ttactcgctc	attcataaaa	caaacctatc	aaaagcaaga	caagagcatt	2040
attcaagact	tgaagaagaa	ttttgaaata	gtgaaattgg	taacgggaag	tttgatcgct	2100
tgtcatcggc	ttgcagctgc	tgtggccggg	cctggagggt	taagtggctc	gacactagtg	2160
gatggccggg	acacttaccg	ggagtattta	gaggcacatc	taaaattttc	agcgtttttc	2220
ttgcaagaag	ctactctgta	tctgggctgg	aatcgtgcca	aggagatctg	ggagtgtctt	2280
gtaactggcc	aggatgtttg	tgaattagat	agagagatgt	gttttgaatg	gtttacaaaa	2340
ggacagcatg	atcttgagag	tgatgttcag	cagcagctct	tcaaggagaa	aattcttaaa	2400
ttggagtcat	atgaaatcac	tatgaatggg	tttaacttat	ttaaaacttt	ttttgaaaat	2460
gtgaatcttt	gtgatcatcg	attgaaaaga	caaggagctc	agttgtatgt	agaaaagctg	2520
gaattgatag	gaatggattt	cattttggaaa	atagccatgg	aatcacctga	tgaagaaatt	2580
gctaataag	ctattcagct	aatcataaac	tatagttaca	ttaatctaaa	tcctagatta	2640
aagaaggatt	cagtatcttt	acataagaaa	ttcattgctg	attgctacac	aagattagaa	2700
gcagccagtt	cagcacttgg	tggccccact	ctaacacatg	ctgtgaccag	agcaacaaaa	2760
atgcttacag	caactgccat	gccaactgta	gcaacctcag	ttcagtcctc	ttatagatct	2820
actaaacttg	taataattga	gagattgctg	cttctggcag	agcgtatgt	gatcactata	2880
gaggattttt	actctgttcc	acgaactatt	ctacctcatg	gtgcctcatt	tcatggacat	2940
cttttaaccc	ttaatgttac	ctatgagtct	accaaagata	ccttcactgt	cgaggctcac	3000
agtaatgaaa	ccatagggag	tgtccgggtg	aaaatagcca	agcagttgtg	ctctcctgtg	3060
gataatatat	agatattttac	aaatgatagc	ctgctgacag	tgaataaaga	tcaaaagcta	3120
ctccaccaac	tgggcttttc	tgatgaacaa	atccttacag	tgaagacttc	tggcagtggg	3180
accccatctg	ggagttcagc	agattcttca	accagctcca	gcagcagcag	cagtgggggt	3240
tttagttctt	catatgccat	ggagcaggag	aaatccctcc	ctgggtgtagt	gatggctctc	3300
gtatgtaacg	tatttgacat	gctttatcag	ctcgccaatc	tggagagacc	aaggataact	3360
ctacgagtac	ggaagcttct	gctcttgata	cccactgatc	cagccattca	ggaagccctt	3420
gatcaacttg	attcttttagg	aagaaagaaa	acattgctgt	ctgaatcaag	ttctcagtc	3480
tcaaaatctc	catccctgtc	atcaaagcaa	cagcaccagc	caagtgccag	ttcaatttta	3540
gaaagtctgt	ttcgatcttt	tgccccggga	atgtctacct	tcagagtgtc	ctacaactta	3600
gaagttctaa	gctccaaact	catgccaaac	gctgatgatg	acatggccag	aagctgtgcc	3660
aatgccttct	ctgaaaactt	cctgaaaagt	ggggatttga	gtttgattgt	aatgtctatg	3720

tctggaagtg	aaggagaacc	agtagccctg	catgcgggaa	tctgtgttcg	acaacagtct	4260
gtatccacca	aagactcgct	gattgcggga	gaggctttgt	ctcttcttgt	tacgtgccta	4320
cagcttcgga	gccagcaact	ggcatctttc	tataacttgc	cctgtgttgc	tgatttcata	4380
attgatattc	tgctcggatc	accaagtgtc	gagattcgcc	gggttgccctg	tgatcagctg	4440
tacactctta	gtcagacaga	cacatcagcg	catccagatg	tgcagaagcc	aaatcagttt	4500
cttctaggcg	taatcctcac	ggctcagctg	cctctctggg	ctccaactag	tattatgaga	4560
ggagtcaatc	agagactggt	atctcagtg	atggagtatt	ttgatttgag	atgccagtta	4620
ttagatgatc	tgacaacttc	agaaatggag	cagttaagga	tcagcccagc	tacgatgctt	4680
gaagatgaga	ttacttggct	ggataacttt	gaacctaatc	gtacagctga	atgtgagacc	4740
agtgaagcgg	acaacatctt	actggcaggg	cacttacgcc	tcatcaagac	ccttctttca	4800
ctctgtgggg	cagaaaagga	aatgcttggg	tcatcactca	ttaaaccatt	gttagatgac	4860
ttccttttcc	gagcttctag	aattatttta	aatagtcatt	ctccagctgg	cagtgccgcc	4920
atcagtcaac	aggactttca	tccaaagtgt	agtacagcga	atagccgatt	ggcagcctat	4980
gaagtccttg	tgatgttggc	tgatagtcca	ccttcaaata	ttcaaattat	tataaaagaa	5040
ctgcttttcta	tgcatcacca	gcctgaccct	gctcttacca	aggagtgttg	ttaccttccc	5100
ccagtggata	gcaggtccag	ttcagggttt	gtggggctga	gaaatgggtg	tgcaacttgt	5160
tatatgaatg	cagtcttcca	gcagctgtat	atgcaacctg	ggctccctga	gtcattactt	5220
tcagtggatg	atgacacaga	caatccagat	gatagcgtgt	tttaccaggt	gcagtctctc	5280
tttggacatt	taatggaaag	caagctgcag	tactatgtac	ctgagaattt	ttggaagatt	5340
ttcaagatgt	ggaataaaga	actttatgtg	agagaacagc	aggatgcata	tgaattcttt	5400
actagtctca	ttgatcagat	ggatgaatac	ctcaagaaaa	tggggagaga	ccaaattttt	5460
aagaatacat	ttcagggtcat	ctactctgat	cagaagatct	gtaaagactg	tcctcacaga	5520
tatgagcgtg	aagaagcttt	catggctctc	aatctaggag	tgacttcttg	tcagagtttg	5580
gaaatttctt	tggaccaatt	tgttagagga	gaagttctag	aaggaagtaa	tgcgtactac	5640
tgtgaaaagt	gtaaagaaaa	gagaataaca	gtgaaaagga	cctgtattaa	atctttacct	5700
agcgtcttgg	taattcacct	aatgagattt	gggtttgact	gggaaagcgg	acgctccatt	5760
aaatatgatg	aacaaataag	gtttccctgg	atgctaaaca	tggagcctta	cacagtttca	5820
ggaatggctc	gccaaagattc	ttcttctgaa	gttggggaaa	atgggcgaag	tgtggatcag	5880
ggcgggtggag	gatccccacg	aaaaaagggt	gccctcacag	aaaactatga	acttgtcggg	5940
gtcatcgtac	acagtgggca	ggcacacgca	ggccactact	attccttcat	taaggacagg	6000
cgaggggtgtg	gaaaaggaaa	gtggtataaa	tttaatgaca	cagttataga	agaatttgac	6060
ctaaatgacg	agaccctgga	gtatgaatgc	tttggaggag	aatatagacc	aaaagtttat	6120
gatcaaacaa	accatacac	tgatgtgcgc	cgaagatact	ggaatgccta	tatgcttttc	6180
taccaaaggg	tgtctgatca	gaactcccca	gtattaccaa	agaaaagtcg	agtcagcgtt	6240
gtacggcagg	aagctgagga	tctctctctg	tcagctccat	cttcaccaga	aatttcacct	6300
cagtcatccc	ctcggcccca	taggccgaac	aatgaccggc	tgtctattct	taccaagctg	6360
gttaaaaaag	gcgagaagaa	aggactgttt	gtggagaaaa	tgcttgctcg	aatataaccag	6420
atggtgagag	atgagaacct	caagtttatg	aagaatagag	atgtatacag	tagtgattat	6480
ttcagttttg	ttttgtcttt	agcttcattg	aatgctacta	aattaaagca	tccatattat	6540
ccttgcatgg	caaagggtgag	cttacagctt	gctattcaat	tcctttttca	aacttatcta	6600
cggacaaaga	agaaactcag	ggttgatact	gaagaatgga	ttgctaccat	tgaagcattg	6660
ctttcaaaaa	gttttgatgc	ttgtcagtg	ttagttgaat	attttattag	ttctgaagga	6720
cgagaattga	taaagatttt	cttactggag	tgcaatgtga	gagaagtacg	agttgctgtg	6780
gccaccattc	tggagaaaac	cctagacagt	gccttgtttt	atcaggataa	gttaaaaagc	6840
cttcatcagt	tactggaggt	actacttgct	ctgttgagca	aagacgtccc	agaaaattgt	6900
aaaaactgtg	ctcagtactt	tttcctgttc	aacacttttg	tacaaaagca	aggaattagg	6960
gctggagatc	ttcttctgag	gcattcagct	ctgcggcaca	tgatcagctt	cctcctaggg	7020
gccagtcggc	aaaacaatca	gatacgtcga	tggagttcag	cacaagcacg	agaatttggg	7080
aatcttcaca	atacagtggc	gttacttggt	ttgcattcag	atgtctcatc	ccaaaggaat	7140
gttgctcctg	gcataattta	gcaacgacca	ccatttagca	ttgctccctc	aagccctctg	7200
ttgccctcc	atgaggaggt	agaagccttg	ttgttcatgt	ctgaaggga	accttacctg	7260
ttagaggtaa	tgtttgcttt	gcgggagctg	acaggctcgc	tcttggcact	cattgagatg	7320
gtagtgact	gctgtttctg	taatgagcat	ttttccttca	caatgctgca	tttcattaag	7380
aaccaactag	aaacggctcc	acctcatgag	ttaaagaata	cgttccaact	acttcatgaa	7440
atattgggtta	ttgaagatcc	tatacaagca	gagcgagtca	aatttggtgt	tgagacagaa	7500
aatggattac	tagctttgat	gcaccacagt	aatcatgtgg	acagtagtcg	ctgctaccag	7560
tgtgtcaaat	ttcttgtcac	tcttgctcaa	aagtgtcctg	cagctaagga	gtacttcaag	7620
gagaattccc	accactggag	ctgggctgtg	cagtggctac	agaagaagat	gtcagaacat	7680
tactggacac	cacagagtaa	tgtctcta	gaaacatcaa	ctggaaaaac	ctttcagcga	7740
accatttcag	ctcaggacac	gttagcgtat	gccacagctt	tggtgaatga	aaaagagcaa	7800
tcaggaagca	gtaatgggtc	ggagagtagt	cctgccaatg	agaacggaga	caggcatcta	7860
cagcagggtt	cagaatctcc	catgatgatt	ggtgagttga	gaagtgcctt	tgatgatgtt	7920
gatccctaga	ggaacatgcc	cagcctgaga	ggagtcaaga	cacaatactg	gatgctcagc	7980

```

accttcttgg aatcagaatc tcgaaccctt tggaagagcc tggagattgg actgggaaag 8040
ctgctgtgac ttgggcgat cgtgtatttc tcaaggaaag catttttaag ccactagaag 8100
gtttgggagc tgtttggcag tgggagaact ccggcatgtg gatcagctgt cccgggagcg 8160
tgggtctatat gtggattcac atttctgtgg agattttcgg aaatagagcc agtggcagac 8220
ttttttgtta cacgaacata caagagtggag cataaagctg ttgctttctc tacgatgcta 8280
caaag 8285

```

```

<210> 20
<211> 2767
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 5928830CB1

```

```

<400> 20
ggatcattgg ctgtactgtg gagctgtcac atctccatac tgaaaaggta cactgcaagt 60
gaatttaaac cgtttttggc tttctattgc atattgccaa gagctttgag actgatgaac 120
caagtaaatg ctctatttag ggctaagtga gatacagatt cctccaatct gatagcgttt 180
cagcctccgg agtgaggaag cagcagaaac agaagcagca gaagcaacag cagtagcagc 240
ggcagcagca acagcagcag cccctactga agtccaatag aggagacttg atctctagtt 300
cattctggaa ctccgcctgg gattgtgcac tgtccagggt cctgaaacat gaaccaaact 360
gccagcgtgt cccatcacat caagtgtcaa ccctcaaaaa caatcaagga actgggaagt 420
aacagccctc cacagagaaa ctggaaggga attgctattg ctctgctggg gatttttagtt 480
gtatgctcac tcatcactat gtcagtcac ctcttaacc cagatgaact cacaattcg 540
tcagaaacca gattgtcttt ggaagacctc tttaggaaag actttgtgct tcacgatcca 600
gaggctcggg ggatcaatga tacagatgtg gtgtataaaa gcgagaatgg acatgtcatt 660
aaactgaata tagaaacaaa tgctaccaca ttattattgg aaaacacaac ttttgtaacc 720
ttcaaagcat caagacattc agtttcacca gatttaaaat atgtccttct ggcatatgat 780
gtcaaacaga tttttcatta ttcgtatact gcttcatatg tgatttacia catacacact 840
aggggaagttt gggagttaaa tcctccagaa gtagaggact ccgtcttgca gtacgcggcc 900
tgggggtgtc aagggcagca gctgatttat atttttgaaa ataatatcta ctatcaacct 960
gatataaaga gcagttcatt gcgactgaca tcttctggaa aagaagaaat aatttttaat 1020
gggattgctg actggttata tgaagaggaa ctctgctcatt ctacatcgc ccactgggtg 1080
tcaccagatg gagaaagact tgccttcctg atgataaatg actctttggg acccaccatg 1140
gttatccctc ggtttactgg agcgttgtat cccaaaggaa agcagtatcc gtatcctaag 1200
gcagggtcaag tgaacccaac aataaaaatta tatgttgtaa acctgtatgg accaactcac 1260
actttggagc tcatgccacc tgacagcttt aaatcaagag aatactatat cactatgggt 1320
aaatgggtaa gcaataccaa gactgtggta agatgggttaa accgacctca gaacatctcc 1380
atcctcacag tctgtgagac cactacaggt gcttgtagta aaaaatatga gatgacatca 1440
gatacgtggc tctctcagca gaatgaggag ccgtgtttt ctagagacgg cagcaaattc 1500
tttatgacag tgccgtgtaa gcaaggggga cgtggagaat ttcaccacat agctatgttc 1560
ctcatccaga gtaaaagtga gcaaattacc gtgcggcatc tgacatcagg aaactgggaa 1620
gtgataaaga tcttggcata cgatgaaact actcaaaaaa tttactttct gagcactgaa 1680
tcttctccca gaggaaggca gctgtacagt gcttctactg aaggattatt gaatcgccaa 1740
tgcatttcat gtaatttcat gaaagaacaa tgtacatatt ttgatgccag ttttagtccc 1800
atgaatcaac atttcttatt attctgtgaa ggtccaaggg tcccagtggg cagcctacat 1860
agtacggaca acccagcaaa atattttata ttggaaagca attctatgct gaaggaagct 1920
atcctgaaga agaagatagg aaagccagaa attaaaatcc ttcataattga cgactatgaa 1980
cttcctttac agttgtccct tcccaaagat tttatggacc gaaaccagta tgctcttctg 2040
ttaataatgg atgaagaacc aggaggccag ctgggttacag ataagttcca tattgactgg 2100
gattccgtac tcattgacat ggataatgtc attgtagcaa gatttgatgg cagaggaagt 2160
ggattccagg gtctgaaaat tttgcaggag attcatcgaa gattaggttc agtagaagta 2220
aaggaccaa taacagctgt gaaatttttg ctgaaactgc cttacattga ctccaaaaga 2280
ttaagcattt ttggaaaggg ttatgggtggc tatattgcat caatgatctt aaaatcagat 2340
gaaaagcttt ttaaatgtgg atccgtggtt gcacctatca cagacttgaa attgtatgcc 2400
tcagctttct ctgaaagata ccttgggatg ccatctaagg aagaaagcac ttaccaggca 2460
gccagtgtgc tacataatgt tcatggcttg aaagaagaaa atatattaat aattcatgga 2520
actgctgaca caaaagttca tttccaacac tcagcagaat taatcaagca cctaataaaa 2580
gctggagtga attatactat gcaggtctac ccagatgaag gtcataacgt atctgagaag 2640
agcaagtatc atctctacag cacaatctc aaattcttca gtgattgttt gaaggaagaa 2700
atatctgtgc taccacagga accagaagaa gatgaataat ggaccgtatt tatacagaac 2760

```

tgaaggg

2767

<210> 21

<211> 5266

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7473607CB1

<400> 21

gaggagagag	gctgcacttt	gaggggtgga	aagacaaggt	gaatccccct	gctgggtcatc	60
agcttgtgcg	gctctgtggg	tgtaaaagag	tggtttggag	catgtgaagt	gagtcttcca	120
ggagatgaag	gggattgcca	ggccgtttgt	gatgatgctg	agatctgggtg	ccgtgcagcc	180
tgcttctgcg	actctcctca	tcaggcgcag	gcacagagta	ggtggagagt	tgagccagaa	240
ccacgatgtc	tttggcacag	cctctcatct	gtcagatggg	agcggggacc	ccggagaggg	300
agtcagccga	ggtcctggca	ttccttgtga	acccccgtct	gtgggtttct	ggtccagtgt	360
cccttctcca	gattagatgg	cttaggcctc	ctctaagggg	gtgggcgtgc	acatccggag	420
agctgtctgg	tgtgcaggac	tgggctgcag	gttaccctga	actgcaacca	tcttagagca	480
aggcccagct	tgcagcagga	ggagctgcag	gccgcccacc	ctagccacgg	cccctgccct	540
ggcaggaagc	ttccaagagt	aaacactgcc	taatcgtccc	gccagtagt	gagcaggcct	600
gtcccattcc	atactgacca	gattcccagt	caccaaggcc	ccctctcact	ccgctccact	660
cctcgggctg	gctctcctga	ggatgcacca	gcgtcacccc	cgggcaagat	gccctccctt	720
ctgtgtggcc	ggaatccttg	cctgtggcct	tctcctgggc	tgctggggac	cctcccattt	780
ccagcagagt	tgtcttcagg	ctttggagcc	acaggccgtg	tcttcttact	tgagccctgg	840
tgctccctta	aaagaaccat	cgccctctgc	tctccctctc	cccctccagg	ccgcccctct	900
tcccctggct	tccagaggca	gaggcagagg	cagaggcggg	ctgcaggcgg	catcctacac	960
ctggagctgc	tggtggccgt	gggccccgat	gtcttccagg	ctcaccagga	ggacacagag	1020
cgctatgtgc	tcaccaacct	caacatcggg	gcagaactgc	ttcgggaccc	gtccctgggg	1080
gctcagtttc	gggtgcacct	ggtgaagatg	gtcattctga	cagagcctga	gggtgctcca	1140
aatatcacag	ccaacctcac	ctcgtccctg	ctgagcgtct	gtgggtggag	ccagaccatc	1200
aaccctgagg	acgacacgga	tcctggccat	gctgacctgg	tcctctatat	cactagggtt	1260
gacctggagt	tgcctgatgg	taaccggcag	gtgcggggcg	tcacccagct	gggcgggtgc	1320
tgctcccca	cctggagctg	cctcattacc	gaggacactg	gcttcgacct	gggagtcacc	1380
attgcccattg	agattgggca	cagcttcggc	ctggagcacg	acggcgcgcc	cggcagcggc	1440
tgcgggccca	gcggacacgt	gatggcttcg	gacggcgccg	cgccccgcgc	cggcctcgcc	1500
tggtccccc	gcagccgccc	gcagctgctg	agcctgctca	gcgcaggacg	ggcgcgctgc	1560
gtgtggggacc	cgccgcggcc	tcaacccggg	tccgcggggc	acccgcccga	tgcgcagcct	1620
ggcctctact	acagcgccaa	cgagcagtgc	cgcgtggcct	tcggcccca	ggctgtcgcc	1680
tgcaccttcg	ccagggagca	cctggatatg	tgccaggccc	tctcctgcca	cacagaccgc	1740
ctggaccaaa	gcagctgcag	ccgcctcctc	gttcctctcc	tggatgggac	agaatgtggc	1800
gtggagaagt	ggtgctccaa	gggtcgctgc	cgctccctgg	tggagctgac	ccccatagca	1860
gcagtgcattg	ggcgctggtc	tagctggggg	ccccgaagtc	cttgctcccg	ctcctgcgga	1920
ggaggtgtgg	tcaccaggag	gcggcagtgc	aacaacccca	gacctgcctt	tggggggcgt	1980
gcatgtgttg	gtgctgacct	ccaggccgag	atgtgcaaca	ctcaggcctg	cgagaagacc	2040
cagctggagt	tcatgtcgca	acagtgcgcc	aggaccgacg	gccagccgct	gcgctcctcc	2100
cctggcgggc	cctccttcta	ccactggggg	gctgctgtac	cacacagcca	aggggatgct	2160
ctgtgcagac	acatgtgcgg	ggccattggc	gagagcttca	tcatgaagcg	tggagacagc	2220
ttcctcgatg	ggacccgggtg	tatgccaaagt	ggcccccggg	aggacgggac	cctgagcctg	2280
tgtgtgtcgg	gcagctgcag	gacatttggc	tgtgatggta	ggatggactc	ccagcaggta	2340
tgggacaggt	gccaggtgtg	tgggtggggac	aacagcacgt	gcagcccacg	gaagggctct	2400
ttcacagctg	gcagagcgag	agaatatgtc	acatttctga	cagttacccc	caacctgacc	2460
agtgtctaca	ttgccaacca	caggcctctc	ttcacacact	tggcggtgag	gatcggaggg	2520
cgctatgtcg	tggctgggaa	gatgagcatc	tcccctaaca	ccacctacgc	ctccctcctg	2580
gaggatggtc	gtgtcgagta	cagagtggcc	ctcaccgagg	accggctgcc	ccgcctggag	2640
gagatccgca	tctggggacc	cctccaggaa	gatgctgaca	tccaggttta	caggcggtat	2700
ggcgaggagt	atggcaacct	cacccgcccc	gacatcacct	tcacctactt	ccagcctaag	2760
ccacggcagg	cctgggtgtg	ggccgctgtg	cgtggggcct	gctcggtgag	ctgtggggca	2820
gggctgcgct	gggtaaacta	cagctgcctg	gaccaggcca	ggaaggagtt	ggtggagact	2880
gtccagtgcc	aagggagcca	gcagccacca	gcgtggccag	aggcctgcgt	gctcgaaccc	2940
tgccctccct	actgggcggt	gggagacttc	ggcccatgca	gcgcctcctg	tgggggcggc	3000
ctgcggggagc	ggccagtgcg	ctgcgtggag	gcccagggca	gcctcctgaa	gacattgccc	3060

ccagccccggt	gcagagcagg	ggccccagcag	ccagctgtgg	cgctggaaac	ctgcaacccc	3120
cagccctgcc	ctgccagggtg	ggagggtgtca	gagcccagct	catgcacatc	agctgggtgga	3180
gcaggcctgg	ccttgagaa	cgagacctgt	gtgccagggg	cagatggcct	ggaggctcca	3240
gtgactgagg	ggcctggctc	cgtagatgag	aagctgcctg	cccctgagcc	ctgtgtcggg	3300
atgtcatgtc	ctccaggctg	gggccatctg	gatgccacct	ctgcagggga	gaaggctccc	3360
tccccatggg	gcagcatcag	gacggggggct	caagctgcac	acgtgtggac	ccctgcggca	3420
gggtcgtgct	ccgtctcctg	cgggcgagggt	ctgatggagc	tgcgtttcct	gtgcatggac	3480
tctgccctca	gggtgcctgt	ccaggaagag	ctgtgtggcc	tggcaagcaa	gcctgggagc	3540
cggcgggagg	tctgccaggc	tgtcccgtgc	cctgctcggg	ggcagtacaa	gctggcggcc	3600
tgcagcgtga	gctgtgggag	aggggtcgtg	cggaggatcc	tgtattgtgc	ccgggccccat	3660
ggggaggacg	atgggtgagga	gatcctgttg	gacaccacgt	gccagggggt	gcctcgccccg	3720
gaaccccagg	aggcctgcag	cctggagccc	tgccaccta	ggtggaaagt	catgtccctt	3780
ggcccatgtt	cgccagctg	tggccttggc	actgctagac	gctcgggtggc	ctgtgtgcag	3840
ctcgaccaag	gccaggacgt	ggagggtggac	gaggcggcct	gtgcggcgct	ggtgcggccc	3900
gaggccagtg	tcccctgtct	cattgccgac	tgcacctacc	gctggcatgt	tggcacctgg	3960
atggagtgct	ctgtttcctg	tggggatggc	atccagcgcc	ggcgtgacac	ctgcctcgga	4020
ccccaggccc	aggcgcctgt	gccagctgat	ttctgccagc	acttgcccaa	gccggtgact	4080
gtgcgtggct	gctgggctgg	gccctgtgtg	ggacagggtg	cgcccagcct	ggtgccccac	4140
gaagaagccg	ctgctccagg	acggaccaca	gccaccctg	ctgggtgcctc	cctggagtgg	4200
tcccaggccc	ggggcctgct	cttctccccg	gctccccagc	ctcggcggct	cctgccccgg	4260
ccccaggaaa	actcagtgc	gtccagttat	gtcctgtcct	ccttcctgtc	aggcagctgc	4320
tgcaggaggg	gtgcctgtgg	caggcagcac	cttgagccaa	caggaaccat	tgacatgcga	4380
ggcccagggc	aggcagactg	tgcagtggcc	attgggcggc	ccctcgggga	ggtggtgacc	4440
ctccgcgtcc	ttgagagttc	tctcaactgc	agtgcggggg	acatgttgct	gctttggggc	4500
cggctcacct	ggaggaagat	gtgcaggaag	ctgttgga	tgactttcag	ctccaagacc	4560
aacacgctgg	tgggtaggca	gcgctgcggg	cggccaggag	gtgggggtgct	gctgcggtat	4620
gggagccagc	ttgctcctga	aaccttctac	agagaatgtg	acatgcagct	ctttggggcc	4680
tgggggtgaa	tcgtgagccc	ctcgtctagt	ccagccacga	gtaatgcagg	gggctgccgg	4740
ctcttcatta	atgtggctcc	gcacgcacgg	attgccatcc	atgccctggc	caccaacatg	4800
ggcgtctggg	ccgaggggagc	caatgccagc	tacatcttga	tccgggacac	ccacagcttg	4860
aggaccacag	cgttccatgg	gcagcagggtg	ctctactggg	agtcagagag	cagccaggct	4920
gagatggagt	tcagcgaggg	cttcctgaag	gctcaggcca	gcctgcgggg	ccagtactgg	4980
accctccaat	catgggtacc	ggagatgcag	gaccctcagt	cctggaaggg	aaaggaagga	5040
acctgagggt	cattgaacat	ttgttccgtg	tctggccagc	cctggagggt	tgacccttgg	5100
tctcagtgtc	ttccaattcg	aactttttcc	aatcttaggt	atctacttta	gagtcttctc	5160
caatgtccaa	aaggctaggg	ggttgagggt	ggggactctg	gaaaagcagc	ccccatttcc	5220
tcgggtacca	ataaataaaa	catgcaggct	caaaaaaaaa	aaaaaa		5266

<210> 22

<211> 1779

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7481673CB1

<400> 22

ctaagcctgt	ttgccggaat	taagagcaga	gctagggcca	gaaacgctag	tctgggcggt	60
taggtcagaa	ctaccccggg	agcctgacag	caggagctcg	agagaagcat	ggctcagcgg	120
tgcgttttgcg	tgtctggccct	ggtggctatg	ctgctcctag	ttttccctac	cgtctccaga	180
tcgatggggcc	cgaggagcgg	ggagtatcaa	agggcgctgc	gaatcccttc	tcagtccagc	240
aaagaggaac	gcgtcgcgat	gaaagaggca	ctgaaagggtg	ccatccagat	tccaacagtg	300
acttttagct	ctgagaagtc	caatactaca	gccctggctg	agttcggaaa	atacattcgt	360
aaagtctttc	ctacagtggg	cagcaccagc	tttatccagc	atgaagtcgt	ggaagagtat	420
agccacctgt	tcactatcca	aggctcggac	cccagcttgc	agccctacct	gctgatggct	480
cactttgatg	tgggtgcctgc	ccctgaagaa	ggctgggagg	tgccccatt	ctctgggttg	540
gagcgtgatg	gcgtcatcta	tggctcggggc	acactggacg	acaagaactc	tgtgatggca	600
ttactgcagg	ccttgagact	cctgctgac	aggaagtaca	tccccgaag	atctttcttc	660
atctctctgg	gccatgatga	ggagtcatca	gggacagggg	ctcagaggat	ctcagccctg	720
ctacagtcaa	ggggcgctcca	gctagccttc	attgtggacg	aggggggctt	catcttggat	780
gatttcatte	ctaacttcaa	gaagcccatc	gccttgattg	cagtctcaga	gaagggttcc	840
atgaacctca	tgttgcaagt	aaacatgact	tcaggccact	cttcagctcc	tccaaaggag	900

acaagcattg	gcateccttgc	agctgctgtc	agccgattgg	agcagacacc	aatgcctatc	960
atatttggaa	gcgggacagt	ggtgactgta	ttgcagcaac	tggaacatga	ggtttatgga	1020
gagaaatccc	ttaaccaatg	caataatcag	gaccaccacg	gcactcacca	tattcaaagc	1080
aggggtggccc	aggccacagt	caacttccgg	attcaccctg	gacagacagt	ccaagaggtc	1140
ctagaactca	cgaagaacat	tgtggctgat	aacagagtcc	agttccatgt	gttgagtgcc	1200
tttgaccccc	tccccgtcag	cccttctgat	gacaaggcct	tgggctacca	gctgctccgc	1260
cagaccgtac	agtccgtctt	cccgggaagtc	aataattactg	ccccagttac	ttctattggc	1320
aacacagaca	gccgattctt	tacaaacctc	accactggca	tctacagggt	ctacccccatc	1380
tacatacagc	ctgaagactt	caaacgcac	catggagtca	acgagaaaat	ctcagtccaa	1440
gcctatgaga	cccaagtga	attcatcttt	gagttgattc	agaatgctga	cacagaccag	1500
gagccagttt	ctcacctgca	caaactgtga	ggtcaagggg	cctgctgggt	taggcatgcc	1560
cgaccccggg	acaggactaa	cccaaggggg	aaagctagt	ttgatgaaac	ttttgatcaa	1620
aaccacattg	taaaacattg	cccatctgtc	ttgctcactc	ttaaactctc	ccaagaacaa	1680
ggccggggta	aggtaaagtc	agcagaaatc	tggcttctcc	cttcctccc	acatctgcat	1740
cccttgatcc	actggcattt	gctgccctct	tgtccctta			1779

<210> 23

<211> 5187

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7484316CB1

<400> 23

ggctccctgg	tagctatagc	agccgcggcg	gttaagtatg	cggcgccagg	agctgctaaa	60
tgtgaacaat	aatgtcttgg	aagagaaatt	acttttcagg	gggtcgtggg	agtgtacaag	120
ggatgtttgc	acctcgaagc	tcaacctcca	tagccccag	caaaggcctc	agcaatgagc	180
cagggcaaaa	cagctgcttc	ctcaacagtg	ccctgcagg	tttgtggcac	ttggatatct	240
tccgacgtag	ctttaggcag	cttacaactc	acaagtgc	gggagattcc	tgcattcttt	300
gcgctctcaa	gggaatcttt	aaccagtttc	agtgtagtag	tgaaaaagt	cttccatctg	360
acactctccg	cagtgtctctg	gcaaagactt	tccaggatga	acaacgtttc	cagctgggaa	420
ttatggatga	tgctgcagag	tgctttgaaa	acctcctgat	gagaattcac	ttccacattg	480
ctgatgaaac	caaagaggat	atatgtactg	cccaacactg	catttcccat	cagaaatttg	540
caatgacatt	gtttgagcag	tgtgtatgta	ctagctgtgg	tgccacttct	gatccgctgc	600
ctttcatcca	gatggtacat	tatatctcca	ccacttccct	ttgcaatcag	gctatttgta	660
tgctggaaag	acgagagaaa	ccttcaccaa	gcatgtttgg	tgagctgctg	cagaatgcca	720
gcaccatggg	ggatctgcgg	aactgtccaa	gcaactgtgg	agagaggatc	aggattcgcc	780
gtgtgttgat	gaatgctcca	cagattatca	cgattgggct	ggatggggac	tcagaccact	840
cagacttagc	agaagatgtt	atccacagcc	tgggaacctg	ccttaagctg	ggatgatctg	900
ttttcagagt	gacggatgac	cgggccaagc	aatctgaact	gtacttagtt	ggaatgatct	960
gttactatgg	caaacattat	tctacattct	tttttcaaac	aaagattcgc	aaatggatgt	1020
attttgatga	tgctcatgtc	aaggagattg	ggcccaaattg	gaaggatgtg	gtgaccaa	1080
gcatcaaggg	gcattatcag	cccctgctgc	tgctttatgc	agatccccag	ggtagccccag	1140
tttccaccca	ggacctgcct	ccccagctg	agttccagtc	atacagcagg	acatgctacg	1200
acagtgaaga	ttcaggacac	ctgactgata	gtgaatgtaa	tcagaaacac	acatccaaga	1260
aagggtcact	gatagagcgc	aagaggagct	ctgggtcggg	taggaggaaa	ggcgatgagc	1320
cccaggcctc	gggataccac	agtgaaggag	aaacactgaa	agagaagcag	gctcctagaa	1380
atgcctccaa	accatccagc	agcaccaaca	ggctgagaga	ttttaaaagag	acagtcagca	1440
atatgatcca	taacagacca	tccctggctt	ctcagaccaa	tgtaggctct	cactgcaggg	1500
gcagaggagg	agaccagcct	gacaaaaaac	ctcctaggac	cctgccttta	cactctcgtg	1560
actgggaaat	agagagtacc	agcagtgagt	caaaatccag	ttcttccagc	aagtatcgct	1620
ccacatggag	acccaaacga	gaatctctga	atattgacag	tatcttttagt	aaggacaaaa	1680
ggaagcactg	tggctatacc	cagcttagcc	ccttttctga	ggattcagct	aaagaattta	1740
taccagatga	accaagcaag	ccaccttctt	acgacattaa	atgttggtgga	ccaagcccc	1800
agtacaagcg	ctggggccca	gcacggccag	gctctcacct	tttagagcag	cacccccgac	1860
taatccagcg	aatggaatct	ggctatgaaa	gcagtgagag	gaacagcagc	agccctgtca	1920
gcctggatgc	agccctgcct	gagagctcaa	atgtctacag	ggatccaagt	gctaagagat	1980
cagctgggtt	ggttccttcc	tggcgctcata	tcccaaagtc	gcacagcagt	agcatcctgg	2040
aggtagactc	cacagcatcc	atgggtggct	ggacaaagag	tcagcctttc	tctgggtgagg	2100
agatatcttc	taaaagtga	ctggatgaat	tgcaggaaga	ggtagccagg	agggcgagc	2160
aacaggaact	tcgaagaaaa	cgggagaagg	agttagaggc	agcgaaagg	tttaaccctc	2220

```

atcctagccg cttcatggac ttggatgaac tgcagaatca ggggaggagt gacggctttg 2280
agaggtcctt gcaagaggca gagtcagtgt ttgaagagtc actacatctg gaacagaaag 2340
gagactgtgc tgcagctttg gctctctgta atgaagctat ctctaaacta agacttgccc 2400
tgcattggtg cagctgtagc acgcacagca gagccctagt cgataagaag ttgcaaatac 2460
gtattcgaaa agcacggagc ctgacaggatc gcatgcagca gcagcaatca ccacagcagc 2520
cgtcgcagcc ctacagcctg ctcccaacac aggcggggac tctctctcag ccaacaagtg 2580
aacagcctat cccgctccaa gtattgttaa gccaaaggag ccaactggaa tccggcatgg 2640
atacagagtt tggggccagt tctttcttcc attcacctgc ttcctgccat gaggcacact 2700
catcactatc tccagagtca tctgccccac agcacagctc cccagtaga tctgccttga 2760
agcttctgac ttcggttgaa gtagacaaca ttgaaccctc tgcattccac aggcaagggt 2820
tacctaaagc accagggtgg actgagaaga attctcatca tagttgggag ccattggatg 2880
ccccagaggg taagctgcaa ggctctaggt gtgacaacag cagttgcagc aagctccctc 2940
cacaagaagg aagaggcatt gctcaagaac agctgttcca agaaaagaag gatcctgcta 3000
accctcccc ggtgatgcct ggaatagcca cctctgagag gggatgatga cacagcctag 3060
gctgtagtcc ttcaaattca tcagctcagc ccagccttcc cctgtataga acctgccacc 3120
ccataatgcc tgttgcttct tcatttctgc ttcactgtcc tgatcctgtg cagaaaacta 3180
accaatgcct ccaaggccaa agcctcaaaa cttcattgac tttaaaagtg gacagaggca 3240
gtgaggagac ctataggcca gagtttccca gcacaaaggg gcttgtccgt tctctggctg 3300
agcagttcca gaggatgcag ggtgtctcca tgagggatag tacaggtttc aaggatagaa 3360
gtttgtcagg tagtctaagg aagaactctt ccccttctga ttctaagcct ctttctcac 3420
aggggtcaaga gaaaggccac tggccatggg caaagcaaca atcctctctg gaggggtggg 3480
atagaccact ttcctgggaa gagtccactg aacattcttc tcttgcccta aactctgggc 3540
tgcctaattg tgaaacttct agcggaggac agcccagggt ggcagagcca gacatatacc 3600
aagagaagct gtcccaagtg agagatgtta ggtctaagga tctgggcagc agtactgact 3660
tggggacttc cttgcctttg gattcctggg tgaatatcac aagggtctgt gattctcagc 3720
ttaagcatgg ggcacctagg ccaggaatga agtcctcccc tcatgattcc catacgtgtg 3780
taacctatcc agagagaaat cacatccttt tgcattccca ttggaaccaa gacacagagc 3840
aggagacctc agaattggag tctctgtatc aggcagctc tcaggcttct caagctggct 3900
gttctggatg ggggcagcag gataccgctt ggcacccact tagccaaaca ggctctgcag 3960
atggcatggg gaggaggttg cactcagccc atgatcctgg tctctcaaag acttcaacag 4020
cagaaatgga gcatggtctc catgaagcca gaacagtgcg tacttctcag gctacacctt 4080
gccgaggcct cagcaggagc tgtggggagg atgagcagta cagtgcagag aatttacgtc 4140
gcatctcagc cagtctcagt ggcaccgttg tctcagagag ggaggaagct ccggtttctt 4200
cccacagttt tgattcatca aacgtgagga agcctttgga aaccgggcac cgttggtcca 4260
gctcctcttc cctccctgtc atccatgacc cttctgtgtt tctcctcggg ccccaactct 4320
accttcccc accacagttc ctgtccccag atgtcctgat gccaccatg gcaggggagc 4380
ccaatagact cccaggaact tcaaggagtg tccagcagtt tctggctatg tgtgacaggg 4440
gtgaaacttc ccaaggggac aagtacacag gaaggacttt gaactaccag agcctcccc 4500
atcgctccag aacagacaac tcctgggcac cctggtcaga gaccaaccag catattggga 4560
ccagattcct gactactcca ggggtgcaat ctcaactaac ctacactgcc aactaccag 4620
aaagaagcaa gggccttcag gttcctcaca ctacagctct gagtgatctt ttccattcac 4680
cctccacccc tccattgtt catcctgtgt acccaccatc tagcagctct catgtacccc 4740
tgaggtcagc ttggaattca gatcctgttc cagggtcccg aaccctgggt cctcgaagag 4800
tagatatgcc cccagatgat gactggaggc aaagcagtta tgcctccac tctggacaca 4860
ggagaacagt gggagagggg tttctgtttg ttctatcaga tgctcccaga agagagcaga 4920
tcagggttag agtcctgcag cacagtcaat ggttaaagggt attcctttcc tttcctggag 4980
ctacaccttt ctttgtaaaa ctgtactgtg ggccgggcgc ggtggctcac acctgtaatc 5040
ccagcacttt gggaggctga ggcgggtgga tcacgaggtc aggagattga gaccatcctg 5100
gccaacatgg tgaaaccccg tctctaccaa aatacaaaaa attagccagg cgtgacgggtg 5160
cgtgcctgta gtcccaacta ctcggaa 5187

```

<210> 24

<211> 3165

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7485008CB1

<400> 24

```

atgcgcctga ctcacatctg ctgctgctgc ctcccttacc agctgggggt cctgtcgaat 60
gggatcgttt cagagctgca gttcgcccc gaccgcgagg agtgggaagt cgtgtttcct 120

```

```

gcgctctggc gccgggagcc ggtggacccg gctggcggca gcggggggcag cgcggacccg 180
ggctgggtgc gcggcggttg gggcgggcga agcgcccggg cgagggtgc cggcagctca 240
cgcgaggtgc gctctgtggc tccggtgcct ttggaggagc ccgtggaggg ccgatcagag 300
tcccggctcc ggcccccgcc gccgtcggag ggtgaggagg acgaggagct cgagtcgcag 360
gagctgccgc ggggatccag cggggctgcc gccttgctcc cgggcgcccc ggcctcgtgg 420
cagccgccgc ctcccccgca gccgcccccg tccccgcccc cggcccagca tgccgagccg 480
gatggcgacg aagtgttgct gcggatcccc gccttctctc gggacctgta cctgctgctc 540
cggagagacg gccgcttcct ggcgccgcgc ttgcagtggt aacagcggcc aaatccccggc 600
cccgggccca cgggggcagc atccgccccg caacctcccg cgccaccaga cgcaggctgc 660
ttctacaccg gagctgtgct gcggcacccct ggctcgctgg cttctttcag cacctgtgga 720
ggtggcctgg tatttaacct ttccaacac aagagtctgg gtgtgcaggt caatcttcgt 780
gtgataaagc ttattctgct ccatgaaact ccaccagaac tatatattgg gcatcatgga 840
gaaaaaatgc tagagagttt ttgtaagtgg caacatgaag aatttgga aaagaatgat 900
atacatttag agatgtcaac aaactggggg gaagacatga cttcagtgga tgcagctata 960
cttataacaa ggaaagattt ctgtgtgcac aaagatgaac catgtgatac tgttggtata 1020
gcttacttga gtggaatgtg tagtgaaaag agaaaatgta ttattgctga agacaatggc 1080
ttgaatcttg cttttacaat tgctcatgaa atgggtcaca acatgggcat taaccatgac 1140
aatgaccacc catcgtgtgc tgatggctct catatcatgt ctgggtgaatg gattaaagga 1200
cagaatcttg gtgacgtttc atggctctga tgtagcaagg aagatttgga aagatttctc 1260
aggtcaaagg ccagtaactg cttgctacaa acaaatccgc agagtgtcaa ttctgtgatg 1320
gttccctcca agctgccagg gatgacatac actgctgatg aacaatgcca gatccttttt 1380
gggccatttg cttctttttg tcaggagatg cagcatgtta tttgcacagg attatggtgc 1440
aaggtagaag gtgagaaaga atgcagaacc aagctagacc caccaatgga tggaactgac 1500
tgtgaccttg gtaagtgggt taaggctgga gaatgtacca gcaggacctc agcacctgaa 1560
catctggccg gagagtggag cctgtggagt ccttgtagcc gaacctgcag tgctgggatc 1620
agcagtcgag agcgcaaatg tcttgggcta gattctgaag caagggattg taatggtccc 1680
agaaaacaat acagaatatg tgagaatcca ccttgtcctg caggtttgcc tggattcaga 1740
gactggcaat gtcaggctta tagtgtaga acttcccccc caaagcatat acttcagtgg 1800
caagctgtcc tggatgaaga aaaaccatgt gccttgtttt gctctcctgt tggaaaagaa 1860
cagcctattc ttctatcaga aaaagtgatg gatggaactt cttgtggcta tcagggatta 1920
gatattctgt caaatggcag gtgccagaaa gttggctgtg atggtttatt agggctctct 1980
gcaagagaag atcattgtgg tgtatgcaat ggcaatggaa aatcatgcaa gatcattaaa 2040
ggggatttta atcacaccag aggagcagggt tatgtagaag tgctgggtgat acctgctgga 2100
gcaagaagaa tcaaagtgtg ggaggaaaag ccggcacata gctatttagc tctccgagat 2160
gctggcaaac agtctattaa tagtgactgg aagattgaac actctggagc cttcaatttg 2220
gctggaacta ccgttcatta tgtaagacga ggcctctggg agaagatctc tgccaaaggt 2280
cctactacag cacctttaca tcttctgggt ctctgtttc aggatcagaa ttatggtcct 2340
cactatgaat acactatccc atcagaccct cttccagaaa accagagctc taaagcacct 2400
gagccccctc tcatgtggac acacacaagc tgggaagatt gcgatgccac ttgtggagga 2460
ggagaaagga agacaacagt gtcctgcaca aaaatcatga gcaaaaatat cagcattgtg 2520
gacaatgaga aatgcaaata cttaaccaag ccagagccac agattcgaaa gtgcaatgag 2580
caaccatgtc aaacaagggt gatgatgaca gaatggaccc cttgttcacg aacttgtgga 2640
aaaggaatgc agagcagaca agtggcctgt acccaacaac tgagcaatgg aacactgatt 2700
agagccccgag agagggactg cattggggcc aagccccgct ctgccagcg ctgtgagggc 2760
caggactgca tgaccgtgtg ggaggcgga gtgtggtctg agtggtcagt caagtgtggc 2820
aaaggcatac gtcacgagc cgtagatgt accaacccaa gaaagaagtg tgtcctctct 2880
accagaccca gggaggctga agactgtgag gattattcaa aatgctatgt gtggcgaatg 2940
ggtgactggt ctaagtgtc aattacctgt ggcaaaggaa tgcagtcccg tgtaatccaa 3000
tgcatgcata agatcacagg aagacatgga aatgaatgtt tttcctcaga aaaacctgca 3060
gcatacaggc catgccatct tcaacctgca atgagaaaat taatgtaa ataccataacat 3120
caccagact ggctgctctg actttcaagt gcctgggaga tcagt 3165

```

<210> 25

<211> 1567

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 4820375CB1

<400> 25

```

ggaaaatgat gttgccca gccacgtgat ccaggctctc actccaaact caattcaaaa 60

```

```

gtatgttctg gaaagctggg agggagctaa ggtgggatct aacctgtcta acatcactgc 120
ctccaaaaca ctaaaactat attaatagca aggagagttt agaagttatc agttactcca 180
ggatacgccc gggaaaagga aacaagatca gccaaatact agaagctgga aaacagatgg 240
aagaatggga actggcacag ccagagaaag ccaggatgtg agcggcactg ggggaggctg 300
tggctcttca ggtttgtgag gcagagagag tcctcagaga cccaggaatt ggtgtcccca 360
gaactgaagg tgaaggggac cgtaggggctg aaatcagttt gattgggtag aaaagcagag 420
aagcagtgag acccctcgct ctcttgtct tccaccaga ggcctcacc tcctcagcag 480
aggtctggga gttcattccc tggagagggg aacagaaatc tctggaggaa aagctgccag 540
gcctgggagg ggttatggct cccaccctat aagtgactgt gcacactgca tgctgagggt 600
ataaaaacag aagtgagggg tgaaggacct cagatgtaca cagaacacgg tcctcagggt 660
tacacagaag tgagcacaga tgccaggaga ataggcacct gtggcccggg ggaagggggg 720
cattgaagca gagaccacg tggggataga gccattttgg ttctagggtg tgttgctgag 780
agtgagtggg ccagtgagg ggaagatctt tggagatgtg gatggggggc caggaattga 840
gacagcacag tgttgatgga ccaccacct ggctctgcag atcaccccaa caactgcagg 900
attgtgaaga ggaagattga gctctattac caggttttta acttcgccat gatcgtgtct 960
tctgcactca tgatatggaa aggcttgatc gtgctcacag gcagtgagag ccccatcgtg 1020
gtgggtgctg gtggcagtat ggagccggcc ttccacagag gagacctcct gttcctcaca 1080
aatttccggg aagacccaat cagagctggg gaaatagttg tttttaaagt tgaaggacga 1140
gacattccaa tagttcacag agtaatcaaa gttcatgaaa aagataatgg agacatcaaa 1200
tttctgacta aaggagataa taatgaagtt gatgatagag gcttgtaaa agaaggccag 1260
aactggctgg aaaagaagga cgtgggtggg agagcaagag ggtttttacc atatgttggt 1320
atgggtacca taataatgaa tgactatcca aaattcaagt atgctctttt ggctgtaatg 1380
ggtgcatatg tgttactaaa acgtgaatcc taaaatgaga agcagttcct gggaccagat 1440
tgaaatgaat tctgttgaaa aagagaaaaa ctaatatatt tgagatgttc catthttctgt 1500
ataaaaggga acagtgtgga gatgtttttg tcttgtccaa ataaaagatt caccagtaaa 1560
aaaaaaa

```

<210> 26
<211> 3308
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7483698CB1

```

<400> 26
cacatatgaa ccggttcccc cgagttaacc cgccgtgatt gtaacatcac tatggcgaat 60
gccccttaga caagttcgac gccccaggt gatggaatcg ccagatcgcc ttgtaaacga 120
ctcactatag ccgaattggc ctctagagca gctcgacgcc ccagtggtgc tggaaaccgat 180
cgggggtccc aggggtggaac catgtgggtg gcaagtggct gactgggctg ctctaccatc 240
tctcgctctt catcaccagg tcttgggaaag ttgacttcca cccaggcaa gaagccctgg 300
tgaggacact gacctcctac gaagtagtga tccccgagcg ggtcaatgag tttggagaag 360
tgttccctca gagccaccac ttccagccggc agaaacgcag ctccgaggcg ctggaaccca 420
tgccgttccg aaccactat cgcttccact cctacgggca gctcttccag ctgaacctga 480
ccgcccgatg atcctttctg gccgccggct acaccgaggt gcacttggga accccggagc 540
gcggggcctg ggagagcgac gcagggccct cggacctgcg ccactgcttc taccgcggcc 600
aggtcaactc acaggaggat tacaaggccg tcgtcagctt atgcggaggc ctgacgggaa 660
catttaaagg acagaacggt gaatatttct tagaacctat aatgaaggca gatgggaatg 720
aatatgaaga tggtcacaac aagccacatc ttatatacag acaagactta aataactctt 780
ttctgcagac tctgaagtat tgcagtgtgt cagaaagtca aataaaggaa accagtttac 840
cctttcatac ctacagcaac atgaatgaag atcttaatgt aatgaaagaa agagtthttag 900
gacacacatc aaaaaatgta ccattgaaag atgaaagaag acattccagg aaaaaacgtc 960
ttatatcata tccaagatac attgaaatta tggttacagc tgatgctaaa gtggtttctg 1020
ctcatggatc gaatttgcaa aactatatac tgactcta atgtcaattgtt gcaacaatct 1080
acaaagatcc aagtattgga aatttgatac acatagtagt ggtaaaatta gttatgattc 1140
accgtgagga ggaaggacca gtcattaatt ttgatgggtc taccacatta aagaactttt 1200
gttcatggca acaaactcag aatgaccttg atgatgttca cccttcccac catgacactg 1260
ctgttcttat cactagggaa gacatttgtt catctaaaga gaaatgtaac atgttaggtt 1320
tatcataatt aggtaccata tgtgatcctt tacaagctg ctttattaat gaagaaaaag 1380
gactcatttc tgcttttact atagcccatg agcttgggca cacacttggg gttcaacatg 1440
atgataatcc tagatgtaaa gaaatgaaag ttacaaagta tcatgtaatg gccctgctt 1500
taagttttca catgagtcct tggagctggg caaactgtag tcggaaatat gttactgaat 1560

```

```

tcctagatac tgggttacggg gaatgtcttc ttgacaaacc agatgaagaa atatataatc 1620
tgccttcaga acttcctgga tcacgatatg atggaaacaa gcagtgtgag cttgcgtttg 1680
gtcctgggtc acaaatgtgt ccccatatag agaatatatg catgcatctg tgggtgcacaa 1740
gcacagaaaa gcttcacaaa ggctgtttca ctcaacacgt gccaccagca gatggaacag 1800
actgcggtcc tggaaatgcat tgccgtcatg ggctatgtgt aaacaaagaa acggaaacac 1860
gtcctgtaaa tgggtgaatgg ggaccatggg aaccttacag ttcttgttca agaacatgtg 1920
gaggcggaat cgaaagtgca accaggcgct gtaatcgtcc tgagccaaga aacggaggaa 1980
attactgtgt gggccgcagg atgaaatttc gatcatgtaa tactgattca tgtccaaaag 2040
gcacacaaga ctttcgagag aagcagtgtc ctgattttta tggtaaacad ttggacatca 2100
gtggcattcc ctctaattgt aggtggcttc caagatacag tggcattggc acaaaggatc 2160
gttgtaaaact ctattgtcag gttgctggaa ccaattatct ctacctattg aaggatatgg 2220
ttgaagatgg tactccttgt ggaactgaaa ctcatgacat ctgtgttcaa ggccagtgtg 2280
tggcagctgg ttgtgatcac gtgttaaact ccagtgccaa gatagacaaa tgtggagtgt 2340
gtgggtgggga caactcttca tgcaagacaa taacagggtg cttcaacagt tctcattatg 2400
gttataaatgt tgttgtaaag attcccgcag gagcaacaaa cgttgacatt cgtcagtaca 2460
gctatttctgg acaaccagat gacagtacc ttgcattatc tgacgctgaa gggaaattttc 2520
ttttcaatgg aaattttctt ctaagtacgt caaaaaaaga aatcaatgtg caaggaacaa 2580
gaactgttat tgaatacagt ggatcaaata acgcagttga aagaattaat agtactaatc 2640
gacaagagaa agaacttatt ttgcagggtg tgtgtgtggg taatttatac aaccctgatg 2700
tacattattc cttcaatatc cctttggaag agaggagtga catgttcaca tgggaccctc 2760
atggaccatg ggaaggctgt accaaaatgt gtcaaggctc tcagcgaaga aacataactt 2820
gcatacataa gagtgatcat agtggtgtgt ctgataaaga atgtgaccac ttgccacttc 2880
catcatttgt tactcaaagt tgcaatacag actgtgaact aagggtggcat gttattggca 2940
aaagtgaatg ttcattccaa tgtgggtcaag gatatagaac cttggacatc cattgcatga 3000
agtattccat tcatgaagga cagactgttc aagttgatga ccactactgt ggtgaccagc 3060
ttaaacctcc tacccaagaa ctatgccatg gtaactgtgt cttcacaaga tggcattatt 3120
cagaatggtc tcagtgttcc aggagtgtg gagggaggga aagggtctcg gaatcttatt 3180
gtatgaataa ctttggccat cgtcttgctg acaatgaatg ccaagaactg tcccagatga 3240
cgagagagaa ttgcaatgaa ttttcctgtc ccagttgggc tgctagtga tggagcgagg 3300
tacattaa                                     3308

```

<210> 27

<211> 2207

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7485421CB1

<400> 27

```

gcctgtttct tcttctgaca ttgcttgatt catgtctccc tcccttttct tgttggcttc 60
taaatttgct gtggtcctgg aacctggatt tgcaaatctg aaatgatgtc cagggttaagg 120
gtgactaggt gtgagttgct ggcctgcagt ggggatggga gctgggcttg ggggtgcccac 180
caggatcagg agcaggaggt gggggaacat ccccgtagct tacagggttt tggcagttga 240
agttgtggac ggagttttcc tatgggggagc tagttttcat ctgaatgaag tcctgttttg 300
gggaactgaa aggattaaac tgatcctcta gagtagtaag ccttgaagat ggaattccct 360
gtcctttctt ccagtagctg tctcgggggg atgctttgct tgactgttct ctctgagcac 420
ccctgtctta tcacacagcg ttcactcctc ctcttctctg aatttcaggc caagtcctgt 480
atctgccatg tctgtggcgt ccacctcaac aggcctgcatt cctgcctcta ctgtgtcttc 540
ttcggctgtt tcacaaagaa gcatattcac gagcatgcga aggcgaagcg gcacaacctg 600
gccattgatc tgatgtatgg aggcattctac tgttttctgt gccaggacta catctatgac 660
aaagacatgg aaataatcgc caaggaggag cagcgaaaag cttggaaaat gcaaggcggt 720
ggagagaagt tttcaacttg ggaaccaacc aaacgggagc ttgaactgct gaagcacaaac 780
ccgaaaagga gaaagatcac ctcgaaactgc accatagggtc tgcgtgggct gatcaacctt 840
gggaacacat gcttcatgaa ctgcatcgtg caggccccga cccacacgcc acttctgcgg 900
gacttcttcc tgtctgacag gcaccgctgt gagatgcaga gcccagctc ctgtctgggtc 960
tgtgagatgt cctcactgtt tcaggagttt tactctggac accggtcccc tcacatcccg 1020
tataagttgc tgcacctggt gtggaccac gcgaggcacc tagcaggcta cgagcagcag 1080
gacgcccacg agttcctcat cgcggccctg gacgtgctcc accgacactg caaaggatgat 1140
gacaatggga agaaggccaa caaccccaac cactgcaact gcatcataga ccagatcttc 1200
acaggcgggt tgcagtcaga cgtcacctgc caagtctgcc atggagtctc caccaccatc 1260
gaccccttct gggacatcag cttggatatc cccggctctt ccacccatt ctggcccctg 1320

```

```

agcccagggga gcgaggggcaa cgtgggtaaac ggggaaagcc acgtgtcggg aaccaccacg 1380
ctcacggact gcctgcgacg attcaccaga ccagagcact tgggcagcag cgccaagatc 1440
aagtgcagcg gttgccatag ctaccaggag tccacaaagc agctcactat gaagaaactg 1500
cccatcgtag cctgttttca tctcaaacga tttgaacact cagccaagct gcggcggaag 1560
atcaccacgt atgtgtcctt ccccttgagg ctggacatga cccctttcat ggcctccagc 1620
aaagagagca ggatgaatgg acagtaccag cagcccacgg acagtctcaa caatgacaac 1680
aagtattccc tgtttgctgt tgtaaacat caagggacct tggagagtgg ccactacacc 1740
agctttatcc ggcagcacia agaccagtgg ttcaagtgtg acgatgccat catcaccaag 1800
gccagcatca aggacgtcct ggacagcgaa gggacttgc tgttctatca caaacagttc 1860
ctggaatacg agtagcctta tctgcagctg gtcagaaaaa caaaggcaat gcattggcaa 1920
gcctcaciaa gtgatcctcc ctggccccc cctcccccaa gcctcccacc gcctccccgg 1980
cctggtgaca ccacctccca tgcagatgtg gccctctgc acctgggacc catcgggtcg 2040
ggatggacca cacggacggg gaggtcctg gaggtgcttt gaagatggat gagatgaggg 2100
gtgtgctctg ggtgggagga gcagcgtaca cccgtcacca gaacatctct tgtgtcatga 2160
catgggggtg caacgggggc ctcacagcac agagtgaccg ctgcctg 2207

```

<210> 28
 <211> 986
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 7485720CB1

```

<400> 28
cctgggtggc agagtgtatc atgaggctgg acctgggtgg gcagaagggt gtagtccacc 60
cccaggtact gctcaatgtg gtggatcatt tcaacagaat cagcaagggt ggaaaccaga 120
aatgcattct tcatgtgctt ttgcggtcat ggcaaatgaa agtacttgat gtatccagca 180
gttttacagt cccttttaat gaagatgaca aagataattg ttttttagcc cacgattatt 240
tgaaaaacac atacagaatg tttaagaggg tgaatgccag ggaaagaata gttgagtggg 300
accacatagg cctaaacta cacaagaatg aactgcctt caatgaaatc atgaaaagat 360
actgccgtaa ctcagtattg gtcactagt acatgaagcc aaaggactta gggctgccta 420
cagaagcata tatttcagta gaagtctatg aagatggaac ttcagccttg aaaacatttg 480
agcatgtgac cagtgaact gcagcagagg aagctaagga aattggagtt aaacacttgt 540
tacaagacat caaagacact acagtgggca ctctttccca gtgtatcaca aaccagggtcc 600
tggaattgaa gggactgaac tccaagcttc tgggtaccag aagctacctg gaaaaagttg 660
ccacaggcaa actgtccacc aaccaccaat tcatctatca gctgcaggtc ttcaagctgc 720
tgccagatgt cagcctgcag gagttcgtca aggcctttta cctgaagacc aatgaccaga 780
tggtggtagt gtacttggcc tcgctgatcc gttccgtggg cgccctgcac aacctcatca 840
acaacaagat tgccaaccgg gatgcagaga agaaagaagg gcaggagaaa gaagagagca 900
aaaaggatag gaaagaggac aaggagaaag ataaagataa ggaaaagagt gatgtaaaga 960
aagaggagaa aaaggagaaa aagtaa 986

```

<210> 29
 <211> 3492
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 7485896CB1

```

<400> 29
atggacctgg gccccgggga cgcggcagga gggggaccgc tcgcgccccg gccccgccgc 60
cgccgctccc tgcgcgcct gttcagccgc ttctgtctgg cgctgggcag ccgctcacgc 120
cccggggact caccgccccg gcccagccg ggacactgtg atggcgacgg tgaggggggc 180
ttcgctgctg ccccgggccc agttccagcg gccccgggga gccccgggga ggaacgcccg 240
cccggacccc agccccagct ccagctcccc gccggcgatg gggcgcgggc gccgggcgct 300
cagggcttga agaaccacgg caacacctgt ttcatgaacg cgggtggtgca gtgtctcagc 360
aacaccgacc tgctggccga gttcctggcg ctggggcgct accgggcggc tccggggccg 420
gccgaggtca ccgagcagct ggcggcgctg gtgcgcgcgc tctggactcg cgaatacacg 480
ccccaacttt ccgcggagtt caagaatgca gtttccaagt acggctctca gttccaaggc 540

```

```

aattcccagc acgacgccct ggaattcctg ctctgggttg tggatcgtgt acatgaggac 600
ctggaggggt catcccagag gccgggtgtc gagaagcttc cgcctgaagc cactaaaacc 660
tctgagaact gcctgtcacc atcagctcag cttcctctag gtcaaagctt tgtgcaaagc 720
cactttcaag cacaatatag atcttccttg acttgtcccc actgcctgaa acagagcaac 780
acctttgatc ctttcctgtg tgtgtcccta cctatcccct tgcgccagac gaggttcttg 840
agtgtcacct tgggtcttccc ctctaagagc cagcgggttc tgcgggttg cctggccctg 900
ccgatcctca gcacagtggc agccctgagg aagatgggtg cagaggaagg aggcgtccct 960
gcagatgagg tgatcttggg tgaactgtat ccagtggtat tccagcggtc tttctttgat 1020
gaagaggacc tgaataccat cgcagagggg gataatgtgt atgcctttca agttcctccc 1080
tcacccagcc aggggactct ctcagctcat ccactgggtc tgtcggcctc cccacgcctg 1140
gcagcccctg agggccagcg attctccctc tctctccaca gtgagagcaa ggtgctaata 1200
ctcttctgta acttgggtgg gtcagggcag caggctagca ggtttgggac acccttcctg 1260
ataaggggag acagagctgt ttcctgggac cagctccagc agtctatcct cagcaagggtc 1320
cgccatctta tgaagagtga ggcccctgta cagaacctgg ggtctctgtt ctccatccgt 1380
gttgtgggac tctctgtggc ctgcagctat ttgtctccga aggacagtcg gccctctgt 1440
cactgggcag ttgacagggg tttgcatctc aggaggccag gaggccctcc acatgtcaag 1500
ctggcgggtg agtgggtag ctctgtcaag gagcgcctgt tcgggagcct ccaggaggag 1560
cgagcgcagg atgccgacag tgtgtggcag cagcagcagg cgcacagca gcacagctgt 1620
accttggatg aatgttttca gttctacacc aaggaggagc agctggccca ggatgacgcc 1680
tggaagtgtc ctactgcca agtcctgcag caggggatgg tgaagctgag tttgtggacg 1740
ctgcctgaca tcctcatcat ccacctcaaa aggttctgcc aggtgggcca gagaagaaac 1800
aagctctcca cgctgggtgaa gtttccgctc tctggactca acatggctcc ccatgtggcc 1860
cagagaagca ccagccctga ggcaggactg ggcccctggc cttcctggaa gcagccggac 1920
tgctgcccc cagttaccc gctggacttc ctgtacgacc tgtatgccgt ctgcaaccac 1980
catggcaacc tgcaagggtg gcattacaca gcctactgcc ggaactctct ggatggccag 2040
tggtacagtt atgatgacag cacgggtggaa ccgcttcgag aagatgaggt caacaccaga 2100
ggggccttata tcctgttcta tcagaagcgg aacagcatcc ctccctggtc agccagcagc 2160
tccatgagag gctctaccag ctccctccctg tctgatcact ggctcttacg gctcgggagc 2220
cacgtgggca gcacaagggg aagcctgctg tcctggagct ctgccccctg cccctccctg 2280
ccccagggtc ctgactctcc catcttcacc aacagcctct gcaatcagga aaaggaggag 2340
ttggagccca ggcgtttggg acggggcgtg aaaggcagaa gcattagcat gaaggcacc 2400
accacttccc gagccaagca gggaccattc aagaccatgc ctctgcggtg gtcctttgga 2460
tccaaggaga aaccaccagg tgcctccgtc gagttgggtg agtacttgga atccagacga 2520
agacctcggg ccacgagcca gtccattgtg tcgctgttga cgggcaactg ggtgaggat 2580
gagaagttag catcgccgag gtccaacgtc gcccttcctg ctaacagcga agatgggtgg 2640
cgggccattg aaagagggtc agccgggggtg ccctgtccct cggctcaacc caaccactgt 2700
ctggccccct gaaactcaga tgggtccaaac acagcaagga aactcaagga aaatgcagg 2760
caggacatca agcttcccag aaagtttgac ctgcctctca ctgtgatgcc ttcagtggag 2820
catgagaaac cagctcgacc ggaggggccag aaggccatga actggaagga gagcttccag 2880
atgggaagca aaagcagccc accctcccc tatatgggat tctctggaaa cagcaaagac 2940
agtcgcccag gcacctctga gctagacaga cccctgcagg ggacactcac ccttctgagg 3000
tccgtgtttc ggaagaagga gaacaggagg aatgagaggg cagaggtctc tccacagggt 3060
cccccggtc ccctgggtgag tggcgggctg agccctgcca tggacgggca ggctccaggc 3120
tcacctcctg ccctcaggat cccagagggc ctggccaggg gcctgggagc ccggctcgag 3180
agggatgtct ggtcagcccc cagctctctc cgcctccctc gtaaagccag cagggccccc 3240
agaggcagtg cactgggcat gtcacaaagg actgttccag gggagcaggc ttcttatggc 3300
acctttcaga gagtcaaata tcacactctt tctttagggtc gaaagaaaac cttaccggag 3360
tccagctttt gatggagcgt gtcagtattg tgtgacgctg gcattcttgg gactttgcca 3420
agcaactgta ggcagctcat gttgagaatg ggtttccagg aaaccggtg tcttgtaata 3480
tctaaaaaaa aa 3492

```

<210> 30

<211> 3716

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7972712CB1

<400> 30

```

ggtgcctgag ccggcgggtc ccctgtgtcc gccgcggctg tcgtcccccg ctcccggcac 60
ttccgggggtc gcagtcccgg gcatggagcc gcgaccgtga ggcgcgctg gaccggggac 120

```

```

gacctgccca gtccggccgc cgcacacgt cccggctctgt gtccacgcgc tgcagctgga 180
atggaggctc tctggaccct ttagaaggca cccctgccct cctgaggtca gctgagcggg 240
taatgcggaa ggttaagaaa ctgcgcctgg acaaggagaa caccggaagt tggagaagct 300
tctcgctgaa ttccgagggg gctgagagga tggccaccac cgggaccca acggccgacc 360
gaggcgacgc agccgccaca gatgaccggg cgcggcgctt ccagggtgcag aagcactcgt 420
gggacgggct ccggagcatc atccacggca gccgcaagta ctcgggcctc attgtcaaca 480
aggcgcccca cgacttccag tttgtgcaga agacggatga gtctggggcc cactcccacc 540
gcctctacta cctgggaatg ccataatggca gccgagagaa ctccctcctc tactctgaga 600
ttcccaagaa ggtccggaaa gaggtctctgc tgctcctgtc ctggaagcag atgctggatc 660
atttccaggc cagcggccac catgggggtct actctcgga ggaggagctg ctgagggagc 720
ggaaacgcct ggggggtctt gccatcacct cctacgactt ccacagcgag agtggcctct 780
tcctcttcca ggccagcaac agcctcttcc actgccgcga cggcggaag aacggcttca 840
tggtgtcccc tatgaaaccg ctggaaatca agaccagtg ctcaggggcc cggatggacc 900
ccaaaatctg ccctgccgac cctgccttct tctccttcat caataacagc gacctgtggg 960
tgccaacat cgagacaggc gaggagcggc ggctgacctt ctgccacca ggtttatcca 1020
atgtcctgga tgacccaag tctgcgggtg tggccacctt cgtcatacag gaagagtctg 1080
accgcttcac tgggtactgg tggtgcccc cagcctcctg ggaagggttca gagggcctca 1140
agacgctgcg aatcctgtat gaggaagtgc atgagtccga ggtggaggtc attcacgtcc 1200
cctctcctgc gctagaagaa aggaagacgg actcgtatcg gtaccccgag acaggcagca 1260
agaatcccaa gattgccttg aaactggctg agttccagac tgacagccag ggcaagatcg 1320
tctcgacca ggagaaggag ctggtgcagc ccttcagctc gctgttccc aaggtggagt 1380
acatcgccag ggccgggtgg acccgggatg gcaaatacgc ctggggccatg ttcctggacc 1440
ggccccagca gtggctccag ctgcctcctc tcccccggc cctgttcac ccgagcacag 1500
agaatgagga gcagcggtta gcctctgcca gagctgtccc caggaatgtc cagccgtatg 1560
tggtgtacga ggaggtcacc aacgtctgga tcaatgttca tgacatcttc tatcccttcc 1620
cccaatcaga gggagaggac gagctctgct ttctccgcgc caatgaatgc aagaccggct 1680
tctgccattt gtacaaagtc accgccgttt taaaatccca gggctacgat tggagtgagc 1740
ccttcagccc cggggaagat gaatttaagt gcccattaa ggaagagatt gctctgacca 1800
gcgggtgaatg ggaggttttg gcgaggcacg gctccaagat ctgggtcaat gaggagacca 1860
agctgggtgta cttccagggc accaaggaca cgcgcgtgga gcaccacctc tacgtgggtca 1920
gctatgaggc ggccggcgag atcgtacgcc tcaccacgcc cggcttctcc catagctgct 1980
ccatgagcca gaacttcgac atgttcgtca gccactacag cagcgtgagc acgccgccct 2040
gcgtgcacgt ctacaagctg agcgcccccg acgacgaccc cctgcacaag cagccccgct 2100
tctgggctag catgatggag gcagccagct gcccccgga ttatgttctt ccagagatct 2160
tccatttcca cagcgctcg gatgtgcggc tctacggcat gatctacaag cccacgcct 2220
tgcagccagg gaagaagcac cccaccgtcc tctttgtata tggaggcccc caggtgcagc 2280
tggtgaataa ctcttcaaa ggcataagt acttgcggt caacacactg gcctccctgg 2340
gctacgccgt ggttgtgatt gacggcaggg gctcctgtca gcgagggtt cggttcgaag 2400
gggccctgaa aaaccaaagt ggccagggtg agatcgagga ccagggtggag ggcctgcagt 2460
tcgtggccga gaagtatggc ttcacgacc tgagccgagt tgccatccat ggctgggtcct 2520
acgggggctt cctctcgctc atggggctaa tccacaagcc ccagggtgtt caggtgacac 2580
tcgcgggtgc cccggtcacc gtctggatgg cctacgacac agggtagact gagcgctaca 2640
tggaagctcc tgagaacaac cagcacggct atgaggcggg ttccgtggcc ctgcacgtgg 2700
agaagctgcc caatgagccc aaccgcttgc ttatcctcca cggcttcttg gacgaaaacg 2760
tgactttttt ccacacaaac ttcctcgtct cccaactgat ccgagcaggg aaaccttacc 2820
agctccagat ctacccaac gagagacaca gtattcgttg ccccgagtcg ggcgagcact 2880
atgaagtcac gttgctgcac tttctacagg aatacctctg agcctgcccc ccgggagccg 2940
ccacatcaca gcacaagtgg ctgcagcctc cgcggggaac caggcgggag ggactgagt 3000
gcccgcgggc ccagtgagg cactttgtcc cgcacgcgc tggccagccc cgaggagccg 3060
ctgccttcac cgcggcgacg ccttttatcc ttttttaaac gctcttgggt tttatgtccg 3120
ctgcttcttg gttgccgaga cagagagatg gtggtctcgg gccagccctt cctctccccg 3180
ccttctggga ggaggaggtc acacgtgat gggcactgga gaggccagaa gagactcaga 3240
ggagcgggct gccttcggcc tggggctccc tgtgacctct cagtcctctg gccgggccag 3300
ccacgctccc cagcaccaca gcatgcaatt gcctgtcccc cccggccagc ctccccact 3360
tgatgtttgt gttttgtttg ggggatatt tttcataatt atttaaaaga caggccgggc 3420
gcgggtggctc acgtctgtaa tcccagcact ttgggagggt gaggcgggag gatcacctga 3480
ggttgggagt tcaagaccag cctggccaac atggggaaac cccgtctcta ctaaaatac 3540
aaaaaattag ccgggtgtgg tggcgctgc ctataatccc agctactcgg gaggctgagg 3600
caggagaatc gcttgaaccc gggagggtga ggttgcgggt agccaagatc gcaccattgc 3660
actccagcct gggcaacaag agcgaaactc tgtctcaaaa taataaaaa ataaaa 3716

```

<210> 31

<211> 2681

WO 02/46383

PCT/US01/46964

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2751509CB1

<400> 31

```

atggccccggc acctgctcct ccccttgtg atgcttgtca tcagtcccat cccaggagcc 60
ttccaggact cagctctcag tcctaccag gaagaacctg aagatctgga ctgcgggccc 120
cctgagccct cggcccgcac cgtggggggc tcaaacgcgc agccggggcac ctggccttgg 180
caagtgagcc tgcaccatgg aggtggccac atctgcgggg gctccctcat cgccccctcc 240
tgggtcctct ccgctgctca ctgtttcatg acgaatggga cgctggagcc cgcggccgag 300
tggtcgggtac tgctgggctg gactccagc gacggggccc tggacggcgc gcacaccgac 360
gcagtggccg ccctcgtggt gccggccaac tacagccaag tggagctggg cgcggacctg 420
gccctgctgc gcctggcctc acccgccagc ctggggcccc ccgtgtggcc tgtctgcctg 480
ccccgcgcct cacaccgctt cgtgcacggc accgcctgct gggccaccgg ctggggagac 540
gtccaggagg cagatcctct gcctctcccc tgggtgctac aggaagtgga gctaaggctg 600
ctgggagagg ccacctgtca atgtctctac agccagcccg gtcccttcaa cctcactctc 660
cagatattgc cagggatgct gtgtgctggc taccagagg gccgcaggga cacctgccag 720
ggtgactctg gggggccctt ggtctgtgag gaaggcggcc gctggttcca ggcaggaatc 780
accagctttg gctttggctg tggacggaga aaccgccctg gagttttcac tgctgtggct 840
acctatgagg catggatacg ggagcagggt atgggttcag agcctgggac tgcctttccc 900
accagcccc agaagacca gtcagatccc caggagccca gggaggagaa ctgcaccatt 960
gccctgcctg agtgcgggaa ggccccgcgc ccaggggcct ggccctggga ggcccagggt 1020
atgggtgccag gatccagacc ctgccatggg gcgctggtgt ctgaaagctg ggtcttggca 1080
cctgccagct gctttctgga cccgaacagc tccgacagcc caccgcgca cctcgacgcc 1140
tggcgcgtgc tgctgccctc gcgcccgcgc gcggagcggg tggcgcgcct ggtgcagcac 1200
gagaacgctt cgtgggacaa cgcctcggac ctggcgctgc tgcagctgcg cacgcccgtg 1260
aacctgagcg cggcttcgcg gcccgctgtg ctaccccacc cggaacacta cttcctgccc 1320
gggagccgct gccgcctggc ccgctggggc gcgggggaac ccgcgcttgg cccaggcgcg 1380
ctgctggagg cggagctggt aggcggctgg tgggtgccact gcctgtacgg ccgcccagggg 1440
gcggcagtag cgctgcccg agaccgcgc cacgcgctct gccctgccta ccaggaaaag 1500
gaggaggtgg gcagctgctg gactcatggc ccatggatca gccatgtgac tcggggagcc 1560
tacctggagg accagctagc ctgggattgg ggccctgatg gggaggagac tgagacacag 1620
acttgtcccc cacacacaga gcatggtgcc tgtggcctgc ggctggaggc tgctccagt 1680
ggggtcctgt ggccctggct ggcagagggt catgtggctg gtgatcgagt ctgcactggg 1740
atcctcctgg ccccaggctg ggtcctggca gccactcact gtgtcctcag gccaggctct 1800
acaacagtgc cttacattga agtgtatctg ggccgggcag gggccagctc cctcccacag 1860
ggccaccagg tatcccgctt ggtcatcagc atccggctgc ccagcacct gggactcagg 1920
ccccccctgg ccctcctgga gctgagctcc cgggtggagc cctccccatc agccctgccc 1980
atctgtctcc acccggcggg tatccccccg ggggccagct gctgggtgtt gggctggaaa 2040
gaaccccagg accgagctcc tgtggctgct gctgtctcca tcttgacaca acgaatctgt 2100
gactgcctct atcagggcat cctgcccccc ggaacctct gtgtcctgta tgcagagggg 2160
caggagaaca ggtgtgagat gacctcagca ccgcccctcc tgtgccagat gacggaagg 2220
tcctggatcc tcgtgggcat ggctgttcaa gggagccggg agctgtttgc tgccattgg 2280
cctgaagagg cctggatctc ccagacagt ggagaggcca acttcctgcc cccagtggt 2340
tccccacact ggccactgg aggcagcaat ctctgcccc cagaactggc caaggcctcg 2400
ggatccccgc atgcagtcta cttcctgctc ctgctgactc tcctgatcca gagctgagg 2460
gctaggggtc cagcaccact tcccccttct ccacctcta cttcccgccc agtggggctg 2520
gaatgtggcc cagccggctg gaacctcaag ggccccaccc accgagattg cagcggctct 2580
ggctaattgg gcctcagtgc ccgggctatt ttgaaccag gaatccttgg ggggtgggg 2640
aggagcggac aataaagggt taaacacaaa aaaaaaaaaa a 2681

```

<210> 32

<211> 1293

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7480192CB1

<400> 32

```

cagaaccaag actagggata agacagctgc ccatgggtgtc cgcggcgggt ctctctgggg 60
atggcaagat gcgaggggtg ctcttggtgc tgctcggcct tctctactct tccaccagaa 120
ggacattgtc gttatagtgg gtataagtaa catggatcct agcaagattg ctcacacaga 180
gtatccagtc aataccatca tcatccatga ggactttgat aacaactcca tgagcaacaa 240
catagccctc ctgaagacag acacagcgat gcatttttggc aacctgggtcc agtccatctg 300
cttcctcggc agaatgctgc atacaccacc agtcttgagc aactgctggg tgtcaggatg 360
gaatcccaca tctgcaacag gaaatcacat gacgatgagt gtcctgagga aaatcttcgt 420
gaaagatctt gacatgtgtc ccctatacaa actccagaag acagaatgcg gcagccacac 480
gaaagaggaa accaagactg cctgcttggg ggacccagga agcccaatga tgtgccagct 540
acagcagttc gatctgtggg ttctgagagg agtcctgaac ttcgggtgggt agacgtgccc 600
tggcctgttt ctgtacacca aggtggaaga ctacagcaaa tggatcacat ccaaggctga 660
gagggccggc cctcccctgt cctcactcca ccactgggaa aagttgattt ctttctccca 720
ccatggacca aatgccacca tgacacagaa gacatattct gattctgaac tgggccatgt 780
tggatcatac ttgcagggac aaagaaggac catcacgcat tcacgactag gaaacagctc 840
tagagatagt ctagatgtta gggagaagga tgtaaaggaa tcaggcaggt ctctgagggc 900
gtctgtacaa cccttatact atgactatta cgggtggggag gtggggggaag gtaggatttt 960
tgcaggtcag aacaggttgt atcagcccga agaaatcatc ttgggtttcct tcgtgcttgt 1020
tttcttttgc agcagtatct agtccaggag ctacccacc aaactgaaga gtaaactgag 1080
aatgctgagt gccaggcatt caccatgctg ttttgatgtc tgtttttgat agttgcacac 1140
tggggctgcc acggataagc ccatggcata cactgggctg gctctccctc ctctatccct 1200
ctcccaggtg tgggaaggtc actttcacta tgcttgtgaa ctaaattgctg gctaacaagt 1260
gtcaaaccaa aaaaaaaaaa aaaaaaaaaa ttg 1293

```

<210> 33

<211> 1579

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 55047465CB1

<400> 33

```

cctgggtcgg tgtctgcgcg ctgggtgtctg agggccaggc tgaggcctcc gctactgctg 60
gagcgcaggc ggcggagagg atgactgccg ctgccattct ctcttgagct agcgagccgc 120
cgccaccctc caccctcccc cggcagggcg gagaggagcg gccggagtca gcgatgggtgc 180
cgggcgagga gaaccaactg gtcccgaag agatagaaaa tgctgctgaa gaacctagag 240
tcttatgtat tatacaagat actactaatt caaagacagt gaatcaacgg atcactttaa 300
atttaccagc atctactcca gtcagaaagc tctttgaaga tgtggccaac aaagtaggct 360
acataaatgg aacctttgac ttgggtgtggg gaaatggaat caatactgct gatgtggcac 420
cactggatca taccagtgcg aagtcacttc tcgacgctaa ttttgagcca ggaaagaaga 480
actttctgca tttgacagat aaagatgggtg aacaacctca aatactgctg gaggattcca 540
gtgctgggga agacagtgtt catgacaggt ttatagggtc gcttccaaga gaaggttctg 600
tgggttctac cagtgattat gtcagccgaa gctactccta ctcatctatt ttgaataaat 660
cagaaactgg atatgtggga ctagtaaacc aagcaatgac ttgctatttg aatagccttt 720
tgcaaact ttttatgact cctgaattta ggaatgcatt atataagtgg gaatttgaag 780
aatctgaaga agatccagtg acaagtattc cataccaact tcaaaggctt tttgttttgt 840
tacaaccag caaaaagaga gcaattgaaa ccacagatgt tacaaggagc tttggatggg 900
atagtagtga ggcttggcag cagcatgatg tacaagaact atgcagagtc atgtttgatg 960
ctttggaaca gaaatggaag caaacagaa aggctgatct tataaatgag ctatatcaag 1020
gcaagctgaa ggactacgtg agatgtctgg aatgtgggtt tgagggtctg cgaatcgaca 1080
catatcttga tattccattg gtcacccgac cttatgggtc cagccaagca tttgctagt 1140
tggaagaagc attgcatgca tttattcagc cagagattct ggatggccca aatcagtatt 1200
tttgtgaacg ttgtaagaag aagtgtgatg cacggaagg ccttcggttt ttgcattttc 1260
cttatctgct gaccttacag ctgaaaagat tcgattttga ttatacaacc atgcatagga 1320
ttaaactgaa tgatcgaatg acatttcccg aggaactaga tatgagtact tttattgatg 1380
ttgaagatga ggtaaatatt tgttatttta aagtattttt cattaatcca cattagtcag 1440
tcatggtttt gaaagagctg tttgaggtaa taacactgat gtcattttgga ctgtcttctt 1500
gatggaaact ttttctttct tattttctaa tgtggaattc aataaaatga tttacctttg 1560
taaaaaaaaa aaaaaagg 1579

```

<210> 34

WO 02/46383

PCT/US01/46964

<211> 2591

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 55063036CB1

<400> 34

```

tggagggcgct catggttgca ggcggggcgcc gccgttcagt tcaggggtctg agcctggagg 60
agtgagccag gcagtgagac tggctcgggc gggccgggac gcgtcgttgc agcagcggct 120
cccagctccc agccaggatt ccgcgcgccc cttcacgcgc cctgctcctg aacttcagct 180
cctgcacagt cctccccacc gcaaggctca aggcgcgcgc ggcgtggacc gcgcacggcc 240
tctaggtctc ctgccagga cagcaacctc tcccctggcc ctcatgggca ccgtcagctc 300
caggcgggtcc tgggtggccgc tgccactgct gctgctgctg ctgctgctgc tcctgggtcc 360
cgcgggcgcc cgtgcgcagg aggacgagga cggcgactac gaggagctgg tgctagcctt 420
gcgttccgag gaggacggcc tggccgaagc acccgagcac ggaaccacag ccaccttcca 480
ccgctgcgcc aaggccttga agttgccccca tgtcgactac atcgaggagg actcctctgt 540
ctttgcccag agcatcccgt ggaacctgga gcggattacc cctccacggg accgggcgga 600
tgaataccag cccccgcagc gaggcagcct ggtggagggtg tatctcctag acaccagcat 660
acagagtgcac caccgggaaa tcgaggggcag ggtcatgggtc accgacttcg agaatgtgcc 720
cgaggaggac gggaccccgct tccacagaca ggccagcaag tgtgacagtc atggcaccca 780
cctggcaggg gtggtcagcg gccgggatgc cggcgtggcc aagggtgcca gcatgcgcag 840
cctgcgcgtg ctcaactgcc aagggaaggg cacgggttagc ggcaccctca taggcctgga 900
gtttatttcgg aaaagccagc tgggtccagcc tgtggggcca ctggtggtgc tgctgcccct 960
ggcgggtggg tacagccgcg tcctcaacgc cgcctgccag cgcctggcga gggctggggg 1020
cgtgctgggtc accgctgccg gcaacttccg ggacgatgcc tgcctctact cccagcctc 1080
agctcccagc gtcattcacg ttggggccac caatgccccg gaccagccgg tgaccctggg 1140
gacttttggg accaactttg gccgctgtgt ggacctcttt gcccagggg aggacatcat 1200
tgggtgcctcc agcgactgca gcacctgctt tgtgtcacag agtgggacat cacaggctgc 1260
tgcccacgtg gctggcattg cagccatgat gctgtctgcc gagccggagc tcaccctggc 1320
cgagttgagg cagagactga tccacttctc tgccaaagat gtcattcaatg aggcctgggt 1380
ccctgaggac cagcgggtac tgacccccaa cctgggtggcc gccctgcccc ccagcaccca 1440
tggggcaggt tggcagctgt tttgcaggac tgtgtggtca gcacactcgg ggcctacacg 1500
gatggccaca gccatcgccc gctgcgcccc agatgaggag ctgctgagct gctccagttt 1560
ctccaggagt ggggaagcggc ggggcgagcg catggaggcc caagggggca agctggtctg 1620
ccgggcccac aacgcttttg ggggtgaggg tgtctacgcc attgccaggt gctgcctgct 1680
accccaggcc aactgcagcg tccacacagc tccaccagct gaggccagca tggggacccg 1740
tgtccactgc caccaacagg gccacgtcct cacaggctgc agctcccact gggagggtgga 1800
ggaccttggc accacaagc cgcctgtgct gaggccacga ggtcagccca accagtgcgt 1860
gggccacagg gaggccagca tccacgcttc ctgctgccat gcccaggtc tggaatgcaa 1920
agtcaaggag catggaatcc cggccccctca ggagcaggtg accgtggcct gcgaggaggg 1980
ctggaccctg actggctgca gtgccctccc tgggacctcc cacgtcctgg gggcctacgc 2040
cgtagacaac acgtgtgtag tcaggagccg ggacgtcagc actacaggca gcaccagcga 2100
agaggccgtg acagccgttg ccatctgctg ccggagccgg cacctggcgc aggcctcca 2160
ggagctccag tgacagcccc atcccaggat ggggtgtctg ggaggggtcaa gggctggggc 2220
tgagctttta aatgggttccg acttgtccct ctctcagccc tccatggcct ggcacgaggg 2280
gatggggatg cttccgcctt tccggggctg ctggcctggc ccttgagtgg ggcagcctcc 2340
ttgcctggaa ctcaactact ctgggtgcct cctccccagg tggagggtgc aggaagctcc 2400
ctccctcact gtggggcatt tcaccattca aacaggctga gctgtgctcg ggtgctgcca 2460
gctgctccca atgtgccgat gtccgtgggc agaattgact ttattgagct cttgttccgt 2520
gccaggcatt caatcctcag gtctccacca ggaggcagga ttcttcccat ggatagggga 2580
gggggcggtg g
2591

```

<210> 35

<211> 1197

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 6178623CB1

<400> 35

```

atgtccccgaa agcagggcggc gaagagccgg cggggcagcg gcagccggaa agccgaggcc 60
gagcgcaagc gggacgagcg ggcggcgcg cgggccctgg ccaaggagcg gcggaatcgg 120
ccggagtctg gcggcgggcg cggctgcgag gaggagtctg tcagcttcgc caaccagctg 180
cagggccctgg ggctgaagct gcgggaggtg cggggggacg gcaattgctt gttcagagct 240
cttggtgatc aattggaggg aactcacga aatcatctca agcacagaca ggagacagtg 300
gactacatga taaagcagcg ggaagatttt gaaccctttg tagaagatga cattcctttt 360
gagaagcatg tggccagttt ggcaaagcct ggtacttttg ctggcaatga tgcaattgta 420
gcctttgcaa gaaatcatca gttgaatgta gtgattcatc aacttaatgc ccctttgtgg 480
cagattcgtg gtacagagaa aagcagcgtg agggagtac acatcgcata tcggtatgga 540
gagcactacg acagtgttcg gaggatcaat gacaactcag aggcacctgc acatctccag 600
acggattttc agatgcttca tcaagatgaa tcaaataaaa gagaaaagat caagacaaag 660
ggaatggact ctgaagacga cctgagagat gaagtagagg atgctgtcca gaaagtttgt 720
aatgcaactg gatgttcaga ttttaattta atagtccaga acctggaagc tgaaaattat 780
aatattgaat ctgcaataat tgccgtgctt cggatgaacc aagggaagag aaataatgca 840
gaagagaatc ttgagcccag tggctcgagt ctgaagcagt gtggcccttt gtgggaggag 900
ggtggcagtg gtgccagaat ctttggaat cagggcttaa atgaaggcag gaccgaaaac 960
aataaggcac aggccagccc tagtgaagaa aacaaagcaa ataaaaacca gctcgcaaag 1020
gtcacaaaca aacagaggcg agaacagcag tggatggaga agaagaagcg gcaggaggag 1080
aggcaccgcc acaaagccct ggagagcaga ggtagccaca gggacaataa cagaagcgaa 1140
gcagaggcga acacgcaggt caccttggtg aagaccttcg ccgctctcaa catctga 1197

```

<210> 36

<211> 2627

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7484157CB1

<400> 36

```

ggagggagga cgtgcgaggg cgggtgcgtgg aggtatggg cctgctggcc agtgctgggt 60
tgttgctgtt gctggtcatc ggccacccca gaagcctagg actgaagtgt ggaattcgca 120
tgggtcaacat gaaaagtaag gaacctgccg tgggatctag attcttctct agaattagta 180
gttggagaaa ttcaacagtg actggacatc catggcaggt ctccctaaaa tcagatgagc 240
accacttctg tggaggaagc ttgattcaag aagatcgggt tgttacagca gcacactgcc 300
tggacagcct cagtgagaag cagctgaaga atataactgt gacttctggg gagtacagcc 360
tctttcagaa ggataagcaa gaacagaata ttcctgtctc aaaaattatt acccatcctg 420
aatacaacag ccgtgaatat atgagtcctg atattgcact gctgtatcta aaacacaaag 480
tcaagtttgg aaatgctgtt cagccaatct gtcttcctga cagcgatgat aaagttgaac 540
caggaattct ttgcttatcc agtggatggg gcaagatttc caaaacatca gaatattcaa 600
atgtcctaca agaaatggaa ctcccatca tggatgacag agcgtgtaat actgtgctca 660
agagcatgaa cctccctccc ctgggaagga ccatgctgtg tgctggcttc cctgattggg 720
gaatggacgc ctgccagggg gactctggag gaccactggg ttgtagaaga ggtggtggaa 780
tctggattct tgctgggata acttcctggg tagctgggtg tgctggaggt tcagttcccg 840
taagaaacaa ccatgtgaag gcatcacttg gcattttctc caaagtgtct gagttgatgg 900
atthttatcac tcaaaacctg ttcacaggtt tggatcgggg ccaacccttc tcaaaagtgg 960
gctcaaggta tataacaaag gccctgaggt ctgtccaaga agtgaatgga agccagagag 1020
gaaagggtaa ggtctgtgga aaaatattgc cttcaccatt gctggcagag accagtgagg 1080
ccatggttcc atttgtttct gatacagaag acagtggcag tggctttgag cttaccgtta 1140
ctgctgtaca gaagtcagaa gcagggctcag gttgtgggag tctggctata ttggtagaag 1200
aagggacaaa tcaactctgc aagtatcctg atttgtatcc cagtaacaca aggtgtcatt 1260
ggttcatttg tgctccagag aagcacatta taaagttgac atttgaggac tttgctgtca 1320
aatttagtcc aaactgtatt tatgatgctg ttgtgattta cggtgattct gaagaaaagc 1380
acaagttagc taaactttgt ggaatgttga ccatcacttc aatattcagt tctagtaaca 1440
tgacgggtgat atacttttaa agtgatggta aaaatcgttt acaaggcttc aaggccagat 1500
ttaccatttt gccctcagag tcttttaaaca aatttgaacc aaagttacct ccccaaaaca 1560
atcctgtatc taccgtaaaa gctattctgc atgatgtctg tggcatccct ccatttagtc 1620
cccagtggct ttccagaaga atcgcaggag gggaagaagc ctgccccac tgttggccat 1680
ggcaggtggg tctgaggttt ctaggcgatt accaatgtgg aggtgccatc atcaaccag 1740
tgtggattct gaccgcagcc cactgtgtgc aattgaagaa taatccactc tcttgacta 1800
ttattgctgg ggaccatgac agaaacctga aggaatcaac agagcagaat tctacatcag 1860

```

ctcaggccaa	actgaatgac	ttcagctatg	ttggtacaga	gctacatctg	aacttaaata	1920
catttctaac	aacactttct	gcttacttca	tcatcgagct	ctctctgaat	gtttcttcat	1980
tagatggtgg	cctagcaagt	cgccacagc	agattcaagt	gcatgtgtta	gaaagagagg	2040
tctgtgaaca	cacttactat	tctgcccac	caggagggat	cacagagaag	atgatctgtg	2100
ctggcctttg	agcatctgga	gagaaagatt	tctgccaggg	agactctggg	gggccactag	2160
tatgtagaca	tgaaaatggg	ccctttgtcc	tctatggcat	tgtcagctgg	ggagctggct	2220
gtgtccagcc	atggaagccg	ggtgtatttg	ccagagtgat	gatcttcttg	gactggatcc	2280
aatcaaaaat	caatggtcct	gcttcacttc	agacaaaata	taaatgcaa	accttaaaac	2340
aacaattgcc	accacccaca	ccttcaccag	acagtgcac	ttggccaggt	ccaaaggaca	2400
gtaaaataac	cagactttcc	caaagttcaa	acagagagca	cttggtcctt	tgtgaggatg	2460
ttcttctgac	caagccagaa	gggatcatgc	agatccaag	aaattctcac	agaactacta	2520
tgggacacat	gaggattatg	gaagctacaa	ttcaaggatg	tccagtattg	gatttgattc	2580
cagtgacttc	tgttgagatc	acatctcttg	attatcctaa	cagttaa		2627

What is claimed is:

1. An isolated polypeptide selected from the group consisting of:
 - a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-18,
 - b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-18,
 - c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-18, and
 - d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-18.
2. An isolated polypeptide of claim 1 comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-18.
3. An isolated polynucleotide encoding a polypeptide of claim 1.
4. An isolated polynucleotide encoding a polypeptide of claim 2.
5. An isolated polynucleotide of claim 4 comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:19-36.
6. A recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide of claim 3.
7. A cell transformed with a recombinant polynucleotide of claim 6.
8. A transgenic organism comprising a recombinant polynucleotide of claim 6.
9. A method of producing a polypeptide of claim 1, the method comprising:
 - a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide, and said recombinant polynucleotide comprises a promoter sequence operably linked to a polynucleotide encoding the polypeptide of claim 1, and

- b) recovering the polypeptide so expressed.

10. A method of claim 9, wherein the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-18.

5

11. An isolated antibody which specifically binds to a polypeptide of claim 1.

12. An isolated polynucleotide selected from the group consisting of:

- 10 a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:19-36,
- b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:19-36,
- c) a polynucleotide complementary to a polynucleotide of a),
- 15 d) a polynucleotide complementary to a polynucleotide of b), and
- e) an RNA equivalent of a)-d).

13. An isolated polynucleotide comprising at least 60 contiguous nucleotides of a polynucleotide of claim 12.

20

14. A method of detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 12, the method comprising:

- 25 a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and
- b) detecting the presence or absence of said hybridization complex, and, optionally, if present, the amount thereof.

30

15. A method of claim 14, wherein the probe comprises at least 60 contiguous nucleotides.

16. A method of detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 12, the method comprising:

- a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and
- b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.

5

17. A composition comprising a polypeptide of claim 1 and a pharmaceutically acceptable excipient.

18. A composition of claim 17, wherein the polypeptide comprises an amino acid sequence
10 selected from the group consisting of SEQ ID NO:1-18.

19. A method for treating a disease or condition associated with decreased expression of functional PMMM, comprising administering to a patient in need of such treatment the composition of claim 17.

15

20. A method of screening a compound for effectiveness as an agonist of a polypeptide of claim 1, the method comprising:

- a) exposing a sample comprising a polypeptide of claim 1 to a compound, and
- b) detecting agonist activity in the sample.

20

21. A composition comprising an agonist compound identified by a method of claim 20 and a pharmaceutically acceptable excipient.

22. A method for treating a disease or condition associated with decreased expression of
25 functional PMMM, comprising administering to a patient in need of such treatment a composition of claim 21.

23. A method of screening a compound for effectiveness as an antagonist of a polypeptide of claim 1, the method comprising:

- 30 a) exposing a sample comprising a polypeptide of claim 1 to a compound, and
- b) detecting antagonist activity in the sample.

24. A composition comprising an antagonist compound identified by a method of claim 23 and a pharmaceutically acceptable excipient.

35

25. A method for treating a disease or condition associated with overexpression of functional PMMM, comprising administering to a patient in need of such treatment a composition of claim 24.

5 26. A method of screening for a compound that specifically binds to the polypeptide of claim 1, the method comprising:

- a) combining the polypeptide of claim 1 with at least one test compound under suitable conditions, and
- b) detecting binding of the polypeptide of claim 1 to the test compound, thereby
10 identifying a compound that specifically binds to the polypeptide of claim 1.

27. A method of screening for a compound that modulates the activity of the polypeptide of claim 1, the method comprising:

- a) combining the polypeptide of claim 1 with at least one test compound under
15 conditions permissive for the activity of the polypeptide of claim 1,
- b) assessing the activity of the polypeptide of claim 1 in the presence of the test compound, and
- c) comparing the activity of the polypeptide of claim 1 in the presence of the test
20 compound with the activity of the polypeptide of claim 1 in the absence of the test compound, wherein a change in the activity of the polypeptide of claim 1 in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide of claim 1.

28. A method of screening a compound for effectiveness in altering expression of a target
25 polynucleotide, wherein said target polynucleotide comprises a sequence of claim 5, the method comprising:

- a) exposing a sample comprising the target polynucleotide to a compound, under conditions suitable for the expression of the target polynucleotide,
- b) detecting altered expression of the target polynucleotide, and
- c) comparing the expression of the target polynucleotide in the presence of varying
30 amounts of the compound and in the absence of the compound.

29. A method of assessing toxicity of a test compound, the method comprising:

- a) treating a biological sample containing nucleic acids with the test compound,

- 5 b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide of claim 12 under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide comprising a polynucleotide sequence of a polynucleotide of claim 12 or fragment thereof,
- c) quantifying the amount of hybridization complex, and
- 10 d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.

30. A diagnostic test for a condition or disease associated with the expression of PMMM in a biological sample, the method comprising:

- 15 a) combining the biological sample with an antibody of claim 11, under conditions suitable for the antibody to bind the polypeptide and form an antibody:polypeptide complex, and
- b) detecting the complex, wherein the presence of the complex correlates with the presence of the polypeptide in the biological sample.

20

31. The antibody of claim 11, wherein the antibody is:

- a) a chimeric antibody,
- b) a single chain antibody,
- c) a Fab fragment,
- 25 d) a F(ab')₂ fragment, or
- e) a humanized antibody.

32. A composition comprising an antibody of claim 11 and an acceptable excipient.

30 33. A method of diagnosing a condition or disease associated with the expression of PMMM in a subject, comprising administering to said subject an effective amount of the composition of claim 32.

34. A composition of claim 32, wherein the antibody is labeled.

35

35. A method of diagnosing a condition or disease associated with the expression of PMMM in a subject, comprising administering to said subject an effective amount of the composition of claim 34.

5 36. A method of preparing a polyclonal antibody with the specificity of the antibody of claim 11, the method comprising:

- a) immunizing an animal with a polypeptide consisting of an amino acid sequence selected from the group consisting of SEQ ID NO:1-18, or an immunogenic fragment thereof, under conditions to elicit an antibody response,
- 10 b) isolating antibodies from said animal, and
- c) screening the isolated antibodies with the polypeptide, thereby identifying a polyclonal antibody which binds specifically to a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-18.

15 37. A polyclonal antibody produced by a method of claim 36.

38. A composition comprising the polyclonal antibody of claim 37 and a suitable carrier.

20 39. A method of making a monoclonal antibody with the specificity of the antibody of claim 11, the method comprising:

- a) immunizing an animal with a polypeptide consisting of an amino acid sequence selected from the group consisting of SEQ ID NO:1-18, or an immunogenic fragment thereof, under conditions to elicit an antibody response,
- b) isolating antibody producing cells from the animal,
- 25 c) fusing the antibody producing cells with immortalized cells to form monoclonal antibody-producing hybridoma cells,
- d) culturing the hybridoma cells, and
- e) isolating from the culture monoclonal antibody which binds specifically to a polypeptide comprising an amino acid sequence selected from the group consisting
- 30 of SEQ ID NO:1-18.

40. A monoclonal antibody produced by a method of claim 39.

41. A composition comprising the monoclonal antibody of claim 40 and a suitable carrier.

42. The antibody of claim 11, wherein the antibody is produced by screening a Fab expression library.

43. The antibody of claim 11, wherein the antibody is produced by screening a recombinant
5 immunoglobulin library.

44. A method of detecting a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-18 in a sample, the method comprising:

- 10 a) incubating the antibody of claim 11 with a sample under conditions to allow specific binding of the antibody and the polypeptide, and
- b) detecting specific binding, wherein specific binding indicates the presence of a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-18 in the sample.

15 45. A method of purifying a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-18 from a sample, the method comprising:

- a) incubating the antibody of claim 11 with a sample under conditions to allow specific binding of the antibody and the polypeptide, and
- 20 b) separating the antibody from the sample and obtaining the purified polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-18.

46. A microarray wherein at least one element of the microarray is a polynucleotide of claim 13.

25

47. A method of generating an expression profile of a sample which contains polynucleotides, the method comprising:

- a) labeling the polynucleotides of the sample,
- 30 b) contacting the elements of the microarray of claim 46 with the labeled polynucleotides of the sample under conditions suitable for the formation of a hybridization complex, and
- c) quantifying the expression of the polynucleotides in the sample.

48. An array comprising different nucleotide molecules affixed in distinct physical
35 locations on a solid substrate, wherein at least one of said nucleotide molecules comprises a first

oligonucleotide or polynucleotide sequence specifically hybridizable with at least 30 contiguous nucleotides of a target polynucleotide, and wherein said target polynucleotide is a polynucleotide of claim 12.

5 49. An array of claim 48, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to at least 30 contiguous nucleotides of said target polynucleotide.

 50. An array of claim 48, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to at least 60 contiguous nucleotides of said target polynucleotide.

10

 51. An array of claim 48, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to said target polynucleotide.

 52. An array of claim 48, which is a microarray.

15

 53. An array of claim 48, further comprising said target polynucleotide hybridized to a nucleotide molecule comprising said first oligonucleotide or polynucleotide sequence.

 54. An array of claim 48, wherein a linker joins at least one of said nucleotide molecules to
20 said solid substrate.

 55. An array of claim 48, wherein each distinct physical location on the substrate contains multiple nucleotide molecules, and the multiple nucleotide molecules at any single distinct physical location have the same sequence, and each distinct physical location on the substrate contains
25 nucleotide molecules having a sequence which differs from the sequence of nucleotide molecules at another distinct physical location on the substrate.

 56. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:1.

30 57. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:2.

 58. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:3.

 59. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:4.

35

60. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:5.
61. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:6.
- 5 62. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:7.
63. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:8.
64. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:9.
- 10 65. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:10.
66. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:11.
- 15 67. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:12.
68. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:13.
69. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:14.
- 20 70. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:15.
71. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:16.
- 25 72. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:17.
73. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:18.
74. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
30 NO:19.
75. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
NO:20.

76. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
NO:21.

77. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
5 NO:22.

78. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
NO:23.

10 79. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
NO:24.

80. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
NO:25.
15

81. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
NO:26.

82. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
20 NO:27.

83. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
NO:28.

25 84. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
NO:29.

85. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
NO:30.
30

86. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
NO:31.

87. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
35 NO:32.

88. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:33.

89. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
5 NO:34.

90. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
NO:35.

10 91. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
NO:36.