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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
12 December 2002 (12.12.2002)

PCT

(10) International Publication Number
WO 02/099138 A2

(51) International Patent Classification[7]: C12Q 1/68

(21) International Application Number: PCT/US02/17409

(22) International Filing Date: 3 June 2002 (03.06.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/296,076	5 June 2001 (05.06.2001)	US
60/328,605	10 October 2001 (10.10.2001)	US
60/357,253	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). **BELVIN, Marcia** [US/US]; 921 Santa Fe Avenue, Albany, CA 94706 (US). **FRANCIS-LANG, Helen** [GB/US]; 1782 Pacific Avenue, Apt. 2, San Francisco, CA 94109 (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).

(74) Agents: **BRUNELLE, Jan** et al.; Exelixis, Inc., 170 Harbor Way, P.O. Box 511, 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, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

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

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

WO 02/099138 A2

(54) Title: HS2STs AS MODIFIERS OF THE P53 PATHWAY AND METHODS OF USE

(57) Abstract: Human HS2ST 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 HS2ST are provided.

HS2STs AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE

REFERENCE TO RELATED APPLICATIONS

This application claims priority to U.S. provisional patent applications 60/296,076 filed 6/5/2001, 60/328,605 filed 10/10/2001, and 60/357,253 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).

Heparan sulphate (HS) proteoglycans, are ubiquitous on cell surfaces and in the extracellular matrix, are composed of extended polysaccharide (glycosaminoglycan) chains covalently bound to various core proteins. HS has been associated in a variety of

biological processes, such as assembly of extracellular matrices, control of cellular growth and differentiation, regulation of blood coagulation, viral infection, etc. (Lindahl, U. et al. (1994) Thromb. Res. 75, 1–32) (Lindahl, U. et al. (1998) J. Biol. Chem. 273, 24979–24982) (Salmivirta, M., et al. (1996) FASEB J. 10, 1270–1279)(Rosenberg, R, et al. (1997) J. Clin. Invest. 99, 2062–2070). The functional roles of HS appear to depend on interactions of specific polysaccharide structures with specific proteins. Such structures are generated in HS biosynthesis, through the coordinated action of several enzymes. A precursor polysaccharide composed of alternating D-glucuronic acid (GlcA) and *N*-acetyl-D-glucosamine (GlcNAc; 2-deoxy-2-acetoamido-D-glucose) units is therefore modified through a series of reactions that include, in consecutive order, N-deacetylation and N-sulphation of GlcN residues, C-5 epimerization of GlcA to L-iduronic acid (IdoA), 2-O-sulphation of uronic acid residues, and finally 6-O- and 3-O-sulphation of GlcN residues (Lindahl, U. et al. (1998) *supra*) (Salmivirta, M., et al. (1996) *supra*) (Rosenberg, R, et al. (1997) *supra*).

Heparan sulfate 2-O-sulfotransferase (HS2OST) has been noted as a putative sulfotransferase enzyme that may play a role in heparan sulfate proteoglycan biosynthesis. Uronyl 2-sulfotransferase (UST or DS2ST) is a closely related enzyme that contains sulfates iduronyl and glucuronyl that residues in dermatan/chondroitin sulfate (Kobayashi, M. et al (1997) J. Biol. Chem. 272, 13980-13985).

UST has ubiquitous expression of messages in a number of human tissues and in several human cancer cell lines (Kobayashi, M. et al. (1999) J Biol Chem 274, 10474-80).

The ability to manipulate the genomes of model organisms such as *Drosophila* provides a powerful means to analyze biochemical processes that, due to significant evolutionary conservation, has 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, Mechler BM et al., 1985 EMBO J 4:1551-1557; Gateff E. 1982 Adv. Cancer Res. 37: 33-74; Watson KL., et al., 1994 J Cell Sci. 18: 19-33; Miklos GL, and Rubin GM. 1996 Cell 86:521-529; Wassarman DA, et al., 1995 Curr Opin Gen Dev 5: 44-50; and Booth DR. 1999 Cancer Metastasis Rev. 18: 261-284). 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 *Drosophila*, and identified their human orthologs, hereinafter referred to as HS2ST. The invention provides methods for utilizing these p53 modifier genes and polypeptides to identify candidate therapeutic agents that can be used in the treatment of disorders associated with defective p53 function. Preferred HS2ST-modulating agents specifically bind to HS2ST polypeptides and restore p53 function. Other preferred HS2ST-modulating agents are nucleic acid modulators such as antisense oligomers and RNAi that repress HS2ST gene expression or product activity by, for example, binding to and inhibiting the respective nucleic acid (i.e. DNA or mRNA).

HS2ST-specific modulating agents may be evaluated by any convenient *in vitro* or *in vivo* assay for molecular interaction with an HS2ST polypeptide or nucleic acid. In one embodiment, candidate p53 modulating agents are tested with an assay system comprising a HS2ST polypeptide or nucleic acid. Candidate 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. HS2ST-modulating agents include HS2ST related proteins (e.g. dominant negative mutants, and biotherapeutics); HS2ST-specific antibodies; HS2ST-specific antisense oligomers and other nucleic acid modulators; and chemical agents that specifically bind HS2ST or compete with HS2ST binding target. In one specific embodiment, a small molecule modulator is identified using a transferase 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 p53 pathway in a mammalian cell by contacting the mammalian cell with an agent that specifically binds a HS2ST 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 *Drosophila* in which p53 was overexpressed in the wing (Ollmann M, et al., Cell 2000 101: 91-101). The pipe gene was identified as a modifier of the p53 pathway. Accordingly, vertebrate orthologs of these modifiers, and preferably the human orthologs, HS2ST 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 HS2ST function are provided herein. Modulation of the HS2ST 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. HS2ST-modulating agents that act by inhibiting or enhancing HS2ST expression, directly or indirectly, for example, by affecting an HS2ST function such as enzymatic (e.g., catalytic) or binding activity, can be identified using methods provided herein. HS2ST modulating agents are useful in diagnosis, therapy and pharmaceutical development.

Nucleic acids and polypeptides of the invention

Sequences related to HS2ST nucleic acids and polypeptides that can be used in the invention are disclosed in Genbank (referenced by Genbank identifier (GI) number) as GI#s 6683563 (SEQ ID NO:1), 12545388 (SEQ ID NO:2), and 4803734 (SEQ ID NO:5)

for nucleic acid, and GI#s 6683564 (SEQ ID NO:6), 6912420 (SEQ ID NO:7), 4803735 (SEQ ID NO:8), and 5032219 (SEQ ID NO:9) for polypeptides. Additionally, newly identified nucleic acid sequences of SEQ ID NOs:3 and 4 can also be used in the invention.

HS2STs are sulfotransferase proteins with transferase domains. The term "HS2ST polypeptide" refers to a full-length HS2ST protein or a functionally active fragment or derivative thereof. A "functionally active" HS2ST fragment or derivative exhibits one or more functional activities associated with a full-length, wild-type HS2ST protein, such as antigenic or immunogenic activity, enzymatic activity, ability to bind natural cellular substrates, etc. The functional activity of HS2ST 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 HS2ST, such as a binding domain. Protein domains can be identified using the PFAM program (Bateman A., et al., Nucleic Acids Res, 1999, 27:260-2; http://pfam.wustl.edu). Methods for obtaining HS2ST 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:6, 7, 8, or 9 (an HS2ST). In further preferred embodiments, the fragment comprises the entire transferase (functionally active) domain.

The term "HS2ST nucleic acid" refers to a DNA or RNA molecule that encodes a HS2ST polypeptide. Preferably, the HS2ST 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 HS2ST. 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 *Drosophila*, 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; http://blast.wustl.edu/blast/README.html) 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 http://www.ebi.ac.uk/MPsrch/; 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 any of SEQ ID NOs:1, 2, 3, 4, or 5. 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, or 5 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 HS2ST Nucleic Acids and Polypeptides

HS2ST nucleic acids and polypeptides, useful for identifying and testing agents that modulate HS2ST function and for other applications related to the involvement of HS2ST in the p53 pathway. HS2ST 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 HS2ST protein for assays used to assess HS2ST 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 HS2ST 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 HS2ST polypeptide can be inserted into any appropriate expression vector. The necessary transcriptional and translational signals, including promoter/enhancer element, can derive from the native HS2ST 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 HS2ST gene product, the expression vector can comprise a promoter operably linked to an HS2ST 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 HS2ST gene product based on the physical or functional properties of the HS2ST protein in *in vitro* assay systems (*e.g.* immunoassays).

The HS2ST 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 HS2ST 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, cite purification reference). Alternatively, native HS2ST 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 HS2ST 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 HS2ST 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 HS2ST in a p53 pathway process such as apoptosis or cell proliferation. Preferably, the altered HS2ST expression results in a detectable phenotype, such as decreased or increased levels of cell proliferation, angiogenesis, or apoptosis compared to control animals having normal HS2ST 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 HS2ST gene that results in a decrease of HS2ST function, preferably such that HS2ST 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 HS2ST gene is used to construct a homologous recombination vector suitable for altering an endogenous HS2ST 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 HS2ST gene, e.g., by introduction of additional copies of HS2ST, or by operatively inserting a regulatory sequence that provides for altered expression of an endogenous copy of the HS2ST 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 HS2ST function and phenotypic changes are compared with appropriate control animals such as genetically modified animals that receive placebo treatment, and/or animals with unaltered HS2ST expression that receive candidate therapeutic agent.

In addition to the above-described genetically modified animals having altered HS2ST function, animal models having defective p53 function (and otherwise normal HS2ST 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 HS2ST and/or the p53 pathway. 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 HS2ST 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 HS2ST activity by administering a HS2ST-interacting or -modulating agent.

In a preferred embodiment, HS2ST-modulating agents inhibit or enhance HS2ST activity or otherwise affect normal HS2ST function, including transcription, protein expression, protein localization, and cellular or extra-cellular activity. In a further preferred embodiment, the candidate p53 pathway- modulating agent specifically modulates the function of the HS2ST. The phrases "specific modulating agent", "specifically modulates", etc., are used herein to refer to modulating agents that directly bind to the HS2ST polypeptide or nucleic acid, and preferably inhibit, enhance, or otherwise alter, the function of the HS2ST. The term also encompasses modulating agents that alter the interaction of the HS2ST with a binding partner or substrate (e.g. by binding to a binding partner of an HS2ST, or to a protein/binding partner complex, and inhibiting function).

Preferred HS2ST-modulating agents include small molecule compounds; HS2ST-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 HS2ST 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 HS2ST–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 HS2ST-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 HS2ST-modulating agents. In a preferred embodiment, HS2ST-interacting proteins affect normal HS2ST function, including transcription, protein expression, protein localization, and cellular or extra-cellular activity. In another embodiment, HS2ST-interacting proteins are useful in detecting and providing information about the function of HS2ST proteins, as is relevant to p53 related disorders, such as cancer (e.g., for diagnostic means).

An HS2ST-interacting protein may be endogenous, i.e. one that naturally interacts genetically or biochemically with an HS2ST, such as a member of the HS2ST pathway that modulates HS2ST expression, localization, and/or activity. HS2ST-modulators include dominant negative forms of HS2ST-interacting proteins and of HS2ST proteins themselves. Yeast two-hybrid and variant screens offer preferred methods for identifying endogenous HS2ST-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 HS2ST-interacting protein may be an exogenous protein, such as an HS2ST-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). HS2ST antibodies are further discussed below.

In preferred embodiments, an HS2ST-interacting protein specifically binds an HS2ST protein. In alternative preferred embodiments, an HS2ST-modulating agent binds an HS2ST substrate, binding partner, or cofactor.

Antibodies

In another embodiment, the protein modulator is an HS2ST specific antibody agonist or antagonist. The antibodies have therapeutic and diagnostic utilities, and can be used in screening assays to identify HS2ST modulators. The antibodies can also be used in dissecting the portions of the HS2ST pathway responsible for various cellular responses and in the general processing and maturation of the HS2ST.

Antibodies that specifically bind HS2ST polypeptides can be generated using known methods. Preferably the antibody is specific to a mammalian ortholog of HS2ST polypeptide, and more preferably, to human HS2ST. 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 HS2ST which are particularly antigenic can be selected, for example, by routine screening of HS2ST 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:6, ,7 8, or 9. 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 HS2ST or substantially purified fragments thereof. If HS2ST fragments are used, they preferably comprise at least 10, and more preferably, at least 20 contiguous amino acids of an HS2ST protein. In a particular embodiment, HS2ST-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 HS2ST-specific antibodies is assayed by an appropriate assay such as a solid phase enzyme-linked immunosorbant assay (ELISA) using immobilized corresponding HS2ST polypeptides. Other assays, such as radioimmunoassays or fluorescent assays might also be used.

Chimeric antibodies specific to HS2ST 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).

HS2ST-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 HS2ST-modulating agents comprise nucleic acid molecules, such as antisense oligomers or double stranded RNA (dsRNA), which generally inhibit HS2ST activity. Preferred nucleic acid modulators interfere with the function of the HS2ST nucleic acid such as DNA replication, transcription, translocation of the HS2ST RNA to the site of protein translation, translation of protein from the HS2ST RNA, splicing of the HS2ST RNA to yield one or more mRNA species, or catalytic activity which may be engaged in or facilitated by the HS2ST RNA.

In one embodiment, the antisense oligomer is an oligonucleotide that is sufficiently complementary to an HS2ST mRNA to bind to and prevent translation, preferably by binding to the 5′ untranslated region. HS2ST-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 HS2ST 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 HS2ST-specific nucleic acid modulator is used in an assay to further elucidate the role of the HS2ST in the p53 pathway, and/or its relationship to other members of the pathway. In another aspect of the invention, an HS2ST-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 HS2ST 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 HS2ST nucleic acid or protein. In general, secondary assays further assess the activity of a HS2ST modulating agent identified by a primary assay and may confirm that the modulating agent

affects HS2ST in a manner relevant to the p53 pathway. In some cases, HS2ST modulators will be directly tested in a secondary assay.

In a preferred embodiment, the screening method comprises contacting a suitable assay system comprising an HS2ST polypeptide with a candidate agent under conditions whereby, but for the presence of the agent, the system provides a reference activity (e.g. transferase 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 HS2ST activity, and hence the p53 pathway.

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 HS2ST 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 HS2ST-interacting proteins are used in screens to identify small molecule modulators, the binding specificity of the interacting protein to the HS2ST protein may be assayed by various known methods such as substrate processing (e.g. ability of the candidate HS2ST-specific binding agents to function as negative effectors in HS2ST-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 HS2ST 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 HS2ST polypeptide, a fusion protein thereof, or to cells or membranes bearing the polypeptide or fusion protein. The HS2ST polypeptide can be full length or a fragment thereof that retains functional HS2ST activity. The HS2ST polypeptide may be fused to another polypeptide, such as a peptide tag for detection or anchoring, or to another tag. The HS2ST polypeptide is preferably human HS2ST, or is an ortholog or derivative thereof as described above. In a preferred embodiment, the screening assay detects candidate agent-based modulation of HS2ST interaction with a binding target, such as an endogenous or exogenous protein or other substrate that has HS2ST –specific binding activity, and can be used to assess normal HS2ST gene function.

Suitable assay formats that may be adapted to screen for HS2ST 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 HS2ST and p53 pathway modulators (e.g. U.S. Pat. Nos. 5,550,019 and 6,133,437 (apoptosis assays); and

U.S. Pat. No. 6,020,135 (p53 modulation), among others). Specific preferred assays are described in more detail below.

Sulfotransferase assays. Assays for sulfotransferase activity are known in the art. An example of a high-throughput method is a continuous coupled enzyme assay for the spectrophotometric analysis of sulfotransferases using aryl sulfotransferase IV (Burkart MD, and Wong CH. (1999) Anal Biochem. 274:131-7). This assay is based on the regeneration of 3'-phosphoadenosine-5'-phosphosulfate (PAPS) from the desulfated 3'-phosphoadenosine-5'-phosphate (PAP) by a recombinant aryl sulfotransferase using p-nitrophenyl sulfate as the sulfate donor and visible spectrophotometric indicator of enzyme turnover.

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 HS2ST, 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 HS2ST function plays a direct role in apoptosis. For example, an apoptosis assay may be performed on cells that over- or under-express HS2ST relative to wild type cells. Differences in apoptotic response compared to wild type cells suggests that the HS2ST 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 HS2ST 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 HS2ST 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 HS2ST, 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 HS2ST 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 HS2ST relative to wild type cells. Differences in proliferation or cell cycle compared to wild type cells suggests that the HS2ST 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 HS2ST, 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 HS2ST function plays a direct role in cell proliferation. For example, an angiogenesis assay may be performed on cells that over- or under-express HS2ST relative to wild type cells. Differences in angiogenesis compared to wild type cells suggests that the HS2ST 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 HS2ST 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 HS2ST, 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 HS2ST 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 HS2ST relative to wild type cells. Differences in hypoxic response compared to wild type cells suggests that the HS2ST 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).

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 HS2ST 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 HS2ST-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 HS2ST gene expression, preferably mRNA expression. In general, expression analysis comprises comparing HS2ST expression in like populations of cells (*e.g.*, two pools of cells that endogenously or recombinantly express HS2ST) 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 HS2ST 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 HS2ST 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 HS2ST-modulating agent identified by any of the above methods to confirm that the modulating agent affects HS2ST in a manner relevant to the p53 pathway. As used herein, HS2ST-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 HS2ST.

Secondary assays generally compare like populations of cells or animals (*e.g.*, two pools of cells or animals that endogenously or recombinantly express HS2ST) in the presence and absence of the candidate modulator. In general, such assays test whether treatment of cells or animals with a candidate HS2ST–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 HS2ST 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 HS2ST 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 HS2ST. 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 HS2ST 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 HS2ST 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 HS2ST-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 p53 function, comprising the step of administering an agent to the cell that specifically modulates HS2ST activity. Preferably, the modulating agent produces a detectable phenotypic change in the cell indicating that the p53 function is restored, i.e., for example, the cell undergoes normal proliferation or progression through the cell cycle.

The discovery that HS2ST 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 HS2ST 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 HS2ST, are identified as amenable to treatment with an HS2ST modulating agent. In a preferred application, the p53 defective tissue overexpresses an HS2ST 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 HS2ST cDNA sequences as probes, can determine whether particular tumors express or overexpress HS2ST. Alternatively, the TaqMan® is used for quantitative RT-PCR analysis of HS2ST 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 HS2ST oligonucleotides, and antibodies directed against an HS2ST, as described above for: (1) the detection of the presence of HS2ST gene mutations, or the detection of either over- or under-expression of HS2ST mRNA relative to the non-disorder state; (2) the detection of either an over- or an under-abundance of HS2ST gene product relative to the non-disorder state; and (3) the detection of perturbations or abnormalities in the signal transduction pathway mediated by HS2ST.

Thus, in a specific embodiment, the invention is drawn to a method for diagnosing a disease in a patient, the method comprising: a) obtaining a biological sample from the patient; b) contacting the sample with a probe for HS2ST expression; c) comparing results from step (b) with a control; and d) determining whether step (c) indicates a likelihood of disease. 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. Drosophila p53 screen

The Drosophila p53 gene was overexpressed specifically in the wing using the vestigial margin quadrant enhancer. Increasing quantities of Drosophila p53 (titrated using different strength transgenic inserts in 1 or 2 copies) caused deterioration of normal wing morphology from mild to strong, with phenotypes including disruption of pattern and polarity of wing hairs, shortening and thickening of wing veins, progressive crumpling of the wing and appearance of dark "death" inclusions in wing blade. In a screen designed to identify enhancers and suppressors of Drosophila p53, homozygous females carrying two copies of p53 were crossed to 5663 males carrying random insertions of a piggyBac transposon (Fraser M *et al.*, Virology (1985) 145:356-361). Progeny containing insertions were compared to non-insertion-bearing sibling progeny for enhancement or suppression of the p53 phenotypes. Sequence information surrounding the piggyBac insertion site was used to identify the modifier genes. Modifiers of the wing phenotype were identified as members of the p53 pathway. *Drosophila*.pipe was an enhancer of the wing phenotype. Human orthologs of the modifiers, are referred to herein as HS2ST.

BLAST analysis (Altschul et al., *supra*) was employed to identify Targets from Drosophila modifiers. For example, representative sequences from HS2ST (GI# 6912420, SEQ ID NO:7) and HS2ST (GI#5032219, SEQ ID NO:9) share 28% and 30% amino acid identity, respectively, with the *Drosophila*.pipe amino acid.

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; http://pfam.wustl.edu), 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.

II. High-Throughput In Vitro Fluorescence Polarization Assay

Fluorescently-labeled HS2ST 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 HS2ST activity.

III. High-Throughput In Vitro Binding Assay.

^{33}P-labeled HS2ST 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 HS2ST 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 × 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. 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, http://www.appliedbiosystems.com/).

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

	breast		colon		lung		ovary	
HS2ST (SEQ ID NO: 3)	2	12	12	30	5	14	1	7
GI#4803743 (SEQ ID NO: 5)	1	12	2	30	1	14	0	7

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 HS2ST 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 HS2ST 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 HS2ST polypeptide, and the candidate test agent is a small molecule modulator.

5. The method of Claim 4 wherein the assay is a transferase 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 HS2ST 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 HS2ST 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 HS2ST polypeptide comprising an amino acid sequence selected from group consisting of SEQ ID NOs:6, 7, 8, and 9, 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 HS2ST ,

(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 HS2ST 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 HS2ST 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> HS2STs AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE

<130> EX02-091C-PC

<150> US 60/296,076
<151> 2001-06-05

<150> US 60/328,605
<151> 2001-10-10

<150> US 60/357,253
<151> 2002-02-15

<160> 9

<170> PatentIn version 3.1

<210> 1
<211> 2185
<212> DNA
<213> Homo sapiens

<400> 1
gggaaggaag gaagagaggg aggcgggcaa gcaggcgggc gcgggggtcg gagactgagg 60

cagtagaggg aggcgagagc ccggcagccg cttcgcgctg tttgctggcg cgggttttgg 120

agggggcggc cgtttagtcg gctgaggaga agcggacacc agcggcgttg gtgatagcgc 180

ctgggggagg gggactggag aggcgagaag gggggttcgc tgcggtggtt ctctcgctgt 240

cgctctctct ttgcctcgct cccggctcgg cgggctcctc ccggcgtctc tctcgcctcc 300

ggggtcccgc tccccgcccc ccgcggtatg tcttgatccc gagcagcggg tttcatgggg 360

ctcctcagga ttatgatgcc gcccaagttg cagctgctgg cggtggtggc cttcgcggtg 420

gcgatgctct tcttggaaaa ccagatccag aaactggagg agtcccgctc gaagctagaa 480

agggctattg caagacacga agtccgagaa attgagcagc gacatacaat ggatggccct 540

cggcaagatg ccactttaga tgaggaagag gacatggtga tcatttataa cagagttccc 600

aaaacggcaa gcacttcatt taccaatatc gcctatgacc tgtgtgcaaa gaataaatac 660

catgtccttc atatcaacac taccaaaaat aatccagtga tgtcattgca agatcaggtg 720

cgctttgtaa agaatataac ttcctggaaa gagatgaaac caggatttta tcatggacac 780

gtttcttact tggattttgc aaaatttggt gtgaagaaga aaccaattta cattaatgtc 840

ataagggatc ctattgagag gctagtttct tattattact ttctgagatt tggagatgat 900

tatagaccag ggttacggag acgaaaacaa ggagacaaaa agacctttga tgaatgtgta 960

gcagaaggtg gctcagactg tgctccagag aagctctggc ttcaaatccc gttcttctgt 1020

ggccatagct ccgaatgctg gaatgtggga agcaggtggg ctatggatca agccaagtat 1080

aacctaatta atgaatattt tctggtggga gttactgaag aacttgaaga ttttatcatg 1140

ttattggagg cagcattgcc ccggtttttc aggggtgcta ctgaactcta tcgcacagga 1200

aagaaatctc atcttaggaa aaccacagag aagaaactcc ccactaaaca aaccattgca 1260

aaactacagc aatctgatat ttggaaaatg gagaatgagt tctatgaatt tgcactagag 1320

cagttccaat tcatcagagc ccatgccgtt cgagaaaaag atggagacct ctacatcctc 1380

gcacaaaact ttttctatga aaagatttac cctaagtcga actgagtata aggtgtgact 1440

attagattct tgaactaaaa tttgaccctg tcttcacctt tgttctcagc tccacagtct 1500

ggattgctga cagtaggtgt atatgacaat ttgtattgag ccaaattagg aaacagacag 1560

taacgtcaag gaagtagata ctggctggca ttgtcagtgt tctaagtttc aggcattttt 1620

atttttcct ggctaaacgt tggtgaaagt tataacctcc tgcctgggag aaaatataca 1680

tcacctaaaa tgaacttatg gcaggtctaa tcaaaaggct aaatacaatt tcagaaaagg 1740

ttctgatact cttgtttttg ataaagcatt ttttcaacta accatgaatt aagatgagtc 1800

catttgcctc ttctgccttc actgagggtt tgggttatac acctctactg aattgtgtta 1860

ataactgttt ggcagtgtgt actttgtttt tgtgagtcat gtctcatgaa atttattgga 1920

atgtttaatc atatttgcta agaaatgttt ctgctgtagt tggatttgcc catatttatg 1980

taggtggttt taatttttta aatggtgatt agtgttaaaa atcaatttaa atcatgacta 2040

atatggtaaa aagataaagc atcaaagcag tatttctcat tcctgcctcc tcaatatcta 2100

atactgggaa gatacttcaa agaatattga gattgtctga agttttagtt aagattttca 2160

cacattaata tcaaaaaaaa aaaaa 2185

<210> 2
<211> 6708
<212> DNA
<213> Homo sapiens

<400> 2
agggagggaa ggaaggaaga gagggaggcg ggcaagcagg cgggcgcggg ggtcggggac 60

tgaggcagta gagggaggcg agagcccggc agccgcttcg cgctgtttgc tgcgcgggct 120

tttggagggg gcggccgttt agtcggctga ggagaagcgg acaccagcgg cgttggtgat 180

agcgcctggg ggaggggggac tggagaggcg agaaggggggg tcgctgcggt ggttctctcg 240

ctgtcgctct ctctttgcct cgctcccggc tcggcgggct cctcccggcg tctctctcgc 300

ctccggggtc ccgctccccg ccccccgcgg tatgtcttga tcccgagcag cgggtttcat 360

ggggctcctc aggattatga tgccgcccaa gttgcagctg ctggcggtgg tggccttcgc 420

ggtggcgatg ctcttcttgg aaaaccagat ccagaaactg gaggagtccc gctcgaagct 480

agaaagggct attgcaagac acgaagtccg agaaattgag cagcgacata caatggatgg 540

ccctcggcaa gatgccactt tagatgagga agaggacatg gtgatcattt ataacagagt 600

tcccaaaacg gcaagcactt catttaccaa tatcgcctat gacctgtgtg caaagaataa 660

ataccatgtc cttcatatca acactaccaa aaataatcca gtgatgtcat tgcaagatca 720

ggtgcgcttt gtaaagaata taacttcctg gaaagagatg aaaccaggat tttatcatgg 780

acacgtttct tacttggatt ttgcaaaatt tggtgtgaag aagaaaccaa tttacattaa 840

tgtcataagg gatcctattg agaggctagt ttcttattat tactttctga gatttggaga 900

tgattataga ccagggttac ggagacgaaa acaaggagac aaaaagacct ttgatgaatg 960

tgtagcagaa ggtggctcag actgtgctcc agagaagctc tggcttcaaa tcccgttctt 1020

ctgtggccat agctccgaat gctggaatgt gggaagcagg tgggctatgg atcaagccaa 1080

gtataaccta attaatgaat attttctggt gggagttact gaagaacttg aagattttat 1140

catgttattg gaggcagcat tgccccggtt tttcaggggt gctactgaac tctatcgcac 1200

aggaaagaaa tctcatctta ggaaaaccac agagaagaaa ctccccacta aacaaaccat 1260

tgcaaaacta cagcaatctg atatttggaa aatggagaat gagttctatg aatttgcact 1320

agagcagttc caattcatca gagcccatgc cgttcgagaa aaagatggag acctctacat 1380

cctcgcacaa aacttttttct atgaaaagat ttaccctaag tcgaactgag tataaggtgt 1440

gactattaga ttcttgaact aaaatttgac cctgtcttca cctttgttct cagctccaca 1500

gtctggattg ctgacagtag gtgtatatga caatttgtat tgagccaaat taggaaacag 1560

acagtaacgt caaggaagta gatactggct ggcattgtca gtgttctaag tttcaggcat 1620

ttttatttttt cctggctaaa cgttggtgaa agttataacc tcctgcctgg gagaaaatat 1680

acatcaccta aaatgaactt atggcaggtc taatcaaaag gctaaataca atttcagaaa 1740

aggttctgat actcttgttt ttgataaagc atttttttcaa ctaaccatga attaagatga 1800

gtccatttgc ctcttctgcc ttcactgagg gtttgggtta tacacctcta ctgaattgtg 1860

ttaataactg tttggcagtg tgtactttgt ttttgtgagt catgtctcat gaaatttatt 1920

ggaatgttta atcatatttg ctaagaaatg tttctgctgt agttggattt gcccatattt 1980

atgtaggtgg ttttaatttt ttaaatggtg attagtgtta aaaatcaatt taaatcatga 2040

ctaatatggt aaaaagataa agcatcaaag cagtatttct cattcctgcc tcctcaatat 2100

ctaatactgg gaagatactt caaagaatat tgagattgtc tgaagtttta gttaagattt 2160

tcacacatta atatcaaaaa agtaagttta gtatttgttt ctccatgggt tatttgtaaa 2220

gctgtaaact gagatatcgg tgactccgta ttatgactcc attagtgagc tgtggtatgg 2280

gtaggatttt cctacttctt ctgtacttttt acctgtagac tatttttact aaggtgcttt 2340

ataatgtgtt ttaaagcatt gcatttacaa aacaaggaaa atgctgtaaa tattgcatat 2400

tttatgtatt tggaccaaaa ggttacaagt aattagacaa aagtggtttt gcaccaattt 2460

tatgtcaagt aaaaccatca gacctactgt tcttgtattt ctcatttaac tttactgtta 2520

agacatcact gaaatgaact tcagtaagct ttcaattttg atacacagtt cattattcat 2580

aacttgaggc agtaattaca gtggaatgag tactggacaa ggagtcaaaa aacttgattt 2640

caggtcctag ctctagcact tacagctgtg tgatcttggg caagtcactt aacctctctt 2700

tgcctcaatt tcctcatctt gaaatgagga taataatacc tgctgtacct acctcacagg 2760

gctgttgtga ggattaaatg agatggcatg tgaaagcact ttgaaaattg taaagcgcta 2820

tgtaaatgta aggtattata gaaacatctt taacatatag tttcatacca ttcatttttt 2880

aacaaagaaa gggaaaagtc tgcttgtaag ctggttgaaa aagttaatct tgatataaat 2940

ttgtgtttga taaatatcct ctcagtgttt tatcttccat gtttcaacaa ctattgaaat 3000

atgaaatgcc tgtgaactct taaagcttca tgagcagctg cttgagttca ggaagttcac 3060

tgttagaaat aggctttgtt agctgactag ggtcagggaa acttttctct tcaaatttga 3120

aagctgtttc tgttttcatt ttacattatt attcagaaat ggtagctatt ctatacctat 3180

ggtttaagta aatatttctg aataaggctt caccatactg taagcatttt aggtagattg 3240

ccttaaaggt tatgggaggg catgagggaa cacttcttat gagaaaacat ttataaacaa 3300

aagaaacatt tataaactaa agaaaaacta aaagaatgac agaacaatca tcttagcacc 3360

ctttcctcac aataatataa aaatattaaa agaacatagg caggcttttt ttaaatttgg 3420

ctttttttctt tccttttttc aaattgactt ttataggtat ttcctgaaag tgtatacaaa 3480

ttatttcctc gcccaaaata aagcaccact tcaaggtgtg gtttgacatt acatgctaat 3540

gaacaaaccc agtatgcaag ttattcttgc accacatgct caaatcttct tgaggtgcat 3600

taactctttt aggtaactag agcagtactt ggtgaactag atcaggaggt cagtaaactt 3660

tctgtggaag ggccagagag taaatatttt aggctttgca gcccatacgg tctctgtcac 3720

agctagtcaa ccctgccatt ttaccacaaa agcagcaata gacattatgt aaacaaatga 3780

gcacagttat gttccaataa aactttattt acaaaaacag atgacatccc agatgcagac 3840

catgggcaac caaccattgc actggctaaa tcattattta tggagaaatc ctctttgtgt 3900

ctctactcta gatgcctaaa agagtttata tacttctaaa agctcctaac ttatatccaa 3960

agaattgctt tctgattcgt gtagtctctc ccacagattc ataaactttt atgacttata 4020

ttgtttccag gtgggcatgg tttatttccc agtttaacag ttcagaatag gggcatttat 4080

tttatcatat tttagggtgg gttaggagta tcctttctgg agactgagaa aggggtgtat 4140

ttaattccat caggtccagt acagtactag gagtcataat actttataat caattaaata 4200

aatagaacca ctgagacaat aatgtatttt tttaaagtgg caaatgtggt tttcttttt 4260

cagcctttgc gctttttcag tattttgacc atagggagat aatttttta taatacaaaa 4320

gtaaccactt ggaattttaa agataatgtt atgtgtgtat gtgaaatata tatacatata 4380

tatatatatt tcctaaaaga agaaaagata cctttctgtt caacttgtat caactcctct 4440

tttctaattg ctgtgaaatg gcaactgttg ataaattatt gtgattgttt taaaatctaa 4500

tgggaagtaa aatatatttt gattttaccc agcttaatct gtaaagtagc acttaaatat 4560

atctgatagc aacacttaag atattgcatg gggattactt tcctatcatc catatgcatt 4620

tgtgcaactt caaacatatt gggtgcttct gaattcctga tgattggatt taagctattg 4680

aaaattggat aatttaaact taatgatttt tataattttc tgatcttaaa atttggttaa 4740

tgcctataat ctgttgcttt ttctcaatat gtgtcctatt ggaaattcct caaatcgttg 4800

gtgccatcag tgatttacaa acaatatttt gatattgcag atgacttgct tactgtattt 4860

gcattgttag aaaacagttt gtagacaatg attctttttt aataaaatca aataattcta 4920

aaagtgctag agaatttaac taaaagctgg ttcccaaatg catagctggc attttaattt 4980

aaattcaaat ctacatagag aacatccgtg taaatcatct aactggattt tcccattggt 5040

cattcccaaa cacacctatg gtcctagaat ccttaagaga agcaccctgt aacctttttat 5100

gtggtttgcc tttaagaggc ccaggtgctt ctcctttatg atttgagttg gcctcttcat 5160

aaattagtgc tgtttacttt cagaggaagc agagaagttg ctgttatgtt tttgcatccg 5220

tttaccctat gcaaagttgc tgtatgatgc caactaaact gctctttagg cagccttctg 5280

aggagaaaag caaccctgtt tcaaatccac tgccaattca gctcctctgg agtggagctt 5340

tctgatttct tggagcagga attttagaga ttgaaatgaa tgatcattta gtcagattta 5400

tcctgtaatt tcatgcagct ttgtggcctt tgcagtacta tttataaaat ggaccctgat 5460

ggtgatgaac tctttagaac gcattactgt taagcctgtg ttgagacatt gatgctgtct 5520

atctcatttt ttagacagtt tttgtagctt tctattgaga gtcaggtatg tgagcatctc 5580

tgaagcagtg ttgaatgtaa ttttcggaaa catggattgt gtattttgac ttttatttta 5640

taaatacaca gctcaacagt gccttttttt tccctcatag tcctgttgga agatgctcac 5700

tactttctct cttctctctc cctgccctcc cccactccat tcagttgatt catttatgca 5760

aattctgttt ccaacttgaa accattttgt cacatctgtt ggagagataa tcactccttt 5820

tccttaacat tctgccagct ttctgatgtt gaagtgtttc agttgactac ctgatgcaaa 5880

agctataaaa taaacagtgg gaaggggaaa aattggtgtc ctgttttaat attttctttt 5940

gtagccttga cactgatgga cattttccaa gctgactcag tgttcagtgt caacttaact 6000

ctcagatagt gttgccatca agaaagcatg caacatcatt ggtttctaat gattttatgg 6060

cttgtgacaa tattttatct ggactgacat gcctctgctg cttttgcttt gtacttcatt 6120

gctggtaata aaatttcaga tggaaaactt acaaaatata tacttaatta gaagaaaaaa 6180

atagagaaag ggctattaga attaaaaaaa tttgaaagta acttaatcta acatttatgg 6240

cacagtttgg acatatccat aatttttttt gggaacacac atttctgatt ttttttttcc 6300

cccttaaaga agaaagtctc aattccattg attttcaatt cttagccact ggctcattgc 6360

tttgagcaat gcttgattga ttctatttat attatatgat attgggttga taaaatacca 6420

gttcaatgat gagttttctt aacagaattt ggtttgtact tgcagtggct gaacaaagag 6480

catggcttga gaatcaaagg gatctgcatt tagcaatgtg atgtcagtaa atggacataa 6540

caggattgtt gtaaaggttg ggcatgatgt atgcaaagta ctggccaggg tagactaata 6600

actgatggca tttatatgct gtgctggaat attgttacca agctgatgtg ccgttctcac 6660

cctgcagaat actggttttg tcatttcata aatgatattt ttataaat 6708

<210> 3
<211> 2185
<212> DNA
<213> Homo sapiens

<400> 3
gggaaggaag gaagagaggg aggcgggcaa gcaggcgggc gcgggggtcg gagactgagg 60

cagtagaggg aggcgagagc ccggcagccg cttcgcgctg tttgctggcg cgggttttgg 120

agggggcggc cgtttagtcg gctgaggaga agcggacacc agcggcgttg gtgatagcgc 180

ctgggggagg gggactggag aggcgagaag gggggttcgc tgcggtggtt ctctcgctgt 240

cgctctctct ttgcctcgct cccggctcgg cgggctcctc ccggcgtctc tctcgcctcc 300

ggggtcccgc tccccgcccc ccgcggtatg tcttgatccc gagcagcggg tttcatgggg 360

ctcctcagga ttatgatgcc gcccaagttg cagctgctgg cggtggtggc cttcgcggtg 420

gcgatgctct tcttggaaaa ccagatccag aaactggagg agtcccgctc gaagctagaa 480

agggctattg caagacacga agtccgagaa attgagcagc gacatacaat ggatggccct 540

cggcaagatg ccactttaga tgaggaagag gacatggtga tcatttataa cagagttccc 600

aaaacggcaa gcacttcatt taccaatatc gcctatgacc tgtgtgcaaa gaataaatac 660

catgtccttc atatcaacac taccaaaaat aatccagtga tgtcattgca agatcaggtg 720

cgctttgtaa agaatataac ttcctggaaa gagatgaaac caggatttta tcatggacac 780

gtttcttact tggattttgc aaaatttggt gtgaagaaga aaccaattta cattaatgtc 840

ataagggatc ctattgagag gctagtttct tattattact ttctgagatt tggagatgat 900

tatagaccag ggttacggag acgaaaacaa ggagacaaaa agacctttga tgaatgtgta 960

gcagaaggtg gctcagactg tgctccagag aagctctggc ttcaaatccc gttcttctgt 1020

ggccatagct ccgaatgctg gaatgtggga agcaggtggg ctatggatca agccaagtat 1080

aacctaatta atgaatattt tctggtggga gttactgaag aacttgaaga ttttatcatg 1140

ttattggagg cagcattgcc ccggtttttc aggggtgcta ctgaactcta tcgcacagga 1200

aagaaatctc atcttaggaa aaccacagag aagaaactcc ccactaaaca aaccattgca 1260

aaactacagc aatctgatat ttggaaaatg gagaatgagt tctatgaatt tgcactagag 1320

cagttccaat tcatcagagc ccatgccgtt cgagaaaaag atggagacct ctacatcctc 1380

gcacaaaact ttttctatga aaagatttac cctaagtcga actgagtata aggtgtgact 1440

attagattct tgaactaaaa tttgaccctg tcttcacctt tgttctcagc tccacagtct 1500

ggattgctga cagtaggtgt atatgacaat ttgtattgag ccaaattagg aaacagacag 1560

taacgtcaag gaagtagata ctggctggca ttgtcagtgt tctaagtttc aggcattttt 1620

atttttcct ggctaaacgt tggtgaaagt tataacctcc tgcctgggag aaaatataca 1680

tcacctaaaa tgaacttatg gcaggtctaa tcaaaaggct aaatacaatt tcagaaaagg 1740

ttctgatact cttgtttttg ataaagcatt ttttcaacta accatgaatt aagatgagtc 1800

catttgcctc ttctgccttc actgagggtt tgggttatac acctctactg aattgtgtta 1860

ataactgttt ggcagtgtgt actttgtttt tgtgagtcat gtctcatgaa atttattgga 1920

atgtttaatc atatttgcta agaaatgttt ctgctgtagt tggatttgcc catatttatg 1980

taggtggttt taatttttta aatggtgatt agtgttaaaa atcaatttaa atcatgacta 2040

atatggtaaa aagataaagc atcaaagcag tatttctcat tcctgcctcc tcaatatcta 2100

atactgggaa gatacttcaa agaatattga gattgtctga agttttagtt aagattttca 2160

cacattaata tcaaaaaaaa aaaaa 2185

<210> 4
<211> 1147
<212> DNA
<213> Homo sapiens

<400> 4
cgggtttcat ggggctcctc aggattatga tgccgcccaa gttgcagctg ctggcggtgg 60

tggccttcgc ggtggcgatg ctcttcttgg aaaaccagat ccagaaactg gaggagtccc 120

gctcgaagct agaaagggct attgcaagac acgaagtccg agaaattgag cagcgacata 180

caatggatgg ccctcggcaa gatgccactt tagatgagga agaggacatg gtgatcattt 240

ataacagagt tcccaaaacg gcaagcactt catttaccaa tatcgcctat gacctgtgtg 300

caaagaataa ataccatgtc cttcatatca acactaccaa aaataatcca gtgatgtcat 360

tgcaagatca ggtgcgcttt gtaaagaata taacttcctg gaaagagatg aaaccaggat 420

tttatcatgg acacgtttct tacttggatt ttgcaaaatt tggtgtgaag aagaaaccaa 480

tttacattaa tgtcataagg gatcctattg agaggctagt ttcttattat tactttctga 540

gatttggaga tgattataga ccagggttac ggagacgaaa acaaggagac aaaaagacct 600

ttgatgaatg tgtagcagaa ggtggctcag actgtgctcc agagaagctc tggcttcaaa 660

tcccgttctt ctgtggccat agctccgaat gctggaatgt gggaagcagg tgggctatgg 720

atcaagccaa gtataaccta attaatgaat attttctggt gggagttact gaagaacttg 780

aagattttat catgttattg gaggcagcat tgccccggtt tttcaggggt gctactgaac 840

tctatcgcac aggaaagaaa tctcatctta ggaaaaccac agagaagaaa ctccccacta 900

aacaaaccat tgcaaaacta cagcaatctg atatttggaa aatggagaat gagttctatg 960

aatttgcact agagcagttc caattcatca gagcccatgc cgttcgagaa aaagatggag 1020

acctctacat cctcgcacaa aacttttttct atgaaaagat ttaccctaag tcgaactgag 1080

tataaggtgt gactattaga ttcttgaact aaaatttgac cctgtcttca cctttgttct 1140

cagctcc 1147

<210> 5
<211> 4196
<212> DNA
<213> Homo sapiens

<400> 5
cggccctccc atgtgcagcc cggccagccg ggctctcctc ctcgcggcgg atgggtgacc 60

ttttcctggc acgggcaggc tgtgggaggc agcggagcag gcgatgaaga agaagcagca 120

gcatcccggc ggcggcgcgg atccctggcc ccatggggcc cctatggggg gcgcccctcc 180

gggcctgggc agctggaagc gtcgggtgcc cctgctgcct ttcctgcgct tctccctccg 240

ggactacggc ttctgcatgg ccaccctgct ggtcttctgc ctgggctccc tcctctatca 300

gctcagcggg ggacccctc gcttcctgct cgacctgcgg cagtacttgg gaaattccac 360

ttacttggat gaccatggac cacctcctag taaggtacta cctttcccaa gccaggtggt 420

gtacaacagg gtaggcaagt gtgggagccg tactgtggtc ttgcttctga gaatcttgtc 480

ggagaagcac ggatttaatt tggtcacatc agacattcac aacaaaacca ggcttactaa 540

aaatgaacaa atggaactga ttaaaaatat aagtactgcc gaacaaccct atttattcac 600

tcgacatgtt catttcctca acttctcaag gtttggagga gaccagcctg tctacatcaa 660

catcattaga gaccccgtca accggttctt atccaactat tttttccgtc gctttggaga 720

ctggagaggg gaacaaaatc acatgatccg caccccagc atgaggcagg aggagcgcta 780

cctggatatc aatgagtgta ttcttgaaaa ctatcccgag tgctccaacc ccaggttatt 840

ttacatcatt ccgtactttt gtggacagca tcccagatgc agggagcctg gtgaatgggc 900

ccttgagaga gcaaagctga acgtgaatga aaacttcctg ctcgtgggga ttcttgaaga 960

gttggaagat gtgctgctgt tactggaaag atttttacct cattacttca agggcgtgct 1020

cagtatctac aaagacccag agcacaggaa gcttggaaac atgactgtga cggtgaagaa 1080

gactgtcccc tctcctgagg ctgtgcagat cctctaccag cggatgagat acgagtacga 1140

gttttaccac tacgtcaaag agcagttcca cctgctgaag cgcaagtttg gacttaagtc 1200

tcacgtcagc aagccccccc tgaggccaca cttctttatc ccaactccac tggaaaccga 1260

ggagccaatc gacgatgaag aacaggatga tgaaaagtgg ctggaagata tttataagag 1320

gtgatgtgac tgtgttgcct ctatggcttt atctcccttt tccagaaagt tctttgtttg 1380

gggaagtaaa atccttaagg gactaaatta atgcttgggt gcattaaaaa gaacaaaaca 1440

ttcccacatg ttggggtcat tgggagatgc ccggttttgc gggttttatt tgtttaattt 1500

tattctgtgt tttctcttgg ctctttgggt ctttcccggg tacactagat ggctccatcc 1560

caaggcatct tgtcataaaa cagctttccc ccaccccata tcatgggaaa agggggagaa 1620

atatagcccc tagcctaata acttatcatt tgtaaaatga cttataaaaa tattacctca 1680

atggtaggag acatccagac ttgtatattt cagtggaaat acaaaaccac ttcagagacc 1740

agggtatctc ctctggaagg atctaagaga aggtaagaca gattaggaca tcgaaaagga 1800

ggatggagcc aggtgccatg gcttgagcct ataatccgag gctgaggtgg gaggatcact 1860

tgagcccagg agtttgaggt tgcagtgagc tgtgatcaca ccactgcact ccagcctggg 1920

tgacagagtg agactctgtc tcaattaatt ttttttttt aaaggaggag gatctccatg 1980

ggtaagtggt ttctacccgc atgggtagag ttctgcctct ggtccttctc agggggcact 2040

ttcaccaaga gcagtgtaat tatctctgaa agagcaagtc agcttgtgcc gcatccccaa 2100

ccaatccaca gcctggagta cctttcaagg tcaaagtgca tggccagctc cattgagaca 2160

ttccatttca aagcaccgtg ctgacagata tcaaagtact ctagcaggga aaataatttg 2220

tttgctgtgt aaggaagaat gtagacaaga cagataaatc tgaaggtcat gtggcatcag 2280

ggaaagggca tggctgtgtc ttttgcaccc aatatgaaac atcttctccc aacactgctt 2340

taatggaagt tctaggaacc aatttagctc aggcatttga ctcctacagc agaagttctg 2400

agcctgacca cagatggtgt gtaatctatc aaacacaccc ctggccaagt tgggtcctat 2460

aggacctggt actatgtact attgtaactt ctagttccct aagaggtacc tgttttcagt 2520

aaaaaggggt cctgagttct gtgcaggtgg aagagctacc cgagaactac ctgagttctg 2580

tgcaggtaga gtcccatttc ttatgggacc tgtgtgctcc tgagaactct tacttgagac 2640

atcaaaaaga agcagcaaga gcttctggga cagagactgc ttggccagct ttgtaagtaa 2700

gtggctgcct ccaatgtgat gtgagtacat gttgggcagt ctcactgtcc taaggtatgt 2760

cttctttcca cctcccactg cccctcccct gccacctatc aatgatgcct tggttcagtc 2820

attagaaatc tgttgctttg agttctgaaa tattttcacc ttaaaaaaaa tgctgaaaat 2880

acacattctc ctgggaagac gataaacagc tagctaagaa gccgaggttc agtggtggca 2940

gcaggaagga cactgccaca aattttgtct atttcatatt tgtcccctag agccagccct 3000

agcaaatgtg tgagttggga gtagttaata gtaaataaga ctctgacttt acacaagcta 3060

cacattttat acttttcata aaccacaaag tctctctaga attttttctg ccttcactaa 3120

aattggactg tagccaagat ataaagcaag tcatttggaa cctgccgagt gagcactgaa 3180

gctactttat catgagatgt gtgttaagaa ggctgcagcc cacaggagtc cagggaaggc 3240

ggggaccaca gaggcacaga gtccagcact tggccgctca tgggccttct ttctgcctca 3300

gaggacgggg gcagagaagt gatgaaggga aatgttctta gaggaggaaa tatcctttgt 3360

cctgttcaga gagaccaggg ccctaccatt aggcatactt tcagaagcaa cctggagaac 3420

agctatcaat catattcaaa accagtacaa gaactgctgc ctggtaccct gtgagtcatt 3480

tctatgaaat tccatataaa gaatgatgat aagtttacac actgtgcaat ctcacaatct 3540

gaaaataaag ttgagttggc tgtgttttct ctgctcttgt cagaacattg ggacaattgg 3600

tcgttcaaaa acattcatcc tcttactgca agtttatctg ggtactttta cctgtgtgtt 3660

caaaggcatt tcttttcagc agtgatcatt ataacttcac aaaaaaagat gctgacggat 3720

ttacttacag ggccttaatg ttattttgtc ccagccaaca ccctctaggt cctaaaagtc 3780

aaggtacttc agtttatttg gcaaacatga caacattttt ttggccctgg gcccaacagt 3840

ttgtacttca tgaaacatat tgtacatttt acatagttta atttaaaaaa taccttttaa 3900

gctagttgat ctttgactgt cttatttatt ataacctttc agcacattcc aaggttttag 3960

ttactcagga aggagttaat taaaatgatt ttattttggt ctgatggatg ttttttaaaa 4020

ggaaaattat tattatgaac cttcagccta ctttcttgag tgccgtaaaa gtgcttgtaa 4080

atctttttt ttttttaaga agaaagaaaa aaatggtgtt tgacgttgat ggaaattcaa 4140

aaatatatat ggaactgaaa cattaactta gctaaaataa aagcaatctg tgtttg 4196

<210> 6
<211> 356
<212> PRT
<213> Homo sapiens

<400> 6

Met Gly Leu Leu Arg Ile Met Met Pro Pro Lys Leu Gln Leu Leu Ala
1 5 10 15

Val Val Ala Phe Ala Val Ala Met Leu Phe Leu Glu Asn Gln Ile Gln

```
            20                  25                  30

Lys Leu Glu Glu Ser Arg Ser Lys Leu Glu Arg Ala Ile Ala Arg His
        35                  40                  45

Glu Val Arg Glu Ile Glu Gln Arg His Thr Met Asp Gly Pro Arg Gln
    50                  55                  60

Asp Ala Thr Leu Asp Glu Glu Glu Asp Met Val Ile Ile Tyr Asn Arg
65                  70                  75                  80

Val Pro Lys Thr Ala Ser Thr Ser Phe Thr Asn Ile Ala Tyr Asp Leu
                85                  90                  95

Cys Ala Lys Asn Lys Tyr His Val Leu His Ile Asn Thr Thr Lys Asn
            100                 105                 110

Asn Pro Val Met Ser Leu Gln Asp Gln Val Arg Phe Val Lys Asn Ile
        115                 120                 125

Thr Ser Trp Lys Glu Met Lys Pro Gly Phe Tyr His Gly His Val Ser
    130                 135                 140

Tyr Leu Asp Phe Ala Lys Phe Gly Val Lys Lys Lys Pro Ile Tyr Ile
145                 150                 155                 160

Asn Val Ile Arg Asp Pro Ile Glu Arg Leu Val Ser Tyr Tyr Tyr Phe
                165                 170                 175

Leu Arg Phe Gly Asp Asp Tyr Arg Pro Gly Leu Arg Arg Arg Lys Gln
            180                 185                 190

Gly Asp Lys Lys Thr Phe Asp Glu Cys Val Ala Glu Gly Gly Ser Asp
        195                 200                 205

Cys Ala Pro Glu Lys Leu Trp Leu Gln Ile Pro Phe Phe Cys Gly His
    210                 215                 220

Ser Ser Glu Cys Trp Asn Val Gly Ser Arg Trp Ala Met Asp Gln Ala
225                 230                 235                 240

Lys Tyr Asn Leu Ile Asn Glu Tyr Phe Leu Val Gly Val Thr Glu Glu
                245                 250                 255

Leu Glu Asp Phe Ile Met Leu Leu Glu Ala Ala Leu Pro Arg Phe Phe
            260                 265                 270
```

```
Arg Gly Ala Thr Glu Leu Tyr Arg Thr Gly Lys Lys Ser His Leu Arg
        275                 280                 285

Lys Thr Thr Glu Lys Lys Leu Pro Thr Lys Gln Thr Ile Ala Lys Leu
    290                 295                 300

Gln Gln Ser Asp Ile Trp Lys Met Glu Asn Glu Phe Tyr Glu Phe Ala
305                 310                 315                 320

Leu Glu Gln Phe Gln Phe Ile Arg Ala His Ala Val Arg Glu Lys Asp
                325                 330                 335

Gly Asp Leu Tyr Ile Leu Ala Gln Asn Phe Phe Tyr Glu Lys Ile Tyr
            340                 345                 350

Pro Lys Ser Asn
        355
```

<210> 7
<211> 356
<212> PRT
<213> Homo sapiens

<400> 7

```
Met Gly Leu Leu Arg Ile Met Met Pro Pro Lys Leu Gln Leu Leu Ala
1               5                   10                  15

Val Val Ala Phe Ala Val Ala Met Leu Phe Leu Glu Asn Gln Ile Gln
            20                  25                  30

Lys Leu Glu Glu Ser Arg Ser Lys Leu Glu Arg Ala Ile Ala Arg His
        35                  40                  45

Glu Val Arg Glu Ile Glu Gln Arg His Thr Met Asp Gly Pro Arg Gln
    50                  55                  60

Asp Ala Thr Leu Asp Glu Glu Glu Asp Met Val Ile Ile Tyr Asn Arg
65                  70                  75                  80

Val Pro Lys Thr Ala Ser Thr Ser Phe Thr Asn Ile Ala Tyr Asp Leu
                85                  90                  95

Cys Ala Lys Asn Lys Tyr His Val Leu His Ile Asn Thr Thr Lys Asn
            100                 105                 110

Asn Pro Val Met Ser Leu Gln Asp Gln Val Arg Phe Val Lys Asn Ile
        115                 120                 125
```

```
Thr Ser Trp Lys Glu Met Lys Pro Gly Phe Tyr His Gly His Val Ser
  130                 135                 140

Tyr Leu Asp Phe Ala Lys Phe Gly Val Lys Lys Lys Pro Ile Tyr Ile
145                 150                 155                 160

Asn Val Ile Arg Asp Pro Ile Glu Arg Leu Val Ser Tyr Tyr Tyr Phe
                165                 170                 175

Leu Arg Phe Gly Asp Asp Tyr Arg Pro Gly Leu Arg Arg Arg Lys Gln
            180                 185                 190

Gly Asp Lys Lys Thr Phe Asp Glu Cys Val Ala Glu Gly Gly Ser Asp
        195                 200                 205

Cys Ala Pro Glu Lys Leu Trp Leu Gln Ile Pro Phe Phe Cys Gly His
    210                 215                 220

Ser Ser Glu Cys Trp Asn Val Gly Ser Arg Trp Ala Met Asp Gln Ala
225                 230                 235                 240

Lys Tyr Asn Leu Ile Asn Glu Tyr Phe Leu Val Gly Val Thr Glu Glu
                245                 250                 255

Leu Glu Asp Phe Ile Met Leu Leu Glu Ala Ala Leu Pro Arg Phe Phe
            260                 265                 270

Arg Gly Ala Thr Glu Leu Tyr Arg Thr Gly Lys Lys Ser His Leu Arg
        275                 280                 285

Lys Thr Thr Glu Lys Lys Leu Pro Thr Lys Gln Thr Ile Ala Lys Leu
    290                 295                 300

Gln Gln Ser Asp Ile Trp Lys Met Glu Asn Glu Phe Tyr Glu Phe Ala
305                 310                 315                 320

Leu Glu Gln Phe Gln Phe Ile Arg Ala His Ala Val Arg Glu Lys Asp
                325                 330                 335

Gly Asp Leu Tyr Ile Leu Ala Gln Asn Phe Phe Tyr Glu Lys Ile Tyr
            340                 345                 350

Pro Lys Ser Asn
        355
```

<210> 8
<211> 406
<212> PRT
<213> Homo sapiens

<400> 8

Met Lys Lys Lys Gln Gln His Pro Gly Gly Gly Ala Asp Pro Trp Pro
1 5 10 15

His Gly Ala Pro Met Gly Gly Ala Pro Pro Gly Leu Gly Ser Trp Lys
20 25 30

Arg Arg Val Pro Leu Leu Pro Phe Leu Arg Phe Ser Leu Arg Asp Tyr
35 40 45

Gly Phe Cys Met Ala Thr Leu Leu Val Phe Cys Leu Gly Ser Leu Leu
50 55 60

Tyr Gln Leu Ser Gly Gly Pro Pro Arg Phe Leu Leu Asp Leu Arg Gln
65 70 75 80

Tyr Leu Gly Asn Ser Thr Tyr Leu Asp Asp His Gly Pro Pro Pro Ser
85 90 95

Lys Val Leu Pro Phe Pro Ser Gln Val Val Tyr Asn Arg Val Gly Lys
100 105 110

Cys Gly Ser Arg Thr Val Val Leu Leu Leu Arg Ile Leu Ser Glu Lys
115 120 125

His Gly Phe Asn Leu Val Thr Ser Asp Ile His Asn Lys Thr Arg Leu
130 135 140

Thr Lys Asn Glu Gln Met Glu Leu Ile Lys Asn Ile Ser Thr Ala Glu
145 150 155 160

Gln Pro Tyr Leu Phe Thr Arg His Val His Phe Leu Asn Phe Ser Arg
165 170 175

Phe Gly Gly Asp Gln Pro Val Tyr Ile Asn Ile Ile Arg Asp Pro Val
180 185 190

Asn Arg Phe Leu Ser Asn Tyr Phe Phe Arg Arg Phe Gly Asp Trp Arg
195 200 205

Gly Glu Gln Asn His Met Ile Arg Thr Pro Ser Met Arg Gln Glu Glu
210 215 220

```
Arg Tyr Leu Asp Ile Asn Glu Cys Ile Leu Glu Asn Tyr Pro Glu Cys
225                 230                 235                 240

Ser Asn Pro Arg Leu Phe Tyr Ile Ile Pro Tyr Phe Cys Gly Gln His
                245                 250                 255

Pro Arg Cys Arg Glu Pro Gly Glu Trp Ala Leu Glu Arg Ala Lys Leu
            260                 265                 270

Asn Val Asn Glu Asn Phe Leu Leu Val Gly Ile Leu Glu Glu Leu Glu
        275                 280                 285

Asp Val Leu Leu Leu Leu Glu Arg Phe Leu Pro His Tyr Phe Lys Gly
    290                 295                 300

Val Leu Ser Ile Tyr Lys Asp Pro Glu His Arg Lys Leu Gly Asn Met
305                 310                 315                 320

Thr Val Thr Val Lys Lys Thr Val Pro Ser Pro Glu Ala Val Gln Ile
                325                 330                 335

Leu Tyr Gln Arg Met Arg Tyr Glu Tyr Glu Phe Tyr His Tyr Val Lys
            340                 345                 350

Glu Gln Phe His Leu Leu Lys Arg Lys Phe Gly Leu Lys Ser His Val
        355                 360                 365

Ser Lys Pro Pro Leu Arg Pro His Phe Phe Ile Pro Thr Pro Leu Glu
    370                 375                 380

Thr Glu Glu Pro Ile Asp Asp Glu Glu Gln Asp Asp Glu Lys Trp Leu
385                 390                 395                 400

Glu Asp Ile Tyr Lys Arg
                405
```

<210> 9
<211> 406
<212> PRT
<213> Homo sapiens

<400> 9

```
Met Lys Lys Lys Gln Gln His Pro Gly Gly Gly Ala Asp Pro Trp Pro
1               5                   10                  15

His Gly Ala Pro Met Gly Gly Ala Pro Pro Gly Leu Gly Ser Trp Lys
            20                  25                  30
```

```
Arg Arg Val Pro Leu Leu Pro Phe Leu Arg Phe Ser Leu Arg Asp Tyr
        35                  40                  45

Gly Phe Cys Met Ala Thr Leu Leu Val Phe Cys Leu Gly Ser Leu Leu
    50                  55                  60

Tyr Gln Leu Ser Gly Gly Pro Pro Arg Phe Leu Leu Asp Leu Arg Gln
65                  70                  75                  80

Tyr Leu Gly Asn Ser Thr Tyr Leu Asp Asp His Gly Pro Pro Pro Ser
                85                  90                  95

Lys Val Leu Pro Phe Pro Ser Gln Val Val Tyr Asn Arg Val Gly Lys
            100                 105                 110

Cys Gly Ser Arg Thr Val Val Leu Leu Leu Arg Ile Leu Ser Glu Lys
        115                 120                 125

His Gly Phe Asn Leu Val Thr Ser Asp Ile His Asn Lys Thr Arg Leu
    130                 135                 140

Thr Lys Asn Glu Gln Met Glu Leu Ile Lys Asn Ile Ser Thr Ala Glu
145                 150                 155                 160

Gln Pro Tyr Leu Phe Thr Arg His Val His Phe Leu Asn Phe Ser Arg
                165                 170                 175

Phe Gly Gly Asp Gln Pro Val Tyr Ile Asn Ile Ile Arg Asp Pro Val
            180                 185                 190

Asn Arg Phe Leu Ser Asn Tyr Phe Phe Arg Arg Phe Gly Asp Trp Arg
        195                 200                 205

Gly Glu Gln Asn His Met Ile Arg Thr Pro Ser Met Arg Gln Glu Glu
    210                 215                 220

Arg Tyr Leu Asp Ile Asn Glu Cys Ile Leu Glu Asn Tyr Pro Glu Cys
225                 230                 235                 240

Ser Asn Pro Arg Leu Phe Tyr Ile Ile Pro Tyr Phe Cys Gly Gln His
                245                 250                 255

Pro Arg Cys Arg Glu Pro Gly Glu Trp Ala Leu Glu Arg Ala Lys Leu
            260                 265                 270
```

```
Asn Val Asn Glu Asn Phe Leu Leu Val Gly Ile Leu Glu Glu Leu Glu
        275                 280                 285

Asp Val Leu Leu Leu Leu Glu Arg Phe Leu Pro His Tyr Phe Lys Gly
    290                 295                 300

Val Leu Ser Ile Tyr Lys Asp Pro Glu His Arg Lys Leu Gly Asn Met
305                 310                 315                 320

Thr Val Thr Val Lys Lys Thr Val Pro Ser Pro Glu Ala Val Gln Ile
                325                 330                 335

Leu Tyr Gln Arg Met Arg Tyr Glu Tyr Glu Phe Tyr His Tyr Val Lys
            340                 345                 350

Glu Gln Phe His Leu Leu Lys Arg Lys Phe Gly Leu Lys Ser His Val
        355                 360                 365

Ser Lys Pro Pro Leu Arg Pro His Phe Phe Ile Pro Thr Pro Leu Glu
    370                 375                 380

Thr Glu Glu Pro Ile Asp Asp Glu Glu Gln Asp Asp Glu Lys Trp Leu
385                 390                 395                 400

Glu Asp Ile Tyr Lys Arg
                405
```