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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
20 February 2003 (20.02.2003)

PCT

(10) International Publication Number
WO 03/014299 A2

(51) International Patent Classification[7]: C12N

(21) International Application Number: PCT/US02/24623

(22) International Filing Date: 2 August 2002 (02.08.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/310,362 6 August 2001 (06.08.2001) US
60/357,501 15 February 2002 (15.02.2002) US

(71) Applicant *(for all designated States except US)*: **EXELIXIS, INC.** [US/US]; P.O. Box 511, 170 Harbor Way, South San Francisco, CA 94083-0511 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **FRIEDMAN, Lori** [US/US]; One Bayside Village Place, Unit 212, San Francisco, CA 94107 (US). **PLOWMAN, Gregory, D.** [US/US]; 35 Winding Way, San Carlos, CA 94070 (US). **COSTA, Michael, A.** [US/US]; 18 Hazelwood Avenue, San Francisco, CA 94112 (US). **LI, Danxi** [CN/US]; 90 Behr Avenue # 302, San Francisco, CA 94131 (US). **FUNKE, Roel, P.** [NL/US]; 343 California Avenue, South San Francisco, CA 94080 (US). **HUNG, Tak** [CN/US]; 1072 Gull Avenue, Foster City, CA 94404 (US).

(74) Agents: **BRUNELLE, Jan** et al.; Exelixis, Inc., P. O. Box 511,, 170 Harbor Way, South San Francisco, CA 94083-0511 (US).

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

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

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

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

WO 03/014299 A2

(54) Title: SPHKs AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE

(57) Abstract: Human SPHK genes are identified as modulators of the p53 pathway, and thus are therapeutic targets for disorders associated with defective p53 function. Methods for identifying modulators of p53, comprising screening for agents that modulate the activity of SPHK are provided.

SPHKs AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE

REFERENCE TO RELATED APPLICATIONS

This application claims priority to U.S. provisional patent applications 60/310,362 filed 8/6/2001, and 60/357,501 filed 2/15/2002. The contents of the prior applications are hereby incorporated in their entirety.

BACKGROUND OF THE INVENTION

The p53 gene is mutated in over 50 different types of human cancers, including familial and spontaneous cancers, and is believed to be the most commonly mutated gene in human cancer (Zambetti and Levine, FASEB (1993) 7:855-865; Hollstein, *et al.*, Nucleic Acids Res. (1994) 22:3551-3555). Greater than 90% of mutations in the p53 gene are missense mutations that alter a single amino acid that inactivates p53 function. Aberrant forms of human p53 are associated with poor prognosis, more aggressive tumors, metastasis, and short survival rates (Mitsudomi et al., Clin Cancer Res 2000 Oct; 6(10):4055-63; Koshland, Science (1993) 262:1953).

The human p53 protein normally functions as a central integrator of signals including DNA damage, hypoxia, nucleotide deprivation, and oncogene activation (Prives, Cell (1998) 95:5-8). In response to these signals, p53 protein levels are greatly increased with the result that the accumulated p53 activates cell cycle arrest or apoptosis depending on the nature and strength of these signals. Indeed, multiple lines of experimental evidence have pointed to a key role for p53 as a tumor suppressor (Levine, Cell (1997) 88:323-331). For example, homozygous p53 "knockout" mice are developmentally normal but exhibit nearly 100% incidence of neoplasia in the first year of life (Donehower *et al.*, Nature (1992) 356:215-221).

The biochemical mechanisms and pathways through which p53 functions in normal and cancerous cells are not fully understood, but one clearly important aspect of p53 function is its activity as a gene-specific transcriptional activator. Among the genes with known p53-response elements are several with well-characterized roles in either regulation of the cell cycle or apoptosis, including GADD45, p21/Waf1/Cip1, cyclin G, Bax, IGF-BP3, and MDM2 (Levine, Cell (1997) 88:323-331).

Sphingosine-1-phosphate (SPP) is a lipid messenger containing both intracellular and extracellular functions. Intracellularly, it mediates proliferation and survival, and extracellularly, it is a ligand for EDG1 (Meyer zu Heringdorf, D. et al. (2001) Eur J

Pharmacol 414, 145-54; Ancellin, N. et al. (2002) J Biol Chem 277, 6667-75). A variety of stimuli can increase cellular levels of SPP by activating sphingosine kinase (SPHK), the enzyme that initiates the phosphorylation of sphingosine. Inhibitors of SPHK can block formation of SPP and inhibit cellular proliferation induced by a variety of factors, including platelet-derived growth factor and serum (Banno, Y. et al. (1998) Biochem J 335, 301-4).

Sphingosine kinase-1 (SPHK1) phosphorylates sphingosine to form sphingosine 1-phosphate, a lipid messenger mediating signaling pathways involved in diverse functions like regulating intracellular Ca^{2+} mobilization and cell proliferation and cell survival (Meyer zu Heringdorf, D. et al. (1998) Embo Journal 17, 2830-7; Nava, V. et al. (2000) FEBS Lett 473, 81-4).

Sphingosine kinase 2 (SPHK2) phosphorylates sphingosine to form sphingosine 1-phosphate, which functions as a second messenger regulating proliferation and survival, as well as a ligand for selected G protein-coupled receptors (Meyer zu Heringdorf, D. et al. (1998) *supra*; Nava, V. et al. (2000) *supra*; Liu, H. et al. (2000) J Biol Chem; 275(26): 19513-20). Expression of SPHK2 in HEK 293 cells elevates SPP levels (Liu, H. et al. (2000) *supra*).

SPHK sequences are highly conserved among evolutionarily diverse organisms such as human, *Caenorhabditis elegans*, and yeast (Kohama, T. et al. (1998) J. Biol. Chem. 273: 23722-23728).

The ability to manipulate the genomes of model organisms such as *C. elegans* provides a powerful means to analyze biochemical processes that, due to significant evolutionary conservation, have direct relevance to more complex vertebrate organisms. Due to a high level of gene and pathway conservation, the strong similarity of cellular processes, and the functional conservation of genes between these model organisms and mammals, identification of the involvement of novel genes in particular pathways and their functions in such model organisms can directly contribute to the understanding of the correlative pathways and methods of modulating them in mammals (see, for example, Dulubova I, et al, J Neurochem 2001 Apr;77(1):229-38; Cai T, et al., Diabetologia 2001 Jan;44(1):81-8; Pasquinelli AE, et al., Nature. 2000 Nov 2;408(6808):37-8; Ivanov IP, et al., EMBO J 2000 Apr 17;19(8):1907-17; Vajo Z et al., Mamm Genome 1999 Oct;10(10):1000-4). For example, a genetic screen can be carried out in an invertebrate model organism having underexpression (e.g. knockout) or overexpression of a gene (referred to as a "genetic entry point") that yields a visible phenotype. Additional genes are mutated in a random or

targeted manner. When a gene mutation changes the original phenotype caused by the mutation in the genetic entry point, the gene is identified as a "modifier" involved in the same or overlapping pathway as the genetic entry point. When the genetic entry point is an ortholog of a human gene implicated in a disease pathway, such as p53, modifier genes can be identified that may be attractive candidate targets for novel therapeutics.

All references cited herein, including sequence information in referenced Genbank identifier numbers and website references, are incorporated herein in their entireties.

SUMMARY OF THE INVENTION

We have discovered genes that modify the p53 pathway in *C. elegans*, and identified their human orthologs, hereinafter referred to as SPHK. The invention provides methods for utilizing these p53 modifier genes and polypeptides to identify SPHK-modulating agents that are candidate therapeutic agents that can be used in the treatment of disorders associated with defective or impaired p53 function and/or SPHK function. Preferred SPHK-modulating agents specifically bind to SPHK polypeptides and restore p53 function. Other preferred SPHK-modulating agents are nucleic acid modulators such as antisense oligomers and RNAi that repress SPHK gene expression or product activity by, for example, binding to and inhibiting the respective nucleic acid (i.e. DNA or mRNA).

SPHK modulating agents may be evaluated by any convenient *in vitro* or *in vivo* assay for molecular interaction with an SPHK polypeptide or nucleic acid. In one embodiment, candidate SPHK modulating agents are tested with an assay system comprising a SPHK polypeptide or nucleic acid. Agents that produce a change in the activity of the assay system relative to controls are identified as candidate p53 modulating agents. The assay system may be cell-based or cell-free. SPHK-modulating agents include SPHK related proteins (e.g. dominant negative mutants, and biotherapeutics); SPHK-specific antibodies; SPHK-specific antisense oligomers and other nucleic acid modulators; and chemical agents that specifically bind to or interact with SPHK or compete with SPHK binding partner (e.g. by binding to an SPHK binding partner). In one specific embodiment, a small molecule modulator is identified using a kinase assay. In specific embodiments, the screening assay system is selected from a binding assay, an apoptosis assay, a cell proliferation assay, an angiogenesis assay, and a hypoxic induction assay.

In another embodiment, candidate p53 pathway modulating agents are further tested using a second assay system that detects changes in the p53 pathway, such as angiogenic, apoptotic, or cell proliferation changes produced by the originally identified candidate

agent or an agent derived from the original agent. The second assay system may use cultured cells or non-human animals. In specific embodiments, the secondary assay system uses non-human animals, including animals predetermined to have a disease or disorder implicating the p53 pathway, such as an angiogenic, apoptotic, or cell proliferation disorder (e.g. cancer).

The invention further provides methods for modulating the SPHK function and/or the p53 pathway in a mammalian cell by contacting the mammalian cell with an agent that specifically binds a SPHK polypeptide or nucleic acid. The agent may be a small molecule modulator, a nucleic acid modulator, or an antibody and may be administered to a mammalian animal predetermined to have a pathology associated the p53 pathway.

DETAILED DESCRIPTION OF THE INVENTION

Genetic screens were designed to identify modifiers of the p53 pathway in *C. elegans,* where a homozygous p53 deletion mutant was used. Various specific genes were silenced by RNA inhibition (RNAi). Methods for using RNAi to silence genes in *C. elegans* are known in the art (Fire A, et al., 1998 Nature 391:806-811; Fire, A. Trends Genet. 15, 358-363 (1999); WO9932619). Genes causing altered phenotypes in the worms were identified as modifiers of the p53 pathway. A modifier of particular interest, C34C6.5 was identified followed by identification of its human orthologs. Accordingly, vertebrate orthologs of these modifiers, and preferably the human orthologs, sphingosine kinase (SPHK) genes (i.e., nucleic acids and polypeptides) are attractive drug targets for the treatment of pathologies associated with a defective p53 signaling pathway, such as cancer.

In vitro and in vivo methods of assessing SPHK function are provided herein. Modulation of the SPHK or their respective binding partners is useful for understanding the association of the p53 pathway and its members in normal and disease conditions and for developing diagnostics and therapeutic modalities for p53 related pathologies. SPHK-modulating agents that act by inhibiting or enhancing SPHK expression, directly or indirectly, for example, by affecting an SPHK function such as enzymatic (e.g., catalytic) or binding activity, can be identified using methods provided herein. SPHK modulating agents are useful in diagnosis, therapy and pharmaceutical development.

Nucleic acids and polypeptides of the invention

Sequences related to SPHK nucleic acids and polypeptides that can be used in the invention are disclosed in Genbank (referenced by Genbank identifier (GI) number) as GI#s 11464966 (SEQ ID NO:1), 9910561 (SEQ ID NO:3), 19923819 (SEQ ID NO:4), 12052923 (SEQ ID NO:6), 13359166 (SEQ ID NO:7), 18594429 (SEQ ID NO:8), and 12232440 (SEQ ID NO:9) for nucleic acid, and GI#s 11464967 (SEQ ID NO:10), 9910562 (SEQ ID NO:11), 13359167 (SEQ ID NO:12), and 20336726 (SEQ ID NO:13) for polypeptides. Additionally, sequences of clones N18H08 (SEQ ID NO:2) and N31F03 (SEQ ID NO:5) can also be used in the methods of invention.

SPHKs are kinase proteins with kinase domains. The term "SPHK polypeptide" refers to a full-length SPHK protein or a functionally active fragment or derivative thereof. A "functionally active" SPHK fragment or derivative exhibits one or more functional activities associated with a full-length, wild-type SPHK protein, such as antigenic or immunogenic activity, enzymatic activity, ability to bind natural cellular substrates, etc. The functional activity of SPHK proteins, derivatives and fragments can be assayed by various methods known to one skilled in the art (Current Protocols in Protein Science (1998) Coligan *et al.*, eds., John Wiley & Sons, Inc., Somerset, New Jersey) and as further discussed below. For purposes herein, functionally active fragments also include those fragments that comprise one or more structural domains of an SPHK, such as a kinase domain or a binding domain. Protein domains can be identified using the PFAM program (Bateman A., et al., Nucleic Acids Res, 1999, 27:260-2). For example, the kinase domains of SPHKs from GI#s 11461967, 9910562, and 20336726 (SEQ ID NOs:10, 11, and 13, respectively) are located respectively at approximately amino acid residues 16-153, 146 to 283, and 132 to 278. Methods for obtaining SPHK polypeptides are also further described below. In some embodiments, preferred fragments are functionally active, domain-containing fragments comprising at least 25 contiguous amino acids, preferably at least 50, more preferably 75, and most preferably at least 100 contiguous amino acids of any one of SEQ ID NOs:10, 11, 12, or 13 (an SPHK). In further preferred embodiments, the fragment comprises the entire kinase (functionally active) domain.

The term "SPHK nucleic acid" refers to a DNA or RNA molecule that encodes a SPHK polypeptide. Preferably, the SPHK polypeptide or nucleic acid or fragment thereof is from a human, but can also be an ortholog, or derivative thereof with at least 70% sequence identity, preferably at least 80%, more preferably 85%, still more preferably 90%, and most preferably at least 95% sequence identity with SPHK. Normally, orthologs

in different species retain the same function, due to presence of one or more protein motifs and/or 3-dimensional structures. Orthologs are generally identified by sequence homology analysis, such as BLAST analysis, usually using protein bait sequences. Sequences are assigned as a potential ortholog if the best hit sequence from the forward BLAST result retrieves the original query sequence in the reverse BLAST (Huynen MA and Bork P, Proc Natl Acad Sci (1998) 95:5849-5856; Huynen MA *et al.*, Genome Research (2000) 10:1204-1210). Programs for multiple sequence alignment, such as CLUSTAL (Thompson JD et al, 1994, Nucleic Acids Res 22:4673-4680) may be used to highlight conserved regions and/or residues of orthologous proteins and to generate phylogenetic trees. In a phylogenetic tree representing multiple homologous sequences from diverse species (e.g., retrieved through BLAST analysis), orthologous sequences from two species generally appear closest on the tree with respect to all other sequences from these two species. Structural threading or other analysis of protein folding (e.g., using software by ProCeryon, Biosciences, Salzburg, Austria) may also identify potential orthologs. In evolution, when a gene duplication event follows speciation, a single gene in one species, such as *C. elegans*, may correspond to multiple genes (paralogs) in another, such as human. As used herein, the term "orthologs" encompasses paralogs. As used herein, "percent (%) sequence identity" with respect to a subject sequence, or a specified portion of a subject sequence, is defined as the percentage of nucleotides or amino acids in the candidate derivative sequence identical with the nucleotides or amino acids in the subject sequence (or specified portion thereof), after aligning the sequences and introducing gaps, if necessary to achieve the maximum percent sequence identity, as generated by the program WU-BLAST-2.0a19 (Altschul *et al.*, J. Mol. Biol. (1997) 215:403-410) with all the search parameters set to default values. The HSP S and HSP S2 parameters are dynamic values and are established by the program itself depending upon the composition of the particular sequence and composition of the particular database against which the sequence of interest is being searched. A % identity value is determined by the number of matching identical nucleotides or amino acids divided by the sequence length for which the percent identity is being reported. "Percent (%) amino acid sequence similarity" is determined by doing the same calculation as for determining % amino acid sequence identity, but including conservative amino acid substitutions in addition to identical amino acids in the computation.

A conservative amino acid substitution is one in which an amino acid is substituted for another amino acid having similar properties such that the folding or activity of the protein

is not significantly affected. Aromatic amino acids that can be substituted for each other are phenylalanine, tryptophan, and tyrosine; interchangeable hydrophobic amino acids are leucine, isoleucine, methionine, and valine; interchangeable polar amino acids are glutamine and asparagine; interchangeable basic amino acids are arginine, lysine and histidine; interchangeable acidic amino acids are aspartic acid and glutamic acid; and interchangeable small amino acids are alanine, serine, threonine, cysteine and glycine.

Alternatively, an alignment for nucleic acid sequences is provided by the local homology algorithm of Smith and Waterman (Smith and Waterman, 1981, Advances in Applied Mathematics 2:482-489; database: European Bioinformatics Institute; Smith and Waterman, 1981, J. of Molec.Biol., 147:195-197; Nicholas et al., 1998, "A Tutorial on Searching Sequence Databases and Sequence Scoring Methods" (www.psc.edu) and references cited therein.; W.R. Pearson, 1991, Genomics 11:635-650). This algorithm can be applied to amino acid sequences by using the scoring matrix developed by Dayhoff (Dayhoff: Atlas of Protein Sequences and Structure, M. O. Dayhoff ed., 5 suppl. 3:353-358, National Biomedical Research Foundation, Washington, D.C., USA), and normalized by Gribskov (Gribskov 1986 Nucl. Acids Res. 14(6):6745-6763). The Smith-Waterman algorithm may be employed where default parameters are used for scoring (for example, gap open penalty of 12, gap extension penalty of two). From the data generated, the "Match" value reflects "sequence identity."

Derivative nucleic acid molecules of the subject nucleic acid molecules include sequences that hybridize to the nucleic acid sequence of SEQ ID NOs:1, 2, 3, 4, 5, 6, 7, 8, or 9. The stringency of hybridization can be controlled by temperature, ionic strength, pH, and the presence of denaturing agents such as formamide during hybridization and washing. Conditions routinely used are set out in readily available procedure texts (*e.g.*, Current Protocol in Molecular Biology, Vol. 1, Chap. 2.10, John Wiley & Sons, Publishers (1994); Sambrook *et al.*, Molecular Cloning, Cold Spring Harbor (1989)). In some embodiments, a nucleic acid molecule of the invention is capable of hybridizing to a nucleic acid molecule containing the nucleotide sequence of any one of SEQ ID NOs:1, 2, 3, 4, 5, 6, 7, 8, or 9 under stringent hybridization conditions that comprise: prehybridization of filters containing nucleic acid for 8 hours to overnight at 65° C in a solution comprising 6X single strength citrate (SSC) (1X SSC is 0.15 M NaCl, 0.015 M Na citrate; pH 7.0), 5X Denhardt's solution, 0.05% sodium pyrophosphate and 100 μg/ml herring sperm DNA; hybridization for 18-20 hours at 65° C in a solution containing 6X SSC, 1X Denhardt's solution, 100 μg/ml yeast tRNA and 0.05% sodium pyrophosphate;

and washing of filters at 65° C for 1h in a solution containing 0.2X SSC and 0.1% SDS (sodium dodecyl sulfate).

In other embodiments, moderately stringent hybridization conditions are used that comprise: pretreatment of filters containing nucleic acid for 6 h at 40° C in a solution containing 35% formamide, 5X SSC, 50 mM Tris-HCl (pH7.5), 5mM EDTA, 0.1% PVP, 0.1% Ficoll, 1% BSA, and 500 μg/ml denatured salmon sperm DNA; hybridization for 18-20h at 40° C in a solution containing 35% formamide, 5X SSC, 50 mM Tris-HCl (pH7.5), 5mM EDTA, 0.02% PVP, 0.02% Ficoll, 0.2% BSA, 100 μg/ml salmon sperm DNA, and 10% (wt/vol) dextran sulfate; followed by washing twice for 1 hour at 55° C in a solution containing 2X SSC and 0.1% SDS.

Alternatively, low stringency conditions can be used that comprise: incubation for 8 hours to overnight at 37° C in a solution comprising 20% formamide, 5 x SSC, 50 mM sodium phosphate (pH 7.6), 5X Denhardt's solution, 10% dextran sulfate, and 20 μg/ml denatured sheared salmon sperm DNA; hybridization in the same buffer for 18 to 20 hours; and washing of filters in 1 x SSC at about 37° C for 1 hour.

Isolation, Production, Expression, and Mis-expression of SPHK Nucleic Acids and Polypeptides

SPHK nucleic acids and polypeptides, useful for identifying and testing agents that modulate SPHK function and for other applications related to the involvement of SPHK in the p53 pathway. SPHK nucleic acids and derivatives and orthologs thereof may be obtained using any available method. For instance, techniques for isolating cDNA or genomic DNA sequences of interest by screening DNA libraries or by using polymerase chain reaction (PCR) are well known in the art. In general, the particular use for the protein will dictate the particulars of expression, production, and purification methods. For instance, production of proteins for use in screening for modulating agents may require methods that preserve specific biological activities of these proteins, whereas production of proteins for antibody generation may require structural integrity of particular epitopes. Expression of proteins to be purified for screening or antibody production may require the addition of specific tags (*e.g.*, generation of fusion proteins). Overexpression of an SPHK protein for assays used to assess SPHK function, such as involvement in cell cycle regulation or hypoxic response, may require expression in eukaryotic cell lines capable of these cellular activities. Techniques for the expression, production, and purification of proteins are well known in the art; any suitable means therefore may be

used (e.g., Higgins SJ and Hames BD (eds.) Protein Expression: A Practical Approach, Oxford University Press Inc., New York 1999; Stanbury PF et al., Principles of Fermentation Technology, 2nd edition, Elsevier Science, New York, 1995; Doonan S (ed.) Protein Purification Protocols, Humana Press, New Jersey, 1996; Coligan JE et al, Current Protocols in Protein Science (eds.), 1999, John Wiley & Sons, New York). In particular embodiments, recombinant SPHK is expressed in a cell line known to have defective p53 function (e.g. SAOS-2 osteoblasts, H1299 lung cancer cells, C33A and HT3 cervical cancer cells, HT-29 and DLD-1 colon cancer cells, among others, available from American Type Culture Collection (ATCC), Manassas, VA). The recombinant cells are used in cell-based screening assay systems of the invention, as described further below.

The nucleotide sequence encoding an SPHK polypeptide can be inserted into any appropriate expression vector. The necessary transcriptional and translational signals, including promoter/enhancer element, can derive from the native SPHK gene and/or its flanking regions or can be heterologous. A variety of host-vector expression systems may be utilized, such as mammalian cell systems infected with virus (*e.g.* vaccinia virus, adenovirus, *etc.*); insect cell systems infected with virus (*e.g.* baculovirus); microorganisms such as yeast containing yeast vectors, or bacteria transformed with bacteriophage, plasmid, or cosmid DNA. A host cell strain that modulates the expression of, modifies, and/or specifically processes the gene product may be used.

To detect expression of the SPHK gene product, the expression vector can comprise a promoter operably linked to an SPHK gene nucleic acid, one or more origins of replication, and, one or more selectable markers (*e.g.* thymidine kinase activity, resistance to antibiotics, *etc.*). Alternatively, recombinant expression vectors can be identified by assaying for the expression of the SPHK gene product based on the physical or functional properties of the SPHK protein in *in vitro* assay systems (*e.g.* immunoassays).

The SPHK protein, fragment, or derivative may be optionally expressed as a fusion, or chimeric protein product (i.e. it is joined via a peptide bond to a heterologous protein sequence of a different protein), for example to facilitate purification or detection. A chimeric product can be made by ligating the appropriate nucleic acid sequences encoding the desired amino acid sequences to each other using standard methods and expressing the chimeric product. A chimeric product may also be made by protein synthetic techniques, *e.g.* by use of a peptide synthesizer (Hunkapiller *et al.*, Nature (1984) 310:105-111).

Once a recombinant cell that expresses the SPHK gene sequence is identified, the gene product can be isolated and purified using standard methods (*e.g.* ion exchange, affinity,

and gel exclusion chromatography; centrifugation; differential solubility; electrophoresis). Alternatively, native SPHK proteins can be purified from natural sources, by standard methods (*e.g.* immunoaffinity purification). Once a protein is obtained, it may be quantified and its activity measured by appropriate methods, such as immunoassay, bioassay, or other measurements of physical properties, such as crystallography.

The methods of this invention may also use cells that have been engineered for altered expression (mis-expression) of SPHK or other genes associated with the p53 pathway. As used herein, mis-expression encompasses ectopic expression, over-expression, under-expression, and non-expression (*e.g.* by gene knock-out or blocking expression that would otherwise normally occur).

Genetically modified animals

Animal models that have been genetically modified to alter SPHK expression may be used in *in vivo* assays to test for activity of a candidate p53 modulating agent, or to further assess the role of SPHK in a p53 pathway process such as apoptosis or cell proliferation. Preferably, the altered SPHK expression results in a detectable phenotype, such as decreased or increased levels of cell proliferation, angiogenesis, or apoptosis compared to control animals having normal SPHK expression. The genetically modified animal may additionally have altered p53 expression (e.g. p53 knockout). Preferred genetically modified animals are mammals such as primates, rodents (preferably mice), cows, horses, goats, sheep, pigs, dogs and cats. Preferred non-mammalian species include zebrafish, *C. elegans*, and *Drosophila*. Preferred genetically modified animals are transgenic animals having a heterologous nucleic acid sequence present as an extrachromosomal element in a portion of its cells, i.e. mosaic animals (see, for example, techniques described by Jakobovits, 1994, Curr. Biol. 4:761-763.) or stably integrated into its germ line DNA (i.e., in the genomic sequence of most or all of its cells). Heterologous nucleic acid is introduced into the germ line of such transgenic animals by genetic manipulation of, for example, embryos or embryonic stem cells of the host animal.

Methods of making transgenic animals are well-known in the art (for transgenic mice see Brinster et al., Proc. Nat. Acad. Sci. USA 82: 4438-4442 (1985), U.S. Pat. Nos. 4,736,866 and 4,870,009, both by Leder et al., U.S. Pat. No. 4,873,191 by Wagner et al., and Hogan, B., Manipulating the Mouse Embryo, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., (1986); for particle bombardment see U.S. Pat. No., 4,945,050, by Sandford *et al.*; for transgenic *Drosophila* see Rubin and Spradling, Science (1982)

218:348-53 and U.S. Pat. No. 4,670,388; for transgenic insects see Berghammer A.J. *et al.*, A Universal Marker for Transgenic Insects (1999) Nature 402:370-371; for transgenic Zebrafish see Lin S., Transgenic Zebrafish, Methods Mol Biol. (2000);136:375-3830); for microinjection procedures for fish, amphibian eggs and birds see Houdebine and Chourrout, Experientia (1991) 47:897-905; for transgenic rats see Hammer *et al.*, Cell (1990) 63:1099-1112; and for culturing of embryonic stem (ES) cells and the subsequent production of transgenic animals by the introduction of DNA into ES cells using methods such as electroporation, calcium phosphate/DNA precipitation and direct injection see, e.g., Teratocarcinomas and Embryonic Stem Cells, A Practical Approach, E. J. Robertson, ed., IRL Press (1987)). Clones of the nonhuman transgenic animals can be produced according to available methods (see Wilmut, I. *et al.* (1997) Nature 385:810-813; and PCT International Publication Nos. WO 97/07668 and WO 97/07669).

In one embodiment, the transgenic animal is a "knock-out" animal having a heterozygous or homozygous alteration in the sequence of an endogenous SPHK gene that results in a decrease of SPHK function, preferably such that SPHK expression is undetectable or insignificant. Knock-out animals are typically generated by homologous recombination with a vector comprising a transgene having at least a portion of the gene to be knocked out. Typically a deletion, addition or substitution has been introduced into the transgene to functionally disrupt it. The transgene can be a human gene (e.g., from a human genomic clone) but more preferably is an ortholog of the human gene derived from the transgenic host species. For example, a mouse SPHK gene is used to construct a homologous recombination vector suitable for altering an endogenous SPHK gene in the mouse genome. Detailed methodologies for homologous recombination in mice are available (see Capecchi, Science (1989) 244:1288-1292; Joyner *et al.*, Nature (1989) 338:153-156). Procedures for the production of non-rodent transgenic mammals and other animals are also available (Houdebine and Chourrout, *supra*; Pursel *et al.*, Science (1989) 244:1281-1288; Simms *et al.*, Bio/Technology (1988) 6:179-183). In a preferred embodiment, knock-out animals, such as mice harboring a knockout of a specific gene, may be used to produce antibodies against the human counterpart of the gene that has been knocked out (Claesson MH et al., (1994) Scan J Immunol 40:257-264; Declerck PJ et al., (1995) J Biol Chem. 270:8397-400).

In another embodiment, the transgenic animal is a "knock-in" animal having an alteration in its genome that results in altered expression (e.g., increased (including ectopic) or decreased expression) of the SPHK gene, e.g., by introduction of additional

copies of SPHK, or by operatively inserting a regulatory sequence that provides for altered expression of an endogenous copy of the SPHK gene. Such regulatory sequences include inducible, tissue-specific, and constitutive promoters and enhancer elements. The knock-in can be homozygous or heterozygous.

Transgenic nonhuman animals can also be produced that contain selected systems allowing for regulated expression of the transgene. One example of such a system that may be produced is the cre/loxP recombinase system of bacteriophage P1 (Lakso *et al.*, PNAS (1992) 89:6232-6236; U.S. Pat. No. 4,959,317). If a cre/loxP recombinase system is used to regulate expression of the transgene, animals containing transgenes encoding both the Cre recombinase and a selected protein are required. Such animals can be provided through the construction of "double" transgenic animals, e.g., by mating two transgenic animals, one containing a transgene encoding a selected protein and the other containing a transgene encoding a recombinase. Another example of a recombinase system is the FLP recombinase system of Saccharomyces cerevisiae (O'Gorman et al. (1991) Science 251:1351-1355; U.S. Pat. No. 5,654,182). In a preferred embodiment, both Cre-LoxP and Flp-Frt are used in the same system to regulate expression of the transgene, and for sequential deletion of vector sequences in the same cell (Sun X et al (2000) Nat Genet 25:83-6).

The genetically modified animals can be used in genetic studies to further elucidate the p53 pathway, as animal models of disease and disorders implicating defective p53 function, and for *in vivo* testing of candidate therapeutic agents, such as those identified in screens described below. The candidate therapeutic agents are administered to a genetically modified animal having altered SPHK function and phenotypic changes are compared with appropriate control animals such as genetically modified animals that receive placebo treatment, and/or animals with unaltered SPHK expression that receive candidate therapeutic agent.

In addition to the above-described genetically modified animals having altered SPHK function, animal models having defective p53 function (and otherwise normal SPHK function), can be used in the methods of the present invention. For example, a p53 knockout mouse can be used to assess, *in vivo*, the activity of a candidate p53 modulating agent identified in one of the *in vitro* assays described below. p53 knockout mice are described in the literature (Jacks et al., Nature 2001;410:1111-1116, 1043-1044; Donehower *et al.*, supra). Preferably, the candidate p53 modulating agent when administered to a model system with cells defective in p53 function, produces a detectable

phenotypic change in the model system indicating that the p53 function is restored, i.e., the cells exhibit normal cell cycle progression.

Modulating Agents

The invention provides methods to identify agents that interact with and/or modulate the function of SPHK and/or the p53 pathway. Modulating agents identified by the methods are also part of the invention. Such agents are useful in a variety of diagnostic and therapeutic applications associated with the p53 pathway, as well as in further analysis of the SPHK protein and its contribution to the p53 pathway. Accordingly, the invention also provides methods for modulating the p53 pathway comprising the step of specifically modulating SPHK activity by administering a SPHK-interacting or -modulating agent.

As used herein, an "SPHK-modulating agent" is any agent that modulated SPHK function, for example, an agent that interacts with SPHK to inhibit or enhance SPHK activity or otherwise affect normal SPHK function. SPHK function can be affected at any level, including transcription, protein expression, protein localization, and cellular or extra-cellular activity. In a preferred embodiment, the SPHK - modulating agent specifically modulates the function of the SPHK. The phrases "specific modulating agent", "specifically modulates", etc., are used herein to refer to modulating agents that directly bind to the SPHK polypeptide or nucleic acid, and preferably inhibit, enhance, or otherwise alter, the function of the SPHK. These phrases also encompasses modulating agents that alter the interaction of the SPHK with a binding partner, substrate, or cofactor (e.g. by binding to a binding partner of an SPHK, or to a protein/binding partner complex, and altering SPHK function). In a further preferred embodiment, the SPHK- modulating agent is a modulator of the p53 pathway (e.g. it restores and/or upregulates p53 function) and thus is also a p53-modulating agent.

Preferred SPHK-modulating agents include small molecule compounds; SPHK-interacting proteins, including antibodies and other biotherapeutics; and nucleic acid modulators such as antisense and RNA inhibitors. The modulating agents may be formulated in pharmaceutical compositions, for example, as compositions that may comprise other active ingredients, as in combination therapy, and/or suitable carriers or excipients. Techniques for formulation and administration of the compounds may be found in "Remington's Pharmaceutical Sciences" Mack Publishing Co., Easton, PA, 19th edition.

Small molecule modulators

Small molecules, are often preferred to modulate function of proteins with enzymatic function, and/or containing protein interaction domains. Chemical agents, referred to in the art as "small molecule" compounds are typically organic, non-peptide molecules, having a molecular weight less than 10,000, preferably less than 5,000, more preferably less than 1,000, and most preferably less than 500. This class of modulators includes chemically synthesized molecules, for instance, compounds from combinatorial chemical libraries. Synthetic compounds may be rationally designed or identified based on known or inferred properties of the SPHK protein or may be identified by screening compound libraries. Alternative appropriate modulators of this class are natural products, particularly secondary metabolites from organisms such as plants or fungi, which can also be identified by screening compound libraries for SPHK–modulating activity. Methods for generating and obtaining compounds are well known in the art (Schreiber SL, Science (2000) 151: 1964-1969; Radmann J and Gunther J, Science (2000) 151:1947-1948).

Small molecule modulators identified from screening assays, as described below, can be used as lead compounds from which candidate clinical compounds may be designed, optimized, and synthesized. Such clinical compounds may have utility in treating pathologies associated with the p53 pathway. The activity of candidate small molecule modulating agents may be improved several-fold through iterative secondary functional validation, as further described below, structure determination, and candidate modulator modification and testing. Additionally, candidate clinical compounds are generated with specific regard to clinical and pharmacological properties. For example, the reagents may be derivatized and re-screened using *in vitro* and *in vivo* assays to optimize activity and minimize toxicity for pharmaceutical development.

Protein Modulators

Specific SPHK-interacting proteins are useful in a variety of diagnostic and therapeutic applications related to the p53 pathway and related disorders, as well as in validation assays for other SPHK-modulating agents. In a preferred embodiment, SPHK-interacting proteins affect normal SPHK function, including transcription, protein expression, protein localization, and cellular or extra-cellular activity. In another embodiment, SPHK-interacting proteins are useful in detecting and providing information about the function of SPHK proteins, as is relevant to p53 related disorders, such as cancer (e.g., for diagnostic means).

An SPHK-interacting protein may be endogenous, i.e. one that naturally interacts genetically or biochemically with an SPHK, such as a member of the SPHK pathway that modulates SPHK expression, localization, and/or activity. SPHK-modulators include dominant negative forms of SPHK-interacting proteins and of SPHK proteins themselves. Yeast two-hybrid and variant screens offer preferred methods for identifying endogenous SPHK-interacting proteins (Finley, R. L. et al. (1996) in DNA Cloning-Expression Systems: A Practical Approach, eds. Glover D. & Hames B. D (Oxford University Press, Oxford, England), pp. 169-203; Fashema SF et al., Gene (2000) 250:1-14; Drees BL Curr Opin Chem Biol (1999) 3:64-70; Vidal M and Legrain P Nucleic Acids Res (1999) 27:919-29; and U.S. Pat. No. 5,928,868). Mass spectrometry is an alternative preferred method for the elucidation of protein complexes (reviewed in, e.g., Pandley A and Mann M, Nature (2000) 405:837-846; Yates JR 3rd, Trends Genet (2000) 16:5-8).

An SPHK-interacting protein may be an exogenous protein, such as an SPHK-specific antibody or a T-cell antigen receptor (see, e.g., Harlow and Lane (1988) Antibodies, A Laboratory Manual, Cold Spring Harbor Laboratory; Harlow and Lane (1999) Using antibodies: a laboratory manual. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press). SPHK antibodies are further discussed below.

In preferred embodiments, an SPHK-interacting protein specifically binds an SPHK protein. In alternative preferred embodiments, an SPHK-modulating agent binds an SPHK substrate, binding partner, or cofactor.

Antibodies

In another embodiment, the protein modulator is an SPHK specific antibody agonist or antagonist. The antibodies have therapeutic and diagnostic utilities, and can be used in screening assays to identify SPHK modulators. The antibodies can also be used in dissecting the portions of the SPHK pathway responsible for various cellular responses and in the general processing and maturation of the SPHK.

Antibodies that specifically bind SPHK polypeptides can be generated using known methods. Preferably the antibody is specific to a mammalian ortholog of SPHK polypeptide, and more preferably, to human SPHK. Antibodies may be polyclonal, monoclonal (mAbs), humanized or chimeric antibodies, single chain antibodies, Fab fragments, F(ab').sub.2 fragments, fragments produced by a FAb expression library, anti-idiotypic (anti-Id) antibodies, and epitope-binding fragments of any of the above. Epitopes of SPHK which are particularly antigenic can be selected, for example, by

routine screening of SPHK polypeptides for antigenicity or by applying a theoretical method for selecting antigenic regions of a protein (Hopp and Wood (1981), Proc. Nati. Acad. Sci. U.S.A. 78:3824-28; Hopp and Wood, (1983) Mol. Immunol. 20:483-89; Sutcliffe et al., (1983) Science 219:660-66) to the amino acid sequence shown in any of SEQ ID NOs:10, 11, 12, or 13. Monoclonal antibodies with affinities of 10^8 M^{-1} preferably 10^9 M^{-1} to 10^{10} M^{-1}, or stronger can be made by standard procedures as described (Harlow and Lane, *supra*; Goding (1986) Monoclonal Antibodies: Principles and Practice (2d ed) Academic Press, New York; and U.S. Pat. Nos. 4,381,292; 4,451,570; and 4,618,577). Antibodies may be generated against crude cell extracts of SPHK or substantially purified fragments thereof. If SPHK fragments are used, they preferably comprise at least 10, and more preferably, at least 20 contiguous amino acids of an SPHK protein. In a particular embodiment, SPHK-specific antigens and/or immunogens are coupled to carrier proteins that stimulate the immune response. For example, the subject polypeptides are covalently coupled to the keyhole limpet hemocyanin (KLH) carrier, and the conjugate is emulsified in Freund's complete adjuvant, which enhances the immune response. An appropriate immune system such as a laboratory rabbit or mouse is immunized according to conventional protocols.

The presence of SPHK-specific antibodies is assayed by an appropriate assay such as a solid phase enzyme-linked immunosorbant assay (ELISA) using immobilized corresponding SPHK polypeptides. Other assays, such as radioimmunoassays or fluorescent assays might also be used.

Chimeric antibodies specific to SPHK polypeptides can be made that contain different portions from different animal species. For instance, a human immunoglobulin constant region may be linked to a variable region of a murine mAb, such that the antibody derives its biological activity from the human antibody, and its binding specificity from the murine fragment. Chimeric antibodies are produced by splicing together genes that encode the appropriate regions from each species (Morrison et al., Proc. Natl. Acad. Sci. (1984) 81:6851-6855; Neuberger et al., Nature (1984) 312:604-608; Takeda et al., Nature (1985) 31:452-454). Humanized antibodies, which are a form of chimeric antibodies, can be generated by grafting complementary-determining regions (CDRs) (Carlos, T. M., J. M. Harlan. 1994. Blood 84:2068-2101) of mouse antibodies into a background of human framework regions and constant regions by recombinant DNA technology (Riechmann LM, et al., 1988 Nature 323: 323-327). Humanized antibodies contain ~10% murine sequences and ~90% human sequences, and thus further reduce or eliminate

immunogenicity, while retaining the antibody specificities (Co MS, and Queen C. 1991 Nature 351: 501-501; Morrison SL. 1992 Ann. Rev. Immun. 10:239-265). Humanized antibodies and methods of their production are well-known in the art (U.S. Pat. Nos. 5,530,101, 5,585,089, 5,693,762, and 6,180,370).

SPHK-specific single chain antibodies which are recombinant, single chain polypeptides formed by linking the heavy and light chain fragments of the Fv regions via an amino acid bridge, can be produced by methods known in the art (U.S. Pat. No. 4,946,778; Bird, Science (1988) 242:423-426; Huston et al., Proc. Natl. Acad. Sci. USA (1988) 85:5879-5883; and Ward et al., Nature (1989) 334:544-546).

Other suitable techniques for antibody production involve in vitro exposure of lymphocytes to the antigenic polypeptides or alternatively to selection of libraries of antibodies in phage or similar vectors (Huse et al., Science (1989) 246:1275-1281). As used herein, T-cell antigen receptors are included within the scope of antibody modulators (Harlow and Lane, 1988, *supra*).

The polypeptides and antibodies of the present invention may be used with or without modification. Frequently, antibodies will be labeled by joining, either covalently or non-covalently, a substance that provides for a detectable signal, or that is toxic to cells that express the targeted protein (Menard S, et al., Int J. Biol Markers (1989) 4:131-134). A wide variety of labels and conjugation techniques are known and are reported extensively in both the scientific and patent literature. Suitable labels include radionuclides, enzymes, substrates, cofactors, inhibitors, fluorescent moieties, fluorescent emitting lanthanide metals, chemiluminescent moieties, bioluminescent moieties, magnetic particles, and the like (U.S. Pat. Nos. 3,817,837; 3,850,752; 3,939,350; 3,996,345; 4,277,437; 4,275,149; and 4,366,241). Also, recombinant immunoglobulins may be produced (U.S. Pat. No. 4,816,567). Antibodies to cytoplasmic polypeptides may be delivered and reach their targets by conjugation with membrane-penetrating toxin proteins (U.S. Pat. No. 6,086,900).

When used therapeutically in a patient, the antibodies of the subject invention are typically administered parenterally, when possible at the target site, or intravenously. The therapeutically effective dose and dosage regimen is determined by clinical studies. Typically, the amount of antibody administered is in the range of about 0.1 mg/kg –to about 10 mg/kg of patient weight. For parenteral administration, the antibodies are formulated in a unit dosage injectable form (e.g., solution, suspension, emulsion) in association with a pharmaceutically acceptable vehicle. Such vehicles are inherently

nontoxic and non-therapeutic. Examples are water, saline, Ringer's solution, dextrose solution, and 5% human serum albumin. Nonaqueous vehicles such as fixed oils, ethyl oleate, or liposome carriers may also be used. The vehicle may contain minor amounts of additives, such as buffers and preservatives, which enhance isotonicity and chemical stability or otherwise enhance therapeutic potential. The antibodies' concentrations in such vehicles are typically in the range of about 1 mg/ml to about10 mg/ml. Immunotherapeutic methods are further described in the literature (US Pat. No. 5,859,206; WO0073469).

Nucleic Acid Modulators

Other preferred SPHK-modulating agents comprise nucleic acid molecules, such as antisense oligomers or double stranded RNA (dsRNA), which generally inhibit SPHK activity. Preferred nucleic acid modulators interfere with the function of the SPHK nucleic acid such as DNA replication, transcription, translocation of the SPHK RNA to the site of protein translation, translation of protein from the SPHK RNA, splicing of the SPHK RNA to yield one or more mRNA species, or catalytic activity which may be engaged in or facilitated by the SPHK RNA.

In one embodiment, the antisense oligomer is an oligonucleotide that is sufficiently complementary to an SPHK mRNA to bind to and prevent translation, preferably by binding to the 5′ untranslated region. SPHK-specific antisense oligonucleotides, preferably range from at least 6 to about 200 nucleotides. In some embodiments the oligonucleotide is preferably at least 10, 15, or 20 nucleotides in length. In other embodiments, the oligonucleotide is preferably less than 50, 40, or 30 nucleotides in length. The oligonucleotide can be DNA or RNA or a chimeric mixture or derivatives or modified versions thereof, single-stranded or double-stranded. The oligonucleotide can be modified at the base moiety, sugar moiety, or phosphate backbone. The oligonucleotide may include other appending groups such as peptides, agents that facilitate transport across the cell membrane, hybridization-triggered cleavage agents, and intercalating agents.

In another embodiment, the antisense oligomer is a phosphothioate morpholino oligomer (PMO). PMOs are assembled from four different morpholino subunits, each of which contain one of four genetic bases (A, C, G, or T) linked to a six-membered morpholine ring. Polymers of these subunits are joined by non-ionic phosphodiamidate intersubunit linkages. Details of how to make and use PMOs and other antisense

oligomers are well known in the art (e.g. see WO99/18193; Probst JC, Antisense Oligodeoxynucleotide and Ribozyme Design, Methods. (2000) 22(3):271-281; Summerton J, and Weller D. 1997 Antisense Nucleic Acid Drug Dev. :7:187-95; US Pat. No. 5,235,033; and US Pat No. 5,378,841).

Alternative preferred SPHK nucleic acid modulators are double-stranded RNA species mediating RNA interference (RNAi). RNAi is the process of sequence-specific, post-transcriptional gene silencing in animals and plants, initiated by double-stranded RNA (dsRNA) that is homologous in sequence to the silenced gene. Methods relating to the use of RNAi to silence genes in *C. elegans*, *Drosophila*, plants, and humans are known in the art (Fire A, et al., 1998 Nature 391:806-811; Fire, A. Trends Genet. 15, 358-363 (1999); Sharp, P. A. RNA interference 2001. Genes Dev. 15, 485-490 (2001); Hammond, S. M., et al., Nature Rev. Genet. 2, 110-1119 (2001); Tuschl, T. Chem. Biochem. 2, 239-245 (2001); Hamilton, A. et al., Science 286, 950-952 (1999); Hammond, S. M., et al., Nature 404, 293-296 (2000); Zamore, P. D., et al., Cell 101, 25-33 (2000); Bernstein, E., et al., Nature 409, 363-366 (2001); Elbashir, S. M., et al., Genes Dev. 15, 188-200 (2001); WO0129058; WO9932619; Elbashir SM, et al., 2001 Nature 411:494-498).

Nucleic acid modulators are commonly used as research reagents, diagnostics, and therapeutics. For example, antisense oligonucleotides, which are able to inhibit gene expression with exquisite specificity, are often used to elucidate the function of particular genes (see, for example, U.S. Pat. No. 6,165,790). Nucleic acid modulators are also used, for example, to distinguish between functions of various members of a biological pathway. For example, antisense oligomers have been employed as therapeutic moieties in the treatment of disease states in animals and man and have been demonstrated in numerous clinical trials to be safe and effective (Milligan JF, *et al*, Current Concepts in Antisense Drug Design, J Med Chem. (1993) 36:1923-1937; Tonkinson JL *et al.*, Antisense Oligodeoxynucleotides as Clinical Therapeutic Agents, Cancer Invest. (1996) 14:54-65). Accordingly, in one aspect of the invention, an SPHK-specific nucleic acid modulator is used in an assay to further elucidate the role of the SPHK in the p53 pathway, and/or its relationship to other members of the pathway. In another aspect of the invention, an SPHK-specific antisense oligomer is used as a therapeutic agent for treatment of p53-related disease states.

Assay Systems

The invention provides assay systems and screening methods for identifying specific modulators of SPHK activity. As used herein, an "assay system" encompasses all the components required for performing and analyzing results of an assay that detects and/or measures a particular event. In general, primary assays are used to identify or confirm a modulator's specific biochemical or molecular effect with respect to the SPHK nucleic acid or protein. In general, secondary assays further assess the activity of a SPHK modulating agent identified by a primary assay and may confirm that the modulating agent affects SPHK in a manner relevant to the p53 pathway. In some cases, SPHK modulators will be directly tested in a secondary assay.

In a preferred embodiment, the screening method comprises contacting a suitable assay system comprising an SPHK polypeptide or nucleic acid with a candidate agent under conditions whereby, but for the presence of the agent, the system provides a reference activity (e.g. kinase activity), which is based on the particular molecular event the screening method detects. A statistically significant difference between the agent-biased activity and the reference activity indicates that the candidate agent modulates SPHK activity, and hence the p53 pathway. The SPHK polypeptide or nucleic acid used in the assay may comprise any of the nucleic acids or polypeptides described above.

Primary Assays

The type of modulator tested generally determines the type of primary assay.

Primary assays for small molecule modulators

For small molecule modulators, screening assays are used to identify candidate modulators. Screening assays may be cell-based or may use a cell-free system that recreates or retains the relevant biochemical reaction of the target protein (reviewed in Sittampalam GS *et al.*, Curr Opin Chem Biol (1997) 1:384-91 and accompanying references). As used herein the term "cell-based" refers to assays using live cells, dead cells, or a particular cellular fraction, such as a membrane, endoplasmic reticulum, or mitochondrial fraction. The term "cell free" encompasses assays using substantially purified protein (either endogenous or recombinantly produced), partially purified or crude cellular extracts. Screening assays may detect a variety of molecular events, including protein-DNA interactions, protein-protein interactions (*e.g.*, receptor-ligand binding), transcriptional activity (*e.g.*, using a reporter gene), enzymatic activity (*e.g.*, via a property

of the substrate), activity of second messengers, immunogenicty and changes in cellular morphology or other cellular characteristics. Appropriate screening assays may use a wide range of detection methods including fluorescent, radioactive, colorimetric, spectrophotometric, and amperometric methods, to provide a read-out for the particular molecular event detected.

Cell-based screening assays usually require systems for recombinant expression of SPHK and any auxiliary proteins demanded by the particular assay. Appropriate methods for generating recombinant proteins produce sufficient quantities of proteins that retain their relevant biological activities and are of sufficient purity to optimize activity and assure assay reproducibility. Yeast two-hybrid and variant screens, and mass spectrometry provide preferred methods for determining protein-protein interactions and elucidation of protein complexes. In certain applications, when SPHK-interacting proteins are used in screens to identify small molecule modulators, the binding specificity of the interacting protein to the SPHK protein may be assayed by various known methods such as substrate processing (e.g. ability of the candidate SPHK-specific binding agents to function as negative effectors in SPHK-expressing cells), binding equilibrium constants (usually at least about 10^7 M^{-1}, preferably at least about 10^8 M^{-1}, more preferably at least about 10^9 M^{-1}), and immunogenicity (e.g. ability to elicit SPHK specific antibody in a heterologous host such as a mouse, rat, goat or rabbit). For enzymes and receptors, binding may be assayed by, respectively, substrate and ligand processing.

The screening assay may measure a candidate agent's ability to specifically bind to or modulate activity of a SPHK polypeptide, a fusion protein thereof, or to cells or membranes bearing the polypeptide or fusion protein. The SPHK polypeptide can be full length or a fragment thereof that retains functional SPHK activity. The SPHK polypeptide may be fused to another polypeptide, such as a peptide tag for detection or anchoring, or to another tag. The SPHK polypeptide is preferably human SPHK, or is an ortholog or derivative thereof as described above. In a preferred embodiment, the screening assay detects candidate agent-based modulation of SPHK interaction with a binding target, such as an endogenous or exogenous protein or other substrate that has SPHK –specific binding activity, and can be used to assess normal SPHK gene function.

Suitable assay formats that may be adapted to screen for SPHK modulators are known in the art. Preferred screening assays are high throughput or ultra high throughput and thus provide automated, cost-effective means of screening compound libraries for lead compounds (Fernandes PB, Curr Opin Chem Biol (1998) 2:597-603; Sundberg SA, Curr

Opin Biotechnol 2000, 11:47-53). In one preferred embodiment, screening assays uses fluorescence technologies, including fluorescence polarization, time-resolved fluorescence, and fluorescence resonance energy transfer. These systems offer means to monitor protein-protein or DNA-protein interactions in which the intensity of the signal emitted from dye-labeled molecules depends upon their interactions with partner molecules (*e.g.*, Selvin PR, Nat Struct Biol (2000) 7:730-4; Fernandes PB, *supra*; Hertzberg RP and Pope AJ, Curr Opin Chem Biol (2000) 4:445-451).

A variety of suitable assay systems may be used to identify candidate SPHK and p53 pathway modulators (e.g. U.S. Pat. No. 6,165,992 (kinase assays); U.S. Pat. Nos. 5,550,019 and 6,133,437 (apoptosis assays); U.S. Pat. No. 6,020,135 (p53 modulation), among others). Specific preferred assays are described in more detail below.

Kinase assays. In some preferred embodiments the screening assay detects the ability of the test agent to modulate the kinase activity of an SPHK polypeptide. In further embodiments, a cell-free kinase assay system is used to identify a candidate p53 modulating agent, and a secondary, cell-based assay, such as an apoptosis or hypoxic induction assay (described below), may be used to further characterize the candidate p53 modulating agent. Many different assays for kinases have been reported in the literature and are well known to those skilled in the art (e.g. U.S. Pat. No. 6,165,992; Zhu et al., Nature Genetics (2000) 26:283-289; and WO0073469). Radioassays, which monitor the transfer of a gamma phosphate are frequently used. For instance, a scintillation assay for p56 (lck) kinase activity monitors the transfer of the gamma phosphate from gamma -^{33}P ATP to a biotinylated peptide substrate; the substrate is captured on a streptavidin coated bead that transmits the signal (Beveridge M *et al.*, J Biomol Screen (2000) 5:205-212). This assay uses the scintillation proximity assay (SPA), in which only radio-ligand bound to receptors tethered to the surface of an SPA bead are detected by the scintillant immobilized within it, allowing binding to be measured without separation of bound from free ligand.

Various sphingosine kinase assays, appropriate for identifying modulators of hSPHK, have been described in the literature (Meyer zu Heringsdorf et al. (1998) EMBO J. 17, 2830-7; Edsal and Spiegel, Analytical Biochemistry 272, 80-86 (1999); and Tolan et al., Cell. Signal. Vol. 11, No. 5, pp. 349-354, 1999))

Other assays for protein kinase activity may use antibodies that specifically recognize phosphorylated substrates. For instance, the kinase receptor activation (KIRA) assay

measures receptor tyrosine kinase activity by ligand stimulating the intact receptor in cultured cells, then capturing solubilized receptor with specific antibodies and quantifying phosphorylation via phosphotyrosine ELISA (Sadick MD, Dev Biol Stand (1999) 97:121-133).

Another example of antibody based assays for protein kinase activity is TRF (time-resolved fluorometry). This method utilizes europium chelate-labeled anti-phosphotyrosine antibodies to detect phosphate transfer to a polymeric substrate coated onto microtiter plate wells. The amount of phosphorylation is then detected using time-resolved, dissociation-enhanced fluorescence (Braunwalder AF, et al., Anal Biochem 1996 Jul 1;238(2):159-64).

Apoptosis assays. Assays for apoptosis may be performed by terminal deoxynucleotidyl transferase-mediated digoxigenin-11-dUTP nick end labeling (TUNEL) assay. The TUNEL assay is used to measure nuclear DNA fragmentation characteristic of apoptosis (Lazebnik *et al.*, 1994, Nature 371, 346), by following the incorporation of fluorescein-dUTP (Yonehara *et al.*, 1989, J. Exp. Med. 169, 1747). Apoptosis may further be assayed by acridine orange staining of tissue culture cells (Lucas, R., et al., 1998, Blood 15:4730-41). An apoptosis assay system may comprise a cell that expresses an SPHK, and that optionally has defective p53 function (e.g. p53 is over-expressed or under-expressed relative to wild-type cells). A test agent can be added to the apoptosis assay system and changes in induction of apoptosis relative to controls where no test agent is added, identify candidate p53 modulating agents. In some embodiments of the invention, an apoptosis assay may be used as a secondary assay to test a candidate p53 modulating agents that is initially identified using a cell-free assay system. An apoptosis assay may also be used to test whether SPHK function plays a direct role in apoptosis. For example, an apoptosis assay may be performed on cells that over- or under-express SPHK relative to wild type cells. Differences in apoptotic response compared to wild type cells suggests that the SPHK plays a direct role in the apoptotic response. Apoptosis assays are described further in US Pat. No. 6,133,437.

Cell proliferation and cell cycle assays. Cell proliferation may be assayed via bromodeoxyuridine (BRDU) incorporation. This assay identifies a cell population undergoing DNA synthesis by incorporation of BRDU into newly-synthesized DNA. Newly-synthesized DNA may then be detected using an anti-BRDU antibody (Hoshino *et*

al., 1986, Int. J. Cancer 38, 369; Campana *et al.*, 1988, J. Immunol. Meth. 107, 79), or by other means.

Cell Proliferation may also be examined using [^{3}H]-thymidine incorporation (Chen, J., 1996, Oncogene 13:1395-403; Jeoung, J., 1995, J. Biol. Chem. 270:18367-73). This assay allows for quantitative characterization of S-phase DNA syntheses. In this assay, cells synthesizing DNA will incorporate [^{3}H]-thymidine into newly synthesized DNA. Incorporation can then be measured by standard techniques such as by counting of radioisotope in a scintillation counter (e.g., Beckman LS 3800 Liquid Scintillation Counter).

Cell proliferation may also be assayed by colony formation in soft agar (Sambrook et al., Molecular Cloning, Cold Spring Harbor (1989)). For example, cells transformed with SPHK are seeded in soft agar plates, and colonies are measured and counted after two weeks incubation.

Involvement of a gene in the cell cycle may be assayed by flow cytometry (Gray JW et al. (1986) Int J Radiat Biol Relat Stud Phys Chem Med 49:237-55). Cells transfected with an SPHK may be stained with propidium iodide and evaluated in a flow cytometer (available from Becton Dickinson).

Accordingly, a cell proliferation or cell cycle assay system may comprise a cell that expresses an SPHK, and that optionally has defective p53 function (e.g. p53 is over-expressed or under-expressed relative to wild-type cells). A test agent can be added to the assay system and changes in cell proliferation or cell cycle relative to controls where no test agent is added, identify candidate p53 modulating agents. In some embodiments of the invention, the cell proliferation or cell cycle assay may be used as a secondary assay to test a candidate p53 modulating agents that is initially identified using another assay system such as a cell-free kinase assay system. A cell proliferation assay may also be used to test whether SPHK function plays a direct role in cell proliferation or cell cycle. For example, a cell proliferation or cell cycle assay may be performed on cells that over- or under-express SPHK relative to wild type cells. Differences in proliferation or cell cycle compared to wild type cells suggests that the SPHK plays a direct role in cell proliferation or cell cycle.

Angiogenesis. Angiogenesis may be assayed using various human endothelial cell systems, such as umbilical vein, coronary artery, or dermal cells. Suitable assays include Alamar Blue based assays (available from Biosource International) to measure

proliferation; migration assays using fluorescent molecules, such as the use of Becton Dickinson Falcon HTS FluoroBlock cell culture inserts to measure migration of cells through membranes in presence or absence of angiogenesis enhancer or suppressors; and tubule formation assays based on the formation of tubular structures by endothelial cells on Matrigel® (Becton Dickinson). Accordingly, an angiogenesis assay system may comprise a cell that expresses an SPHK, and that optionally has defective p53 function (e.g. p53 is over-expressed or under-expressed relative to wild-type cells). A test agent can be added to the angiogenesis assay system and changes in angiogenesis relative to controls where no test agent is added, identify candidate p53 modulating agents. In some embodiments of the invention, the angiogenesis assay may be used as a secondary assay to test a candidate p53 modulating agents that is initially identified using another assay system. An angiogenesis assay may also be used to test whether SPHK function plays a direct role in cell proliferation. For example, an angiogenesis assay may be performed on cells that over- or under-express SPHK relative to wild type cells. Differences in angiogenesis compared to wild type cells suggests that the SPHK plays a direct role in angiogenesis.

Hypoxic induction. The alpha subunit of the transcription factor, hypoxia inducible factor-1 (HIF-1), is upregulated in tumor cells following exposure to hypoxia in vitro. Under hypoxic conditions, HIF-1 stimulates the expression of genes known to be important in tumour cell survival, such as those encoding glyolytic enzymes and VEGF. Induction of such genes by hypoxic conditions may be assayed by growing cells transfected with SPHK in hypoxic conditions (such as with 0.1% O2, 5% CO2, and balance N2, generated in a Napco 7001 incubator (Precision Scientific)) and normoxic conditions, followed by assessment of gene activity or expression by Taqman®. For example, a hypoxic induction assay system may comprise a cell that expresses an SPHK, and that optionally has a mutated p53 (e.g. p53 is over-expressed or under-expressed relative to wild-type cells). A test agent can be added to the hypoxic induction assay system and changes in hypoxic response relative to controls where no test agent is added, identify candidate p53 modulating agents. In some embodiments of the invention, the hypoxic induction assay may be used as a secondary assay to test a candidate p53 modulating agents that is initially identified using another assay system. A hypoxic induction assay may also be used to test whether SPHK function plays a direct role in the hypoxic response. For example, a hypoxic induction assay may be performed on cells that

over- or under-express SPHK relative to wild type cells. Differences in hypoxic response compared to wild type cells suggests that the SPHK plays a direct role in hypoxic induction.

Cell adhesion. Cell adhesion assays measure adhesion of cells to purified adhesion proteins, or adhesion of cells to each other, in presence or absence of candidate modulating agents. Cell-protein adhesion assays measure the ability of agents to modulate the adhesion of cells to purified proteins. For example, recombinant proteins are produced, diluted to 2.5g/mL in PBS, and used to coat the wells of a microtiter plate. The wells used for negative control are not coated. Coated wells are then washed, blocked with 1% BSA, and washed again. Compounds are diluted to 2× final test concentration and added to the blocked, coated wells. Cells are then added to the wells, and the unbound cells are washed off. Retained cells are labeled directly on the plate by adding a membrane-permeable fluorescent dye, such as calcein-AM, and the signal is quantified in a fluorescent microplate reader.

Cell-cell adhesion assays measure the ability of agents to modulate binding of cell adhesion proteins with their native ligands. These assays use cells that naturally or recombinantly express the adhesion protein of choice. In an exemplary assay, cells expressing the cell adhesion protein are plated in wells of a multiwell plate. Cells expressing the ligand are labeled with a membrane-permeable fluorescent dye, such as BCECF , and allowed to adhere to the monolayers in the presence of candidate agents. Unbound cells are washed off, and bound cells are detected using a fluorescence plate reader.

High-throughput cell adhesion assays have also been described. In one such assay, small molecule ligands and peptides are bound to the surface of microscope slides using a microarray spotter, intact cells are then contacted with the slides, and unbound cells are washed off. In this assay, not only the binding specificity of the peptides and modulators against cell lines are determined, but also the functional cell signaling of attached cells using immunofluorescence techniques in situ on the microchip is measured (Falsey JR et al., Bioconjug Chem. 2001 May-Jun;12(3):346-53).

Cell Migration. An invasion/migration assay (also called a migration assay) tests the ability of cells to overcome a physical barrier and to migrate towards pro-angiogenic signals. Migration assays are known in the art (e.g., Paik JH et al., 2001, J Biol Chem

276:11830-11837). In a typical experimental set-up, cultured endothelial cells are seeded onto a matrix-coated porous lamina, with pore sizes generally smaller than typical cell size. The matrix generally simulates the environment of the extracellular matrix, as described above. The lamina is typically a membrane, such as the transwell polycarbonate membrane (Corning Costar Corporation, Cambridge, MA), and is generally part of an upper chamber that is in fluid contact with a lower chamber containing pro-angiogenic stimuli. Migration is generally assayed after an overnight incubation with stimuli, but longer or shorter time frames may also be used. Migration is assessed as the number of cells that crossed the lamina, and may be detected by staining cells with hemotoxylin solution (VWR Scientific, South San Francisco, CA), or by any other method for determining cell number. In another exemplary set up, cells are fluorescently labeled and migration is detected using fluorescent readings, for instance using the Falcon HTS FluoroBlok (Becton Dickinson). While some migration is observed in the absence of stimulus, migration is greatly increased in response to pro-angiogenic factors. As described above, a preferred assay system for migration/invasion assays comprises testing an SPHK's response to a variety of pro-angiogenic factors, including tumor angiogenic and inflammatory angiogenic agents, and culturing the cells in serum free medium.

Sprouting assay. A sprouting assay is a three-dimensional *in vitro* angiogenesis assay that uses a cell-number defined spheroid aggregation of endothelial cells ("spheroid"), embedded in a collagen gel-based matrix. The spheroid can serve as a starting point for the sprouting of capillary-like structures by invasion into the extracellular matrix (termed "cell sprouting") and the subsequent formation of complex anastomosing networks (Korff and Augustin, 1999, J Cell Sci 112:3249-58). In an exemplary experimental set-up, spheroids are prepared by pipetting 400 human umbilical vein endothelial cells into individual wells of a nonadhesive 96-well plates to allow overnight spheroidal aggregation (Korff and Augustin: J Cell Biol 143: 1341-52, 1998). Spheroids are harvested and seeded in 900μl of methocel-collagen solution and pipetted into individual wells of a 24 well plate to allow collagen gel polymerization. Test agents are added after 30 min by pipetting 100 μl of 10-fold concentrated working dilution of the test substances on top of the gel. Plates are incubated at 37°C for 24h. Dishes are fixed at the end of the experimental incubation period by addition of paraformaldehyde. Sprouting intensity of endothelial cells can be quantitated by an automated image analysis system to determine the cumulative sprout length per spheroid.

Primary assays for antibody modulators

For antibody modulators, appropriate primary assays test is a binding assay that tests the antibody's affinity to and specificity for the SPHK protein. Methods for testing antibody affinity and specificity are well known in the art (Harlow and Lane, 1988, 1999, *supra*). The enzyme-linked immunosorbant assay (ELISA) is a preferred method for detecting SPHK-specific antibodies; others include FACS assays, radioimmunoassays, and fluorescent assays.

Primary assays for nucleic acid modulators

For nucleic acid modulators, primary assays may test the ability of the nucleic acid modulator to inhibit or enhance SPHK gene expression, preferably mRNA expression. In general, expression analysis comprises comparing SPHK expression in like populations of cells (*e.g.*, two pools of cells that endogenously or recombinantly express SPHK) in the presence and absence of the nucleic acid modulator. Methods for analyzing mRNA and protein expression are well known in the art. For instance, Northern blotting, slot blotting, ribonuclease protection, quantitative RT-PCR (*e.g.*, using the TaqMan®, PE Applied Biosystems), or microarray analysis may be used to confirm that SPHK mRNA expression is reduced in cells treated with the nucleic acid modulator (*e.g.*, Current Protocols in Molecular Biology (1994) Ausubel FM *et al.*, *eds.*, John Wiley & Sons, Inc., chapter 4; Freeman WM *et al.*, Biotechniques (1999) 26:112-125; Kallioniemi OP, Ann Med 2001, 33:142-147; Blohm DH and Guiseppi-Elie, A Curr Opin Biotechnol 2001, 12:41-47). Protein expression may also be monitored. Proteins are most commonly detected with specific antibodies or antisera directed against either the SPHK protein or specific peptides. A variety of means including Western blotting, ELISA, or in situ detection, are available (Harlow E and Lane D, 1988 and 1999, *supra*).

Secondary Assays

Secondary assays may be used to further assess the activity of SPHK-modulating agent identified by any of the above methods to confirm that the modulating agent affects SPHK in a manner relevant to the p53 pathway. As used herein, SPHK-modulating agents encompass candidate clinical compounds or other agents derived from previously identified modulating agent. Secondary assays can also be used to test the activity of a modulating agent on a particular genetic or biochemical pathway or to test the specificity of the modulating agent's interaction with SPHK.

Secondary assays generally compare like populations of cells or animals (*e.g.*, two pools of cells or animals that endogenously or recombinantly express SPHK) in the presence and absence of the candidate modulator. In general, such assays test whether treatment of cells or animals with a candidate SPHK–modulating agent results in changes in the p53 pathway in comparison to untreated (or mock- or placebo-treated) cells or animals. Certain assays use "sensitized genetic backgrounds", which, as used herein, describe cells or animals engineered for altered expression of genes in the p53 or interacting pathways.

Cell-based assays

Cell based assays may use a variety of mammalian cell lines known to have defective p53 function (e.g. SAOS-2 osteoblasts, H1299 lung cancer cells, C33A and HT3 cervical cancer cells, HT-29 and DLD-1 colon cancer cells, among others, available from American Type Culture Collection (ATCC), Manassas, VA). Cell based assays may detect endogenous p53 pathway activity or may rely on recombinant expression of p53 pathway components. Any of the aforementioned assays may be used in this cell-based format. Candidate modulators are typically added to the cell media but may also be injected into cells or delivered by any other efficacious means.

Animal Assays

A variety of non-human animal models of normal or defective p53 pathway may be used to test candidate SPHK modulators. Models for defective p53 pathway typically use genetically modified animals that have been engineered to mis-express (*e.g.*, over-express or lack expression in) genes involved in the p53 pathway. Assays generally require systemic delivery of the candidate modulators, such as by oral administration, injection, etc.

In a preferred embodiment, p53 pathway activity is assessed by monitoring neovascularization and angiogenesis. Animal models with defective and normal p53 are used to test the candidate modulator's affect on SPHK in Matrigel® assays. Matrigel® is an extract of basement membrane proteins, and is composed primarily of laminin, collagen IV, and heparin sulfate proteoglycan. It is provided as a sterile liquid at 4° C, but rapidly forms a solid gel at 37° C. Liquid Matrigel® is mixed with various angiogenic agents, such as bFGF and VEGF, or with human tumor cells which over-express the SPHK. The mixture is then injected subcutaneously(SC) into female athymic nude mice (Taconic,

Germantown, NY) to support an intense vascular response. Mice with Matrigel® pellets may be dosed via oral (PO), intraperitoneal (IP), or intravenous (IV) routes with the candidate modulator. Mice are euthanized 5 - 12 days post-injection, and the Matrigel® pellet is harvested for hemoglobin analysis (Sigma plasma hemoglobin kit). Hemoglobin content of the gel is found to correlate the degree of neovascularization in the gel.

In another preferred embodiment, the effect of the candidate modulator on SPHK is assessed via tumorigenicity assays. In one example, xenograft human tumors are implanted SC into female athymic mice, 6-7 week old, as single cell suspensions either from a pre-existing tumor or from *in vitro* culture. The tumors which express the SPHK endogenously are injected in the flank, 1 x 10^5 to 1 x 10^7 cells per mouse in a volume of 100 μL using a 27gauge needle. Mice are then ear tagged and tumors are measured twice weekly. Candidate modulator treatment is initiated on the day the mean tumor weight reaches 100 mg. Candidate modulator is delivered IV, SC, IP, or PO by bolus administration. Depending upon the pharmacokinetics of each unique candidate modulator, dosing can be performed multiple times per day. The tumor weight is assessed by measuring perpendicular diameters with a caliper and calculated by multiplying the measurements of diameters in two dimensions. At the end of the experiment, the excised tumors maybe utilized for biomarker identification or further analyses. For immunohistochemistry staining, xenograft tumors are fixed in 4% paraformaldehyde, 0.1M phosphate, pH 7.2, for 6 hours at 4°C, immersed in 30% sucrose in PBS, and rapidly frozen in isopentane cooled with liquid nitrogen.

Diagnostic and therapeutic uses

Specific SPHK-modulating agents are useful in a variety of diagnostic and therapeutic applications where disease or disease prognosis is related to defects in the p53 pathway, such as angiogenic, apoptotic, or cell proliferation disorders. Accordingly, the invention also provides methods for modulating the p53 pathway in a cell, preferably a cell pre-determined to have defective or impaired p53 function (e.g. due to overexpression, underexpression, or misexpression of p53, or due to gene mutations), comprising the step of administering an agent to the cell that specifically modulates SPHK activity. Preferably, the modulating agent produces a detectable phenotypic change in the cell indicating that the p53 function is restored. The phrase "function is restored", and equivalents, as used herein, means that the desired phenotype is achieved, or is brought closer to normal compared to untreated cells. For example, with restored p53 function,

cell proliferation and/or progression through cell cycle may normalize, or be brought closer to normal relative to untreated cells. The invention also provides methods for treating disorders or disease associated with impaired p53 function by administering a therapeutically effective amount of an SPHK -modulating agent that modulates the p53 pathway. The invention further provides methods for modulating SPHK function in a cell, preferably a cell pre-determined to have defective or impaired SPHK function, by administering an SPHK -modulating agent. Additionally, the invention provides a method for treating disorders or disease associated with impaired SPHK function by administering a therapeutically effective amount of an SPHK -modulating agent.

The discovery that SPHK is implicated in p53 pathway provides for a variety of methods that can be employed for the diagnostic and prognostic evaluation of diseases and disorders involving defects in the p53 pathway and for the identification of subjects having a predisposition to such diseases and disorders.

Various expression analysis methods can be used to diagnose whether SPHK expression occurs in a particular sample, including Northern blotting, slot blotting, ribonuclease protection, quantitative RT-PCR, and microarray analysis. (*e.g.*, Current Protocols in Molecular Biology (1994) Ausubel FM *et al.*, *eds.*, John Wiley & Sons, Inc., chapter 4; Freeman WM *et al.*, Biotechniques (1999) 26:112-125; Kallioniemi OP, Ann Med 2001, 33:142-147; Blohm and Guiseppi-Elie, Curr Opin Biotechnol 2001, 12:41-47). Tissues having a disease or disorder implicating defective p53 signaling that express an SPHK, are identified as amenable to treatment with an SPHK modulating agent. In a preferred application, the p53 defective tissue overexpresses an SPHK relative to normal tissue. For example, a Northern blot analysis of mRNA from tumor and normal cell lines, or from tumor and matching normal tissue samples from the same patient, using full or partial SPHK cDNA sequences as probes, can determine whether particular tumors express or overexpress SPHK. Alternatively, the TaqMan® is used for quantitative RT-PCR analysis of SPHK expression in cell lines, normal tissues and tumor samples (PE Applied Biosystems).

Various other diagnostic methods may be performed, for example, utilizing reagents such as the SPHK oligonucleotides, and antibodies directed against an SPHK, as described above for: (1) the detection of the presence of SPHK gene mutations, or the detection of either over- or under-expression of SPHK mRNA relative to the non-disorder state; (2) the detection of either an over- or an under-abundance of SPHK gene product relative to the

non-disorder state; and (3) the detection of perturbations or abnormalities in the signal transduction pathway mediated by SPHK.

Thus, in a specific embodiment, the invention is drawn to a method for diagnosing a disease or disorder in a patient that is associated with alterations in SPHK expression, the method comprising: a) obtaining a biological sample from the patient; b) contacting the sample with a probe for SPHK expression; c) comparing results from step (b) with a control; and d) determining whether step (c) indicates a likelihood of the disease or disorder. Preferably, the disease is cancer, most preferably a cancer as shown in TABLE 1. The probe may be either DNA or protein, including an antibody.

EXAMPLES

The following experimental section and examples are offered by way of illustration and not by way of limitation.

I. C. elegans p53 screen

A systematic RNAi of various genes was carried out in worms homozygous for p53 deletion. p53 (-/-) worms have a normal phenotype, but are defective in germline apoptotic response to ionizing radiation as p53 is involved in the DNA damage response. After silencing of each gene by RNAi, worms were subject to gamma-irradiation, and phenotypes were scored.

Worms subjected to SPHK RNAi had an increase in germline apoptosis in p53 (-/-) mutants compared to non-RNAi control animals, as visualized by acridine orange staining. C34C6.5 was thus a suppressor of the p53 (-/-) phenotype. Human orthologs of the modifiers are referred to herein as SPHK.

BLAST analysis (Altschul et al., *supra*) was employed to identify Targets from C.elegans modifiers. For example, representative sequences from SPHK, GI#s 11464967, 9910562, and 20336726 (SEQ ID NOs:10, 11, and 13, respectively), share 28%, 32%, and 28% amino acid identity, respectively, with the *C. elegans* C34C6.5.

Various domains, signals, and functional subunits in proteins were analyzed using the PSORT (Nakai K., and Horton P., Trends Biochem Sci, 1999, 24:34-6; Kenta Nakai, Protein sorting signals and prediction of subcellular localization, Adv. Protein Chem. 54, 277-344 (2000)), PFAM (Bateman A., et al., Nucleic Acids Res, 1999, 27:260-2), SMART (Ponting CP, et al., SMART: identification and annotation of domains from signaling and extracellular protein sequences. Nucleic Acids Res. 1999 Jan 1;27(1):229-

32), TM-HMM (Erik L.L. Sonnhammer, Gunnar von Heijne, and Anders Krogh: A hidden Markov model for predicting transmembrane helices in protein sequences. In Proc. of Sixth Int. Conf. on Intelligent Systems for Molecular Biology, p 175-182 Ed J. Glasgow, T. Littlejohn, F. Major, R. Lathrop, D. Sankoff, and C. Sensen Menlo Park, CA: AAAI Press, 1998), and clust (Remm M, and Sonnhammer E. Classification of transmembrane protein families in the Caenorhabditis elegans genome and identification of human orthologs. Genome Res. 2000 Nov;10(11):1679-89) programs. For example, the kinase domains of SPHKs from GI#s 11461967, 9910562, and 20336726 (SEQ ID NOs:10, 11, and 13, respectively) are located respectively at approximately amino acid residues 16-153, 146 to 283, and 132 to 278

II. High-Throughput In Vitro Fluorescence Polarization Assay

Fluorescently-labeled SPHK peptide/substrate are added to each well of a 96-well microtiter plate, along with a test agent in a test buffer (10 mM HEPES, 10 mM NaCl, 6 mM magnesium chloride, pH 7.6). Changes in fluorescence polarization, determined by using a Fluorolite FPM-2 Fluorescence Polarization Microtiter System (Dynatech Laboratories, Inc), relative to control values indicates the test compound is a candidate modifier of SPHK activity.

III. High-Throughput In Vitro Binding Assay.

^{33}P-labeled SPHK peptide is added in an assay buffer (100 mM KCl, 20 mM HEPES pH 7.6, 1 mM $MgCl_2$, 1% glycerol, 0.5% NP-40, 50 mM beta-mercaptoethanol, 1 mg/ml BSA, cocktail of protease inhibitors) along with a test agent to the wells of a Neutralite-avidin coated assay plate and incubated at 25°C for 1 hour. Biotinylated substrate is then added to each well and incubated for 1 hour. Reactions are stopped by washing with PBS, and counted in a scintillation counter. Test agents that cause a difference in activity relative to control without test agent are identified as candidate p53 modulating agents.

IV. Immunoprecipitations and Immunoblotting

For coprecipitation of transfected proteins, 3×10^6 appropriate recombinant cells containing the SPHK proteins are plated on 10-cm dishes and transfected on the following day with expression constructs. The total amount of DNA is kept constant in each transfection by adding empty vector. After 24 h, cells are collected, washed once with phosphate-buffered saline and lysed for 20 min on ice in 1 ml of lysis buffer containing 50

mM Hepes, pH 7.9, 250 mM NaCl, 20 mM -glycerophosphate, 1 mM sodium orthovanadate, 5 mM p-nitrophenyl phosphate, 2 mM dithiothreitol, protease inhibitors (complete, Roche Molecular Biochemicals), and 1% Nonidet P-40. Cellular debris is removed by centrifugation twice at 15,000 x g for 15 min. The cell lysate is incubated with 25 μl of M2 beads (Sigma) for 2 h at 4 °C with gentle rocking.

After extensive washing with lysis buffer, proteins bound to the beads are solubilized by boiling in SDS sample buffer, fractionated by SDS-polyacrylamide gel electrophoresis, transferred to polyvinylidene difluoride membrane and blotted with the indicated antibodies. The reactive bands are visualized with horseradish peroxidase coupled to the appropriate secondary antibodies and the enhanced chemiluminescence (ECL) Western blotting detection system (Amersham Pharmacia Biotech).

V. Kinase assay

A purified or partially purified SPHK is diluted in a suitable reaction buffer, e.g., 50 mM Hepes, pH 7.5, containing magnesium chloride or manganese chloride (1-20 mM) and a peptide or polypeptide substrate, such as myelin basic protein or casein (1-10 μg/ml). The final concentration of the kinase is 1-20 nM. The enzyme reaction is conducted in microtiter plates to facilitate optimization of reaction conditions by increasing assay throughput. A 96-well microtiter plate is employed using a final volume 30-100 μl. The reaction is initiated by the addition of ^{33}P-gamma-ATP (0.5 μCi/ml) and incubated for 0.5 to 3 hours at room temperature. Negative controls are provided by the addition of EDTA, which chelates the divalent cation ($Mg2^{+}$ or Mn^{2+}) required for enzymatic activity. Following the incubation, the enzyme reaction is quenched using EDTA. Samples of the reaction are transferred to a 96-well glass fiber filter plate (MultiScreen, Millipore). The filters are subsequently washed with phosphate-buffered saline, dilute phosphoric acid (0.5%) or other suitable medium to remove excess radiolabeled ATP. Scintillation cocktail is added to the filter plate and the incorporated radioactivity is quantitated by scintillation counting (Wallac/Perkin Elmer). Activity is defined by the amount of radioactivity detected following subtraction of the negative control reaction value (EDTA quench).

VI. Expression analysis

All cell lines used in the following experiments are NCI (National Cancer Institute) lines, and are available from ATCC (American Type Culture Collection, Manassas, VA

20110-2209). Normal and tumor tissues were obtained from Impath, UC Davis, Clontech, Stratagene, and Ambion.

TaqMan analysis was used to assess expression levels of the disclosed genes in various samples.

RNA was extracted from each tissue sample using Qiagen (Valencia, CA) RNeasy kits, following manufacturer's protocols, to a final concentration of 50ng/μl. Single stranded cDNA was then synthesized by reverse transcribing the RNA samples using random hexamers and 500ng of total RNA per reaction, following protocol 4304965 of Applied Biosystems (Foster City, CA).

Primers for expression analysis using TaqMan assay (Applied Biosystems, Foster City, CA) were prepared according to the TaqMan protocols, and the following criteria: a) primer pairs were designed to span introns to eliminate genomic contamination, and b) each primer pair produced only one product.

Taqman reactions were carried out following manufacturer's protocols, in 25 μl total volume for 96-well plates and 10 μl total volume for 384-well plates, using 300nM primer and 250 nM probe, and approximately 25ng of cDNA. The standard curve for result analysis was prepared using a universal pool of human cDNA samples, which is a mixture of cDNAs from a wide variety of tissues so that the chance that a target will be present in appreciable amounts is good. The raw data were normalized using 18S rRNA (universally expressed in all tissues and cells).

For each expression analysis, tumor tissue samples were compared with matched normal tissues from the same patient. A gene was considered overexpressed in a tumor when the level of expression of the gene was 2 fold or higher in the tumor compared with its matched normal sample. In cases where normal tissue was not available, a universal pool of cDNA samples was used instead. In these cases, a gene was considered overexpressed in a tumor sample when the difference of expression levels between a tumor sample and the average of all normal samples from the same tissue type was greater than 2 times the standard deviation of all normal samples (i.e., Tumor – average(all normal samples) > 2 x STDEV(all normal samples)).

Results are shown in Table 1. Data presented in bold indicate that greater than 50% of tested tumor samples of the tissue type indicated in row 1 exhibited over expression of the gene listed in column 1, relative to normal samples. Underlined data indicates that between 25% to 49% of tested tumor samples exhibited over expression. A modulator identified by an assay described herein can be further validated for therapeutic effect by

administration to a tumor in which the gene is overexpressed. A decrease in tumor growth confirms therapeutic utility of the modulator. Prior to treating a patient with the modulator, the likelihood that the patient will respond to treatment can be diagnosed by obtaining a tumor sample from the patient, and assaying for expression of the gene targeted by the modulator. The expression data for the gene(s) can also be used as a diagnostic marker for disease progression. The assay can be performed by expression analysis as described above, by antibody directed to the gene target, or by any other available detection method.

Table 1

SEQ ID NO	NA_GI#	breast		colon		lung		ovary	
8	18594429	2	11	4	30	1	13	1	7
1	11464966	4	11	4	30	6	14	1	7
4	9910561	2	11	7	25	4	11	0	3

9. The method of claim 8 wherein the nucleic acid modulator is an antisense oligomer.

10. The method of Claim 8 wherein the nucleic acid modulator is a PMO.

11. The method of Claim 1 additionally comprising:

(d) administering the candidate p53 pathway modulating agent identified in (c) to a model system comprising cells defective in p53 function and, detecting a phenotypic change in the model system that indicates that the p53 function is restored.

12. The method of Claim 11 wherein the model system is a mouse model with defective p53 function.

13. A method for modulating a p53 pathway of a cell comprising contacting a cell defective in p53 function with a candidate modulator that specifically binds to a SPHK polypeptide comprising an amino acid sequence selected from group consisting of SEQ ID NOs:10, 11, 12, and 13, whereby p53 function is restored.

14. The method of claim 13 wherein the candidate modulator is administered to a vertebrate animal predetermined to have a disease or disorder resulting from a defect in p53 function.

15. The method of Claim 13 wherein the candidate modulator is selected from the group consisting of an antibody and a small molecule.

16. The method of Claim 1, comprising the additional steps of:

(d) providing a secondary assay system comprising cultured cells or a non-human animal expressing SPHK ,

(e) contacting the secondary assay system with the test agent of (b) or an agent derived therefrom under conditions whereby, but for the presence of the test agent or agent derived therefrom, the system provides a reference activity; and

(f) detecting an agent-biased activity of the second assay system,

wherein a difference between the agent-biased activity and the reference activity of the second assay system confirms the test agent or agent derived therefrom as a candidate p53 pathway modulating agent,

and wherein the second assay detects an agent-biased change in the p53 pathway.

17. The method of Claim 16 wherein the secondary assay system comprises cultured cells.

18. The method of Claim 16 wherein the secondary assay system comprises a non-human animal.

19. The method of Claim 18 wherein the non-human animal mis-expresses a p53 pathway gene.

20. A method of modulating p53 pathway in a mammalian cell comprising contacting the cell with an agent that specifically binds a SPHK polypeptide or nucleic acid.

21. The method of Claim 20 wherein the agent is administered to a mammalian animal predetermined to have a pathology associated with the p53 pathway.

22. The method of Claim 20 wherein the agent is a small molecule modulator, a nucleic acid modulator, or an antibody.

23. A method for diagnosing a disease in a patient comprising:

(a) obtaining a biological sample from the patient;

(b) contacting the sample with a probe for SPHK expression;

(c) comparing results from step (b) with a control;

(d) determining whether step (c) indicates a likelihood of disease.

24. The method of claim 23 wherein said disease is cancer.

25. The method according to claim 24, wherein said cancer is a cancer as shown in Table 1 as having >25% expression level.

WHAT IS CLAIMED IS:

1. A method of identifying a candidate p53 pathway modulating agent, said method comprising the steps of:

(a) providing an assay system comprising a purified SPHK polypeptide or nucleic acid or a functionally active fragment or derivative thereof;

(b) contacting the assay system with a test agent under conditions whereby, but for the presence of the test agent, the system provides a reference activity; and

(c) detecting a test agent-biased activity of the assay system, wherein a difference between the test agent-biased activity and the reference activity identifies the test agent as a candidate p53 pathway modulating agent.

2. The method of Claim 1 wherein the assay system comprises cultured cells that express the SPHK polypeptide.

3. The method of Claim 2 wherein the cultured cells additionally have defective p53 function.

4. The method of Claim 1 wherein the assay system includes a screening assay comprising a SPHK polypeptide, and the candidate test agent is a small molecule modulator.

5. The method of Claim 4 wherein the assay is a kinase assay.

6. The method of Claim 1 wherein the assay system is selected from the group consisting of an apoptosis assay system, a cell proliferation assay system, an angiogenesis assay system, and a hypoxic induction assay system.

7. The method of Claim 1 wherein the assay system includes a binding assay comprising a SPHK polypeptide and the candidate test agent is an antibody.

8. The method of Claim 1 wherein the assay system includes an expression assay comprising a SPHK nucleic acid and the candidate test agent is a nucleic acid modulator.

9. The method of claim 8 wherein the nucleic acid modulator is an antisense oligomer.

10. The method of Claim 8 wherein the nucleic acid modulator is a PMO.

11. The method of Claim 1 additionally comprising:

(d) administering the candidate p53 pathway modulating agent identified in (c) to a model system comprising cells defective in p53 function and, detecting a phenotypic change in the model system that indicates that the p53 function is restored.

12. The method of Claim 11 wherein the model system is a mouse model with defective p53 function.

13. A method for modulating a p53 pathway of a cell comprising contacting a cell defective in p53 function with a candidate modulator that specifically binds to a SPHK polypeptide comprising an amino acid sequence selected from group consisting of SEQ ID NOs:10, 11, 12, and 13, whereby p53 function is restored.

14. The method of claim 13 wherein the candidate modulator is administered to a vertebrate animal predetermined to have a disease or disorder resulting from a defect in p53 function.

15. The method of Claim 13 wherein the candidate modulator is selected from the group consisting of an antibody and a small molecule.

16. The method of Claim 1, comprising the additional steps of:

(d) providing a secondary assay system comprising cultured cells or a non-human animal expressing SPHK ,

(e) contacting the secondary assay system with the test agent of (b) or an agent derived therefrom under conditions whereby, but for the presence of the test agent or agent derived therefrom, the system provides a reference activity; and

(f) detecting an agent-biased activity of the second assay system,

wherein a difference between the agent-biased activity and the reference activity of the second assay system confirms the test agent or agent derived therefrom as a candidate p53 pathway modulating agent,

and wherein the second assay detects an agent-biased change in the p53 pathway.

17. The method of Claim 16 wherein the secondary assay system comprises cultured cells.

18. The method of Claim 16 wherein the secondary assay system comprises a non-human animal.

19. The method of Claim 18 wherein the non-human animal mis-expresses a p53 pathway gene.

20. A method of modulating p53 pathway in a mammalian cell comprising contacting the cell with an agent that specifically binds a SPHK polypeptide or nucleic acid.

21. The method of Claim 20 wherein the agent is administered to a mammalian animal predetermined to have a pathology associated with the p53 pathway.

22. The method of Claim 20 wherein the agent is a small molecule modulator, a nucleic acid modulator, or an antibody.

23. A method for diagnosing a disease in a patient comprising:

(a) obtaining a biological sample from the patient;

(b) contacting the sample with a probe for SPHK expression;

(c) comparing results from step (b) with a control;

(d) determining whether step (c) indicates a likelihood of disease.

24. The method of claim 23 wherein said disease is cancer.

25. The method according to claim 24, wherein said cancer is a cancer as shown in Table 1 as having >25% expression level.

SEQUENCE LISTING

<110> EXELIXIS, INC.

<120> SPHKs AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE

<130> EX02-104C-PC

<150> US 60/310,362
<151> 2001-08-06

<150> US 60/357,501
<151> 2002-02-15

<160> 13

<170> PatentIn version 3.1

<210> 1
<211> 1783
<212> DNA
<213> Homo sapiens

<400> 1
ggagggagcg aggccgggga gtccgctcca gcggggcgct ccagtccctc agacgtgggc 60

tgagcttggg acgagctgcg ttccgcccca ggccactgta gggaacggcg gtggcgcctc 120

cccagcaaac cggaccgact gggtccagcc gccgcaggga atgacaccgg tgctcctaca 180

gccacggctc cgggcgggga aggcgagccc cacagccggc cctgcgacgc ccgcctgggc 240

agcaccgata aggagctgaa ggcaggagcc gccgccacgg gcagcgcccc cacagcgcca 300

gggacccccc ggcagcggga gccgcgggtc gaggttatgg atccagcggg cggccccccgg 360

ggcgtgctcc cgcggccctg ccgcgtgctg gtgctgctga acccgcgcgg cggcaagggc 420

aaggccttgc agctcttccg gagtcacgtg cagccccttt tggctgaggc tgaaatctcc 480

ttcacgctga tgctcactga gcggcggaac cacgcgcggg agctggtgcg gtcggaggag 540

ctgggccgct gggacgctct ggtggtcatg tctggagacg ggctgatgca cgaggtggtg 600

aacgggctca tggagcggcc tgactgggag accgccatcc agaagccccct gtgtagcctc 660

ccagcaggct ctggcaacgc gctggcagct tccttgaacc attatgctgg ctatgagcag 720

gtcaccaatg aagacctcct gaccaactgc acgctattgc tgtgccgccg gctgctgtca 780

cccatgaacc tgctgtctct gcacacggct tcggggctgc gcctcttctc tgtgctcagc 840

ctggcctggg gcttcattgc tgatgtggac ctagagagtg agaagtatcg gcgtctgggg 900

gagatgcgct tcactctggg caccttcctg cgtctggcag ccctgcgcac ctaccgcggc 960

cgactggcct acctccctgt aggaagagtg ggttccaaga cacctgcctc ccccgttgtg 1020

gtccagcagg gcccggtaga tgcacacctt gtgccactgg aggagccagt gccctctcac 1080

tggacaatgg tgcccgacga ggactttgtg ctaatcctgg cactgctgca ctcgcacctg 1140

ggcagtgaga tgtttgctgc acccatgggc cgctgtgcag ctggcgtcat gcatctgttc 1200

tacgtgcggg cgggagtgtc tcgtgccatg ctgctgcgct tcttcctggc catggagaag 1260

ggcaggcata tggagtatga atgcccctac ttggtatatg tgcccgtggt cgccttccgc 1320

ttggagccca aggatgggaa aggtgtgttt gcagtggatg gggaattgat ggttagcgag 1380

gccgtgcagg gccaggtgca cccaaactac ttctggatgg tcagtggttg cgtggagccc 1440

ccgcccagct ggaagcccca gcagatgcca ccgccagaag agcccttatg acccctgggc 1500

cacgctgtgc cttagtgtct acttgcagga cccttcctcc ttccctaggg ctgcagggcc 1560

tgtccacagt tcctgtgggg gtggaggaga ctcctctgga gaagggtgag aaggtggagg 1620

ttatgctttg gggggacagg ccagaatgaa gtcctgggtc aggagcccag ctggctgggc 1680

ccagctgcct atgtaaggcc ttctagtttg ttctgagacc cccaccccac gaaccaaatc 1740

caaataaagt gacattccca gcctgaaaaa aaaaaaaaaa aaa 1783

<210> 2
<211> 1272
<212> DNA
<213> Homo sapiens

<400> 2
cctgggcagc accgataagg agctgaaggc aggagccgcc gccacgggca gcgcccccac 60

agcgccaggg acccccctggc agcgggagcc gcgggtcgag gttatggatc cagcgggcgg 120

cccccggggc gtgctcccgc ggccctgccg cgtgctggtg ctgctgaacc cgcgcggcgg 180

caagggcaag gccttgcagc tcttccggag tcacgtgcag ccccttttgg ctgaggctga 240

aatctccttc acgctgatgc tcactgagcg gcggaaccac gcgcgggagc tggtgcggtc 300

ggaggagctg ggccgctggg acgctctggt ggtcatgtct ggagacgggc tgatgcacga 360

ggtggtgaac gggctcatgg agcggcctga ctgggagacc gccatccaga agcccctgtg 420

tagcctccca gcaggctctg gcaacgcgct ggcagcttcc ttgaaccatt atgctggcta 480

tgagcaggtc accaatgaag acctcctgac caactgcacg ctattgctgt gccgccggct 540

gctgtcaccc atgaacctgc tgtctctgca cacggcttcg gggctgcgcc tcttctctgt 600

gctcagcctg gcctggggct tcattgctga tgtggaccta gagagtgaga agtatcggcg 660

tctgggggag atgcgcttca ctctgggcac cttcctgcgt ctggcagccc tgcgcaccta 720

ccgcggccga ctggcctacc tccctgtagg aagagtgggt tccaagacac ctgcctcccc 780

cgttgtggtc cagcagggcc cggtagatgc acaccttgtg ccactggagg agccagtgcc 840

ctctcactgg acagtggtgc ccgacgagga ctttgtgcta gtcctggcac tgctgcactc 900

gcacctgggc agtgagatgt ttgctgcacc catgggccgc tgtgcagctg gcgtcatgca 960

tctgttctac gtgcgggcgg gagtgtctcg tgccatgctg ctgcgcctct tcctggccat 1020

ggagaagggc aggcatatgg agtatgaatg cccctacttg gtatatgtgc ccgtggtcgc 1080

cttccgcttg gagcccaagg atgggaaagg tgtgtttgca gtggatgggg aattgatggt 1140

tagcgaggcc gtgcagggcc aggtgcaccc aaactacttc tggatggtca gcggttgcgt 1200

ggagcccccg cccagctgga agccccagca gatgccaccg ccagaagagc ccttatgacc 1260

cctgggccac gc 1272

<210> 3
<211> 1857
<212> DNA
<213> Homo sapiens

<400> 3
atggccccgc ccccaccgcc actggctgcc agcaccccgc tcctccatgg cgagtttggc 60

tcctacccag cccgaggccc acgctttgcc ctcaccctta catcgcaggc cctgcacata 120

cagcggctgc gccccaaacc tgaagccagg ccccggggtg gcctggtccc gttggccgag 180

gtctcaggct gctgcaccct gcgaagccgc agcccctcag actcagcggc ctacttctgc 240

atctacacct accctcgggg ccggcgcggg gcccggcgca gagccactcg caccttccgg 300

gcagatgggg ccgccaccta cgaagagaac cgtgccgagg cccagcgctg ggccactgcc 360

ctcacctgtc tgctccgagg actgccactg cccggggatg gggagatcac ccctgacctg 420

ctacctcggc cgccccggtt gcttctattg gtcaatccct ttgggggtcg gggcctggcc 480

tggcagtggt gtaagaacca cgtgcttccc atgatctctg aagctgggct gtccttcaac 540

ctcatccaga cagaacgaca gaaccacgcc cgggagctgg tccaggggct gagcctgagt 600

gagtgggatg gcatcgtcac ggtctcggga gacgggctgc tccatgaggt gctgaacggg 660

ctcctagatc gccctgactg ggaggaagct gtgaagatgc ctgtgggcat cctcccctgc 720

ggctcgggca acgcgctggc cggagcagtg aaccagcacg ggggatttga gccagccctg 780

ggcctcgacc tgttgctcaa ctgctcactg ttgctgtgcc ggggtggtgg ccacccactg 840

gacctgctct ccgtgacgct ggcctcgggc tcccgctgtt tctccttcct gtctgtggcc 900

tggggcttcg tgtcagatgt ggatatccag agcgagcgct tcagggcctt gggcagtgcc 960

cgcttcacac tgggcacggt gctgggcctc gccacactgc acacctaccg cggacgcctc 1020

tcctacctcc ccgccactgt ggaacctgcc tcgcccaccc ctgcccatag cctgcctcgt 1080

gccaagtcgg agctgaccct aaccccagac ccagccccgc ccatggccca ctcacccctg 1140

catcgttctg tgtctgacct gcctcttccc ctgccccagc ctgccctggc ctctcctggc 1200

tcgccagaac ccctgcccat cctgtccctc aacggtgggg gcccagagct ggctggggac 1260

tggggtgggg ctggggatgc tccgctgtcc ccggacccac tgctgtcttc acctcctggc 1320

tctcccaagg cagctctaca ctcacccgtc tccgaagggg cccccgtaat tcccccatcc 1380

tctgggctcc cacttcccac ccctgatgcc cgggtagggg cctccacctg cggcccgccc 1440

gaccacctgc tgcctccgct gggcaccccg ctgcccccag actgggtgac gctggagggg 1500

gactttgtgc tcatgttggc catctcgccc agccacctag gcgctgacct ggtggcagct 1560

ccgcatgcgc gcttcgacga cggcctggtg cacctgtgct gggtgcgtag cggcatctcg 1620

cgggctgcgc tgctgcgcct ttcttggcc atggagcgtg gtagccactt cagcctgggc 1680

tgtccgcagc tgggctacgc cgcggcccgt gccttccgcc tagagccgct cacaccacgc 1740

ggcgtgctca cagtggacgg ggagcaggtg gagtatgggc cgctacaggc acagatgcac 1800

cctggcatcg gtacactgct cactgggcct cctggctgcc cggggcggga gccctga 1857

<210> 4
<211> 2731
<212> DNA
<213> Homo sapiens

<400> 4
ggggaacaaa ggtgagcgaa aggaggaggc agaatccggg cagagggcag ggagagggcc 60

tgtggggaag ggacctcagt cctgctccca cccgctccct ggagagcagg cggccagaca 120

cccagaggcc agaccaggag ctgaccggga gctggggcca cgggcctagg agcaccctgg 180

tcagggctaa ggccatggcc ccgcccccac cgccactggc tgccagcacc ccgctcctcc 240

atggcgagtt tggctcctac ccagcccgag gcccacgctt tgccctcacc cttacatcgc 300

aggccctgca catacagcgg ctgcgcccca aacctgaagc caggccccgg ggtggcctgg 360

tcccgttggc cgaggtctca ggctgctgca ccctgcgaag ccgcagcccc tcagactcag 420

cggcctactt ctgcatctac acctaccctc ggggccggcg cggggcccgg cgcagagcca 480

ctcgcacctt ccgggcagat ggggccgcca cctacgaaga gaaccgtgcc gaggcccagc 540

gctgggccac tgccctcacc tgtctgctcc gaggactgcc actgcccggg gatggggaga 600

tcacccctga cctgctacct cggccgcccc ggttgcttct attggtcaat ccctttgggg 660

gtcggggcct ggcctggcag tggtgtaaga accacgtgct tcccatgatc tctgaagctg 720

ggctgtcctt caacctcatc cagacagaac gacagaacca cgcccgggag ctggtccagg 780

ggctgagcct gagtgagtgg gatggcatcg tcacggtctc gggagacggg ctgctccatg 840

aggtgctgaa cgggctccta gatcgccctg actgggagga agctgtgaag atgcctgtgg 900

gcatcctccc ctgcggctcg ggcaacgcgc tggccggagc agtgaaccag cacgggggat 960

ttgagccagc cctgggcctc gacctgttgc tcaactgctc actgttgctg tgccggggtg 1020

gtggccaccc actggacctg ctctccgtga cgctggcctc gggctcccgc tgtttctcct 1080

tcctgtctgt ggcctggggc ttcgtgtcag atgtggatat ccagagcgag cgcttcaggg 1140

ccttgggcag tgcccgcttc acactgggca cggtgctggg cctcgccaca ctgcacacct 1200

accgcggacg cctctcctac ctccccgcca ctgtggaacc tgcctcgccc acccctgccc 1260

atagcctgcc tcgtgccaag tcggagctga ccctaacccc agacccagcc ccgcccatgg 1320

cccactcacc cctgcatcgt tctgtgtctg acctgcctct tcccctgccc cagcctgccc 1380

tggcctctcc tggctcgcca gaacccctgc ccatcctgtc cctcaacggt gggggcccag 1440

agctggctgg ggactggggt ggggctgggg atgctccgct gtccccggac ccactgctgt 1500

cttcacctcc tggctctccc aaggcagctc tacactcacc cgtctccgaa ggggcccccg 1560

taattccccc atcctctggg ctcccacttc ccacccctga tgcccgggta ggggcctcca 1620

cctgcggccc gcccgaccac ctgctgcctc cgctgggcac cccgctgccc ccagactggg 1680

tgacgctgga gggggacttt gtgctcatgt tggccatctc gcccagccac ctaggcgctg 1740

acctggtggc agctccgcat gcgcgcttcg acgacggcct ggtgcacctg tgctgggtgc 1800

gtagcggcat ctcgcgggct gcgctgctgc gccttttctt ggccatggag cgtggtagcc 1860

acttcagcct gggctgtccg cagctgggct acgccgcggc ccgtgccttc cgcctagagc 1920

cgctcacacc acgcggcgtg ctcacagtgg acggggagca ggtggagtat gggccgctac 1980

aggcacagat gcaccctggc atcggtacac tgctcactgg gcctcctggc tgcccggggc 2040

gggagccctg aaactaaaca agcttggtac ccgccggggg cggggcctac attccaatgg 2100

ggcggagcct gagctagggg gtgtggcctg gctgctagag ttgtggtggc aggggccctg 2160

gccccgtctc aggattgcgc tcgctttcat gggaccagac gtgatgctgg aaggtgggcg 2220

tcgtcacggt taaagagaaa tgggctcgtc ccgagggtag tgcctgatca atgagggcgg 2280

ggcctggcgt ctgatctggg gccgcccttа cggggcaggg ctcagtcctg acgcttgcca 2340

cctgctccta cccggccagg atggctgagg gcggagtcta ttttacgcgt cgcccaatga 2400

caggacctgg aatgtactgg ctggggtagg cctcagtgag tcggccggtc agggcccgca 2460

gcctcgcccc atccactccg gtgcctccat ttagctggcc aatcagccca ggaggggcag 2520

gttccccggg gccggcgcta ggatttgcac taatgttcct ctccccgcgg gtgggggcgg 2580

ggaaattcat atcccctgtt cgtctcatgc gcgtcctccg tccccaatct aaaaagcaat 2640

tgaaaaggtc tatgcaataa aggcagtcgc ttcattcctc tcaaaaaaaa aaaaaaaaaa 2700

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa a 2731

<210> 5
<211> 1857
<212> DNA
<213> Homo sapiens

<400> 5

```
atggccccgc ccccaccgcc actggctgcc agcaccccgc tcctccatgg cgagtttggc     60
tcctacccag cccgaggccc acgctttgcc ctcaccctta catcgcaggc cctgcacata    120
cagcggctgc gccccaaacc tgaagccagg ccccggggtg gcctggtccc gttggccgag    180
gtctcaggct gctgcaccct gcgaagccgc agcccctcag actcagcggc ctacttctgc    240
atctacacct accctcgggg ccggcgcggg gcccggcgca gagccactcg caccttccgg    300
gcagatgggg ccgccaccta cgaagagaac cgtgccgagg cccagcgctg ggccactgcc    360
ctcacctgtc tgctccgagg actgccactg cccggggatg gggagatcac ccctgacctg    420
ctacctcggc cgccccggtt gcttctattg gtcaatccct ttgggggtcg gggcctggcc    480
tggcagtggt gtaagaacca cgtgcttccc atgatctctg aagctgggct gtccttcaac    540
ctcatccaga cagaacgaca gaaccacgcc cgggagctgg tccaggggct gagcctgagt    600
gagtgggatg gcatcgtcac ggtctcggga gacgggctgc tccatgaggt gctgaacggg    660
ctcctagatc gccctgactg ggaggaagct gtgaagatgc ctgtgggcat cctcccctgc    720
ggctcgggca acgcgctggc cggagcagtg aaccagcacg ggggatttga gccagccctg    780
ggcctcgacc tgttgctcaa ctgctcactg ttgctgtgcc ggggtggtgg ccacccactg    840
gacctgctct ccgtgacgct ggcctcgggc tcccgctgtt tctccttcct gtctgtggcc    900
tggggcttcg tgtcagatgt ggatatccag agcgagcgct tcagggcctt gggcagtgcc    960
cgcttcacac tgggcacggt gctgggcctc gccacactgc acacctaccg cggacgcctc   1020
tcctacctcc ccgccactgt ggaacctgcc tcgcccaccc ctgcccatag cctgcctcgt   1080
gccaagtcgg agctgaccct aaccccagac ccagccccgc ccatggccca ctcacccctg   1140
catcgttctg tgtctgacct gcctcttccc ctgccccagc ctgccctggc ctctcctggc   1200
tcgccagaac ccctgcccat cctgtccctc aacggtgggg gcccagagct ggctggggac   1260
tggggtgggg ctggggatgc tccgctgtcc ccggacccac tgctgtcttc acctcctggc   1320
tctcccaagg cagctctaca ctcacccgtc tccgaagggg cccccgtaat tcccccatcc   1380
tctgggctcc cacttcccac ccctgatgcc cgggtagggg cctccacctg cggcccgccc   1440
gaccacctgc tgcctccgct gggcaccccg ctgcccccag actgggtgac gctggagggg   1500
gactttgtgc tcatgttggc catctcgccc agccacctag gcgctgacct ggtggcagct   1560
ccgcatgcgc gcttcgacga cggcctggtg cacctgtgct gggtgcgtag cggcatctcg   1620
cgggctgcgc tgctgcgcct ttcttggcc atggagcgtg gtagccactt cagtctgggc   1680
tgtccgcagc tgggctacgc cgcggcccgt gccttccgcc tagagccgct cacaccacgc   1740
ggcgtgctca cagtggacgg ggagcaggtg gagtatgggc cgctacaggc acagatgcac   1800
cctggcatcg gtacactgct cactgggcct cctggctgcc cggggcggga gccctga      1857
```

<210> 6
<211> 2875
<212> DNA
<213> Homo sapiens

<400> 6
agtgttggag gtgaggaggc ggggctggca gggctagtcg gggcatctgg aaatttccga 60

ccccacgctt cgggcgtttc cttatcaggt tcaccgctcc ctgatctcgc gctgcacttc 120

gtaggcgcag ccgctgcttg ggaagtccta cttaagagct gaaggtcagg ccaggacagt 180

gagacctgac tccttgctcc taccagccta ctatggctta agacccaggg ccagggtccc 240

gttgatgtaa cagagcagag gaccagcaga tgaatggaca ccttgaagca gaggagcagc 300

aggaccagag gccagaccag gagctgaccg ggagctgggg ccacgggcct aggagcaccc 360

tggtcagggc taaggccatg gccccgcccc caccgccact ggctgccagc acctcgctcc 420

tccatggcga gtttggctcc tacccagccc gaggcccacg ctttgccctc acccttacat 480

cgcaggccct gcacatacag cggctgcgcc ccaaacctga agccaggccc cggggtggcc 540

tggtcccgtt ggccgaggtc tcaggctgct gcaccctgcg aagccgcagc ccctcagact 600

cagcggccta cttctgcatc tacacctacc ctcggggccg gcgcggggcc cggcgcagag 660

ccactcgcac cttccgggca gatggggccg ccacctacga agagaaccgt gccgaggccc 720

agcgctgggc cactgccctc acctgtctgc tccgaggact gccactgccc ggggatgggg 780

agatcacccc tgacctgcta cctcggccgc cccggttgct tctattggtc aatccctttg 840

ggggtcgggg cctggcctgg cagtggtgta agaaccacgt gcttcccatg atctctgaag 900

ctgggctgtc cttcaacctc atccagacag aacgacagaa ccacgcccgg gagctggtcc 960

aggggctgag cctgagtgag tgggatggca tcgtcacggt ctcgggagac gggctgctcc 1020

atgaggtgct gaacgggctc ctagatcgcc ctgactggga ggaagctgtg aagatgcctg 1080

tgggcatcct cccctgcggc tcgggcaacg cgctggccgg agcagtgaac cagcacgggg 1140

gatttgagcc agccctgggc ctcgacctgt tgctcaactg ctcactgttg ctgtgccggg 1200

gtggtggcca cccactggac ctgctctccg tgacgctggc ctcgggctcc cgctgtttct 1260

ccttcctgtc tgtggcctgg ggcttcgtgt cagatgtgga tatccagagc gagcgcttca 1320

gggccttggg cagtgcccgc ttcacactgg gcacggtgct gggcctcgcc acactgcaca 1380

cctaccgcgg acgcctctcc tacctccccg ccactgtgga acctgcctcg cccacccctg 1440

cccatagcct gcctcgtgcc aagtcggagc tgaccctaac cccagaccca gccccgccca 1500

tggcccactc acccctgcat cgttctgtgt ctgacctgcc tcttcccctg ccccagcctg 1560

ccctggcctc tcctggctcg ccagaacccc tgcccatcct gtccctcaac ggtgggggcc 1620

cagagctggc tggggactgg ggtggggctg gggatgctcc gctgtccccg gacccactgc 1680

tgtcttcacc tcctggctct cccaaggcag ctctacactc acccgtctcc gaaggggccc 1740

ccgtaattcc cccatcctct gggctcccac ttcccacccc tgatgcccgg gtaggggcct 1800

ccacctgcgg cccgcccgac cacctgctgc ctccgctagg caccccgctg cccccagact 1860

gggtgacgct ggaggggggac tttgtgctca tgttggccat ctcgcccagc cacctaggcg 1920

ctgacctggt ggcagctccg catgcgcgct tcgacgacgg cctggtgcac ctgtgctggg 1980

tgcgtagcgg catctcgcgg gctgcgctgc tgcgcctttt cttggccatg gagcgtggta 2040

gccacttcag cctgggctgt ccgcagctgg gctacgccgc ggcccgtgcc ttccgcctag 2100

agccgctcac accacgcggc gtgctcacag tggacgggga gcaggtggag tatgggccgc 2160

tacaggcaca gatgcaccct ggcatcggta cactgctcac tgggcctcct ggctgcccgg 2220

ggcgggagcc ctgaaactaa acaagcttgg tacccgccgg gggcggggcc tacattccaa 2280

tggggcggag cctgagctag ggggtgtggc ctggctgcta gagttgtggt ggcaggggcc 2340

ctggccccgt ctcaggattg cgctcgcttt catgggacca gacgtgatgc tggaaggtgg 2400

gcgtcgtcac ggttaaagag aaatgggctc gtcccgaggg tagtgcctga tcaatgaggg 2460

cggggcctgg cgtctgatct ggggccgccc ttacggggca gggctcagtc ctgacgcttg 2520

ccacctgctc ctacccggcc aggatggctg agggcggagt ctattttacg cgtcgcccaa 2580

tgacaggacc tggaatgtac tggctggggt aggcctcagt gagtcggccg gtcagggccc 2640

gcagcctcgc cccatccact ccggtgcctc catttagctg gccaatcagc ccaggagggg 2700

caggttcccc ggggccggcg ctaggatttg cactaatgtt cctctccccg cgggtggggg 2760

cggggaaatt catatcccct gttcgtctca tgcgcgtcct ccgtccccaa tctaaaaagc 2820

aattgaaaag gtctatgcaa taaaggcagt cgcttcattc ctctcaaaaa aaaaa 2875

<210> 7
<211> 4171
<212> DNA
<213> Homo sapiens

<400> 7
gatcatcgcc gttgaggaaa cagacgttca cgggaaacat caaggcagtg gaaaatggca 60

gaaaatggaa aagccttacg cttttacagt tcactgtgta aagagagcac gacggcaccg 120

ctggaagtgg gcgcaggtga ctttctggtg tccagaggag cagctgtgtc acttgtggct 180

gcagaccctg cgggagatgc tggagaagct gacgtccaga ccaaagcatt tactggtatt 240

tatcaacccg tttggaggaa aaggacaagg caagcggata tatgaaagaa aagtggcacc 300

actgttcacc ttagcctcca tcaccactga catcatcgtt actgaacatg ctaatcaggc 360

caaggagact ctgtatgaga ttaacataga caaatacgac ggcatcgtct gtgtcggcgg 420

agatggtatg ttcagcgagg tgctgcacgg tctgattggg aggacgcaga ggagcgccgg 480

ggtcgaccag aaccaccccc gggctgtgct ggtccccagt agcctccgga ttggaatcat 540

tcccgcaggg tcaacggact gcgtgtgtta ctccaccgtg ggcaccagcg acgcagaaac 600

ctcggcgctg catatcgttg ttggggactc gctggccatg gatgtgtcct cagtccacca 660

caacagcaca ctccttcgct actccgtgtc cctgctgggc tacggcttct acggggacat 720

catcaaggac agtgagaaga aacggtggtt gggtcttgcc agatacgact tttcaggttt 780

aaagaccttc ctctcccacc actgctatga agggacagtg tccttcctcc ctgcacaaca 840

cacggtggga tctccaaggg ataggaagcc ctgccgggca ggatgctttg tttgcaggca 900

aagcaagcag cagctggagg aggagcagaa gaaagcactg tatggtttgg aagctgcgga 960

ggacgtggag gagtggcaag tcgtctgtgg gaagtttctg gccatcaatg ccacaaacat 1020

gtcctgtgct tgtcgccgga gccccagggg cctctccccg gctgcccact tgggagacgg 1080

gtcttctgac ctcatcctca tccggaaatg ctccaggttc aattttctga gatttctcat 1140

caggcacacc aaccagcagg accagtttga cttcactttt gttgaagttt atcgcgtcaa 1200

gaaattccag tttacgtcga agcacatgga ggatgaggac agcgacctca aggaggggg 1260

gaagaagcgc tttgggcaca tttgcagcag ccacccctcc tgctgctgca ccgtctccaa 1320

cagctcctgg aactgcgacg gggaggtcct gcacagccct gccatcgagg tcagagtcca 1380

ctgccagctg gttcgactct ttgcacgagg aattgaagag aatccgaagc cagactcaca 1440

cagctgagaa gccggcgtcc tgctctcgaa ctgggaaagt gtgaaaacta tttaagataa 1500

ttattacaga ccaattatgt tgatatatac atttaaatgt agaaatttat ttttgatagt 1560

taaatcttga ttttagaaga aaaccctttt gtcaacaatt ttgtgtacat atttggcatt 1620

ttcagttctg tacgcatctg cgggttgcag cccacgccgc ttactctcag cggatgcagc 1680

tgctcacttg ggggcactgg cctcttaggt tttaacgatg tcaacagtgt agtttagaaa 1740

atggcccgtt agtggctcta ttgcaataat gttagggaca ttatatgatt tccacgcagg 1800

tcacaccatc tgggcctgag gtagcagtgg gtcactttga tccactttgc aggacttatt 1860

ctgtaacggt ttgtggccaa gttttgggaa gtggttgatt ctctttgcct tcatttcacc 1920

ttcctcttcg tttacggtta ggacatcgct gcttgatcct tacaatactg tgcaactgca 1980

atgcaacgtg gccctgcttc aggtgatccg cgggaggggc ctccacgcca gcgccgggaa 2040

ggctgctggg gcctccacac ctgcctcatc acggcggcga ggctacgaca atccggctgg 2100

gagcatgacc ttggcgtctg ttctgggagc acagatgata agctctggaa gctggcagtg 2160

tgtaaagcac tggcaagttt gttactgtta aaatgtcaaa taccaatgct ttatatcgac 2220

gcgaagtgct taacacagcc gggcttgggg gcagtcagga ggaagctggc catccgtgga 2280

ggaggggccg gtcctggact cccgcaggac tcctctgatg cagggcctga agtctgtaca 2340

cgtggtccag atttgtcctt gtcttttctt cacactgagt tctctatatt tattgaacat 2400

cttgtccttt taagccagag tagtgtaaac tgcgtctcgg atgtctgtct tttgcctcga 2460

agccacgatg gatcgctggt ttcctctgca gcgcgagggc tccggcgacc agaggattct 2520

tcccggaagg cattcctgcc gcgctccccg gggcacccct caattgtgta ctacgtcctt 2580

gtttagtgtg tatccgtgcc cacgtagatg atgtctgtaa cgtagttttg tttgaaatat 2640

gagaatatgc ggcttaaact ttgatctgta aggagcgggg ccgtggccgt ttggagcacg 2700

ctgtagacac cgttcctcat gctgccgggt gggttttgca gaagctccct tagtgatttc 2760

atgtttaaca ggcagcatcc attttcagaa tttcctggca ttgatttata ttttgaagca 2820

tacaggaaac ttctcgtttc ctcgtttagc cccacccaga tcaggtgaaa gggcagcttt 2880

aatggtggtt tttatggacc acattatcag agagcactgt gcaagccaaa tggttcaata 2940

atgaatgaaa attctgggtg taaagagtaa atatgccctg gctctttcta ccaatgtttg 3000

ctcctggttg gaaagaaacc aaagatttaa gacgggctgc tcttccagac tggctgtgcc 3060

tgcctgtgcc cagcaacctg tgcagccggc agtgtgcctg gtgtcacgcc aggaggctgt 3120

ggctgctgtg ggccctctgg aattgtgctc ctcacaaagt ttccccaaaa ggttcttcta 3180

agcctttatt gtccctggta aatgtttccc ggctgggcgc ggtggctcac gcctgtaatc 3240

ccagcacttt gggaggccga ggcgggtgga tcacctaagg tcaggagttt gagatcagcc 3300

tgcccaacat ggtgaaacct cgtctctact aaaaatacac aacttagcca gtcttgttgg 3360

cgcacgcctg taatctcagc tactagggac gctgaggcag gagaatcgct tgaacccaag 3420

aaagaggtgg aggttgcggt gagccaagat tgcgccactg cactccagcc tgggcaaaca 3480

gagggagact ccatcgcccc cccccaacaa aaaaaaaagt ttcccataca ctggcctgcc 3540

ccaaaaccca ctaacaattt tagcaaaaca gtccaggcca aagaggaagc atttcatgtt 3600

caataagaaa cccagccatt ccgcatggct ggttcctgag tggctctggt gatactctcc 3660

agccacctgc tgacattgag aatctcagac ctcgggactg ctgttgcggt accgtgtgtc 3720

tgacacctgc cagcagccct ttgctatctg cgcgcaggat gggggtgact gcccagacat 3780

tcccgctaga taggctctga tttccggggc agcctttcag atgcggcaga catacaacac 3840

ctgtacttta gagttttaag ggaaaaaaaa tcagaagtgc tggttagata gtaaaaactt 3900

aggataactt agaaaggcta gttttagctt cctttgtggc tccctggtgc aaaacaatta 3960

gcagttatgc aatggacctg attctagttt attctaatta agaagtgagg ccgagtttga 4020

cttcgttcct gaatacaatc ttgagtaact gggaaagtct gagtgaaagg atggcctcat 4080

tctctttcta atcttgctgg tttcaagatt agaaaatggc attatttgat ctgaaatgtt 4140

tgagaagaca cgaataaagt tacttgggca g 4171

<210> 8
<211> 4407
<212> DNA
<213> Homo sapiens

<400> 8
ggccgctaac ggtccggcgc ccctcggcgt ccgcgcgccc ccagcctggc ggacgagccc 60

ggcggcggag atgggggcga cgggggcggc ggagccgctg caatccgtgc tgtgggtgaa 120

gcagcagcgc tgcgccgtga gcctggagcc cgcgcgggct ctgctgcgct ggtggcggag 180

cccggggccc ggagccggcg cccccggcgc ggatgcctgc tctgtgcctg tatctgagat 240

catcgccgtt gaggaaacag acgttcacgg gaaacatcaa ggcagtggaa aatggcagaa 300

aatggaaaag ccttacgctt ttacagttca ctgtgtaaag agagcacgac ggcaccgctg 360

gaagtgggcg caggtgactt tctggtgtcc agaggagcag ctgtgtcact tgtggctgca 420

gaccctgcgg gagatgctgg agaagctgac gtccagacca aagcatttac tggtatttat 480

caacccgttt ggaggaaaag gacaaggcaa gcggatatat gaaagaaaag tggcaccact 540

gttcacctta gcctccatca ccactgacat catcgttact gaacatgcta atcaggccaa 600

ggagactctg tatgagatta acatagacaa atacgacggc atcgtctgtg tcggcggaga 660

tggtatgttc agcgaggtgc tgcacggtct gattgggagg acgcagagga gcgccggggt 720

cgaccagaac caccccggg ctgtgctggt ccccagtagc ctccggattg gaatcattcc 780

cgcagggtca acggactgcg tgtgttactc caccgtgggc accagcgacg cagaaacctc 840

ggcgctgcat atcgttgttg gggactcgct ggccatggat gtgtcctcag tccaccacaa 900

cagcacactc cttcgctact ccgtgtccct gctgggctac ggcttctacg gggacatcat 960

caaggacagt gagaagaaac ggtggttggg tcttgccaga tacgactttt caggtttaaa 1020

gaccttcctc tcccaccact gctatgaagg gacagtgtcc ttcctccctg cacaacacac 1080

ggtgggatct ccaagggata ggaagccctg ccgggcagga tgctttgttt gcaggcaaag 1140

caagcagcag ctggaggagg agcagaagaa agcactgtat ggtttggaag ctgcggagga 1200

cgtggaggag tggcaagtcg tctgtgggaa gtttctggcc atcaatgcca caaacatgtc 1260

ctgtgcttgt cgccggagcc ccaggggcct ctccccggct gcccacttgg gagacgggtc 1320

ttctgacctc atcctcatcc ggaaatgctc caggttcaat tttctgagat ttctcatcag 1380

gcacaccaac cagcaggacc agtttgactt cacttttgtt gaagtttatc gcgtcaagaa 1440

attccagttt acgtcgaagc acatggagga tgaggacagc gacctcaagg agggggggaa 1500

gaagcgcttt gggcacattt gcagcagcca cccctcctgc tgctgcaccg tctccaacag 1560

ctcctggaac tgcgacgggg aggtcctgca cagccctgcc atcgaggtca gagtccactg 1620

gcactttggg aggccgaggc gggtggatca cctaaggtca ggagtttgag atcagcctgc 3540

ccaacatggt gaaacctcgt ctctactaaa aatacacaac ttagccagtc ttgttggcgc 3600

acgcctgtaa tctcagctac tagggatgct gaggcaggag aatcgcttga acccaagaaa 3660

gaggtggagg ttgcggtgag ccaagattgc gccactgcac tccagcctgg gcaaacagag 3720

ggagactcca tcgccccccc caacaaaaaa aaaagtttcc catacactgg cctgccccaa 3780

aacccactaa caattttagc aaaacagtcc aggccaaaga ggaagcattt catgttcaat 3840

aagaaaccca gccattccgc atggctggtt cctgagtggc tctggtgata ctctccagcc 3900

acctgctgac attcagaatc tcagacctcg ggactgctgt tgcggtaccg tgtgtctgac 3960

acctgccagc agccctttgc tatctgcgcg caggatgggg gtgactgccc agacattccc 4020

gctagatagg ctctgatttc cggggcagcc tttcagatgc ggcagacata caacacctgt 4080

actttagagt tttaagggaa aaaaaatcag aagtgctggt tagatagtaa aaacttagga 4140

taacttagaa aggctagttt tagcttcctt tgtggctccc tggtgcaaaa caattagcag 4200

ttatgcaatg gacctgattc tagtttattc taattaagaa gtgaggccga gtttgacttc 4260

gttcctgaat acaatcttga gtaactggga aagtctgagt gaaaggatgg cctcattctc 4320

tttctaatct tgctggtttc aagattagaa aatggcatta tttgatctga aatgtttgag 4380

aagacacgaa taaagttact tgggcag 4407

<210> 9
<211> 2619
<212> DNA
<213> Homo sapiens

<400> 9
gattttagaa gaaaaccctt ttgtcaacaa ttttgtgtac atatttggca ttttcagttc 60

tgtacgcatc tgcgggttgc agcccacgcc gcttactctc agcggatgca gctgctcact 120

tgggggcact ggcctcttag gttttaacga tgtcaacagt gtagtttaga aaatggcccg 180

ttagtggctc tattgcaata atgttaggga cattatatga tttccacgca ggtcacacca 240

tctgggcctg aggtagcagt gggtcacttt gatccacttt gcaggactta ttctgtaacg 300

gtttgtggcc aagttttggg aagtggttga ttctctttgc cttcatttca ccttcctctt 360

cgtttacggt taggacatcg ctgcttgatc cttacaatac tgtgcaactg caatgcaacg 420

tggccctgct tcaggtgatc cgcgggaggg gcctccacgc cagcgccggg aaggctgctg 480

gggcctccac acctgcctca tcacggcggc gaggctacga caatccggct gggagcatga 540

ccttggcgtc tgttctggga gcacggatga taagctctgg aagctggcag tgtgtaaagc 600

actggcaagt ttgttactgt taaaatgtca aataccaatg ctttatatcg acgcgaagtg 660

cttaacacag ccgggcttgg gggcagtcag gaggaagctg gccatccgtg gaggaggggc 720

cggtcctgga ctcccgcagg actcctctga ggcagggcct gaagtctgta cacgtggtcc 780

agatttgtcc ttgtcttttc ttcacactga gttctctata tttattgaac atcttgtcct 840

tttaagccag agtagtgtaa actgcgtctc ggatgtctgt cttttgcctc gaagccacga 900

tggatcgctg gtttcctctg cagcgcgagg gctccggcga ccagaggatt cttcccggaa 960

ggcattcctg ccgcgctccc cggggcaccc ctcaattgtg tactacgtcc ttgtttagtg 1020

tgtatccgtg cccacgtaga tgatgtctgt aacgtagttt tgtttgaaat atgagaatat 1080

gcggcttaaa ctttgatctg taaggagcgg ggccgtggcc gtttggagca cgctgtagac 1140

accgttcctc atgctgccgg gtgggttttg cagaagctcc cttagtgatt tcatgtttaa 1200

caggcagcat ccatttccag aatttcctgg cattgattta tattttgaag catacaggaa 1260

acttctcgtt tcctcgttta gccccaccca gatcaggtga aagggcagct ttaatggtgg 1320

tttttatgga ccacattatc agagagcact gtgcaagcca aatggttcaa taatgaatga 1380

aaattctggg tgtaaagagt aaatatgccc tggctctttc taccaatgtt tgctcctggt 1440

tggaaagaaa ccaaagattt aagacgggct gctcttccag actggctgtg cctgcctgtg 1500

cccagcaacc tgtgcagccg gcagtgtgcc tggtgtcacg ccaggaggct gtggctgctg 1560

tgggccctct ggaattgtgc tcctcacaaa gtttccccaa aaggttcttc taagccttta 1620

ttgtccctgg taaatgtttc ccggctgggc gcggtggctc acgcctgtaa tcccagcact 1680

ttgggaggcc gaggcgggtg gatcacctaa ggtcaggagt ttgagatcag cctgcccaac 1740

atggtgaaac ctcgtctcta ctaaaaatac acaacttagc cagtcttgtt ggcgcacgcc 1800

tgtaatctca gctactaggg atgctgaggc aggagaatcg cttgaaccca agaaagaggt 1860

ggaggttgcg gtgagccaag attgcgccac tgcactccag cctgggcaaa cagagggaga 1920

ctccatcgcc ccccccaaca aaaaaaaaag tttcccatac actggcctgc cccaaaaccc 1980

actaacaatt ttagcaaaac agtccaggcc aaagaggaag catttcatgt tcaataagaa 2040

acccagccat tccgcatggc tggttcctga gtggctctgg tgatactctc cagccacctg 2100

ctgacattca gaatctcaga cctcgggact gctgttgcgg taccgtgtgt ctgacacctg 2160

ccagcagccc tttgctatct gcgcgcagga tgggggtgac tgcccagaca ttcccgctag 2220

ataggctctg atttccgggg cagcctttca gatgcggcag acatacaaca cctgtacttt 2280

agagttttaa gggaaaaaaa atcagaagtg ctggttagat agtaaaaact taggataact 2340

tagaaaggct agttttagct tcctttgtgg ctccctggtg caaaacaatt agcagttatg 2400

caatggacct gattctagtt tattctaatt aagaagtgag gccgagtttg acttcgttcc 2460

tgaatacaat cttgagtaac tgggaaagtc tgagtgaaag gatggcctca ttctctttct 2520

ccagctggtt cgactctttg cacgaggaat tgaagagaat ccgaagccag actcacacag 1680

ctgagaagcc ggcgtcctgc tcacaaactg ggaaagtgtg aaaactattt aagataatta 1740

ttacagacca attatgttga tatatacatt taaatgtaga aatttatttt tgatagttaa 1800

atcttgattt tagaagaaaa ccctttttgtc aacaattttg tgtacatatt tggcattttc 1860

agttctgtac gcatctgcgg gttgcagccc acgccgctta ctctcagcgg atgcagctgc 1920

tcacttgggg gcactggcct cttaggtttt aacgatgtca acagtgtagt ttagaaaatg 1980

gcccgttagt ggctctattg caataatgtt agggacatta tatgatttcc acgcaggtca 2040

caccatctgg gcctgaggta gcagtgggtc actttgatcc actttgcagg acttattctg 2100

taacggtttg tggccaagtt ttgggaagtg gttgattctc tttgccttca tttcaccttc 2160

ctcttcgttt acggttagga catcgctgct tgatccttac aatactgtgc aactgcaatg 2220

caacgtggcc ctgcttcagg tgatccgcgg gaggggcctc cacgccagcg ccgggaaggc 2280

tgctggggcc tccacacctg cctcatcacg gcggcgaggc tacgacaatc cggctgggag 2340

catgaccttg gcgtctgttc tgggagcacg gatgataagc tctggaagct ggcagtgtgt 2400

aaagcactgg caagtttgtt actgttaaaa tgtcaaatac caatgcttta tatcgacgcg 2460

aagtgcttaa cacagccggg cttgggggca gtcaggagga agctggccat ccgtggagga 2520

ggggccggtc ctggactccc gcaggactcc tctgaggcag ggcctgaagt ctgtacacgt 2580

ggtccagatt tgtccttgtc ttttcttcac actgagttct ctatatttat tgaacatctt 2640

gtccttttaa gccagagtag tgtaaactgc gtctcggatg tctgtctttt gcctcgaagc 2700

cacgatggat cgctggtttc ctctgcagcg cgagggctcc ggcgaccaga ggattcttcc 2760

cggaaggcat tcctgccgcg ctccccgggg caccctcaa ttgtgtacta cgtccttgtt 2820

tagtgtgtat ccgtgcccac gtagatgatg tctgtaacgt agttttgttt gaaatatgag 2880

aatatgcggc ttaaactttg atctgtaagg agcggggccg tggccgtttg gagcacgctg 2940

tagacaccgt tcctcatgct gccgggtggg ttttgcagaa gctcccttag tgatttcatg 3000

tttaacaggc agcatccatt ttcagaattt cctggcattg atttatattt tgaagcatac 3060

aggaaacttc tcgtttcctc gtttagcccc acccagatca ggtgaaaggg cagctttaat 3120

ggtggttttt atggaccaca ttatcagaga gcactgtgca agccaaatgg ttcaataatg 3180

aatgaaaatt ctgggtgtaa agagtaaata tgccctggct ctttctacca atgtttgctc 3240

ctggttggaa agaaaccaaa gatttaagac gggctgctct tccagactgg ctgtgcctgc 3300

ctgtgcccag caacctgtgc agccggcagt gtgcctggtg tcacgccagg aggctgtggc 3360

tgctgtgggc cctctggaat tgtgctcctc acaaagtttc cccaaaaggt tcttctaagc 3420

ctttattgtc cctggtaaat gtttcccggc tgggcgcggt ggctcacgcc tgtaatccca 3480

gcactttggg aggccgaggc gggtggatca cctaaggtca ggagtttgag atcagcctgc 3540

ccaacatggt gaaacctcgt ctctactaaa aatacacaac ttagccagtc ttgttggcgc 3600

acgcctgtaa tctcagctac tagggatgct gaggcaggag aatcgcttga acccaagaaa 3660

gaggtggagg ttgcggtgag ccaagattgc gccactgcac tccagcctgg gcaaacagag 3720

ggagactcca tcgccccccc caacaaaaaa aaaagtttcc catacactgg cctgccccaa 3780

aacccactaa caattttagc aaaacagtcc aggccaaaga ggaagcattt catgttcaat 3840

aagaaaccca gccattccgc atggctggtt cctgagtggc tctggtgata ctctccagcc 3900

acctgctgac attcagaatc tcagacctcg ggactgctgt tgcggtaccg tgtgtctgac 3960

acctgccagc agccctttgc tatctgcgcg caggatgggg gtgactgccc agacattccc 4020

gctagatagg ctctgatttc cggggcagcc tttcagatgc ggcagacata caacacctgt 4080

actttagagt tttaagggaa aaaaaatcag aagtgctggt tagatagtaa aaacttagga 4140

taacttagaa aggctagttt tagcttcctt tgtggctccc tggtgcaaaa caattagcag 4200

ttatgcaatg gacctgattc tagtttattc taattaagaa gtgaggccga gtttgacttc 4260

gttcctgaat acaatcttga gtaactggga aagtctgagt gaaaggatgg cctcattctc 4320

tttctaatct tgctggtttc aagattagaa aatggcatta tttgatctga aatgtttgag 4380

aagacacgaa taaagttact tgggcag 4407

<210> 9
<211> 2619
<212> DNA
<213> Homo sapiens

<400> 9
gattttagaa gaaaaccctt ttgtcaacaa ttttgtgtac atatttggca ttttcagttc 60

tgtacgcatc tgcgggttgc agcccacgcc gcttactctc agcggatgca gctgctcact 120

tgggggcact ggcctcttag gttttaacga tgtcaacagt gtagtttaga aaatggcccg 180

ttagtggctc tattgcaata atgttaggga cattatatga tttccacgca ggtcacacca 240

tctgggcctg aggtagcagt gggtcacttt gatccacttt gcaggactta ttctgtaacg 300

gtttgtggcc aagttttggg aagtggttga ttctctttgc cttcatttca ccttcctctt 360

cgtttacggt taggacatcg ctgcttgatc cttacaatac tgtgcaactg caatgcaacg 420

tggccctgct tcaggtgatc cgcgggaggg gcctccacgc cagcgccggg aaggctgctg 480

gggcctccac acctgcctca tcacggcggc gaggctacga caatccggct gggagcatga 540

ccttggcgtc tgttctggga gcacggatga taagctctgg aagctggcag tgtgtaaagc 600

actggcaagt ttgttactgt taaaatgtca aataccaatg ctttatatcg acgcgaagtg 660

cttaacacag ccgggcttgg gggcagtcag gaggaagctg gccatccgtg gaggaggggc 720

cggtcctgga ctcccgcagg actcctctga ggcagggcct gaagtctgta cacgtggtcc 780

agatttgtcc ttgtcttttc ttcacactga gttctctata tttattgaac atcttgtcct 840

tttaagccag agtagtgtaa actgcgtctc ggatgtctgt cttttgcctc gaagccacga 900

tggatcgctg gtttcctctg cagcgcgagg gctccggcga ccagaggatt cttcccggaa 960

ggcattcctg ccgcgctccc cggggcaccc ctcaattgtg tactacgtcc ttgtttagtg 1020

tgtatccgtg cccacgtaga tgatgtctgt aacgtagttt tgtttgaaat atgagaatat 1080

gcggcttaaa ctttgatctg taaggagcgg ggccgtggcc gtttggagca cgctgtagac 1140

accgttcctc atgctgccgg gtgggttttg cagaagctcc cttagtgatt tcatgtttaa 1200

caggcagcat ccatttccag aatttcctgg cattgattta tattttgaag catacaggaa 1260

acttctcgtt tcctcgttta gccccaccca gatcaggtga aagggcagct ttaatggtgg 1320

tttttatgga ccacattatc agagagcact gtgcaagcca aatggttcaa taatgaatga 1380

aaattctggg tgtaaagagt aaatatgccc tggctctttc taccaatgtt tgctcctggt 1440

tggaaagaaa ccaaagattt aagacgggct gctcttccag actggctgtg cctgcctgtg 1500

cccagcaacc tgtgcagccg gcagtgtgcc tggtgtcacg ccaggaggct gtggctgctg 1560

tgggccctct ggaattgtgc tcctcacaaa gtttccccaa aaggttcttc taagccttta 1620

ttgtccctgg taaatgtttc ccggctgggc gcggtggctc acgcctgtaa tcccagcact 1680

ttgggaggcc gaggcgggtg gatcacctaa ggtcaggagt ttgagatcag cctgcccaac 1740

atggtgaaac ctcgtctcta ctaaaaatac acaacttagc cagtcttgtt ggcgcacgcc 1800

tgtaatctca gctactaggg atgctgaggc aggagaatcg cttgaaccca agaaagaggt 1860

ggaggttgcg gtgagccaag attgcgccac tgcactccag cctgggcaaa cagagggaga 1920

ctccatcgcc ccccccaaca aaaaaaaaag tttcccatac actggcctgc cccaaaaccc 1980

actaacaatt ttagcaaaac agtccaggcc aaagaggaag catttcatgt tcaataagaa 2040

acccagccat tccgcatggc tggttcctga gtggctctgg tgatactctc cagccacctg 2100

ctgacattca gaatctcaga cctcgggact gctgttgcgg taccgtgtgt ctgacacctg 2160

ccagcagccc tttgctatct gcgcgcagga tgggggtgac tgcccagaca ttcccgctag 2220

ataggctctg atttccgggg cagcctttca gatgcggcag acatacaaca cctgtacttt 2280

agagttttaa gggaaaaaaa atcagaagtg ctggttagat agtaaaaact taggataact 2340

tagaaaggct agttttagct tcctttgtgg ctccctggtg caaaacaatt agcagttatg 2400

caatggacct gattctagtt tattctaatt aagaagtgag gccgagtttg acttcgttcc 2460

tgaatacaat cttgagtaac tgggaaagtc tgagtgaaag gatggcctca ttctctttct 2520

aatcttgctg gtttcaagat tagaaaatgg cattatttga tctgaaatgt ttgagaagac 2580

acgaataaag ttacttgggc agaaaaaaaa aaaaaaaaa 2619

<210> 10
<211> 384
<212> PRT
<213> Homo sapiens

<400> 10

Met Asp Pro Ala Gly Gly Pro Arg Gly Val Leu Pro Arg Pro Cys Arg
1 5 10 15

Val Leu Val Leu Leu Asn Pro Arg Gly Gly Lys Gly Lys Ala Leu Gln
20 25 30

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

Phe Thr Leu Met Leu Thr Glu Arg Arg Asn His Ala Arg Glu Leu Val
50 55 60

Arg Ser Glu Glu Leu Gly Arg Trp Asp Ala Leu Val Val Met Ser Gly
65 70 75 80

Asp Gly Leu Met His Glu Val Val Asn Gly Leu Met Glu Arg Pro Asp
85 90 95

Trp Glu Thr Ala Ile Gln Lys Pro Leu Cys Ser Leu Pro Ala Gly Ser
100 105 110

Gly Asn Ala Leu Ala Ala Ser Leu Asn His Tyr Ala Gly Tyr Glu Gln
115 120 125

Val Thr Asn Glu Asp Leu Leu Thr Asn Cys Thr Leu Leu Leu Cys Arg
130 135 140

Arg Leu Leu Ser Pro Met Asn Leu Leu Ser Leu His Thr Ala Ser Gly
145 150 155 160

Leu Arg Leu Phe Ser Val Leu Ser Leu Ala Trp Gly Phe Ile Ala Asp
165 170 175

Val Asp Leu Glu Ser Glu Lys Tyr Arg Arg Leu Gly Glu Met Arg Phe
180 185 190

Thr Leu Gly Thr Phe Leu Arg Leu Ala Ala Leu Arg Thr Tyr Arg Gly
195 200 205

```
Arg Leu Ala Tyr Leu Pro Val Gly Arg Val Gly Ser Lys Thr Pro Ala
    210                 215                 220

Ser Pro Val Val Val Gln Gln Gly Pro Val Asp Ala His Leu Val Pro
225                 230                 235                 240

Leu Glu Glu Pro Val Pro Ser His Trp Thr Met Val Pro Asp Glu Asp
                245                 250                 255

Phe Val Leu Ile Leu Ala Leu Leu His Ser His Leu Gly Ser Glu Met
            260                 265                 270

Phe Ala Ala Pro Met Gly Arg Cys Ala Ala Gly Val Met His Leu Phe
        275                 280                 285

Tyr Val Arg Ala Gly Val Ser Arg Ala Met Leu Leu Arg Phe Phe Leu
    290                 295                 300

Ala Met Glu Lys Gly Arg His Met Glu Tyr Glu Cys Pro Tyr Leu Val
305                 310                 315                 320

Tyr Val Pro Val Val Ala Phe Arg Leu Glu Pro Lys Asp Gly Lys Gly
                325                 330                 335

Val Phe Ala Val Asp Gly Glu Leu Met Val Ser Glu Ala Val Gln Gly
            340                 345                 350

Gln Val His Pro Asn Tyr Phe Trp Met Val Ser Gly Cys Val Glu Pro
        355                 360                 365

Pro Pro Ser Trp Lys Pro Gln Gln Met Pro Pro Pro Glu Glu Pro Leu
    370                 375                 380
```

<210> 11
<211> 618
<212> PRT
<213> Homo sapiens

<400> 11

```
Met Ala Pro Pro Pro Pro Pro Leu Ala Ala Ser Thr Pro Leu Leu His
1               5                   10                  15

Gly Glu Phe Gly Ser Tyr Pro Ala Arg Gly Pro Arg Phe Ala Leu Thr
            20                  25                  30

Leu Thr Ser Gln Ala Leu His Ile Gln Arg Leu Arg Pro Lys Pro Glu
```

```
        35                  40                  45

Ala Arg Pro Arg Gly Gly Leu Val Pro Leu Ala Glu Val Ser Gly Cys
    50                  55                  60

Cys Thr Leu Arg Ser Arg Ser Pro Ser Asp Ser Ala Ala Tyr Phe Cys
65                  70                  75                  80

Ile Tyr Thr Tyr Pro Arg Gly Arg Arg Gly Ala Arg Arg Arg Ala Thr
                85                  90                  95

Arg Thr Phe Arg Ala Asp Gly Ala Ala Thr Tyr Glu Glu Asn Arg Ala
            100                 105                 110

Glu Ala Gln Arg Trp Ala Thr Ala Leu Thr Cys Leu Leu Arg Gly Leu
        115                 120                 125

Pro Leu Pro Gly Asp Gly Glu Ile Thr Pro Asp Leu Leu Pro Arg Pro
    130                 135                 140

Pro Arg Leu Leu Leu Leu Val Asn Pro Phe Gly Gly Arg Gly Leu Ala
145                 150                 155                 160

Trp Gln Trp Cys Lys Asn His Val Leu Pro Met Ile Ser Glu Ala Gly
                165                 170                 175

Leu Ser Phe Asn Leu Ile Gln Thr Glu Arg Gln Asn His Ala Arg Glu
            180                 185                 190

Leu Val Gln Gly Leu Ser Leu Ser Glu Trp Asp Gly Ile Val Thr Val
        195                 200                 205

Ser Gly Asp Gly Leu Leu His Glu Val Leu Asn Gly Leu Leu Asp Arg
    210                 215                 220

Pro Asp Trp Glu Glu Ala Val Lys Met Pro Val Gly Ile Leu Pro Cys
225                 230                 235                 240

Gly Ser Gly Asn Ala Leu Ala Gly Ala Val Asn Gln His Gly Gly Phe
                245                 250                 255

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

Cys Arg Gly Gly Gly His Pro Leu Asp Leu Leu Ser Val Thr Leu Ala
        275                 280                 285
```

```
Ser Gly Ser Arg Cys Phe Ser Phe Leu Ser Val Ala Trp Gly Phe Val
    290                 295                 300

Ser Asp Val Asp Ile Gln Ser Glu Arg Phe Arg Ala Leu Gly Ser Ala
305                 310                 315                 320

Arg Phe Thr Leu Gly Thr Val Leu Gly Leu Ala Thr Leu His Thr Tyr
                325                 330                 335

Arg Gly Arg Leu Ser Tyr Leu Pro Ala Thr Val Glu Pro Ala Ser Pro
            340                 345                 350

Thr Pro Ala His Ser Leu Pro Arg Ala Lys Ser Glu Leu Thr Leu Thr
        355                 360                 365

Pro Asp Pro Ala Pro Pro Met Ala His Ser Pro Leu His Arg Ser Val
    370                 375                 380

Ser Asp Leu Pro Leu Pro Leu Pro Gln Pro Ala Leu Ala Ser Pro Gly
385                 390                 395                 400

Ser Pro Glu Pro Leu Pro Ile Leu Ser Leu Asn Gly Gly Gly Pro Glu
                405                 410                 415

Leu Ala Gly Asp Trp Gly Gly Ala Gly Asp Ala Pro Leu Ser Pro Asp
            420                 425                 430

Pro Leu Leu Ser Ser Pro Pro Gly Ser Pro Lys Ala Ala Leu His Ser
        435                 440                 445

Pro Val Ser Glu Gly Ala Pro Val Ile Pro Pro Ser Ser Gly Leu Pro
    450                 455                 460

Leu Pro Thr Pro Asp Ala Arg Val Gly Ala Ser Thr Cys Gly Pro Pro
465                 470                 475                 480

Asp His Leu Leu Pro Pro Leu Gly Thr Pro Leu Pro Pro Asp Trp Val
                485                 490                 495

Thr Leu Glu Gly Asp Phe Val Leu Met Leu Ala Ile Ser Pro Ser His
            500                 505                 510

Leu Gly Ala Asp Leu Val Ala Ala Pro His Ala Arg Phe Asp Asp Gly
        515                 520                 525

Leu Val His Leu Cys Trp Val Arg Ser Gly Ile Ser Arg Ala Ala Leu
```

530 535 540

Leu Arg Leu Phe Leu Ala Met Glu Arg Gly Ser His Phe Ser Leu Gly
545 550 555 560

Cys Pro Gln Leu Gly Tyr Ala Ala Ala Arg Ala Phe Arg Leu Glu Pro
565 570 575

Leu Thr Pro Arg Gly Val Leu Thr Val Asp Gly Glu Gln Val Glu Tyr
580 585 590

Gly Pro Leu Gln Ala Gln Met His Pro Gly Ile Gly Thr Leu Leu Thr
595 600 605

Gly Pro Pro Gly Cys Pro Gly Arg Glu Pro
610 615

<210> 12
<211> 481
<212> PRT
<213> Homo sapiens

<400> 12

Ile Ile Ala Val Glu Glu Thr Asp Val His Gly Lys His Gln Gly Ser
1 5 10 15

Gly Lys Trp Gln Lys Met Glu Lys Pro Tyr Ala Phe Thr Val His Cys
20 25 30

Val Lys Arg Ala Arg Arg His Arg Trp Lys Trp Ala Gln Val Thr Phe
35 40 45

Trp Cys Pro Glu Glu Gln Leu Cys His Leu Trp Leu Gln Thr Leu Arg
50 55 60

Glu Met Leu Glu Lys Leu Thr Ser Arg Pro Lys His Leu Leu Val Phe
65 70 75 80

Ile Asn Pro Phe Gly Gly Lys Gly Gln Gly Lys Arg Ile Tyr Glu Arg
85 90 95

Lys Val Ala Pro Leu Phe Thr Leu Ala Ser Ile Thr Thr Asp Ile Ile
100 105 110

Val Thr Glu His Ala Asn Gln Ala Lys Glu Thr Leu Tyr Glu Ile Asn
115 120 125

```
Ile Asp Lys Tyr Asp Gly Ile Val Cys Val Gly Gly Asp Gly Met Phe
    130                 135                 140

Ser Glu Val Leu His Gly Leu Ile Gly Arg Thr Gln Arg Ser Ala Gly
145                 150                 155                 160

Val Asp Gln Asn His Pro Arg Ala Val Leu Val Pro Ser Ser Leu Arg
                165                 170                 175

Ile Gly Ile Ile Pro Ala Gly Ser Thr Asp Cys Val Cys Tyr Ser Thr
            180                 185                 190

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

Asp Ser Leu Ala Met Asp Val Ser Ser Val His His Asn Ser Thr Leu
    210                 215                 220

Leu Arg Tyr Ser Val Ser Leu Leu Gly Tyr Gly Phe Tyr Gly Asp Ile
225                 230                 235                 240

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

Phe Ser Gly Leu Lys Thr Phe Leu Ser His His Cys Tyr Glu Gly Thr
            260                 265                 270

Val Ser Phe Leu Pro Ala Gln His Thr Val Gly Ser Pro Arg Asp Arg
        275                 280                 285

Lys Pro Cys Arg Ala Gly Cys Phe Val Cys Arg Gln Ser Lys Gln Gln
    290                 295                 300

Leu Glu Glu Glu Gln Lys Lys Ala Leu Tyr Gly Leu Glu Ala Ala Glu
305                 310                 315                 320

Asp Val Glu Glu Trp Gln Val Val Cys Gly Lys Phe Leu Ala Ile Asn
                325                 330                 335

Ala Thr Asn Met Ser Cys Ala Cys Arg Arg Ser Pro Arg Gly Leu Ser
            340                 345                 350

Pro Ala Ala His Leu Gly Asp Gly Ser Ser Asp Leu Ile Leu Ile Arg
        355                 360                 365

Lys Cys Ser Arg Phe Asn Phe Leu Arg Phe Leu Ile Arg His Thr Asn
    370                 375                 380
```

```
Gln Gln Asp Gln Phe Asp Phe Thr Phe Val Glu Val Tyr Arg Val Lys
385                 390                 395                 400

Lys Phe Gln Phe Thr Ser Lys His Met Glu Asp Glu Asp Ser Asp Leu
                405                 410                 415

Lys Glu Gly Gly Lys Lys Arg Phe Gly His Ile Cys Ser Ser His Pro
            420                 425                 430

Ser Cys Cys Cys Thr Val Ser Asn Ser Ser Trp Asn Cys Asp Gly Glu
        435                 440                 445

Val Leu His Ser Pro Ala Ile Glu Val Arg Val His Cys Gln Leu Val
    450                 455                 460

Arg Leu Phe Ala Arg Gly Ile Glu Glu Asn Pro Lys Pro Asp Ser His
465                 470                 475                 480

Ser
```

<210> 13
<211> 537
<212> PRT
<213> Homo sapiens

<400> 13

```
Met Gly Ala Thr Gly Ala Ala Glu Pro Leu Gln Ser Val Leu Trp Val
1               5                   10                  15

Lys Gln Gln Arg Cys Ala Val Ser Leu Glu Pro Ala Arg Ala Leu Leu
            20                  25                  30

Arg Trp Trp Arg Ser Pro Gly Pro Gly Ala Gly Ala Pro Gly Ala Asp
        35                  40                  45

Ala Cys Ser Val Pro Val Ser Glu Ile Ile Ala Val Glu Glu Thr Asp
    50                  55                  60

Val His Gly Lys His Gln Gly Ser Gly Lys Trp Gln Lys Met Glu Lys
65                  70                  75                  80

Pro Tyr Ala Phe Thr Val His Cys Val Lys Arg Ala Arg Arg His Arg
                85                  90                  95

Trp Lys Trp Ala Gln Val Thr Phe Trp Cys Pro Glu Glu Gln Leu Cys
```

100 105 110

His Leu Trp Leu Gln Thr Leu Arg Glu Met Leu Glu Lys Leu Thr Ser
115 120 125

Arg Pro Lys His Leu Leu Val Phe Ile Asn Pro Phe Gly Gly Lys Gly
130 135 140

Gln Gly Lys Arg Ile Tyr Glu Arg Lys Val Ala Pro Leu Phe Thr Leu
145 150 155 160

Ala Ser Ile Thr Thr Asp Ile Ile Val Thr Glu His Ala Asn Gln Ala
165 170 175

Lys Glu Thr Leu Tyr Glu Ile Asn Ile Asp Lys Tyr Asp Gly Ile Val
180 185 190

Cys Val Gly Gly Asp Gly Met Phe Ser Glu Val Leu His Gly Leu Ile
195 200 205

Gly Arg Thr Gln Arg Ser Ala Gly Val Asp Gln Asn His Pro Arg Ala
210 215 220

Val Leu Val Pro Ser Ser Leu Arg Ile Gly Ile Ile Pro Ala Gly Ser
225 230 235 240

Thr Asp Cys Val Cys Tyr Ser Thr Val Gly Thr Ser Asp Ala Glu Thr
245 250 255

Ser Ala Leu His Ile Val Val Gly Asp Ser Leu Ala Met Asp Val Ser
260 265 270

Ser Val His His Asn Ser Thr Leu Leu Arg Tyr Ser Val Ser Leu Leu
275 280 285

Gly Tyr Gly Phe Tyr Gly Asp Ile Ile Lys Asp Ser Glu Lys Lys Arg
290 295 300

Trp Leu Gly Leu Ala Arg Tyr Asp Phe Ser Gly Leu Lys Thr Phe Leu
305 310 315 320

Ser His His Cys Tyr Glu Gly Thr Val Ser Phe Leu Pro Ala Gln His
325 330 335

Thr Val Gly Ser Pro Arg Asp Arg Lys Pro Cys Arg Ala Gly Cys Phe
340 345 350

```
Val Cys Arg Gln Ser Lys Gln Gln Leu Glu Glu Glu Gln Lys Lys Ala
        355                 360                 365

Leu Tyr Gly Leu Glu Ala Ala Glu Asp Val Glu Glu Trp Gln Val Val
    370                 375                 380

Cys Gly Lys Phe Leu Ala Ile Asn Ala Thr Asn Met Ser Cys Ala Cys
385                 390                 395                 400

Arg Arg Ser Pro Arg Gly Leu Ser Pro Ala Ala His Leu Gly Asp Gly
                405                 410                 415

Ser Ser Asp Leu Ile Leu Ile Arg Lys Cys Ser Arg Phe Asn Phe Leu
            420                 425                 430

Arg Phe Leu Ile Arg His Thr Asn Gln Gln Asp Gln Phe Asp Phe Thr
        435                 440                 445

Phe Val Glu Val Tyr Arg Val Lys Lys Phe Gln Phe Thr Ser Lys His
    450                 455                 460

Met Glu Asp Glu Asp Ser Asp Leu Lys Glu Gly Gly Lys Lys Arg Phe
465                 470                 475                 480

Gly His Ile Cys Ser Ser His Pro Ser Cys Cys Cys Thr Val Ser Asn
                485                 490                 495

Ser Ser Trp Asn Cys Asp Gly Glu Val Leu His Ser Pro Ala Ile Glu
            500                 505                 510

Val Arg Val His Cys Gln Leu Val Arg Leu Phe Ala Arg Gly Ile Glu
        515                 520                 525

Glu Asn Pro Lys Pro Asp Ser His Ser
    530                 535
```