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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
11 October 2001 (11.10.2001)

PCT

(10) International Publication Number
WO 01/75075 A2

(51) International Patent Classification[7]: C12N 9/00

(21) International Application Number: PCT/EP01/03395

(22) International Filing Date: 26 March 2001 (26.03.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/192,315 27 March 2000 (27.03.2000) US
60/210,735 12 June 2000 (12.06.2000) US

(71) Applicant *(for all designated States except US)*: **BAYER AKTIENGESELLSCHAFT** [DE/DE]; 51368 Leverkusen (DE).

(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: **XIAO, Yonghong** [US/US]; 78 Dana Street, Apt. #1, Cambridge, MA 02138 (US).

(74) Common Representative: **BAYER AKTIENGESELLSCHAFT**; 51368 Leverkusen (DE).

(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, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

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

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

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

WO 01/75075 A2

(54) Title: REGULATION OF HUMAN TYROSINE PHOSPHATASE-LIKE ENZYME

(57) Abstract: Human tyrosine phosphatase-like enzyme can be regulated to treat or prevent diseases such as rheumatoid arthritis and osteoarthritis and carcinomas of the intestine, bladder, prostate, breast, stomach, and brain.

REGULATION OF HUMAN TYROSINE PHOSPHATASE-LIKE ENZYME

TECHNICAL FIELD OF THE INVENTION

The invention relates to the area of human tyrosine phosphatase-like enzymes and their regulation.

BACKGROUND OF THE INVENTION

Tyrosine phosphatase

Phosphatases remove phosphate groups from molecules previously activated by kinases and control most cellular signaling events that regulate cell growth and differentiation, cell-to-cell contacts, the cell cycle and oncogenesis. Protein phosphorylation is the ubiquitous strategy used to control the activities of eukaryotic cells. It is estimated that more than 1000 of the 10,000 proteins active in a typical mammalian cell are phosphorylated. The high energy phosphate which confers activation is transferred from adenosine triphosphate molecules to a protein by protein kinases, and is subsequently removed from the protein by protein phosphatases. *See* U.S. Patent 6,020,179.

There appear to be three, evolutionarily-distinct protein phosphatase gene families (Carbonneau & Tonks, *Ann. Rev. Cell Biol. 8*, 463-93, 1992). They are the protein phosphatases (PP), the protein tyrosine phosphatases (PTP) and the acid/alkaline phosphatases (AP). PTPs remove phosphate groups only from selected phosphotyrosines on particular types of proteins. In so doing, PTPs reverse the effects of protein tyrosine kinases (PTK) and therefore lay a significant role in cell cycle and cell signaling processes. PTPs possess a high specific enzyme activity relative to their PTK counterparts and therefore ensure that tyrosine phosphorylations are very short lived and very uncommon in resting cells. Many PTKs are encoded by oncogenes, and it is well known that oncogenesis is often accompanied by increased tyrosine phos-

phorylation activity. It is therefore possible that PTPs may serve to prevent or reverse cell transformation and the growth of various cancers by controlling the levels of tyrosine phosphorylation in cells. This is supported by studies showing that overexpression of PTP can suppress transformation in cells and, conversely, specific inhibition of PTP can enhance cell transformation.

The PTPs are found in transmembrane, receptor-like, and nontransmembrane, non-receptor forms, and possess a diversity in size (20 kDa to greater than 100 kDa) and structure. All PTPs share homology within a region of 240 residues which delineates the catalytic domain and contain a conserved sequence near the carboxy terminus. The combination of the catalytic domain with a wide variety of structural motifs accounts for the diversity and specificity of these enzymes. In the nonreceptor isoforms, the noncatalytic sequences may also confer different modes of regulation and target PTPs to various intracellular compartments.

Receptor-like PTPs (R-PTPs) are generally large (greater than 100 kDa) and are grouped on the basis of their single transmembrane segment and two, tandem PTP domains within the cytoplasmic tail. In contrast to the similarity within the internal cytoplasmic domains of these molecules, there is considerable diversity among the extracellular segment. Key examples of this type of PTP are CD45, a PTP found on the surface of leucocytes that helps to activate T and B lymphocytes when activated by extracellular antibodies and LAR, a PTP having structural features related to the N-CAM family of cell adhesion molecules on its extracellular domain and which may be involved in cell adhesion processes.

Nonreceptor PTPs (NR-PTP) are generally smaller (about 50 kDa) than the R-PTPs and have single catalytic domains and noncatalytic sequences of variable length positioned at either the N- or C-termini. NR-PTPs are intracellular and may use their noncatalytic sequences to direct them to particular subcellular compartments or to determine their enzyme regulating activity. Some NR-PTPs may be divided into subfamilies based on similarities in their noncatalytic domains. For example, human

PTPH1 and MEG01 contain homologous catalytic domains and N-terminal segments with homology to band 4.1, talin and ezrin. In addition they are thought to be localized between actin stress fibers and the plasma membrane where they modulate cytoskeletal dynamics. T-cell PTP and PTP1β display a high degree of similarity in their catalytic domains and structural similarities in their C-terminal noncatalytic domains that may help direct them to membranes where they regulate enzyme activity.

Recently, a new class of smaller (about 20 kDa) NR-PTPs has been found which have a single catalytic domain and are represented by PRL-1, found in regenerating rat liver and hepatoma cells, and OV-1, found in human ovarian tissue (Diamond *et al.*, *Mol. Cell. Biol. 14*, 3752-62, 1994; Montagna *et al.*, *Hum. Genet. 96*, 532-38, 1995). These PTPs possess homology to other NR-PTPs only within the region of the catalytic active site. Stably transfected cells that overexpress PRL-1 exhibit altered cell growth and morphology and a transformed phenotype. Furthermore, it is postulated that PRL-1 is important in the control of normal cell growth and in the development of tumorigenicity.

It is apparent that PTPs may serve either as positive or negative regulators of cell growth, and that a detailed understanding of phosphatase interaction in signal transduction pathways should reveal many potential mechanisms to provide the means for clinical diagnosis or therapeutic intervention in the progression of cancer, inflammatory illnesses, or oncogenesis. Thus, there is a need in the art to identify new PTPs which can be regulated to prevent or treat PTP-associated diseases.

SUMMARY OF THE INVENTION

It is an object of the invention to provide reagents and methods of regulating a human tyrosine phosphatase-like enzyme. This and other objects of the invention are provided by one or more of the embodiments described below.

One embodiment of the invention is a tyrosine phosphatase-like enzyme polypeptide comprising an amino acid sequence selected from the group consisting of:

amino acid sequences which are at least about 50% identical to the amino acid sequence shown in SEQ ID NO: 2;

the amino acid sequence shown in SEQ ID NO: 2;

amino acid sequences which are at least about 50% identical to the amino acid sequence shown in SEQ ID NO: 4;

the amino acid sequence shown in SEQ ID NO:4;

amino acid sequences which are at least about 50% identical to the amino acid sequence shown in SEQ ID NO: 6;

the amino acid sequence shown in SEQ ID NO: 6;

amino acid sequences which are at least about 50% identical to the amino acid sequence shown in SEQ ID NO: 7;

the amino acid sequence shown in SEQ ID NO: 7;

amino acid sequences which are at least about 50% identical to the amino acid sequence shown in SEQ ID NO: 8;

the amino acid sequence shown in SEQ ID NO:8;

amino acid sequences which are at least about 50% identical to the amino acid sequence shown in SEQ ID NO: 10; and

the amino acid sequence shown in SEQ ID NO: 10.

Yet another embodiment of the invention is a method of screening for agents which decrease the activity of tyrosine phosphatase-like enzyme. A test compound is contacted with a tyrosine phosphatase-like enzyme polypeptide comprising an amino acid sequence selected from the group consisting of:

amino acid sequences which are at least about 50% identical to the amino acid sequence shown in SEQ ID NO: 2;

the amino acid sequence shown in SEQ ID NO: 2;

amino acid sequences which are at least about 50% identical to the amino acid sequence shown in SEQ ID NO: 4;

the amino acid sequence shown in SEQ ID NO:4;

amino acid sequences which are at least about 50% identical to the amino acid sequence shown in SEQ ID NO: 6;

the amino acid sequence shown in SEQ ID NO: 6;

amino acid sequences which are at least about 50% identical to the amino acid sequence shown in SEQ ID NO: 7;

the amino acid sequence shown in SEQ ID NO: 7;

amino acid sequences which are at least about 50% identical to the amino acid sequence shown in SEQ ID NO: 8;

the amino acid sequence shown in SEQ ID NO:8;

amino acid sequences which are at least about 50% identical to the amino acid sequence shown in SEQ ID NO: 10; and

the amino acid sequence shown in SEQ ID NO: 10.

Binding between the test compound and the tyrosine phosphatase-like enzyme polypeptide is detected. A test compound which binds to the tyrosine phosphatase-like enzyme polypeptide is thereby identified as a potential agent for decreasing the activity of tyrosine phosphatase-like enzyme.

Another embodiment of the invention is a method of screening for agents which decrease the activity of tyrosine phosphatase-like enzyme. A test compound is contacted with a polynucleotide encoding a tyrosine phosphatase-like enzyme polypeptide, wherein the polynucleotide comprises a nucleotide sequence selected from the group consisting of:

nucleotide sequences which are at least about 50% identical to the nucleotide sequence shown in SEQ ID NO: 1;

the nucleotide sequence shown in SEQ ID NO: 1;

nucleotide sequences which are at least about 50% identical to the nucleotide sequence shown in SEQ ID NO: 3;

the nucleotide sequence shown in SEQ ID NO: 3;

nucleotide sequences which are at least about 50% identical to the nucleotide sequence shown in SEQ ID NO: 5;

the nucleotide sequence shown in SEQ ID NO:5.

nucleotide sequences which are at least about 50% identical to the nucleotide sequence shown in SEQ ID NO: 9; and

the nucleotide sequence shown in SEQ ID NO: 9.

Binding of the test compound to the polynucleotide is detected. A test compound which binds to the polynucleotide is identified as a potential agent for decreasing the activity of tyrosine phosphatase-like enzyme. The agent can work by decreasing the amount of the tyrosine phosphatase-like enzyme through interacting with the tyrosine phosphatase-like enzyme mRNA.

Another embodiment of the invention is a method of screening for agents which regulate the activity of tyrosine phosphatase-like enzyme. A test compound is contacted with a tyrosine phosphatase-like enzyme polypeptide comprising an amino acid sequence selected from the group consisting of:

amino acid sequences which are at least about 50% identical to the amino acid sequence shown in SEQ ID NO: 2;

the amino acid sequence shown in SEQ ID NO: 2;

amino acid sequences which are at least about 50% identical to the amino acid sequence shown in SEQ ID NO: 4;

the amino acid sequence shown in SEQ ID NO:4;

amino acid sequences which are at least about 50% identical to the amino acid sequence shown in SEQ ID NO: 6;

the amino acid sequence shown in SEQ ID NO: 6;

amino acid sequences which are at least about 50% identical to the amino acid sequence shown in SEQ ID NO: 7;

the amino acid sequence shown in SEQ ID NO: 7;

amino acid sequences which are at least about 50% identical to the amino acid sequence shown in SEQ ID NO: 8;

the amino acid sequence shown in SEQ ID NO:8;

amino acid sequences which are at least about 50% identical to the amino acid sequence shown in SEQ ID NO: 10; and

the amino acid sequence shown in SEQ ID NO: 10.

A tyrosine phosphatase-like enzyme activity of the polypeptide is detected. A test compound which increases tyrosine phosphatase-like enzyme activity of the polypeptide relative to tyrosine phosphatase-like enzyme activity in the absence of the test compound is thereby identified as a potential agent for increasing the activity of tyrosine phosphatase-like enzyme. A test compound which decreases tyrosine phosphatase-like enzyme activity of the polypeptide relative to tyrosine phosphatase-like enzyme activity in the absence of the test compound is thereby identified as a potential agent for decreasing the activity of tyrosine phosphatase-like enzyme.

Even another embodiment of the invention is a method of screening for agents which decrease the activity of tyrosine phosphatase-like enzyme. A test compound is contacted with a tyrosine phosphatase-like enzyme product of a polynucleotide which comprises a nucleotide sequence selected from the group consisting of:

nucleotide sequences which are at least about 50% identical to the nucleotide sequence shown in SEQ ID NO: 1;

the nucleotide sequence shown in SEQ ID NO: 1;

nucleotide sequences which are at least about 50% identical to the nucleotide sequence shown in SEQ ID NO: 3;

the nucleotide sequence shown in SEQ ID NO: 3;

nucleotide sequences which are at least about 50% identical to the nucleotide sequence shown in SEQ ID NO: 5;

the nucleotide sequence shown in SEQ ID NO:5;

nucleotide sequences which are at least about 50% identical to the nucleotide sequence shown in SEQ ID NO: 9; and

the nucleotide sequence shown in SEQ ID NO: 9.

Binding of the test compound to the tyrosine phosphatase-like enzyme product is detected. A test compound which binds to the tyrosine phosphatase-like enzyme product is thereby identified as a potential agent for decreasing the activity of tyrosine phosphatase-like enzyme.

Still another embodiment of the invention is a method of reducing the activity of tyrosine phosphatase-like enzyme. A cell is contacted with a reagent which specifically binds to a polynucleotide encoding a tyrosine phosphatase-like enzyme polypeptide or the product encoded by the polynucleotide, wherein the polynucleotide comprises a nucleotide sequence selected from the group consisting of:

nucleotide sequences which are at least about 50% identical to the nucleotide sequence shown in SEQ ID NO: 1;

the nucleotide sequence shown in SEQ ID NO: 1;

nucleotide sequences which are at least about 50% identical to the nucleotide sequence shown in SEQ ID NO: 3;

the nucleotide sequence shown in SEQ ID NO: 3;

nucleotide sequences which are at least about 50% identical to the nucleotide sequence shown in SEQ ID NO: 5;

the nucleotide sequence shown in SEQ ID NO:5;

nucleotide sequences which are at least about 50% identical to the nucleotide sequence shown in SEQ ID NO: 9; and

the nucleotide sequence shown in SEQ ID NO: 9.

Tyrosine phosphatase-like enzyme activity in the cell is thereby decreased.

The invention thus provides a human tyrosine phosphatase-like enzyme which can be regulated to provide therapeutic effects.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 shows the DNA-sequence encoding a tyrosine phosphatase-like enzyme polypeptide.

Fig. 2 shows the amino acid sequence deduced from the DNA-sequence of Fig.1.

Fig. 3 shows the DNA-sequence encoding a tyrosine phosphatase-like enzyme polypeptide.

Fig. 4 shows the amino acid sequence deduced from the DNA-sequence of Fig. 3.

Fig. 5 shows the DNA-sequence encoding a tyrosine phosphatase-like enzyme polypeptide.

Fig. 6 shows the amino acid sequence deduced from the DNA-sequence of Fig. 5.

Fig. 7 shows a specific part of the amino acid sequence of Fig. 6.

Fig. 8 shows a specific part of the amino acid sequence of Fig. 6.

Fig. 9 shows the DNA-sequence encoding a tyrosine phosphatase-like enzyme polypeptide.

Fig. 10 shows the amino acid sequence deduced from the DNA-sequence of Fig. 9.

DETAILED DESCRIPTION OF THE INVENTION

The invention relates to an isolated polynucleotide encoding a tyrosine phosphatase-like enzyme polypeptide and being selected from the group consisting of:

a) a polynucleotide encoding a haparanase-like enzyme polypeptide comprising an amino acid sequence selected from the group consisting of:

amino acid sequences which are at least about 50% identical to the amino acid sequence shown in SEQ ID NO: 2;
the amino acid sequence shown in SEQ ID NO: 2;
amino acid sequences which are at least about 50% identical to

the amino acid sequence shown in SEQ ID NO: 4;
the amino acid sequence shown in SEQ ID NO:4;
amino acid sequences which are at least about 50% identical to
the amino acid sequence shown in SEQ ID NO: 6;
the amino acid sequence shown in SEQ ID NO: 6;
amino acid sequences which are at least about 50% identical to
the amino acid sequence shown in SEQ ID NO: 7;
the amino acid sequence shown in SEQ ID NO: 7;
amino acid sequences which are at least about 50% identical to
the amino acid sequence shown in SEQ ID NO: 8;
the amino acid sequence shown in SEQ ID NO:8;
amino acid sequences which are at least about 50% identical to
the amino acid sequence shown in SEQ ID NO: 10; and
the amino acid sequence shown in SEQ ID NO: 10.

b) a polynucleotide comprising the sequence of SEQ ID NOS: 1, 3, 5 or 9;

c) a polynucleotide which hybridizes under stringent conditions to a polynucleotide specified in (a) and (b);

d) a polynucleotide the sequence of which deviates from the polynucleotide sequences specified in (a) to (c) due to the degeneration of the genetic code; and

e) a polynucleotide which represents a fragment, derivative or allelic variation of a polynucleotide sequence specified in (a) to (d).

Furthermore, it has been discovered by the present applicant that a tyrosine phosphatase-like enzyme, particularly a human tyrosine phosphatase-like enzyme, can be regulated to treat disorders such as rheumatoid arthritis, osteoarthritis, and various carcinomas. Human tyrosine phosphatase-like enzyme also can be used to screen for tyrosine phosphatase-like enzyme agonists and antagonists.

Tyrosine Phosphatase-Like Enzyme Polypeptides

Tyrosine phosphatase-like enzyme polypeptides according to the invention comprise an amino acid sequence shown in SEQ ID NO:2, 4, 6, 7, or 9, a portion of one of those sequences, or a biologically active variant thereof, as defined below. A tyrosine phosphatase-like enzyme polypeptide therefore can be a portion of a tyrosine phosphatase-like enzyme, a full-length tyrosine phosphatase-like enzyme, or a fusion protein comprising all or a portion of a tyrosine phosphatase-like enzyme. Human tyrosine phosphatase-like enzyme according to the invention comprises a B41 domain at amino acids 33 to 187 (SEQ ID NO:7).

Biologically Active Variants

Tyrosine phosphatase-like enzyme polypeptide variants preferably are biologically active, *i.e.*, retain a tyrosine phosphatase activity. Tyrosine phosphatase activity can be measured, for example, as described in the specific examples, below. Preferably, naturally or non-naturally occurring tyrosine phosphatase-like enzyme polypeptide variants have amino acid sequences which are at least about 50, preferably about 75, 90, 96, or 98% identical to an amino acid sequence shown in SEQ ID NO:2, 4, 6, 7, or 9 or a fragment thereof. Percent identity between a putative tyrosine phosphatase-like enzyme polypeptide variant and an amino acid sequence of SEQ ID NO:2, 4, 6, 7, or 9 is determined using the Blast2 alignment program.

Variations in percent identity can be due, for example, to amino acid substitutions, insertions, or deletions. Amino acid substitutions are defined as one for one amino acid replacements. They are conservative in nature when the substituted amino acid has similar structural and/or chemical properties. Examples of conservative replacements are substitution of a leucine with an isoleucine or valine, an aspartate with a glutamate, or a threonine with a serine.

Amino acid insertions or deletions are changes to or within an amino acid sequence. They typically fall in the range of about 1 to 5 amino acids. Guidance in determining which amino acid residues can be substituted, inserted, or deleted without abolishing biological or immunological activity of a tyrosine phosphatase-like enzyme polypeptide can be found using computer programs well known in the art, such as DNASTAR software. Whether an amino acid change results in a biologically active tyrosine phosphatase-like enzyme polypeptide can readily be determined by assaying for tyrosine phosphatase activity, as described for example, in the specific examples, below.

Fusion Proteins

Fusion proteins are useful for generating antibodies against tyrosine phosphatase-like enzyme polypeptide amino acid sequences and for use in various assay systems. For example, fusion proteins can be used to identify proteins which interact with portions of a tyrosine phosphatase-like enzyme polypeptide. Protein affinity chromatography or library-based assays for protein-protein interactions, such as the yeast two-hybrid or phage display systems, can be used for this purpose. Such methods are well known in the art and also can be used as drug screens.

A tyrosine phosphatase-like enzyme fusion protein comprises two polypeptide segments fused together by means of a peptide bond. The first polypeptide segment comprises at least 5, 6, 8, 10, 25, or 50 or more contiguous amino acids of SEQ ID NO:2, 4, 6, 7, or 9 or from a biologically active variant, such as those described above. The first polypeptide segment also can comprise full-length tyrosine phosphatase-like enzyme protein.

The second polypeptide segment can be a full-length protein or a protein fragment. Proteins commonly used in fusion protein construction include β-galactosidase, β-glucuronidase, green fluorescent protein (GFP), autofluorescent proteins, including blue fluorescent protein (BFP), glutathione-S-transferase (GST), luciferase, horse-

radish peroxidase (HRP), and chloramphenicol acetyltransferase (CAT). Additionally, epitope tags are used in fusion protein constructions, including histidine (His) tags, FLAG tags, influenza hemagglutinin (HA) tags, Myc tags, VSV-G tags, and thioredoxin (Trx) tags. Other fusion constructions can include maltose binding protein (MBP), S-tag, Lex a DNA binding domain (DBD) fusions, GAL4 DNA binding domain fusions, and herpes simplex virus (HSV) BP16 protein fusions. A fusion protein also can be engineered to contain a cleavage site located between the tyrosine phosphatase-like enzyme polypeptide-encoding sequence and the heterologous protein sequence, so that the tyrosine phosphatase-like enzyme polypeptide can be cleaved and purified away from the heterologous moiety.

A fusion protein can be synthesized chemically, as is known in the art. Preferably, a fusion protein is produced by covalently linking two polypeptide segments or by standard procedures in the art of molecular biology. Recombinant DNA methods can be used to prepare fusion proteins, for example, by making a DNA construct which comprises a coding sequence selected from SEQ ID NOS:1, 3, 5, or 7 in proper reading frame with nucleotides encoding the second polypeptide segment and expressing the DNA construct in a host cell, as is known in the art. Many kits for constructing fusion proteins are available from companies such as Promega Corporation (Madison, WI), Stratagene (La Jolla, CA), CLONTECH (Mountain View, CA), Santa Cruz Biotechnology (Santa Cruz, CA), MBL International Corporation (MIC; Watertown, MA), and Quantum Biotechnologies (Montreal, Canada; 1-888-DNA-KITS).

Identification of Species Homologs

Species homologs of human tyrosine phosphatase-like enzyme polypeptide can be obtained using tyrosine phosphatase-like enzyme polypeptide polynucleotides (described below) to make suitable probes or primers for screening cDNA expression libraries from other species, such as mice, monkeys, or yeast, identifying cDNAs

which encode homologs of tyrosine phosphatase-like enzyme polypeptide, and expressing the cDNAs as is known in the art.

Tyrosine Phosphatase-Like Enzyme Polynucleotides

A tyrosine phosphatase-like enzyme polynucleotide can be single- or double-stranded and comprises a coding sequence or the complement of a coding sequence for a tyrosine phosphatase-like enzyme polypeptide. Nucleotide sequences encoding human tyrosine phosphatase-like enzyme polypeptides are shown in SEQ ID NOS:1, 3, and 5.

Degenerate nucleotide sequences encoding human tyrosine phosphatase-like enzyme polypeptides, as well as homologous nucleotide sequences which are at least about 50, preferably about 75, 90, 96, or 98% identical to a nucleotide sequence selected from the group consisting of SEQ ID NOS:1, 3, 5, or 7 also are tyrosine phosphatase-like enzyme polynucleotides. Percent sequence identity between the sequences of two polynucleotides is determined using computer programs such as ALIGN which employ the FASTA algorithm, using an affine gap search with a gap open penalty of -12 and a gap extension penalty of -2. Complementary DNA (cDNA) molecules, species homologs, and variants of tyrosine phosphatase-like enzyme polynucleotides which encode biologically active tyrosine phosphatase-like enzyme polypeptides also are tyrosine phosphatase-like enzyme polynucleotides.

Identification of Variants and Homologs of Tyrosine Phosphatase-Like Enzyme Polynucleotides

Variants and homologs of the tyrosine phosphatase-like enzyme polynucleotides described above also are tyrosine phosphatase-like enzyme polynucleotides. Typically, homologous tyrosine phosphatase-like enzyme polynucleotide sequences can be identified by hybridization of candidate polynucleotides to known tyrosine phos-

phatase-like enzyme polynucleotides under stringent conditions, as is known in the art. For example, using the following wash conditions--2X SSC (0.3 M NaCl, 0.03 M sodium citrate, pH 7.0), 0.1% SDS, room temperature twice, 30 minutes each; then 2X SSC, 0.1% SDS, 50 °C once, 30 minutes; then 2X SSC, room temperature twice, 10 minutes each--homologous sequences can be identified which contain at most about 25-30% basepair mismatches. More preferably, homologous nucleic acid strands contain 15-25% basepair mismatches, even more preferably 5-15% basepair mismatches.

Species homologs of the tyrosine phosphatase-like enzyme polynucleotides disclosed herein also can be identified by making suitable probes or primers and screening cDNA expression libraries from other species, such as mice, monkeys, or yeast. Human variants of tyrosine phosphatase-like enzyme polynucleotides can be identified, for example, by screening human cDNA expression libraries. It is well known that the T_m of a double-stranded DNA decreases by 1-1.5 °C with every 1% decrease in homology (Bonner *et al.*, *J. Mol. Biol. 81*, 123 (1973). Variants of human tyrosine phosphatase-like enzyme polynucleotides or tyrosine phosphatase-like enzyme polynucleotides of other species can therefore be identified by hybridizing a putative homologous tyrosine phosphatase-like enzyme polynucleotide with a polynucleotide having a nucleotide sequence of SEQ ID NO:1, 3, 5, or 7 or the complement thereof to form a test hybrid. The melting temperature of the test hybrid is compared with the melting temperature of a hybrid comprising tyrosine phosphatase polynucleotides having perfectly complementary nucleotide sequences, and the number or percent of basepair mismatches within the test hybrid is calculated.

Nucleotide sequences which hybridize to tyrosine phosphatase polynucleotides or their complements following stringent hybridization and/or wash conditions also are tyrosine phosphatase-like enzyme polynucleotides. Stringent wash conditions are well known and understood in the art and are disclosed, for example, in Sambrook *et al.*, MOLECULAR CLONING: A LABORATORY MANUAL, 2d ed., 1989, at pages 9.50-9.51.

Typically, for stringent hybridization conditions a combination of temperature and salt concentration should be chosen that is approximately 12-20 °C below the calculated T_m of the hybrid under study. The T_m of a hybrid between a tyrosine phosphatase-like enzyme polynucleotide having a nucleotide sequence shown in SEQ ID NO:1, 3, 5, or 7 or the complement thereof and a polynucleotide sequence which is at least about 50, preferably about 75, 90, 96, or 98% identical to one of those nucleotide sequences can be calculated, for example, using the equation of Bolton and McCarthy, *Proc. Natl. Acad. Sci. U.S.A. 48*, 1390 (1962):

$$T_m = 81.5\ ^\circ C - 16.6(\log_{10}[Na^+]) + 0.41(\%G + C) - 0.63(\%\text{formamide}) - 600/l),$$

where l = the length of the hybrid in basepairs.

Stringent wash conditions include, for example, 4X SSC at 65 °C, or 50% formamide, 4X SSC at 42 °C, or 0.5X SSC, 0.1% SDS at 65 °C. Highly stringent wash conditions include, for example, 0.2X SSC at 65 °C.

Preparation of Tyrosine Phosphatase-Like Enzyme Polynucleotides

A naturally occurring tyrosine phosphatase-like enzyme polynucleotide can be isolated free of other cellular components such as membrane components, proteins, and lipids. Polynucleotides can be made by a cell and isolated using standard nucleic acid purification techniques, or synthesized using an amplification technique, such as the polymerase chain reaction (PCR), or by using an automatic synthesizer. Methods for isolating polynucleotides are routine and are known in the art. Any such technique for obtaining a polynucleotide can be used to obtain isolated tyrosine phosphatase-like enzyme polynucleotides. For example, restriction enzymes and probes can be used to isolate polynucleotide fragments which comprises tyrosine phosphatase-like enzyme nucleotide sequences. Isolated polynucleotides are in preparations which are free or at least 70, 80, or 90% free of other molecules.

Tyrosine phosphatase-like enzyme cDNA molecules can be made with standard molecular biology techniques, using tyrosine phosphatase-like enzyme mRNA as a template. tyrosine phosphatase-like enzyme cDNA molecules can thereafter be replicated using molecular biology techniques known in the art and disclosed in manuals such as Sambrook *et al.* (1989). An amplification technique, such as PCR, can be used to obtain additional copies of polynucleotides of the invention, using either human genomic DNA or cDNA as a template.

Alternatively, synthetic chemistry techniques can be used to synthesizes tyrosine phosphatase-like enzyme polynucleotides. The degeneracy of the genetic code allows alternate nucleotide sequences to be synthesized which will encode a tyrosine phosphatase-like enzyme polypeptide having, for example, an amino acid sequence shown in SEQ ID NO:2, 4, 6, 7, or 9 or a biologically active variant thereof.

Extending Tyrosine Phosphatase-Like Enzyme Polynucleotides

Various PCR-based methods can be used to extend the nucleic acid sequences disclosed herein to detect upstream sequences such as promoters and regulatory elements. For example, restriction-site PCR uses universal primers to retrieve unknown sequence adjacent to a known locus (Sarkar, *PCR Methods Applic. 2*, 318-322, 1993). Genomic DNA is first amplified in the presence of a primer to a linker sequence and a primer specific to the known region. The amplified sequences are then subjected to a second round of PCR with the same linker primer and another specific primer internal to the first one. Products of each round of PCR are transcribed with an appropriate RNA polymerase and sequenced using reverse transcriptase.

Inverse PCR also can be used to amplify or extend sequences using divergent primers based on a known region (Triglia *et al.*, *Nucleic Acids Res. 16*, 8186, 1988). Primers can be designed using commercially available software, such as OLIGO 4.06 Primer Analysis software (National Biosciences Inc., Plymouth, Minn.), to be 22-30 nucleotides in length, to have a GC content of 50% or more, and to anneal to the target se-

quence at temperatures about 68-72 °C. The method uses several restriction enzymes to generate a suitable fragment in the known region of a gene. The fragment is then circularized by intramolecular ligation and used as a PCR template.

Another method which can be used is capture PCR, which involves PCR amplification of DNA fragments adjacent to a known sequence in human and yeast artificial chromosome DNA (Lagerstrom *et al.*, *PCR Methods Applic. 1*, 111-119, 1991). In this method, multiple restriction enzyme digestions and ligations also can be used to place an engineered double-stranded sequence into an unknown fragment of the DNA molecule before performing PCR.

Another method which can be used to retrieve unknown sequences is that of Parker *et al.*, *Nucleic Acids Res. 19*, 3055-3060, 1991). Additionally, PCR, nested primers, and PROMOTERFINDER libraries (CLONTECH, Palo Alto, Calif.) can be used to walk genomic DNA (CLONTECH, Palo Alto, Calif.). This process avoids the need to screen libraries and is useful in finding intron/exon junctions.

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

Commercially available capillary electrophoresis systems can be used to analyze the size or confirm the nucleotide sequence of PCR or sequencing products. For example, capillary sequencing can employ flowable polymers for electrophoretic separation, four different fluorescent dyes (one for each nucleotide) which are laser activated, and detection of the emitted wavelengths by a charge coupled device camera. Output/light intensity can be converted to electrical signal using appropriate software (*e.g.* GENOTYPER and Sequence NAVIGATOR, Perkin Elmer), and the entire pro-

cess from loading of samples to computer analysis and electronic data display can be computer controlled. Capillary electrophoresis is especially preferable for the sequencing of small pieces of DNA which might be present in limited amounts in a particular sample.

Obtaining Tyrosine Phosphatase-Like Enzyme Polypeptides

Tyrosine phosphatase-like enzyme polypeptides can be obtained, for example, by purification from human cells, by expression of tyrosine phosphatase-like enzyme polynucleotides, or by direct chemical synthesis.

Protein Purification

Tyrosine phosphatase-like enzyme polypeptides can be purified from any human cell which expresses the receptor, including host cells which have been transfected with tyrosine phosphatase-like enzyme polynucleotides. Eight- to nine-week old fetal cells and melanoma cells are particularly useful sources of tyrosine phosphatase-like enzyme polypeptides. A purified tyrosine phosphatase-like enzyme polypeptide is separated from other compounds which normally associate with the tyrosine phosphatase-like enzyme polypeptide in the cell, such as certain proteins, carbohydrates, or lipids, using methods well-known in the art. Such methods include, but are not limited to, size exclusion chromatography, ammonium sulfate fractionation, ion exchange chromatography, affinity chromatography, and preparative gel electrophoresis. A preparation of purified tyrosine phosphatase-like enzyme polypeptides is at least 80% pure; preferably, the preparations are 90%, 95%, or 99% pure. Purity of the preparations can be assessed by any means known in the art, such as SDS-polyacrylamide gel electrophoresis.

Expression of Tyrosine Phosphatase-Like Enzyme Polynucleotides

To express a tyrosine phosphatase-like enzyme polypeptide, a tyrosine phosphatase-like enzyme polynucleotide can be inserted into an expression vector which contains the necessary elements for the transcription and translation of the inserted coding sequence. Methods which are well known to those skilled in the art can be used to construct expression vectors containing sequences encoding tyrosine phosphatase-like enzyme polypeptides and appropriate transcriptional and translational control elements. These methods include *in vitro* recombinant DNA techniques, synthetic techniques, and *in vivo* genetic recombination. Such techniques are described, for example, in Sambrook *et al.* (1989) and in Ausubel *et al.*, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York, N.Y., 1989.

A variety of expression vector/host systems can be utilized to contain and express sequences encoding a tyrosine phosphatase-like enzyme polypeptide. These include, but are not limited to, microorganisms, such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors, insect cell systems infected with virus expression vectors (*e.g.*, baculovirus), plant cell systems transformed with virus expression vectors (*e.g.*, cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or with bacterial expression vectors (*e.g.*, Ti or pBR322 plasmids), or animal cell systems.

The control elements or regulatory sequences are those non-translated regions of the vector -- enhancers, promoters, 5' and 3' untranslated regions -- which interact with host cellular proteins to carry out transcription and translation. Such elements can vary in their strength and specificity. Depending on the vector system and host utilized, any number of suitable transcription and translation elements, including constitutive and inducible promoters, can be used. For example, when cloning in bacterial systems, inducible promoters such as the hybrid lacZ promoter of the BLUESCRIPT phagemid (Stratagene, LaJolla, Calif.) or pSPORT1 plasmid (Life Technologies) and the like can be used. The baculovirus polyhedrin promoter can

be used in insect cells. Promoters or enhancers derived from the genomes of plant cells (*e.g.*, heat shock, RUBISCO, and storage protein genes) or from plant viruses (*e.g.*, viral promoters or leader sequences) can be cloned into the vector. In mammalian cell systems, promoters from mammalian genes or from mammalian viruses are preferable. If it is necessary to generate a cell line that contains multiple copies of a nucleotide sequence encoding a tyrosine phosphatase-like enzyme polypeptide, vectors based on SV40 or EBV can be used with an appropriate selectable marker.

Bacterial and Yeast Expression Systems

In bacterial systems, a number of expression vectors can be selected depending upon the use intended for the tyrosine phosphatase-like enzyme polypeptide. For example, when a large quantity of a tyrosine phosphatase-like enzyme polypeptide is needed for the induction of antibodies, vectors which direct high level expression of fusion proteins that are readily purified can be used. Such vectors include, but are not limited to, multifunctional *E. coli* cloning and expression vectors such as BLUESCRIPT (Stratagene). In a BLUESCRIPT vector, a sequence encoding the tyrosine phosphatase-like enzyme polypeptide can be ligated into the vector in frame with sequences for the amino-terminal Met and the subsequent 7 residues of β-galactosidase so that a hybrid protein is produced. pIN vectors (Van Heeke & Schuster, *J. Biol. Chem. 264*, 5503-5509, 1989) or pGEX vectors (Promega, Madison, Wis.) also can be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption to glutathione-agarose beads followed by elution in the presence of free glutathione. Proteins made in such systems can be designed to include heparin, thrombin, or factor Xa protease cleavage sites so that the cloned polypeptide of interest can be released from the GST moiety at will.

In the yeast *Saccharomyces cerevisiae*, a number of vectors containing constitutive or inducible promoters such as alpha factor, alcohol oxidase, and PGH can be used.

For reviews, see Ausubel *et al.* (1989) and Grant *et al.*, *Methods Enzymol. 153*, 516-544, 1987.

Plant and Insect Expression Systems

If plant expression vectors are used, the expression of sequences encoding tyrosine phosphatase-like enzyme polypeptides can be driven by any of a number of promoters. For example, viral promoters such as the 35S and 19S promoters of CaMV can be used alone or in combination with the omega leader sequence from TMV (Takamatsu, *EMBO J. 6*, 307-311, 1987). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters can be used (Coruzzi *et al.*, *EMBO J. 3*, 1671-1680, 1984; Broglie *et al.*, *Science 224*, 838-843, 1984; Winter *et al.*, *Results Probl. Cell Differ. 17*, 85-105, 1991). These constructs can be introduced into plant cells by direct DNA transformation or by pathogen-mediated transfection. Such techniques are described in a number of generally available reviews (*e.g.*, Hobbs or Murray, in MCGRAW HILL YEARBOOK OF SCIENCE AND TECHNOLOGY, McGraw Hill, New York, N.Y., pp. 191-196, 1992).

An insect system also can be used to express a tyrosine phosphatase-like enzyme polypeptide. For example, in one such system *Autographa californica* nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign genes in *Spodoptera frugiperda* cells or in *Trichoplusia* larvae. Sequences encoding tyrosine phosphatase-like enzyme polypeptides can be cloned into a non-essential region of the virus, such as the polyhedrin gene, and placed under control of the polyhedrin promoter. Successful insertion of tyrosine phosphatase-like enzyme polypeptides will render the polyhedrin gene inactive and produce recombinant virus lacking coat protein. The recombinant viruses can then be used to infect *S. frugiperda cells* or *Trichoplusia* larvae in which tyrosine phosphatase-like enzyme polypeptides can be expressed (Engelhard *et al.*, *Proc. Nat. Acad. Sci. 91*, 3224-3227, 1994).

Mammalian Expression Systems

A number of viral-based expression systems can be used to express tyrosine phosphatase-like enzyme polypeptides in mammalian host cells. For example, if an adenovirus is used as an expression vector, sequences encoding tyrosine phosphatase-like enzyme polypeptides can be ligated into an adenovirus transcription/translation complex comprising the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome can be used to obtain a viable virus which is capable of expressing a tyrosine phosphatase-like enzyme polypeptide in infected host cells (Logan & Shenk, *Proc. Natl. Acad. Sci. 81*, 3655-3659, 1984). If desired, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, can be used to increase expression in mammalian host cells.

Human artificial chromosomes (HACs) also can be used to deliver larger fragments of DNA than can be contained and expressed in a plasmid. HACs of 6M to 10M are constructed and delivered to cells via conventional delivery methods (*e.g.*, liposomes, polycationic amino polymers, or vesicles).

Specific initiation signals also can be used to achieve more efficient translation of sequences encoding tyrosine phosphatase-like enzyme polypeptides. Such signals include the ATG initiation codon and adjacent sequences. In cases where sequences encoding a tyrosine phosphatase-like enzyme polypeptide, its initiation codon, and upstream 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 (including the ATG initiation codon) should be provided. The initiation codon should be in the correct reading frame to ensure translation of the entire insert. Exogenous translational elements and initiation codons can be of various origins, both natural and synthetic. The efficiency of expression can be enhanced by the inclusion of enhancers which are appropriate for the particular cell system which is used (see Scharf *et al.*, *Results Probl. Cell Differ. 20*, 125-162, 1994).

Host Cells

A host cell strain can be chosen for its ability to modulate the expression of the inserted sequences or to process the expressed tyrosine phosphatase-like enzyme polypeptide 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" form of the polypeptide also can be used to facilitate correct insertion, folding and/or function. 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; 10801 University Boulevard, Manassas, VA 20110-2209) and can be chosen to ensure the correct modification and processing of the foreign protein.

Stable expression is preferred for long-term, high-yield production of recombinant proteins. For example, cell lines which stably express tyrosine phosphatase-like enzyme polypeptides can be transformed using expression vectors which can 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 can be allowed to grow for 1-2 days in an enriched medium before they are switched to a selective medium. The purpose of the selectable marker is to confer resistance to selection, and its presence allows growth and recovery of cells which successfully express the introduced tyrosine phosphatase-like enzyme sequences. Resistant clones of stably transformed cells can be proliferated using tissue culture techniques appropriate to the cell type. See, for example, ANIMAL CELL CULTURE, R.I. Freshney, ed., 1986.

Any number of selection systems can be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase (Wigler *et al.*, *Cell 11*, 223-32, 1977) and adenine phosphoribosyltransferase (Lowy

et al., *Cell 22*, 817-23, 1980) genes which can be employed in *tk*$^-$ or *aprt*$^-$ cells, respectively. Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate (Wigler *et al.*, *Proc. Natl. Acad. Sci. 77*, 3567-70, 1980), *npt* confers resistance to the aminoglycosides, neomycin and G-418 (Colbere-Garapin *et al.*, *J. Mol. Biol. 150*, 1-14, 1981), and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively (Murray, 1992, *supra*). Additional selectable genes have been described. For example, *trpB* allows cells to utilize indole in place of tryptophan, or *hisD*, which allows cells to utilize histinol in place of histidine (Hartman & Mulligan, *Proc. Natl. Acad. Sci. 85*, 8047-51, 1988). Visible markers such as anthocyanins, β-glucuronidase and its substrate GUS, and luciferase and its substrate luciferin, can be used to identify transformants and to quantify the amount of transient or stable protein expression attributable to a specific vector system (Rhodes *et al.*, *Methods Mol. Biol. 55*, 121-131, 1995).

Detecting Expression of Tyrosine Phosphatase-Like Enzyme Polypeptides

Although the presence of marker gene expression suggests that the tyrosine phosphatase-like enzyme polynucleotide is also present, its presence and expression may need to be confirmed. For example, if a sequence encoding a tyrosine phosphatase-like enzyme polypeptide is inserted within a marker gene sequence, transformed cells containing sequences which encode a tyrosine phosphatase-like enzyme polypeptide can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding a tyrosine phosphatase-like enzyme polypeptide under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tyrosine phosphatase-like enzyme polynucleotide.

Alternatively, host cells which contain a tyrosine phosphatase-like enzyme polynucleotide and which express a tyrosine phosphatase-like enzyme polypeptide can be identified by a variety of procedures known to those of skill in the art. These pro-

cedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations 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. For example, the presence of a polynucleotide sequence encoding a tyrosine phosphatase-like enzyme polypeptide can be detected by DNA-DNA or DNA-RNA hybridization or amplification using probes or fragments or fragments of polynucleotides encoding a tyrosine phosphatase-like enzyme polypeptide. Nucleic acid amplification-based assays involve the use of oligonucleotides selected from sequences encoding a tyrosine phosphatase-like enzyme polypeptide to detect transformants which contain a tyrosine phosphatase-like enzyme polynucleotide.

A variety of protocols for detecting and measuring the expression of a tyrosine phosphatase-like enzyme polypeptide, using either polyclonal or monoclonal antibodies specific for the polypeptide, are known in the art. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay using monoclonal antibodies reactive to two non-interfering epitopes on a tyrosine phosphatase-like enzyme polypeptide can be used, or a competitive binding assay can be employed. These and other assays are described in Hampton *et al.*, SEROLOGICAL METHODS: A LABORATORY MANUAL, APS Press, St. Paul, Minn., 1990) and Maddox *et al.*, *J. Exp. Med. 158*, 1211-1216, 1983).

A wide variety of labels and conjugation techniques are known by those skilled in the art and can 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 tyrosine phosphatase-like enzyme polypeptides include oligo-labeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, sequences encoding a tyrosine phosphatase-like enzyme polypeptide can be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and can be used to synthesize RNA probes *in vitro* by addition of labeled nucleotides and an appropriate RNA

polymerase such as T7, T3, or SP6. These procedures can be conducted using a variety of commercially available kits (Amersham Pharmacia Biotech, Promega, and US Biochemical). Suitable reporter molecules or labels which can be used for ease of detection include radionuclides, enzymes, and fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Expression and Purification of Tyrosine Phosphatase-Like Enzyme Polypeptides

Host cells transformed with nucleotide sequences encoding a tyrosine phosphatase-like enzyme polypeptide can be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The polypeptide produced by a transformed cell can be secreted or contained 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 tyrosine phosphatase-like enzyme polypeptides can be designed to contain signal sequences which direct secretion of soluble tyrosine phosphatase-like enzyme polypeptides through a prokaryotic or eukaryotic cell membrane or which direct the membrane insertion of membrane-bound tyrosine phosphatase-like enzyme polypeptide.

As discussed above, other constructions can be used to join a sequence encoding a tyrosine phosphatase-like enzyme polypeptide to a nucleotide sequence encoding a polypeptide domain which will facilitate purification of soluble proteins. Such purification facilitating domains include, but are not limited to, metal chelating peptides such as histidine-tryptophan modules that allow purification on immobilized metals, protein A domains that allow purification on immobilized immunoglobulin, and the domain utilized in the FLAGS extension/affinity purification system (Immunex Corp., Seattle, Wash.). Inclusion of cleavable linker sequences such as those specific for Factor Xa or enterokinase (Invitrogen, San Diego, CA) between the purification domain and the tyrosine phosphatase-like enzyme polypeptide also can be used to facilitate purification. One such expression vector provides for expression of a fu-

sion protein containing a tyrosine phosphatase-like enzyme polypeptide and 6 histidine residues preceding a thioredoxin or an enterokinase cleavage site. The histidine residues facilitate purification by IMAC (immobilized metal ion affinity chromatography, as described in Porath *et al.*, *Prot. Exp. Purif. 3*, 263-281, 1992), while the enterokinase cleavage site provides a means for purifying the tyrosine phosphatase-like enzyme polypeptide from the fusion protein. Vectors which contain fusion proteins are disclosed in Kroll *et al.*, *DNA Cell Biol. 12*, 441-453, 1993.

Chemical Synthesis

Sequences encoding a tyrosine phosphatase-like enzyme polypeptide can be synthesized, in whole or in part, using chemical methods well known in the art (see Caruthers *et al.*, *Nucl. Acids Res. Symp. Ser.* 215-223, 1980; Horn *et al. Nucl. Acids Res. Symp. Ser.* 225-232, 1980). Alternatively, a tyrosine phosphatase-like enzyme polypeptide itself can be produced using chemical methods to synthesize its amino acid sequence, such as by direct peptide synthesis using solid-phase techniques (Merrifield, *J. Am. Chem. Soc. 85*, 2149-2154, 1963; Roberge *et al.*, *Science 269*, 202-204, 1995). Protein synthesis can be performed using manual techniques or by automation. Automated synthesis can be achieved, for example, using Applied Biosystems 431A Peptide Synthesizer (Perkin Elmer). Optionally, fragments of tyrosine phosphatase-like enzyme polypeptides can be separately synthesized and combined using chemical methods to produce a full-length molecule.

The newly synthesized peptide can be substantially purified by preparative high performance liquid chromatography (*e.g.*, Creighton, PROTEINS: STRUCTURES AND MOLECULAR PRINCIPLES, WH Freeman and Co., New York, N.Y., 1983). The composition of a synthetic tyrosine phosphatase-like enzyme polypeptide can be confirmed by amino acid analysis or sequencing (*e.g.*, the Edman degradation procedure; *see* Creighton, *supra*). Additionally, any portion of the amino acid sequence of the tyrosine phosphatase-like enzyme polypeptide can be altered during direct synthesis

and/or combined using chemical methods with sequences from other proteins to produce a variant polypeptide or a fusion protein.

Production of Altered Tyrosine Phosphatase-Like Enzyme Polypeptides

As will be understood by those of skill in the art, it may be advantageous to produce tyrosine phosphatase-like enzyme polypeptide-encoding nucleotide sequences possessing non-naturally occurring codons. For example, codons preferred by a particular prokaryotic or eukaryotic host can be selected to increase the rate of protein expression or to produce an RNA transcript having desirable properties, such as a half-life which is longer than that of a transcript generated from the naturally occurring sequence.

The nucleotide sequences disclosed herein can be engineered using methods generally known in the art to alter tyrosine phosphatase-like enzyme polypeptide-encoding sequences for a variety of reasons, including but not limited to, alterations which modify the cloning, processing, and/or expression of the polypeptide or mRNA product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides can be used to engineer the nucleotide sequences. For example, site-directed mutagenesis can be used to insert new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, introduce mutations, and so forth.

Antibodies

Any type of antibody known in the art can be generated to bind specifically to an epitope of a tyrosine phosphatase-like enzyme polypeptide. "Antibody" as used herein includes intact immunoglobulin molecules, as well as fragments thereof, such as Fab, $F(ab')_2$, and Fv, which are capable of binding an epitope of a tyrosine phosphatase-like enzyme polypeptide. Typically, at least 6, 8, 10, or 12 contiguous amino

acids are required to form an epitope. However, epitopes which involve non-contiguous amino acids may require more, *e.g.*, at least 15, 25, or 50 amino acids.

An antibody which specifically binds to an epitope of a tyrosine phosphatase-like enzyme polypeptide can be used therapeutically, as well as in immunochemical assays, such as Western blots, ELISAs, radioimmunoassays, immunohistochemical assays, immunoprecipitations, or other immunochemical assays known in the art. Various immunoassays can be used to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays are well known in the art. Such immunoassays typically involve the measurement of complex formation between an immunogen and an antibody which specifically binds to the immunogen.

Typically, an antibody which specifically binds to a tyrosine phosphatase-like enzyme polypeptide provides a detection signal at least 5-, 10-, or 20-fold higher than a detection signal provided with other proteins when used in an immunochemical assay. Preferably, antibodies which specifically bind to tyrosine phosphatase-like enzyme polypeptides do not detect other proteins in immunochemical assays and can immunoprecipitate a tyrosine phosphatase-like enzyme polypeptide from solution.

Tyrosine phosphatase-like enzyme polypeptides can be used to immunize a mammal, such as a mouse, rat, rabbit, guinea pig, monkey, or human, to produce polyclonal antibodies. If desired, a tyrosine phosphatase-like enzyme polypeptide can be conjugated to a carrier protein, such as bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin. Depending on the host species, various adjuvants can be used to increase the immunological response. Such adjuvants include, but are not limited to, Freund's adjuvant, mineral gels (*e.g.*, aluminum hydroxide), and surface active substances (*e.g.* lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanin, and dinitrophenol). Among adjuvants used in humans, BCG (*bacilli Calmette-Guerin*) and *Corynebacterium parvum* are especially useful.

Monoclonal antibodies which specifically bind to a tyrosine phosphatase-like enzyme polypeptide can be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These techniques include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique (Kohler *et al.*, *Nature 256*, 495-497, 1985; Kozbor *et al.*, *J. Immunol. Methods 81*, 31-42, 1985; Cote *et al.*, *Proc. Natl. Acad. Sci. 80*, 2026-2030, 1983; Cole *et al.*, *Mol. Cell Biol. 62*, 109-120, 1984).

In addition, techniques developed for the production of "chimeric antibodies," the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used (Morrison *et al.*, *Proc. Natl. Acad. Sci. 81*, 6851-6855, 1984; Neuberger *et al.*, *Nature 312*, 604-608, 1984; Takeda *et al.*, *Nature 314*, 452-454, 1985). Monoclonal and other antibodies also can be "humanized" to prevent a patient from mounting an immune response against the antibody when it is used therapeutically. Such antibodies may be sufficiently similar in sequence to human antibodies to be used directly in therapy or may require alteration of a few key residues. Sequence differences between rodent antibodies and human sequences can be minimized by replacing residues which differ from those in the human sequences by site directed mutagenesis of individual residues or by grating of entire complementarity determining regions. Alternatively, humanized antibodies can be produced using recombinant methods, as described in GB2188638B. Antibodies which specifically bind to a tyrosine phosphatase-like enzyme polypeptide can contain antigen binding sites which are either partially or fully humanized, as disclosed in U.S. 5,565,332.

Alternatively, techniques described for the production of single chain antibodies can be adapted using methods known in the art to produce single chain antibodies which specifically bind to tyrosine phosphatase-like enzyme polypeptides. Antibodies with related specificity, but of distinct idiotypic composition, can be generated by chain shuffling from random combinatorial immunoglobin libraries (Burton, *Proc. Natl. Acad. Sci. 88*, 11120-23, 1991).

Single-chain antibodies also can be constructed using a DNA amplification method, such as PCR, using hybridoma cDNA as a template (Thirion *et al.*, 1996, *Eur. J. Cancer Prev. 5*, 507-11). Single-chain antibodies can be mono- or bispecific, and can be bivalent or tetravalent. Construction of tetravalent, bispecific single-chain antibodies is taught, for example, in Coloma & Morrison, 1997, *Nat. Biotechnol. 15*, 159-63. Construction of bivalent, bispecific single-chain antibodies is taught in Mallender & Voss, 1994, *J. Biol. Chem. 269*, 199-206.

A nucleotide sequence encoding a single-chain antibody can be constructed using manual or automated nucleotide synthesis, cloned into an expression construct using standard recombinant DNA methods, and introduced into a cell to express the coding sequence, as described below. Alternatively, single-chain antibodies can be produced directly using, for example, filamentous phage technology (Verhaar *et al.*, 1995, *Int. J. Cancer 61*, 497-501; Nicholls *et al.*, 1993, *J. Immunol. Meth. 165*, 81-91).

Antibodies which specifically bind to tyrosine phosphatase-like enzyme polypeptides also can be produced by inducing *in vivo* production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature (Orlandi *et al.*, *Proc. Natl. Acad. Sci. 86*, 3833-3837, 1989; Winter *et al.*, *Nature 349*, 293-299, 1991).

Other types of antibodies can be constructed and used therapeutically in methods of the invention. For example, chimeric antibodies can be constructed as disclosed in WO 93/03151. Binding proteins which are derived from immunoglobulins and which are multivalent and multispecific, such as the "diabodies" described in WO 94/13804, also can be prepared.

Antibodies according to the invention can be purified by methods well known in the art. For example, antibodies can be affinity purified by passage over a column to

which a tyrosine phosphatase-like enzyme polypeptide is bound. The bound antibodies can then be eluted from the column using a buffer with a high salt concentration.

Antisense Oligonucleotides

Antisense oligonucleotides are nucleotide sequences which are complementary to a specific DNA or RNA sequence. Once introduced into a cell, the complementary nucleotides combine with natural sequences produced by the cell to form complexes and block either transcription or translation. Preferably, an antisense oligonucleotide is at least 11 nucleotides in length, but can be at least 12, 15, 20, 25, 30, 35, 40, 45, or 50 or more nucleotides long. Longer sequences also can be used. Antisense oligonucleotide molecules can be provided in a DNA construct and introduced into a cell as described above to decrease the level of tyrosine phosphatase-like enzyme gene products in the cell.

Antisense oligonucleotides can be deoxyribonucleotides, ribonucleotides, or a combination of both. Oligonucleotides can be synthesized manually or by an automated synthesizer, by covalently linking the 5' end of one nucleotide with the 3' end of another nucleotide with non-phosphodiester internucleotide linkages such alkylphosphonates, phosphorothioates, phosphorodithioates, alkylphosphonothioates, alkylphosphonates, phosphoramidates, phosphate esters, carbamates, acetamidate, carboxymethyl esters, carbonates, and phosphate triesters. *See* Brown, *Meth. Mol. Biol. 20*, 1-8, 1994; Sonveaux, *Meth. Mol. Biol. 26*, 1-72, 1994; Uhlmann *et al.*, *Chem. Rev. 90*, 543-583, 1990.

Modifications of tyrosine phosphatase-like enzyme gene expression can be obtained by designing antisense oligonucleotides which will form duplexes to the control, 5', or regulatory regions of the tyrosine phosphatase-like enzyme gene. Oligonucleotides derived from the transcription initiation site, *e.g.*, between positions -10 and +10 from the start site, are preferred. 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 chaperons. Therapeutic advances using triplex DNA have been described in the literature (*e.g.*, Gee *et al.*, in Huber & Carr, MOLECULAR AND IMMUNOLOGIC APPROACHES, Futura Publishing Co., Mt. Kisco, N.Y., 1994). An antisense oligonucleotide also can be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Precise complementarity is not required for successful complex formation between an antisense oligonucleotide and the complementary sequence of a tyrosine phosphatase-like enzyme polynucleotide. Antisense oligonucleotides which comprise, for example, 2, 3, 4, or 5 or more stretches of contiguous nucleotides which are precisely complementary to a tyrosine phosphatase-like enzyme polynucleotide, each separated by a stretch of contiguous nucleotides which are not complementary to adjacent tyrosine phosphatase-like enzyme nucleotides, can provide sufficient targeting specificity for tyrosine phosphatase-like enzyme mRNA. Preferably, each stretch of complementary contiguous nucleotides is at least 4, 5, 6, 7, or 8 or more nucleotides in length. Non-complementary intervening sequences are preferably 1, 2, 3, or 4 nucleotides in length. One skilled in the art can easily use the calculated melting point of an antisense-sense pair to determine the degree of mismatching which will be tolerated between a particular antisense oligonucleotide and a particular tyrosine phosphatase-like enzyme polynucleotide sequence.

Antisense oligonucleotides can be modified without affecting their ability to hybridize to a tyrosine phosphatase-like enzyme polynucleotide. These modifications can be internal or at one or both ends of the antisense molecule. For example, internucleoside phosphate linkages can be modified by adding cholesteryl or diamine moieties with varying numbers of carbon residues between the amino groups and terminal ribose. Modified bases and/or sugars, such as arabinose instead of ribose, or a 3', 5'-substituted oligonucleotide in which the 3' hydroxyl group or the 5' phosphate group are substituted, also can be employed in a modified antisense oligonucleotide.

These modified oligonucleotides can be prepared by methods well known in the art. *See, e.g.*, Agrawal *et al.*, *Trends Biotechnol. 10*, 152-158, 1992; Uhlmann *et al.*, *Chem. Rev. 90*, 543-584, 1990; Uhlmann *et al.*, *Tetrahedron. Lett. 215*, 3539-3542, 1987.

Ribozymes

Ribozymes are RNA molecules with catalytic activity. *See, e.g.*, Cech, *Science 236*, 1532-1539; 1987; Cech, *Ann. Rev. Biochem. 59*, 543-568; 1990, Cech, *Curr. Opin. Struct. Biol. 2*, 605-609; 1992, Couture & Stinchcomb, *Trends Genet. 12*, 510-515, 1996. Ribozymes can be used to inhibit gene function by cleaving an RNA sequence, as is known in the art (*e.g.*, Haseloff *et al.*, U.S. Patent 5,641,673). The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. Examples include engineered hammerhead motif ribozyme molecules that can specifically and efficiently catalyze endonucleolytic cleavage of specific nucleotide sequences.

The coding sequence of a tyrosine phosphatase-like enzyme polynucleotide can be used to generate ribozymes which will specifically bind to mRNA transcribed from the tyrosine phosphatase-like enzyme polynucleotide. Methods of designing and constructing ribozymes which can cleave other RNA molecules in trans in a highly sequence specific manner have been developed and described in the art (*see* Haseloff *et al. Nature 334*, 585-591, 1988). For example, the cleavage activity of ribozymes can be targeted to specific RNAs by engineering a discrete "hybridization" region into the ribozyme. The hybridization region contains a sequence complementary to the target RNA and thus specifically hybridizes with the target (see, for example, Gerlach *et al.*, EP 321,201).

Specific ribozyme cleavage sites within a tyrosine phosphatase-like enzyme RNA target can be identified by scanning the target molecule for ribozyme cleavage sites

which include the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides corresponding to the region of the target RNA containing the cleavage site can be evaluated for secondary structural features which may render the target inoperable. Suitability of candidate tyrosine phosphatase-like enzyme RNA targets also can be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays. Longer complementary sequences can be used to increase the affinity of the hybridization sequence for the target. The hybridizing and cleavage regions of the ribozyme can be integrally related such that upon hybridizing to the target RNA through the complementary regions, the catalytic region of the ribozyme can cleave the target.

Ribozymes can be introduced into cells as part of a DNA construct. Mechanical methods, such as microinjection, liposome-mediated transfection, electroporation, or calcium phosphate precipitation, can be used to introduce a ribozyme-containing DNA construct into cells in which it is desired to decrease tyrosine phosphatase-like enzyme expression. Alternatively, if it is desired that the cells stably retain the DNA construct, the construct can be supplied on a plasmid and maintained as a separate element or integrated into the genome of the cells, as is known in the art. A ribozyme-encoding DNA construct can include transcriptional regulatory elements, such as a promoter element, an enhancer or UAS element, and a transcriptional terminator signal, for controlling transcription of ribozymes in the cells.

As taught in Haseloff *et al.*, U.S. Patent 5,641,673, ribozymes can be engineered so that ribozyme expression will occur in response to factors which induce expression of a target gene. Ribozymes also can be engineered to provide an additional level of regulation, so that destruction of mRNA occurs only when both a ribozyme and a target gene are induced in the cells.

Screening Methods

The invention provides assays for screening test compounds which bind to or modulate the activity of a tyrosine phosphatase-like enzyme polypeptide or a tyrosine phosphatase-like enzyme polynucleotide. A test compound preferably binds to a tyrosine phosphatase-like enzyme polypeptide or polynucleotide. More preferably, a test compound decreases or increases the effect of tyrosine phosphatase activity by at least about 10, preferably about 50, more preferably about 75, 90, or 100% relative to the absence of the test compound.

Test Compounds

Test compounds can be pharmacologic agents already known in the art or can be compounds previously unknown to have any pharmacological activity. The compounds can be naturally occurring or designed in the laboratory. They can be isolated from microorganisms, animals, or plants, and can be produced recombinantly, or synthesized by chemical methods known in the art. If desired, test compounds can be obtained using any of the numerous combinatorial library methods known in the art, including but not limited to, biological libraries, spatially addressable parallel solid phase or solution phase libraries, synthetic library methods requiring deconvolution, the "one-bead one-compound" library method, and synthetic library methods using affinity chromatography selection. The biological library approach is limited to polypeptide libraries, while the other four approaches are applicable to polypeptide, non-peptide oligomer, or small molecule libraries of compounds. *See* Lam, *Anticancer Drug Des. 12*, 145, 1997.

Methods for the synthesis of molecular libraries are well known in the art (*see*, for example, DeWitt *et al.*, *Proc. Natl. Acad. Sci. U.S.A. 90*, 6909, 1993; Erb *et al. Proc. Natl. Acad. Sci. U.S.A. 91*, 11422, 1994; Zuckermann *et al.*, *J. Med. Chem. 37*, 2678, 1994; Cho *et al.*, *Science 261*, 1303, 1993; Carell *et al.*, *Angew. Chem. Int. Ed. Engl. 33*, 2059, 1994; Carell *et al.*, *Angew. Chem. Int. Ed. Engl. 33*, 2061; Gallop *et al.*, *J.*

Med. Chem. 37, 1233, 1994). Libraries of compounds can be presented in solution (*see, e.g.*, Houghten, *Biotechniques 13*, 412-421, 1992), or on beads (Lam, *Nature 354*, 82-84, 1991), chips (Fodor, *Nature 364*, 555-556, 1993), bacteria or spores (Ladner, U.S. Patent 5,223,409), plasmids (Cull *et al.*, *Proc. Natl. Acad. Sci. U.S.A. 89*, 1865-1869, 1992), or phage (Scott & Smith, *Science 249*, 386-390, 1990; Devlin, *Science 249*, 404-406, 1990); Cwirla *et al.*, *Proc. Natl. Acad. Sci. 97*, 6378-6382, 1990; Felici, *J. Mol. Biol. 222*, 301-310, 1991; and Ladner, U.S. Patent 5,223,409).

High Throughput Screening

Test compounds can be screened for the ability to bind to tyrosine phosphatase-like enzyme polypeptides or polynucleotides or to affect tyrosine phosphatase-like enzyme activity or tyrosine phosphatase-like enzyme gene expression using high throughput screening. Using high throughput screening, many discrete compounds can be tested in parallel so that large numbers of test compounds can be quickly screened. The most widely established techniques utilize 96-well microtiter plates. The wells of the microtiter plates typically require assay volumes that range from 50 to 500 µl. In addition to the plates, many instruments, materials, pipettors, robotics, plate washers, and plate readers are commercially available to fit the 96-well format.

Alternatively, "free format assays," or assays that have no physical barrier between samples, can be used. For example, an assay using pigment cells (melanocytes) in a simple homogeneous assay for combinatorial peptide libraries is described by Jayawickreme *et al.*, *Proc. Natl. Acad. Sci. U.S.A. 19,* 1614-18 (1994). The cells are placed under agarose in petri dishes, then beads that carry combinatorial compounds are placed on the surface of the agarose. The combinatorial compounds are partially released the compounds from the beads. Active compounds can be visualized as dark pigment areas because, as the compounds diffuse locally into the gel matrix, the active compounds cause the cells to change colors.

Another example of a free format assay is described by Chelsky, "Strategies for Screening Combinatorial Libraries: Novel and Traditional Approaches," reported at the First Annual Conference of The Society for Biomolecular Screening in Philadelphia, Pa. (Nov. 7-10, 1995). Chelsky placed a simple homogenous enzyme assay for carbonic anhydrase inside an agarose gel such that the enzyme in the gel would cause a color change throughout the gel. Thereafter, beads carrying combinatorial compounds via a photolinker were placed inside the gel and the compounds were partially released by UV-light. Compounds that inhibited the enzyme were observed as local zones of inhibition having less color change.

Yet another example is described by Salmon *et al.*, *Molecular Diversity 2*, 57-63 (1996). In this example, combinatorial libraries were screened for compounds that had cytotoxic effects on cancer cells growing in agar.

Another high throughput screening method is described in Beutel *et al.*, U.S. Patent 5,976,813. In this method, test samples are placed in a porous matrix. One or more assay components are then placed within, on top of, or at the bottom of a matrix such as a gel, a plastic sheet, a filter, or other form of easily manipulated solid support. When samples are introduced to the porous matrix they diffuse sufficiently slowly, such that the assays can be performed without the test samples running together.

Binding Assays

For binding assays, the test compound is preferably a small molecule which binds to and occupies the active site of the tyrosine phosphatase-like enzyme polypeptide, thereby making the active site inaccessible to substrate such that normal biological activity is prevented. Examples of such small molecules include, but are not limited to, small peptides or peptide-like molecules.

In binding assays, either the test compound or the tyrosine phosphatase-like enzyme polypeptide can comprise a detectable label, such as a fluorescent, radioisotopic,

chemiluminescent, or enzymatic label, such as horseradish peroxidase, alkaline phosphatase, or luciferase. Detection of a test compound which is bound to the tyrosine phosphatase-like enzyme polypeptide can then be accomplished, for example, by direct counting of radioemmission, by scintillation counting, or by determining conversion of an appropriate substrate to a detectable product.

Alternatively, binding of a test compound to a tyrosine phosphatase-like enzyme polypeptide can be determined without labeling either of the interactants. For example, a microphysiometer can be used to detect binding of a test compound with a tyrosine phosphatase-like enzyme polypeptide. A microphysiometer (*e.g.*, Cytosensor™) is an analytical instrument that measures the rate at which a cell acidifies its environment using a light-addressable potentiometric sensor (LAPS). Changes in this acidification rate can be used as an indicator of the interaction between a test compound and a tyrosine phosphatase-like enzyme polypeptide (McConnell *et al.*, *Science 257*, 1906-1912, 1992).

Determining the ability of a test compound to bind to a tyrosine phosphatase-like enzyme polypeptide also can be accomplished using a technology such as real-time Bimolecular Interaction Analysis (BIA) (Sjolander & Urbaniczky, *Anal. Chem. 63*, 2338-2345, 1991, and Szabo *et al.*, *Curr. Opin. Struct. Biol. 5*, 699-705, 1995). BIA is a technology for studying biospecific interactions in real time, without labeling any of the interactants (*e.g.*, BIAcore™). Changes in the optical phenomenon surface plasmon resonance (SPR) can be used as an indication of real-time reactions between biological molecules.

In yet another aspect of the invention, a tyrosine phosphatase-like enzyme polypeptide can be used as a "bait protein" in a two-hybrid assay or three-hybrid assay (see, *e.g.*, U.S. Patent 5,283,317; Zervos *et al.*, *Cell 72*, 223-232, 1993; Madura *et al.*, *J. Biol. Chem. 268*, 12046-12054, 1993; Bartel *et al.*, *Biotechniques 14*, 920-924, 1993; Iwabuchi *et al.*, *Oncogene 8*, 1693-1696, 1993; and Brent W094/10300), to identify

other proteins which bind to or interact with the tyrosine phosphatase-like enzyme polypeptide and modulate its activity.

The two-hybrid system is based on the modular nature of most transcription factors, which consist of separable DNA-binding and activation domains. Briefly, the assay utilizes two different DNA constructs. For example, in one construct, polynucleotide encoding a tyrosine phosphatase-like enzyme polypeptide can be fused to a polynucleotide encoding the DNA binding domain of a known transcription factor (*e.g.*, GAL-4). In the other construct a DNA sequence that encodes an unidentified protein ("prey" or "sample") can be fused to a polynucleotide that codes for the activation domain of the known transcription factor. If the "bait" and the "prey" proteins are able to interact *in vivo* to form an protein-dependent complex, the DNA-binding and activation domains of the transcription factor are brought into close proximity. This proximity allows transcription of a reporter gene (*e.g.*, LacZ), which is operably linked to a transcriptional regulatory site responsive to the transcription factor. Expression of the reporter gene can be detected, and cell colonies containing the functional transcription factor can be isolated and used to obtain the DNA sequence encoding the protein which interacts with the tyrosine phosphatase-like enzyme polypeptide.

It may be desirable to immobilize either the tyrosine phosphatase-like enzyme polypeptide (or polynucleotide) or the test compound to facilitate separation of bound from unbound forms of one or both of the interactants, as well as to accommodate automation of the assay. Thus, either the tyrosine phosphatase-like enzyme polypeptide (or polynucleotide) or the test compound can be bound to a solid support. Suitable solid supports include, but are not limited to, glass or plastic slides, tissue culture plates, microtiter wells, tubes, silicon chips, or particles such as beads (including, but not limited to, latex, polystyrene, or glass beads). Any method known in the art can be used to attach the tyrosine phosphatase-like enzyme polypeptide (or polynucleotide) or test compound to a solid support, including use of covalent and non-covalent linkages, passive absorption, or pairs of binding moieties attached re-

spectively to the polypeptide (or polynucleotide) or test compound and the solid support. Test compounds are preferably bound to the solid support in an array, so that the location of individual test compounds can be tracked. Binding of a test compound to a tyrosine phosphatase-like enzyme polypeptide (or polynucleotide) can be accomplished in any vessel suitable for containing the reactants. Examples of such vessels include microtiter plates, test tubes, and microcentrifuge tubes.

In one embodiment, the tyrosine phosphatase-like enzyme polypeptide is a fusion protein comprising a domain that allows the tyrosine phosphatase-like enzyme polypeptide to be bound to a solid support. For example, glutathione-S-transferase fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, Mo.) or glutathione derivatized microtiter plates, which are then combined with the test compound or the test compound and the non-adsorbed tyrosine phosphatase-like enzyme polypeptide; the mixture is then incubated under conditions conducive to complex formation (*e.g.*, at physiological conditions for salt and pH). Following incubation, the beads or microtiter plate wells are washed to remove any unbound components. Binding of the interactants can be determined either directly or indirectly, as described above. Alternatively, the complexes can be dissociated from the solid support before binding is determined.

Other techniques for immobilizing proteins or polynucleotides on a solid support also can be used in the screening assays of the invention. For example, either a tyrosine phosphatase-like enzyme polypeptide (or polynucleotide) or a test compound can be immobilized utilizing conjugation of biotin and streptavidin. Biotinylated tyrosine phosphatase-like enzyme polypeptides (or polynucleotides) or test compounds can be prepared from biotin-NHS(N-hydroxysuccinimide) using techniques well known in the art (*e.g.*, biotinylation kit, Pierce Chemicals, Rockford, Ill.) and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical). Alternatively, antibodies which specifically bind to a tyrosine phosphatase-like enzyme polypeptide, polynucleotide, or a test compound, but which do not interfere with a desired binding site, such as the active site of the tyrosine phosphatase-like enzyme polypeptide, can

be derivatized to the wells of the plate. Unbound target or protein can be trapped in the wells by antibody conjugation.

Methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies which specifically bind to the tyrosine phosphatase-like enzyme polypeptide or test compound, enzyme-linked assays which rely on detecting an activity of the tyrosine phosphatase-like enzyme polypeptide, and SDS gel electrophoresis under non-reducing conditions.

Screening for test compounds which bind to a tyrosine phosphatase-like enzyme polypeptide or polynucleotide also can be carried out in an intact cell. Any cell which comprises a tyrosine phosphatase-like enzyme polypeptide or polynucleotide can be used in a cell-based assay system. A tyrosine phosphatase-like enzyme polynucleotide can be naturally occurring in the cell or can be introduced using techniques such as those described above. Binding of the test compound to a tyrosine phosphatase-like enzyme polypeptide or polynucleotide is determined as described above.

Tyrosine Phosphatase Assays

Test compounds can be tested for the ability to increase or decrease tyrosine phosphatase activity of a tyrosine phosphatase-like enzyme polypeptide (see the specific examples, below). Tyrosine phosphatase assays can be carried out after contacting either a purified tyrosine phosphatase-like enzyme polypeptide, a cell membrane preparation, or an intact cell with a test compound. A test compound which decreases a tyrosine phosphatase activity of a tyrosine phosphatase-like enzyme by at least about 10, preferably about 50, more preferably about 75, 90, or 100% is identified as a potential agent for decreasing tyrosine phosphatase-like enzyme activity. A test compound which increases tyrosine phosphatase activity by at least about 10,

preferably about 50, more preferably about 75, 90, or 100% is identified as a potential agent for increasing tyrosine phosphatase-like enzyme activity.

Tyrosine Phosphatase-Like Enzyme Gene Expression

In another embodiment, test compounds which increase or decrease tyrosine phosphatase-like enzyme gene expression are identified. A tyrosine phosphatase-like enzyme polynucleotide is contacted with a test compound, and the expression of an RNA or polypeptide product of the tyrosine phosphatase-like enzyme polynucleotide is determined. The level of expression of appropriate mRNA or polypeptide in the presence of the test compound is compared to the level of expression of mRNA or polypeptide in the absence of the test compound. The test compound can then be identified as a modulator of expression based on this comparison. For example, when expression of mRNA or polypeptide is greater in the presence of the test compound than in its absence, the test compound is identified as a stimulator or enhancer of the mRNA or polypeptide expression. Alternatively, when expression of the mRNA or polypeptide is less in the presence of the test compound than in its absence, the test compound is identified as an inhibitor of the mRNA or polypeptide expression.

The level of tyrosine phosphatase-like enzyme mRNA or polypeptide expression in the cells can be determined by methods well known in the art for detecting mRNA or polypeptide. Either qualitative or quantitative methods can be used. The presence of polypeptide products of a tyrosine phosphatase-like enzyme polynucleotide can be determined, for example, using a variety of techniques known in the art, including immunochemical methods such as radioimmunoassay, Western blotting, and immunohistochemistry. Alternatively, polypeptide synthesis can be determined *in vivo*, in a cell culture, or in an *in vitro* translation system by detecting incorporation of labeled amino acids into a tyrosine phosphatase-like enzyme polypeptide.

Such screening can be carried out either in a cell-free assay system or in an intact cell. Any cell which expresses a tyrosine phosphatase-like enzyme polynucleotide can be used in a cell-based assay system. The tyrosine phosphatase-like enzyme polynucleotide can be naturally occurring in the cell or can be introduced using techniques such as those described above. Either a primary culture or an established cell line, such as CHO or human embryonic kidney 293 cells, can be used.

Pharmaceutical Compositions

The invention also provides pharmaceutical compositions which can be administered to a patient to achieve a therapeutic effect. Pharmaceutical compositions of the invention can comprise, for example, a tyrosine phosphatase-like enzyme polypeptide, tyrosine phosphatase-like enzyme polynucleotide, antibodies which specifically bind to a tyrosine phosphatase-like enzyme polypeptide, or mimetics, agonists, antagonists, or inhibitors of a tyrosine phosphatase-like enzyme polypeptide activity. The compositions can be administered alone or in combination with at least one other agent, such as stabilizing compound, which can be administered in any sterile, biocompatible pharmaceutical carrier, including, but not limited to, saline, buffered saline, dextrose, and water. The compositions can be administered to a patient alone, or in combination with other agents, drugs or hormones.

In addition to the active ingredients, these pharmaceutical compositions can contain suitable pharmaceutically-acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. Pharmaceutical compositions of the invention can be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal, parenteral, topical, sublingual, or rectal means. Pharmaceutical compositions for oral administration can be formulated using pharmaceutically acceptable carriers well known in the art in dosages suitable for oral administration. Such carriers enable the pharmaceutical compositions to be for-

mulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions, and the like, for ingestion by the patient.

Pharmaceutical preparations for oral use can be obtained through combination of active compounds with solid excipient, optionally grinding a resulting mixture, and processing the mixture of granules, after adding suitable auxiliaries, if desired, to obtain tablets or dragee cores. Suitable excipients are carbohydrate or protein fillers, such as sugars, including lactose, sucrose, mannitol, or sorbitol; starch from corn, wheat, rice, potato, or other plants; cellulose, such as methyl cellulose, hydroxy-propylmethyl-cellulose, or sodium carboxymethylcellulose; gums including arabic and tragacanth; and proteins such as gelatin and collagen. If desired, disintegrating or solubilizing agents can be added, such as the cross-linked polyvinyl pyrrolidone, agar, alginic acid, or a salt thereof, such as sodium alginate.

Dragee cores can be used in conjunction with suitable coatings, such as concentrated sugar solutions, which also can contain gum arabic, talc, polyvinylpyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments can be added to the tablets or dragee coatings for product identification or to characterize the quantity of active compound, *i.e.*, dosage.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a coating, such as glycerol or sorbitol. Push-fit capsules can contain active ingredients mixed with a filler or binders, such as lactose or starches, lubricants, such as talc or magnesium stearate, and, optionally, stabilizers. In soft capsules, the active compounds can be dissolved or suspended in suitable liquids, such as fatty oils, liquid, or liquid poly-ethylene glycol with or without stabilizers.

Pharmaceutical formulations suitable for parenteral administration can be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks'

solution, Ringer's solution, or physiologically buffered saline. Aqueous injection suspensions can contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Additionally, suspensions of the active compounds can be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or triglycerides, or liposomes. Non-lipid polycationic amino polymers also can be used for delivery. Optionally, the suspension also can contain suitable stabilizers or agents which increase the solubility of the compounds to allow for the preparation of highly concentrated solutions. For topical or nasal administration, penetrants appropriate to the particular barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

The pharmaceutical compositions of the present invention can be manufactured in a manner that is known in the art, *e.g.*, by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping, or lyophilizing processes. The pharmaceutical composition can be provided as a salt and can be formed with many acids, including but not limited to, hydrochloric, sulfuric, acetic, lactic, tartaric, malic, succinic, etc. Salts tend to be more soluble in aqueous or other protonic solvents than are the corresponding free base forms. In other cases, the preferred preparation can be a lyophilized powder which can contain any or all of the following: 1-50 mM histidine, 0.1%-2% sucrose, and 2-7% mannitol, at a pH range of 4.5 to 5.5, that is combined with buffer prior to use.

Further details on techniques for formulation and administration can be found in the latest edition of REMINGTON'S PHARMACEUTICAL SCIENCES (Maack Publishing Co., Easton, Pa.). After pharmaceutical compositions have been prepared, they can be placed in an appropriate container and labeled for treatment of an indicated condition. Such labeling would include amount, frequency, and method of administration.

Therapeutic Indications and Methods

Human tyrosine phosphatase-like enzyme may have a role in the pathogenesis of diseases such as rheumatoid arthritis and osteoarthritis and carcinomas of the intestine, bladder, prostate, breast, stomach, and brain. Inflammatory diseases and cancer can therefore be treated by increasing human tyrosine phosphatase-like enzyme activity through the use of gene therapy or administration of human tyrosine phosphatase-like enzyme or agonists of the enzyme. Alternatively, in cases where overexpression of human tyrosine phosphatase-like enzyme may be associated with a disease condition, enzyme activity may be decreased, for example, by the use of antisense oligonucleotides, ribozymes, antibodies, or antagonists of the enzyme.

Underexpression or overexpression of protein tyrosine phosphatases may result in tumorigenesis in different instances. Therefore gene therapy, using a nucleotide sequence encoding human tyrosine phosphatase-like enzyme may be useful where increased enzyme activity is needed and, conversely, an antisense molecule to a sequence encoding the enzyme may be administered where decreased expression of the enzyme is needed.

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

A reagent which affects tyrosine phosphatase-like enzyme activity can be administered to a human cell, either *in vitro* or *in vivo*, to reduce tyrosine phosphatase-like enzyme activity. The reagent preferably binds to an expression product of a human tyrosine phosphatase-like enzyme gene. If the expression product is a protein, the reagent is preferably an antibody. For treatment of human cells *ex vivo*, an antibody can be added to a preparation of stem cells which have been removed from the body. The cells can then be replaced in the same or another human body, with or without clonal propagation, as is known in the art.

In one embodiment, the reagent is delivered using a liposome. Preferably, the liposome is stable in the animal into which it has been administered for at least about 30 minutes, more preferably for at least about 1 hour, and even more preferably for at least about 24 hours. A liposome comprises a lipid composition that is capable of targeting a reagent, particularly a polynucleotide, to a particular site in an animal, such as a human. Preferably, the lipid composition of the liposome is capable of targeting to a specific organ of an animal, such as the lung, liver, spleen, heart brain, lymph nodes, and skin.

A liposome useful in the present invention comprises a lipid composition that is capable of fusing with the plasma membrane of the targeted cell to deliver its contents to the cell. Preferably, the transfection efficiency of a liposome is about 0.5 μg of DNA per 16 nmole of liposome delivered to about 10^6 cells, more preferably about 1.0 μg of DNA per 16 nmole of liposome delivered to about 10^6 cells, and even more preferably about 2.0 μg of DNA per 16 nmol of liposome delivered to about 10^6 cells. Preferably, a liposome is between about 100 and 500 nm, more preferably between about 150 and 450 nm, and even more preferably between about 200 and 400 nm in diameter.

Suitable liposomes for use in the present invention include those liposomes standardly used in, for example, gene delivery methods known to those of skill in the art. More preferred liposomes include liposomes having a polycationic lipid composition

and/or liposomes having a cholesterol backbone conjugated to polyethylene glycol. Optionally, a liposome comprises a compound capable of targeting the liposome to a tumor cell, such as a tumor cell ligand exposed on the outer surface of the liposome.

Complexing a liposome with a reagent such as an antisense oligonucleotide or ribozyme can be achieved using methods which are standard in the art (see, for example, U.S. Patent 5,705,151). Preferably, from about 0.1 μg to about 10 μg of polynucleotide is combined with about 8 nmol of liposomes, more preferably from about 0.5 μg to about 5 μg of polynucleotides are combined with about 8 nmol liposomes, and even more preferably about 1.0 μg of polynucleotides is combined with about 8 nmol liposomes.

In another embodiment, antibodies can be delivered to specific tissues *in vivo* using receptor-mediated targeted delivery. Receptor-mediated DNA delivery techniques are taught in, for example, Findeis *et al. Trends in Biotechnol. 11*, 202-05 (1993); Chiou *et al.*, GENE THERAPEUTICS: METHODS AND APPLICATIONS OF DIRECT GENE TRANSFER (J.A. Wolff, ed.) (1994); Wu & Wu, *J. Biol. Chem. 263*, 621-24 (1988); Wu *et al.*, *J. Biol. Chem. 269*, 542-46 (1994); Zenke *et al.*, *Proc. Natl. Acad. Sci. U.S.A. 87*, 3655-59 (1990); Wu *et al.*, *J. Biol. Chem. 266*, 338-42 (1991).

Determination of a Therapeutically Effective Dose

The determination of a therapeutically effective dose is well within the capability of those skilled in the art. A therapeutically effective dose refers to that amount of active ingredient which increases or decreases tyrosine phosphatase-like enzyme activity relative to the tyrosine phosphatase-like enzyme activity which occurs in the absence of the therapeutically effective dose.

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays or in animal models, usually mice, rabbits, dogs, or pigs. The animal model also can 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.

Therapeutic efficacy and toxicity, *e.g.*, ED_{50} (the dose therapeutically effective in 50% of the population) and LD_{50} (the dose lethal to 50% of the population), can be determined by standard pharmaceutical procedures in cell cultures or experimental animals. The dose ratio of toxic to therapeutic effects is the therapeutic index, and it can be expressed as the ratio, LD_{50}/ED_{50}.

Pharmaceutical compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies is used in formulating a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that include the ED_{50} with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, 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 that requires treatment. Dosage and administration are adjusted to provide sufficient levels of the active ingredient or to maintain the desired effect. Factors which can be taken into account include the severity of the disease state, general health of the subject, age, weight, and gender of the subject, diet, time and frequency of administration, drug combination(s), reaction sensitivities, and tolerance/response to therapy. Long-acting pharmaceutical compositions can be administered every 3 to 4 days, every week, or once every two weeks depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts can vary from 0.1 to 100,000 micrograms, up to a total dose of about 1 g, depending upon the route of administration. Guidance as to particular dosages and 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 poly-

nucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

If the reagent is a single-chain antibody, polynucleotides encoding the antibody can be constructed and introduced into a cell either *ex vivo* or *in vivo* using well-established techniques including, but not limited to, transferrin-polycation-mediated DNA transfer, transfection with naked or encapsulated nucleic acids, liposome-mediated cellular fusion, intracellular transportation of DNA-coated latex beads, protoplast fusion, viral infection, electroporation, "gene gun," and DEAE- or calcium phosphate-mediated transfection.

Effective *in vivo* dosages of an antibody are in the range of about 5 μg to about 50 μg/kg, about 50 μg to about 5 mg/kg, about 100 μg to about 500 μg/kg of patient body weight, and about 200 to about 250 μg/kg of patient body weight. For administration of polynucleotides encoding single-chain antibodies, effective *in vivo* dosages are in the range of about 100 ng to about 200 ng, 500 ng to about 50 mg, about 1 μg to about 2 mg, about 5 μg to about 500 μg, and about 20 μg to about 100 μg of DNA.

If the expression product is mRNA, the reagent is preferably an antisense oligonucleotide or a ribozyme. Polynucleotides which express antisense oligonucleotides or ribozymes can be introduced into cells by a variety of methods, as described above.

Preferably, a reagent reduces expression of a tyrosine phosphatase-like enzyme gene or the activity of a tyrosine phosphatase-like enzyme polypeptide by at least about 10, preferably about 50, more preferably about 75, 90, or 100% relative to the absence of the reagent. The effectiveness of the mechanism chosen to decrease the level of expression of a tyrosine phosphatase-like enzyme gene or the activity of a tyrosine phosphatase-like enzyme polypeptide can be assessed using methods well known in the art, such as hybridization of nucleotide probes to tyrosine phosphatase-

like enzyme-specific mRNA, quantitative RT-PCR, immunologic detection of a tyrosine phosphatase-like enzyme polypeptide, or measurement of tyrosine phosphatase-like enzyme activity.

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

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

Diagnostic Methods

Tyrosine phosphatase-like enzymes also can be used in diagnostic assays for detecting diseases and abnormalities or susceptibility to diseases and abnormalities related to the presence of mutations in the nucleic acid sequences which encode a tyrosine phosphatase-like enzyme.

Differences can be determined between the cDNA or genomic sequence encoding a tyrosine phosphatase-like enzyme in individuals afflicted with a disease and in normal individuals. If a mutation is observed in some or all of the afflicted individuals but not in normal individuals, then the mutation is likely to be the causative agent of the disease.

Sequence differences between a reference gene and a gene having mutations can be revealed by the direct DNA sequencing method. In addition, cloned DNA segments can be employed as probes to detect specific DNA segments. The sensitivity of this method is greatly enhanced when combined with PCR. For example, a sequencing primer can be used with a double-stranded PCR product or a single-stranded template molecule generated by a modified PCR. The sequence determination is performed by conventional procedures using radiolabeled nucleotides or by automatic sequencing procedures using fluorescent tags.

Genetic testing based on DNA sequence differences can be carried out by detection of alteration in electrophoretic mobility of DNA fragments in gels with or without denaturing agents. Small sequence deletions and insertions can be visualized, for example, by high resolution gel electrophoresis. DNA fragments of different sequences can be distinguished on denaturing formamide gradient gels in which the mobilities of different DNA fragments are retarded in the gel at different positions according to their specific melting or partial melting temperatures (*see, e.g.*, Myers *et al.*, *Science 230*, 1242, 1985). Sequence changes at specific locations can also be revealed by nuclease protection assays, such as RNase and S 1 protection or the chemical cleavage method (*e.g.*, Cotton *et al.*, *Proc. Natl. Acad. Sci. USA 85*, 4397-4401, 1985). Thus, the detection of a specific DNA sequence can be performed by methods such as hybridization, RNase protection, chemical cleavage, direct DNA sequencing or the use of restriction enzymes and Southern blotting of genomic DNA. In addition to direct methods such as gel-electrophoresis and DNA sequencing, mutations can also be detected by *in situ* analysis.

Altered levels of a tyrosine phosphatase-like enzyme also can be detected in various tissues. Assays used to detect levels of the receptor polypeptides in a body sample, such as blood or a tissue biopsy, derived from a host are well known to those of skill in the art and include radioimmunoassays, competitive binding assays, Western blot analysis, and ELISA assays.

All patents and patent applications cited in this disclosure are expressly incorporated herein by reference. The above disclosure generally describes the present invention. A more complete understanding can be obtained by reference to the following specific examples which are provided for purposes of illustration only and are not intended to limit the scope of the invention.

EXAMPLE 1

Detection of tyrosine phosphatase-like enzyme activity

The polynucleotide of SEQ ID NO: 9 is inserted into pGEX vector and expressed as a fusion protein with glutathione S-transferase. The fusion protein is purified from lysed cells by adsorption by glutathion-agarose-beads followed by elution in the presence of free glutathione. The activity of the fusion protein (tyrosine phosphatase-like enzyme polypeptide of SEQ ID NO: 10) is assessed according to the following procedures:

The fusion protein is incubated at 37°C for 2h in 25 1 of 10mM Tris HCL, pH 7,4, containing 7,5 nM tyrosine phosphopeptide (RRLIEDAEpYAARG), and the reaction is terminated by addition of Malachite green solution (UBI). Phosphate release is measured after 15 min by evaluating absorbance at 605 nm. By comparing the absorbance of the fusion protein to the absorbance of a negative standard such as heat-inactivated enzyme and to a positive standard such as tyrosine phosphatase-like enzyme the tyrosine phosphatase-like enzyme activity of the fusion protein (tyrosine phosphatase-like enzyme polypeptide of SEQ ID NO: 10 is demonstrated).

EXAMPLE 2

Identification of test compounds that bind to tyrosine phosphatase-like enzyme polypeptides

Purified tyrosine phosphatase-like enzyme polypeptides comprising a glutathione-S-transferase protein and absorbed onto glutathione-derivatized wells of 96-well microtiter plates are contacted with test compounds from a small molecule library at pH 7.0 in a physiological buffer solution. Tyrosine phosphatase-like enzyme polypeptides comprise an amino acid sequence shown in SEQ ID NO:2, 4, 6, 7, or 9. The

test compounds comprise a fluorescent tag. The samples are incubated for 5 minutes to one hour. Control samples are incubated in the absence of a test compound.

The buffer solution containing the test compounds is washed from the wells. Binding of a test compound to a tyrosine phosphatase-like enzyme polypeptide is detected by fluorescence measurements of the contents of the wells. A test compound which increases the fluorescence in a well by at least 15% relative to fluorescence of a well in which a test compound is not incubated is identified as a compound which binds to a tyrosine phosphatase-like enzyme polypeptide.

EXAMPLE 3

Identification of a test compound which decreases tyrosine phosphatase-like enzyme gene expression

A test compound is administered to a culture of CHO cells transfected with a tyrosine phosphatase-like enzyme expression construct and incubated at 37 °C for 10 to 45 minutes. A culture of the same type of cells incubated for the same time without the test compound provides a negative control.

RNA is isolated from the two cultures as described in Chirgwin *et al.*, *Biochem. 18*, 5294-99, 1979). Northern blots are prepared using 20 to 30 μg total RNA and hybridized with a ^{32}P-labeled tyrosine phosphatase-like enzyme-specific probe at 65 ° C in Express-hyb (CLONTECH). The probe comprises at least 11 contiguous nucleotides selected from the complement of SEQ ID NO:1, 3, 5, or 7. A test compound which decreases the tyrosine phosphatase-like enzyme-specific signal relative to the signal obtained in the absence of the test compound is identified as an inhibitor of tyrosine phosphatase-like enzyme gene expression.

EXAMPLE 4

Treatment of rheumatoid arthritis with human tyrosine phosphatase-like enzyme

A polynucleotide encoding a human tyrosine phosphatase-like enzyme is administered by direct injection into an inflamed joint in a patient with rheumatoid arthritis. The patient's rheumatoid arthritis is monitored over a period of days or weeks. The severity of the rheumatoid arthritis in the treated joint is decreased.

EXAMPLE 5

Effect of a test compound on tyrosine phosphatase activity

Tyrosine phosphatase activity in the presence or absence of a test compound is measured by the hydrolysis of P-nitrophenyl phosphate (PNPP). Human tyrosine phosphatase-like enzyme is incubated together with PNPP in HEPES buffer pH 7.5, in the presence of 0.1% β-mercaptoethanol at 37 °C for 60 minutes in the presence or absence of a test compound. The reaction is stopped by the addition of 6 µl of 10 N NaOH, and the increase in light absorbance at 410 nm of the hydrolyzed PNPP is measured using a spectrophotometer (Diamond *et al.*, *supra*).

CLAIMS

1. An isolated polynucleotide encoding a tyrosine phosphatase-like enzyme polypeptide and being selected from the group consisting of:

a) a polynucleotide encoding a tyrosine phosphatase-like enzyme polypeptide comprising an amino acid sequence selected from the group consisting of:
amino acid sequences which are at least about 50% identical to the amino acid sequence shown in SEQ ID NO: 2;
the amino acid sequence shown in SEQ ID NO: 2;
amino acid sequences which are at least about 50% identical to the amino acid sequence shown in SEQ ID NO: 4;
the amino acid sequence shown in SEQ ID NO:4;
amino acid sequences which are at least about 50% identical to the amino acid sequence shown in SEQ ID NO: 6;
the amino acid sequence shown in SEQ ID NO: 6;
amino acid sequences which are at least about 50% identical to the amino acid sequence shown in SEQ ID NO: 7;
the amino acid sequence shown in SEQ ID NO: 7;
amino acid sequences which are at least about 50% identical to the amino acid sequence shown in SEQ ID NO: 8;
the amino acid sequence shown in SEQ ID NO:8;
amino acid sequences which are at least about 50% identical to the amino acid sequence shown in SEQ ID NO: 10; and
the amino acid sequence shown in SEQ ID NO: 10.

b) a polynucleotide comprising the sequence of SEQ ID NOS: 1, 3, 5 or 9;

c) a polynucleotide which hybridizes under stringent conditions to a polynucleotide specified in (a) and (b);

d) a polynucleotide the sequence of which deviates from the

polynucleotide sequences specified in (a) to (c) due to the degeneration of the genetic code; and

e) a polynucleotide which represents a fragment, derivative or allelic variation of a polynucleotide sequence specified in (a) to (d).

2. An expression vector containing any polynucleotide of claim 1.

3. A host cell containing the expression vector of claim 2.

4. A substantially purified tyrosine phosphatase-like enzyme polypeptide encoded by a polynucleotide of claim 1.

5. A method for producing a tyrosine phosphatase-like enzyme polypeptide, wherein the method comprises the following steps:

a) culturing the host cell of claim 3 under conditions suitable for the expression of the tyrosine phosphatase-like enzyme polypeptide; and

b) recovering the tyrosine phosphatase-like enzyme polypeptide from the host cell culture.

6. A method for detection of a polynucleotide encoding a tyrosine phosphatase-like enzyme polypetide in a biological sample comprising the following steps:

a) hybridizing any polynucleotide of claim 1 to a nucleic acid material of a biological sample, thereby forming a hybridization complex; and

b) detecting said hybridization complex.

7. The method of claim 6, wherein before hybridization, the nucleic acid material of the biological sample is amplified.

8. A method for the detection of a polynucleotide of claim 1 or a tyrosine phosphatase-like enzyme polypeptide of claim 5 comprising the steps of contact-

ing a biological sample with a reagent which specifically interacts with the polynucleotide or the tyrosine phosphatase-like enzyme polypeptide.

9. A diagnostic kit for conducting the method of any one of claims 6 to 8.

10. A method of screening for agents which decrease the activity of a tyrosine phosphatase-like enzyme, comprising the steps of:

contacting a test compound with any tyrosine phosphatase-like enzyme polypeptide encoded by any polynucleotide of claim 1;

detecting binding of the test compound of the tyrosine phosphatase-like enzyme polypeptide, wherein a test compound which binds to the polypeptide is identified as a potential therapeutic agent for decreasing the activity of a tyrosine phosphatase-like enzyme.

11. A method of screening for agents which regulate the activity of a tyrosine phosphatase-like enzyme, comprising the steps of:

contacting a test compound with a tyrosine phosphatase-like enzyme polypeptide encoded by any polynucleotide of claim 1; and

detecting a tyrosine phosphatase-like enzyme activity of the polypeptide, wherein a test compound which increases the tyrosine phosphatase-like enzyme activity is identified as a potential therapeutic agent for increasing the activity of the tyrosine phosphatase-like enzyme, and wherein a test compound which decreases the tyrosine phosphatase-like enzyme activity of the polypeptide is identified as a potential therapeutic agent for decreasing the activity of the tyrosine phosphatase-like enzyme.

12. A method of screening for agents which decrease the activity of a tyrosine phosphatase-like enzyme, comprising the steps of:

contacting a test compound with any polynucleotide of claim 1 and detecting binding of the test compound to the polynucleotide, wherein a test compound which binds to the polynucleotide is identified as a potential therapeutic agent for decreasing the activity of tyrosine phosphatase-like enzyme.

13. A method of reducing the activity of tyrosine phosphatase-like enzyme, comprising the steps of:

contacting a cell with a reagent which specifically binds to any polynucleotide of claim 1 or any tyrosine phosphatase-like enzyme polypeptide of claim 4, whereby the activity of tyrosine phosphatase-like enzyme is reduced.

14. A reagent that modulates the activity of a tyrosine phosphatase-like enzyme polypeptide or a polynucleotide wherein said reagent is identified by the method of any of the claims 10 to 12.

15. A pharmaceutical composition, comprising:

the expression vector of claim 2 or the reagent of claim 14 and a pharmaceutically acceptable carrier.

16. Use of the pharmaceutical composition of claim 15 for modulating the activity of a tyrosine phosphatase-like enzyme in a disease.

17. Use of claim 16 wherein the disease is rheumatoid arthritis, osteoarthritis, intestine-, bladder-, prostate-, breast-, stomach-, and brain-carcinoma.

Fig. 1

GATGTTGGTTGATGACCCCCGGGATCTGGAGCAGATGAATGAAGAGTCTCTGGAAGTCAGCCCAGACATG
TGCATCTACATCACAGAGGACATGCTCATGTCGCGGAACTNGAATGGACACTCTGGGTTGATTGTGAAAG
AAATTGGGTCTTCCACCTCGAGCTCTTCAGAAACAGTTGTTAAGCTTCGTGGCCAGAGTACTGATTCTCT
TCCACAGACTATATGTCGGAAACCAAAGACCTCCACTGATCGACACAGCTTGAGCCTCGATGACATCAGA
CTTTACCAGAAAGACTTCCTGCGCATTGCAGGTCTGTGTCAGGACACTGCTCAGAGTTACACCTTTGGAT
GTGGCCATGAACTGGATGAGGAAGGCCTCTATTGCAACAGTTGCTTGGCCCAGCAGTGCATCAACATCCA
AGATGCTTTTCCAGTCAAAAGAACCAGCAAATACTTTTCTCTGGATCTCACTCATGATGAAGTTCCAGAG
TTTGTTGTGTAAAGTCCGTCTGTGTGCAGCTGTACAG

Fig. 2

MLVDDPRDLEQMNEESLEVSPDMCIYITEDMLMSRNXNGHSGLIVKEIGS
STSSSSETVVKLRGQSTDSLPQTICRKPKTSTDRHSLSLDDIRLYQKDFL
RIAGLCQDTAQSYTFGCGHELDEEGLYCNSCLAQQCINIQDAFPVKRTSK
YFSLDLTHDEVPEFVV*SPSVCSCT (3'UTR)

Fig. 3

ATACTTTTCTCTGGATCTCACTCATGATGAAGTTCCAGAGTTTGTTGTGTAAAGTCCGTCTGTGTGCAGC
TGTACAGGCAGCTTACTGTTTGCTAGAGGATGCGAAAGTCATAAGTTCTTTACATATTACTTGTGCCATA
TCTTCTTCACCCTAAACATAGCTCTTTCTTTATAATATTTGTGATGATGGAAACAAAAGCCTTGGAACAA
TTGCACTTTAAGTATTACACAGAAGTAAAAGAACTACAGAAAATGTACAGCAAGACAAGTGCCCGGAAGT
TCACTGATCCTTCAGAAGGAAATGCGCTTTACTGATTGCAAAGCCTTCAGAATATTGGAGTGTGGTGTGT
TTGCTCATCTGATGCTTTTTAGTTCAGTTACATGTAACATCACATTTTTTTATCACGTGAAAGATGTTAG
ATTTGTTTGCTTATAAATTTTTTACCACTCCCACATAAAATGCTCATAGTTTGGGAGAGGAAGAGGGAAG
ATTCTCTCTTCTTTTAA

Fig. 4

YFSLDLTHDEVPEFVV*SPSVCSCTGSLLFARGCESHKFFTYYLCHIFFT
LNIALSL*YL**WKQKPWNNCTLSITQK*KNYRKCTARQVPGSSLILQKE
MRFTDCKAFRILECGVFAHLMLFSSVTCNITFFYHVKDVRFVCL*IFYHS
HIKCS*FGRGRGKILSSF (3'UTR)

Fig. 5

tgattactgtgagaaggaatgttctcttttggtaat**atg**gaggggggacgcacgaaccgaa
caaccagatgatccagtaaataaatacagtaacacatactgtatatcaagctctgttcaa
agtgcttcacatatgagtaacccgtttaataaacttcacatgaatctcaattcaaaggtt
aagattctgtgtcaccagttgctggtccaggtttgtgacctgctcaggctaaaggactgc
cacctctttggactcagtgttatacaaatcagactatatgacttggaaatatcgggtcag
tttcttctatattctttgtgaggagaacatgtcgtctcctgaaaaaaaatgtacagataa
tgaacatgtgtatatggagttgtcaaaaaagctttacaaatattgtccaaaagaatggaa
gaaagaggccagcaaggtacgacaatacgaagtcacttgggttgtttccaagaggggaa
ggactacatcctgaagcacattccaaacatgcacaaagatcagtttgcactaacagcttc
cgaagctcatcttaaatatatcaaagaggctgtccgactggatgacgtcgctgttcatta
ctacagattgtataaggataaaagggaaattgaagcatcgctgactcttggattgaccat
gaggggaatacagatttttcagaatttagatgaagagaaacaattactttatgatttccc
ctggacaaatgttggaaaattggtgtttgtgggtaagaaatttgagattttgccagatgg
cttgccttctgcccggaagctcatatactacacggggtgccccatgcgctccagacacct
cctgcaacttctgagcaacagccaccgcctctatatgaatctgcagcctgtcctgcgcca
tatccggaagctggaggaaaacgaagagaagaagcagtaccgggaatcttacatcagtga
caacctggacctcgacatggaccagctggaaaaacggtcgcgggccagcgggagcagtgc
gggcagcatgaaacacaagcgcctgtcccgtcattccaccgccagccacagcagttccca
cacctcgggcattgaggcagacaccaagccccgggacacggggccagaagacagctactc
cagcagtgccatccaccgcaagctgaaaacctgcagctcaatgaccagtcatggcagctc
ccacacctcaggggtggagagtggcggcaaagaccggctggaagaggacttacaggacga
tgaaatagagatgttggttgatgacccccgggatctggagcagatgaatgaagagtctct
ggaagtcagcccagacatgtgcatctacatcacagaggacatgctcatgtcgcggaagct
gaatggacactctgggttgattgtgaaagaaattgggtcttccacctcgagctcttcaga
aacagttgttaagcttcgtggccagagtactgattctcttccacagactatatgtcggaa
accaaagacctccactgatcgacacagcttgagcctcgatgacatcagactttaccagaa
agacttcctgcgcattgcaggtctgtgtcaggacactgctcagagttacacctttggatg
tggccatgaactggatgaggaaggcctctattgcaacagttgcttggcccagcagtgcat
caacatccaagatgcttttccagtcaaaagaaccagcaaatacttttctctggatctcac
tcatgatgaagttccagagtttgttgtgt**taa**agtccgtctgtgtgcagctgtacaggcag

Fig. 6

MEGDARTEQPDDPVNSKYSNTYCISSSVQSASH**MSNPFNKLHMNLNSKVKILCHQLLVQV**
CDLLRLKDCHLFGLSVIQIRLYDLEISGQFLLYSFMSSPEKKCTDNEHVYMELSKKLYKY
CPKEWKKEASKVRQYEVTWVVSKRGKDYILKHIPNMHKDQFALTASEAHLKYIKEAVRLD
DVAVHYYRLYKDKREIEASLTLGLTMRGIQIFQNLDEEKQLLYDFPWTNVGKLVFVGKKF
EILPDGLPSARKLIYYTGCPMRSRHLLQLLSNSHRLYMNLQPVLRHIRKLEENEEKKQYR
ESYISDNLDLDMDQLEKRSRASGSSAGSMKHKRLSRHSTASHSSSHTSGIEADTKPRDTG
PEDSYSSSAIHRKLKTCSSMTSHGSSHTSGVESGGKDRLEEDLQDDEIEMLVDDPRDLEQ
MNEESLEVSPDMCIYITEDMLMSRKLNGHSGLIVKEIGSSTSSSSETVVKLRGQSTDSLP
QTICRKPKTSTDRHSLSLDDIRLYQKDFLRIAGLCQDTAQSYTFGCGHELDEEGLYCNSC
LAQQCINIQDAFPVKRTSKYFSLDLTHDEVPEFVV

B41 or ERM domain

Fig. 7

MSNPFNKLHMNLNSKVKILCHQLLVQV
CDLLRLKDCHLFGLSVIQIRLYDLEISGQFLLYSFMSSPEKKCTDNEHVYMELSKKLYKY
CPKEWKKEASKVRQYEVTWVVSKRGKDYILKHIPNMHKDQFALTASEAHLKYIKEAVRLD
DVAVHYY

Fig. 8

MLVDDPRDLEQ
MNEESLEVSPDMCIYITEDMLMSRKLNGHSGLIVKEIGSSTSSSSETVVKLRGQSTDSLP
QTICRKPKTSTDRHSLSLDDIRLYQKDFLRIAGLCQDTAQSYTFGCGHELDEEGLYCNSC
LAQQCINIQDAFPVKRTSKYFSLDLTHDEVPEFVV

Fig. 9

atggaggggg acgcacgaac cgaacaacca gatgatccag taaatagtaa atacagtaac
acatactgta tatcaagctc tgttcaaagt gcttcacata tgagtaaccc gtttaataaa
cttcacatga atctcaattc aaaggttaag attctgtgtc accagttgct ggtccaggtt
tgtgacctgc tcaggctaaa ggactgccac ctctttggac tcagtgttat acaaaataat
gaacatgtgt atatggagtt gtcacaaaag ctttacaaat attgtccaaa agaatggaag
aaagaggcca gcaaggtacg acaatacgaa gtcacttggg gtatcgacca atttgggcct
cctatgatca tccacttccg tgtgcagtac tatgtggaaa atggcagatt gatcagtgac
agagcagcaa gatactatta ttactggcac ctgagaaaac aagttcttca ttctcagtgt
gtgctccaag aggaggccta cttcctgctg gcagcctttg ccctgcaggc tgatcttggg
aacttcaaaa ggaataagca ctatggaaaa tacttcgagc cagaggctta cttcccatct
tgggttgttt ccaagagggg gaaggactac atcctgaagc acattccaaa catgcacaaa
gatcagtttg cactaacagc ttccgaagct catcttaaat atatcaaaga ggctgtccga
ctggatgacg tcgctgttca ttactacaga ttgtataagg ataaaaggga aattgaagca
tcgctgactc ttggattgac catgagggga atacagattt ttcagaattt agatgaagag
aaacaattac tttatgattt cccctggaca aatgttggaa aattggtgtt tgtgggtaag
aaatttgaga ttttgccaga tggcttgcct tctgcccgga agctcatata ctacacgggg
tgccccatgc gctccagaca cctcctgcaa cttctgagca acagccaccg cctctatatg
aatctgcagc ctgtcctgcg ccatatccgg aagctggagg aaaacgaaga gaagaagcag
taccgggaat cttacatcag tgacaacctg gacctcgaca tggaccagct ggaaaaacgg
tcgcgggcca gcgggagcag tgcgggcagc atgaaacaca agcgcctgtc ccgtcattcc
accgccagcc acagcagttc ccacacctcg ggcattgagg cagacaccaa gccccgggac
acggggccag aagacagcta ctccagcagt gccatccacc gcaagctgaa aacctgcagc
tcaatgacca gtcatggcag ctcccacacc tcaggggtgg agagtggcgg caaagaccgg
ctggaagagg acttacagga cgatgaaata gagatgttgg ttgatgaccc ccgggatctg
gagcagatga atgaagagtc tctggaagtc agcccagaca tgtgcatcta catcacagag
gacatgctca tgtcgcggaa gctgaatgga cactctgggt tgattgtgaa agaaattggg
tcttccacct cgagctcttc agaaacagtt gttaagcttc gtggccagag tactgattct
cttccacaga ctatatgtcg gaaaccaaag acctccactg atcgacacag cttgagcctc
gatgacatca gactttacca gaaagacttc ctgcgcattg caggtctgtg tcaggacact
gctcagagtt acacctttgg atgtggccat gaactggatg aggaaggcct ctattgcaac
agttgcttgg cccagcagtg catcaacatc caagatgctt ttccagtcaa aagaaccagc
aaatactttt ctctggatct cactcatgat gaagttccag agtttgttgt gtaa

Fig. 10

MEGDARTEQP DDPVNSKYSN TYCISSSVQS ASHMSNPFNK LHMNLNSKVK ILCHQLLVQV
CDLLRLKDCH LFGLSVIQNN EHVYMELSQK LYKYCPKEWK KEASKVRQYE VTWGIDQFGP
PMIIHFRVQY YVENGRLISD RAARYYYYWH LRKQVLHSQC VLQEEAYFLL AAFALQADLG
NFKRNKHYGK YFEPEAYFPS WVVSKRGKDY ILKHIPNMHK DQFALTASEA HLKYIKEAVR
LDDVAVHYYR LYKDKREIEA SLTLGLTMRG IQIFQNLDEE KQLLYDFPWT NVGKLVFVGK
KFEILPDGLP SARKLIYYTG CPMRSRHLLQ LLSNSHRLYM NLQPVLRHIR KLEENEEKKQ
YRESYISDNL DLDMDQLEKR SRASGSSAGS MKHKRLSRHS TASHSSSHTS GIEADTKPRD
TGPEDSYSSS AIHRKLKTCS SMTSHGSSHT SGVESGGKDR LEEDLQDDEI EMLVDDPRDL
EQMNEESLEV SPDMCIYITE DMLMSRKLNG HSGLIVKEIG SSTSSSSETV VKLRGQSTDS
LPQTICRKPK TSTDRHSLSL DDIRLYQKDF LRIAGLCQDT AQSYTFGCGH ELDEEGLYCN
SCLAQQCINI QDAFPVKRTS KYFSLDLTHD EVPEFVV

SEQUENCE LISTING

<110> Bayer AG
Bull, Christof

<120> REGULATION OF HUMAN TYROSINE PHOSPHATASE-LIKE ENZYME

<130> Lio043 foreign countries

<150> 60/192,315
<151> 2000-03-27

<160> 10

<170> PatentIn version 3.0

<210> 1
<211> 527
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> ()..()
<223> n = a, t, g, c

<400> 1
gatgttggtt gatgaccccc gggatctgga gcagatgaat gaagagtctc tggaagtcag 60

cccagacatg tgcatctaca tcacagagga catgctcatg tcgcggaact ngaatggaca 120

ctctgggttg attgtgaaag aaattgggtc ttccacctcg agctcttcag aaacagttgt 180

taagcttcgt ggccagagta ctgattctct tccacagact atatgtcgga aaccaaagac 240

ctccactgat cgacacagct tgagcctcga tgacatcaga ctttaccaga aagacttcct 300

gcgcattgca ggtctgtgtc aggacactgc tcagagttac acctttggat gtggccatga 360

actggatgag gaaggcctct attgcaacag ttgcttggcc cagcagtgca tcaacatcca 420

agatgctttt ccagtcaaaa gaaccagcaa atacttttct ctggatctca ctcatgatga 480

agttccagag tttgttgtgt aaagtccgtc tgtgtgcagc tgtacag 527

<210> 2
<211> 174
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<222> ()..()
<223> Xaa = any amino acid

<400> 2

Met Leu Val Asp Asp Pro Arg Asp Leu Glu Gln Met Asn Glu Glu Ser
1 5 10 15

```
Leu Glu Val Ser Pro Asp Met Cys Ile Tyr Ile Thr Glu Asp Met Leu
            20                  25                  30

Met Ser Arg Asn Xaa Asn Gly His Ser Gly Leu Ile Val Lys Glu Ile
        35                  40                  45

Gly Ser Ser Thr Ser Ser Ser Ser Glu Thr Val Val Lys Leu Arg Gly
    50                  55                  60

Gln Ser Thr Asp Ser Leu Pro Gln Thr Ile Cys Arg Lys Pro Lys Thr
65                  70                  75                  80

Ser Thr Asp Arg His Ser Leu Ser Leu Asp Asp Ile Arg Leu Tyr Gln
                85                  90                  95

Lys Asp Phe Leu Arg Ile Ala Gly Leu Cys Gln Asp Thr Ala Gln Ser
            100                 105                 110

Tyr Thr Phe Gly Cys Gly His Glu Leu Asp Glu Glu Gly Leu Tyr Cys
        115                 120                 125

Asn Ser Cys Leu Ala Gln Gln Cys Ile Asn Ile Gln Asp Ala Phe Pro
    130                 135                 140

Val Lys Arg Thr Ser Lys Tyr Phe Ser Leu Asp Leu Thr His Asp Glu
145                 150                 155                 160

Val Pro Glu Phe Val Val Ser Pro Ser Val Cys Ser Cys Thr
                165                 170
```

<210> 3
<211> 507
<212> DNA
<213> Homo sapiens

<400> 3

```
atacttttct ctggatctca ctcatgatga agttccagag tttgttgtgt aaagtccgtc     60

tgtgtgcagc tgtacaggca gcttactgtt tgctagagga tgcgaaagtc ataagttctt    120

tacatattac ttgtgccata tcttcttcac cctaaacata gctctttctt tataatattt    180

gtgatgatgg aaacaaaagc cttggaacaa ttgcacttta agtattacac agaagtaaaa    240

gaactacaga aaatgtacag caagacaagt gcccggaagt tcactgatcc ttcagaagga    300

aatgcgcttt actgattgca aagccttcag aatattggag tgtggtgtgt ttgctcatct    360

gatgcttttt agttcagtta catgtaacat cacatttttt tatcacgtga aagatgttag    420

atttgtttgc ttataaattt tttaccactc ccacataaaa tgctcatagt ttgggagagg    480

aagagggaag attctctctt cttttaa                                        507
```

<210> 4
<211> 161
<212> PRT
<213> Homo sapiens

<400> 4

```
Tyr Phe Ser Leu Asp Leu Thr His Asp Glu Val Pro Glu Phe Val Val
```

1 5 10 15

Ser Pro Ser Val Cys Ser Cys Thr Gly Ser Leu Leu Phe Ala Arg Gly
20 25 30

Cys Glu Ser His Lys Phe Phe Thr Tyr Tyr Leu Cys His Ile Phe Phe
35 40 45

Thr Leu Asn Ile Ala Leu Ser Leu Tyr Leu Trp Lys Gln Lys Pro Trp
50 55 60

Asn Asn Cys Thr Leu Ser Ile Thr Gln Lys Lys Asn Tyr Arg Lys Cys
65 70 75 80

Thr Ala Arg Gln Val Pro Gly Ser Ser Leu Ile Leu Gln Lys Glu Met
85 90 95

Arg Phe Thr Asp Cys Lys Ala Phe Arg Ile Leu Glu Cys Gly Val Phe
100 105 110

Ala His Leu Met Leu Phe Ser Ser Val Thr Cys Asn Ile Thr Phe Phe
115 120 125

Tyr His Val Lys Asp Val Arg Phe Val Cys Leu Ile Phe Tyr His Ser
130 135 140

His Ile Lys Cys Ser Phe Gly Arg Gly Arg Gly Lys Ile Leu Ser Ser
145 150 155 160

Phe

<210> 5
<211> 1800
<212> DNA
<213> Homo sapiens

<400> 5
tgattactgt gagaaggaat gttctctttt ggtaatatgg agggggacgc acgaaccgaa 60

caaccagatg atccagtaaa taaatacagt aacacatact gtatatcaag ctctgttcaa 120

agtgcttcac atatgagtaa cccgtttaat aaacttcaca tgaatctcaa ttcaaaggtt 180

aagattctgt gtcaccagtt gctggtccag gtttgtgacc tgctcaggct aaaggactgc 240

cacctctttg gactcagtgt tatacaaatc agactatatg acttggaaat atcgggtcag 300

tttcttctat attctttgtg aggagaacat gtcgtctcct gaaaaaaaat gtacagataa 360

tgaacatgtg tatatggagt tgtcaaaaaa gctttacaaa tattgtccaa aagaatggaa 420

gaaagaggcc agcaaggtac gacaatacga agtcacttgg gttgtttcca agaggggaa 480

ggactacatc ctgaagcaca ttccaaacat gcacaaagat cagtttgcac taacagcttc 540

cgaagctcat cttaaatata tcaaagaggc tgtccgactg gatgacgtcg ctgttcatta 600

ctacagattg tataaggata aaagggaaat tgaagcatcg ctgactcttg gattgaccat 660

gaggggaata cagatttttc agaatttaga tgaagagaaa caattacttt atgatttccc 720

ctggacaaat gttggaaaat tggtgtttgt gggtaagaaa tttgagattt tgccagatgg 780

cttgccttct gcccggaagc tcatatacta cacggggtgc cccatgcgct ccagacacct 840

cctgcaactt ctgagcaaca gccaccgcct ctatatgaat ctgcagcctg tcctgcgcca 900

tatccggaag ctggaggaaa acgaagagaa gaagcagtac cgggaatctt acatcagtga 960

caacctggac ctcgacatgg accagctgga aaaacggtcg cgggccagcg ggagcagtgc 1020

gggcagcatg aaacacaagc gcctgtcccg tcattccacc gccagccaca gcagttccca 1080

cacctcgggc attgaggcag acaccaagcc ccgggacacg gggccagaag acagctactc 1140

cagcagtgcc atccaccgca agctgaaaac ctgcagctca atgaccagtc atggcagctc 1200

ccacacctca ggggtggaga gtggcggcaa agaccggctg gaagaggact tacaggacga 1260

tgaaatagag atgttggttg atgacccccg ggatctggag cagatgaatg aagagtctct 1320

ggaagtcagc ccagacatgt gcatctacat cacagaggac atgctcatgt cgcggaagct 1380

gaatggacac tctgggttga ttgtgaaaga aattgggtct tccacctcga gctcttcaga 1440

aacagttgtt aagcttcgtg gccagagtac tgattctctt ccacagacta tatgtcggaa 1500

accaaagacc tccactgatc gacacagctt gagcctcgat gacatcagac tttaccagaa 1560

agacttcctg cgcattgcag gtctgtgtca ggacactgct cagagttaca cctttggatg 1620

tggccatgaa ctggatgagg aaggcctcta ttgcaacagt tgcttggccc agcagtgcat 1680

caacatccaa gatgcttttc cagtcaaaag aaccagcaaa tacttttctc tggatctcac 1740

tcatgatgaa gttccagagt ttgttgtgta aagtccgtct gtgtgcagct gtacaggcag 1800

<210> 6
<211> 575
<212> PRT
<213> Homo sapiens

<400> 6

Met Glu Gly Asp Ala Arg Thr Glu Gln Pro Asp Asp Pro Val Asn Ser
1 5 10 15

Lys Tyr Ser Asn Thr Tyr Cys Ile Ser Ser Ser Val Gln Ser Ala Ser
20 25 30

His Met Ser Asn Pro Phe Asn Lys Leu His Met Asn Leu Asn Ser Lys
35 40 45

Val Lys Ile Leu Cys His Gln Leu Leu Val Gln Val Cys Asp Leu Leu
50 55 60

Arg Leu Lys Asp Cys His Leu Phe Gly Leu Ser Val Ile Gln Ile Arg
65 70 75 80

Leu Tyr Asp Leu Glu Ile Ser Gly Gln Phe Leu Leu Tyr Ser Phe Met
85 90 95

Ser Ser Pro Glu Lys Lys Cys Thr Asp Asn Glu His Val Tyr Met Glu
100 105 110

Leu Ser Lys Lys Leu Tyr Lys Tyr Cys Pro Lys Glu Trp Lys Lys Glu
115 120 125

Ala Ser Lys Val Arg Gln Tyr Glu Val Thr Trp Val Val Ser Lys Arg
130 135 140

Gly Lys Asp Tyr Ile Leu Lys His Ile Pro Asn Met His Lys Asp Gln
145 150 155 160

Phe Ala Leu Thr Ala Ser Glu Ala His Leu Lys Tyr Ile Lys Glu Ala
165 170 175

Val Arg Leu Asp Asp Val Ala Val His Tyr Tyr Arg Leu Tyr Lys Asp
180 185 190

Lys Arg Glu Ile Glu Ala Ser Leu Thr Leu Gly Leu Thr Met Arg Gly
195 200 205

Ile Gln Ile Phe Gln Asn Leu Asp Glu Glu Lys Gln Leu Leu Tyr Asp
210 215 220

Phe Pro Trp Thr Asn Val Gly Lys Leu Val Phe Val Gly Lys Lys Phe
225 230 235 240

Glu Ile Leu Pro Asp Gly Leu Pro Ser Ala Arg Lys Leu Ile Tyr Tyr
245 250 255

Thr Gly Cys Pro Met Arg Ser Arg His Leu Leu Gln Leu Leu Ser Asn
260 265 270

Ser His Arg Leu Tyr Met Asn Leu Gln Pro Val Leu Arg His Ile Arg
275 280 285

Lys Leu Glu Glu Asn Glu Glu Lys Lys Gln Tyr Arg Glu Ser Tyr Ile
290 295 300

Ser Asp Asn Leu Asp Leu Asp Met Asp Gln Leu Glu Lys Arg Ser Arg
305 310 315 320

Ala Ser Gly Ser Ser Ala Gly Ser Met Lys His Lys Arg Leu Ser Arg
325 330 335

His Ser Thr Ala Ser His Ser Ser Ser His Thr Ser Gly Ile Glu Ala
340 345 350

Asp Thr Lys Pro Arg Asp Thr Gly Pro Glu Asp Ser Tyr Ser Ser Ser
355 360 365

Ala Ile His Arg Lys Leu Lys Thr Cys Ser Ser Met Thr Ser His Gly
370 375 380

Ser Ser His Thr Ser Gly Val Glu Ser Gly Gly Lys Asp Arg Leu Glu
385 390 395 400

Glu Asp Leu Gln Asp Asp Glu Ile Glu Met Leu Val Asp Asp Pro Arg
405 410 415

Asp Leu Glu Gln Met Asn Glu Glu Ser Leu Glu Val Ser Pro Asp Met
420 425 430

Cys Ile Tyr Ile Thr Glu Asp Met Leu Met Ser Arg Lys Leu Asn Gly
435 440 445

```
His Ser Gly Leu Ile Val Lys Glu Ile Gly Ser Ser Thr Ser Ser Ser
    450                 455                 460

Ser Glu Thr Val Val Lys Leu Arg Gly Gln Ser Thr Asp Ser Leu Pro
465                 470                 475                 480

Gln Thr Ile Cys Arg Lys Pro Lys Thr Ser Thr Asp Arg His Ser Leu
                485                 490                 495

Ser Leu Asp Asp Ile Arg Leu Tyr Gln Lys Asp Phe Leu Arg Ile Ala
            500                 505                 510

Gly Leu Cys Gln Asp Thr Ala Gln Ser Tyr Thr Phe Gly Cys Gly His
        515                 520                 525

Glu Leu Asp Glu Glu Gly Leu Tyr Cys Asn Ser Cys Leu Ala Gln Gln
    530                 535                 540

Cys Ile Asn Ile Gln Asp Ala Phe Pro Val Lys Arg Thr Ser Lys Tyr
545                 550                 555                 560

Phe Ser Leu Asp Leu Thr His Asp Glu Val Pro Glu Phe Val Val
                565                 570                 575
```

<210> 7
<211> 154
<212> PRT
<213> Homo sapiens

<400> 7

```
Met Ser Asn Pro Phe Asn Lys Leu His Met Asn Leu Asn Ser Lys Val
1               5                   10                  15

Lys Ile Leu Cys His Gln Leu Leu Val Gln Val Cys Asp Leu Leu Arg
            20                  25                  30

Leu Lys Asp Cys His Leu Phe Gly Leu Ser Val Ile Gln Ile Arg Leu
        35                  40                  45

Tyr Asp Leu Glu Ile Ser Gly Gln Phe Leu Leu Tyr Ser Phe Met Ser
    50                  55                  60

Ser Pro Glu Lys Lys Cys Thr Asp Asn Glu His Val Tyr Met Glu Leu
65                  70                  75                  80

Ser Lys Lys Leu Tyr Lys Tyr Cys Pro Lys Glu Trp Lys Lys Glu Ala
                85                  90                  95

Ser Lys Val Arg Gln Tyr Glu Val Thr Trp Val Val Ser Lys Arg Gly
            100                 105                 110

Lys Asp Tyr Ile Leu Lys His Ile Pro Asn Met His Lys Asp Gln Phe
        115                 120                 125

Ala Leu Thr Ala Ser Glu Ala His Leu Lys Tyr Ile Lys Glu Ala Val
    130                 135                 140

Arg Leu Asp Asp Val Ala Val His Tyr Tyr
145                 150
```

<210> 8

<211> 166
<212> PRT
<213> Homo sapiens

<400> 8

Met Leu Val Asp Asp Pro Arg Asp Leu Glu Gln Met Asn Glu Glu Ser
1 5 10 15

Leu Glu Val Ser Pro Asp Met Cys Ile Tyr Ile Thr Glu Asp Met Leu
20 25 30

Met Ser Arg Lys Leu Asn Gly His Ser Gly Leu Ile Val Lys Glu Ile
35 40 45

Gly Ser Ser Thr Ser Ser Ser Ser Glu Thr Val Val Lys Leu Arg Gly
50 55 60

Gln Ser Thr Asp Ser Leu Pro Gln Thr Ile Cys Arg Lys Pro Lys Thr
65 70 75 80

Ser Thr Asp Arg His Ser Leu Ser Leu Asp Asp Ile Arg Leu Tyr Gln
85 90 95

Lys Asp Phe Leu Arg Ile Ala Gly Leu Cys Gln Asp Thr Ala Gln Ser
100 105 110

Tyr Thr Phe Gly Cys Gly His Glu Leu Asp Glu Glu Gly Leu Tyr Cys
115 120 125

Asn Ser Cys Leu Ala Gln Gln Cys Ile Asn Ile Gln Asp Ala Phe Pro
130 135 140

Val Lys Arg Thr Ser Lys Tyr Phe Ser Leu Asp Leu Thr His Asp Glu
145 150 155 160

Val Pro Glu Phe Val Val
165

<210> 9
<211> 1914
<212> DNA
<213> Homo sapiens

<400> 9
atggaggggg acgcacgaac cgaacaacca gatgatccag taaatagtaa atacagtaac 60

acatactgta tatcaagctc tgttcaaagt gcttcacata tgagtaaccc gtttaataaa 120

cttcacatga atctcaattc aaaggttaag attctgtgtc accagttgct ggtccaggtt 180

tgtgacctgc tcaggctaaa ggactgccac ctctttggac tcagtgttat acaaaataat 240

gaacatgtgt atatggagtt gtcacaaaag ctttacaaat attgtccaaa agaatggaag 300

aaagaggcca gcaaggtacg acaatacgaa gtcacttggg gtatcgacca atttgggcct 360

cctatgatca tccacttccg tgtgcagtac tatgtggaaa atggcagatt gatcagtgac 420

agagcagcaa gatactatta ttactggcac ctgagaaaac aagttcttca ttctcagtgt 480

gtgctccaag aggaggccta cttcctgctg gcagcctttg ccctgcaggc tgatcttggg 540

aacttcaaaa ggaataagca ctatggaaaa tacttcgagc cagaggctta cttcccatct 600

tgggttgttt ccaagagggg gaaggactac atcctgaagc acattccaaa catgcacaaa 660

gatcagtttg cactaacagc ttccgaagct catcttaaat atatcaaaga ggctgtccga 720

ctggatgacg tcgctgttca ttactacaga ttgtataagg ataaaaggga aattgaagca 780

tcgctgactc ttggattgac catgagggga atacagattt ttcagaattt agatgaagag 840

aaacaattac tttatgattt cccctggaca aatgttggaa aattggtgtt tgtgggtaag 900

aaatttgaga ttttgccaga tggcttgcct tctgcccgga agctcatata ctacacgggg 960

tgccccatgc gctccagaca cctcctgcaa cttctgagca acagccaccg cctctatatg 1020

aatctgcagc ctgtcctgcg ccatatccgg aagctggagg aaaacgaaga gaagaagcag 1080

taccgggaat cttacatcag tgacaacctg gacctcgaca tggaccagct ggaaaaacgg 1140

tcgcgggcca gcgggagcag tgcgggcagc atgaaacaca agcgcctgtc ccgtcattcc 1200

accgccagcc acagcagttc ccacacctcg ggcattgagg cagacaccaa gccccgggac 1260

acggggccag aagacagcta ctccagcagt gccatccacc gcaagctgaa aacctgcagc 1320

tcaatgacca gtcatggcag ctcccacacc tcaggggtgg agagtggcgg caaagaccgg 1380

ctggaagagg acttacagga cgatgaaata gagatgttgg ttgatgaccc ccgggatctg 1440

gagcagatga atgaagagtc tctggaagtc agcccagaca tgtgcatcta catcacagag 1500

gacatgctca tgtcgcggaa gctgaatgga cactctgggt tgattgtgaa agaaattggg 1560

tcttccacct cgagctcttc agaaacagtt gttaagcttc gtggccagag tactgattct 1620

cttccacaga ctatatgtcg gaaaccaaag acctccactg atcgacacag cttgagcctc 1680

gatgacatca gactttacca gaaagacttc ctgcgcattg caggtctgtg tcaggacact 1740

gctcagagtt acacctttgg atgtggccat gaactggatg aggaaggcct ctattgcaac 1800

agttgcttgg cccagcagtg catcaacatc caagatgctt ttccagtcaa aagaaccagc 1860

aaatactttt ctctggatct cactcatgat gaagttccag agtttgttgt gtaa 1914

<210> 10
<211> 637
<212> PRT
<213> Homo sapiens

<400> 10

Met Glu Gly Asp Ala Arg Thr Glu Gln Pro Asp Asp Pro Val Asn Ser
1 5 10 15

Lys Tyr Ser Asn Thr Tyr Cys Ile Ser Ser Ser Val Gln Ser Ala Ser
20 25 30

His Met Ser Asn Pro Phe Asn Lys Leu His Met Asn Leu Asn Ser Lys
35 40 45

```
Val Lys Ile Leu Cys His Gln Leu Leu Val Gln Val Cys Asp Leu Leu
    50                  55                  60

Arg Leu Lys Asp Cys His Leu Phe Gly Leu Ser Val Ile Gln Asn Asn
65                  70                  75                  80

Glu His Val Tyr Met Glu Leu Ser Gln Lys Leu Tyr Lys Tyr Cys Pro
                85                  90                  95

Lys Glu Trp Lys Lys Glu Ala Ser Lys Val Arg Gln Tyr Glu Val Thr
            100                 105                 110

Trp Gly Ile Asp Gln Phe Gly Pro Pro Met Ile Ile His Phe Arg Val
        115                 120                 125

Gln Tyr Tyr Val Glu Asn Gly Arg Leu Ile Ser Asp Arg Ala Ala Arg
    130                 135                 140

Tyr Tyr Tyr Tyr Trp His Leu Arg Lys Gln Val Leu His Ser Gln Cys
145                 150                 155                 160

Val Leu Gln Glu Glu Ala Tyr Phe Leu Leu Ala Ala Phe Ala Leu Gln
                165                 170                 175

Ala Asp Leu Gly Asn Phe Lys Arg Asn Lys His Tyr Gly Lys Tyr Phe
            180                 185                 190

Glu Pro Glu Ala Tyr Phe Pro Ser Trp Val Val Ser Lys Arg Gly Lys
        195                 200                 205

Asp Tyr Ile Leu Lys His Ile Pro Asn Met His Lys Asp Gln Phe Ala
    210                 215                 220

Leu Thr Ala Ser Glu Ala His Leu Lys Tyr Ile Lys Glu Ala Val Arg
225                 230                 235                 240

Leu Asp Asp Val Ala Val His Tyr Tyr Arg Leu Tyr Lys Asp Lys Arg
                245                 250                 255

Glu Ile Glu Ala Ser Leu Thr Leu Gly Leu Thr Met Arg Gly Ile Gln
            260                 265                 270

Ile Phe Gln Asn Leu Asp Glu Glu Lys Gln Leu Leu Tyr Asp Phe Pro
        275                 280                 285

Trp Thr Asn Val Gly Lys Leu Val Phe Val Gly Lys Lys Phe Glu Ile
    290                 295                 300

Leu Pro Asp Gly Leu Pro Ser Ala Arg Lys Leu Ile Tyr Tyr Thr Gly
305                 310                 315                 320

Cys Pro Met Arg Ser Arg His Leu Leu Gln Leu Leu Ser Asn Ser His
                325                 330                 335

Arg Leu Tyr Met Asn Leu Gln Pro Val Leu Arg His Ile Arg Lys Leu
            340                 345                 350

Glu Glu Asn Glu Glu Lys Lys Gln Tyr Arg Glu Ser Tyr Ile Ser Asp
        355                 360                 365

Asn Leu Asp Leu Asp Met Asp Gln Leu Glu Lys Arg Ser Arg Ala Ser
    370                 375                 380
```

```
Gly Ser Ser Ala Gly Ser Met Lys His Lys Arg Leu Ser Arg His Ser
385                 390                 395                 400

Thr Ala Ser His Ser Ser Ser His Thr Ser Gly Ile Glu Ala Asp Thr
                405                 410                 415

Lys Pro Arg Asp Thr Gly Pro Glu Asp Ser Tyr Ser Ser Ser Ala Ile
            420                 425                 430

His Arg Lys Leu Lys Thr Cys Ser Ser Met Thr Ser His Gly Ser Ser
        435                 440                 445

His Thr Ser Gly Val Glu Ser Gly Gly Lys Asp Arg Leu Glu Glu Asp
    450                 455                 460

Leu Gln Asp Asp Glu Ile Glu Met Leu Val Asp Asp Pro Arg Asp Leu
465                 470                 475                 480

Glu Gln Met Asn Glu Glu Ser Leu Glu Val Ser Pro Asp Met Cys Ile
                485                 490                 495

Tyr Ile Thr Glu Asp Met Leu Met Ser Arg Lys Leu Asn Gly His Ser
            500                 505                 510

Gly Leu Ile Val Lys Glu Ile Gly Ser Ser Thr Ser Ser Ser Ser Glu
        515                 520                 525

Thr Val Val Lys Leu Arg Gly Gln Ser Thr Asp Ser Leu Pro Gln Thr
    530                 535                 540

Ile Cys Arg Lys Pro Lys Thr Ser Thr Asp Arg His Ser Leu Ser Leu
545                 550                 555                 560

Asp Asp Ile Arg Leu Tyr Gln Lys Asp Phe Leu Arg Ile Ala Gly Leu
                565                 570                 575

Cys Gln Asp Thr Ala Gln Ser Tyr Thr Phe Gly Cys Gly His Glu Leu
            580                 585                 590

Asp Glu Glu Gly Leu Tyr Cys Asn Ser Cys Leu Ala Gln Gln Cys Ile
        595                 600                 605

Asn Ile Gln Asp Ala Phe Pro Val Lys Arg Thr Ser Lys Tyr Phe Ser
    610                 615                 620

Leu Asp Leu Thr His Asp Glu Val Pro Glu Phe Val Val
625                 630                 635
```