



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<p>(51) International Patent Classification ⁶ : C12N 15/12, C07K 14/47, A61K 38/17, G01N 33/68, C12Q 1/68, C07K 16/18</p>	A2	<p>(11) International Publication Number: WO 99/61614</p> <p>(43) International Publication Date: 2 December 1999 (02.12.99)</p>
<p>(21) International Application Number: PCT/US99/11497</p> <p>(22) International Filing Date: 25 May 1999 (25.05.99)</p> <p>(30) Priority Data: 60/087,104 28 May 1998 (28.05.98) US 09/216,006 17 December 1998 (17.12.98) US</p> <p>(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Applications US 60/087,104 (CIP) Filed on 28 May 1998 (28.05.98) US 09/216,006 (CIP) Filed on 17 December 1998 (17.12.98)</p> <p>(71) Applicant (for all designated States except US): INCYTE PHARMACEUTICALS, INC. [US/US]; 3174 Porter Drive, Palo Alto, CA 94304 (US).</p> <p>(72) Inventors; and (75) Inventors/Applicants (for US only): LAL, Preeti [IN/US]; 2382 Lass Drive, Santa Clara, CA 95054 (US). HILLMAN, Jennifer, L. [US/US]; 230 Monroe Drive #12, Mountain View, CA 94040 (US). GORGONE, Gina [US/US]; 1253 Pinecrest Drive, Boulder Creek, CA 95006 (US). CORLEY,</p>	<p>Neil, C. [US/US]; 1240 Dale Avenue #30, Mountain View, CA 94040 (US). PATTERSON, Chandra [US/US]; 490 Sherwood Way #1, Menlo Park, CA 94025 (US). YUE, Henry [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). TANG, Y., Tom [CN/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). AZIMZAI, Yalda [US/US]; 2045 Rock Springs Drive, Hayward, CA 94547 (US).</p> <p>(74) Agents: BILLINGS, Lucy, J. et al.; Incyte Pharmaceuticals, Inc., 3174 Porter Drive, Palo Alto, CA 94304 (US).</p> <p>(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).</p> <p>Published <i>Without international search report and to be republished upon receipt of that report.</i></p>	
<p>(54) Title: HUMAN SOCS PROTEINS</p>		
<p>(57) Abstract</p> <p>The invention provides human SOCS proteins (HSCOP) and polynucleotides which identify and encode HSCOP. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating, or preventing disorders associated with expression of HSCOP.</p>		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

HUMAN SOCS PROTEINS

TECHNICAL FIELD

5 This invention relates to nucleic acid and amino acid sequences of human SOCS proteins and to the use of these sequences in the diagnosis, treatment, and prevention of cancer, immune and neurological disorders, and infectious diseases.

BACKGROUND OF THE INVENTION

10

Signal transduction is a general process in which cells respond to extracellular signals (hormones, neurotransmitters, growth and differentiation factors, etc.) through a cascade of biochemical reactions beginning with the binding of the signal molecule to a cell membrane receptor and ending with an effect on an intracellular target molecule.

15 Intermediate steps in this process involve the activation of various cytoplasmic proteins by phosphorylation via protein kinases and the translocation of some of these activated proteins to the cell nucleus, where the transcription of specific genes is affected. The signal transduction process regulates all types of cell functions, including cell proliferation, differentiation, and gene transcription.

20 Cytokines are a specific class of extracellular signaling molecules that control growth, differentiation, and various functions of hemopoietic and immune cells. Cytokines include interleukins (ILs), colony-stimulating factors (G-CSF and GM-CSF), erythropoietin (EPO), and various growth factors (EGF, PDGF, TGF, and FGF; Callard, R. and Gearing, A. (1994) The Cytokine Facts Book, pp 2-6, Academic Press, San Diego, CA).

25 Many of the cytokine receptors, including those for the growth factors EGF, PDGF, and FGF exhibit intrinsic protein kinase activity. Binding of the cytokine to its receptor triggers the autophosphorylation of a tyrosine residue on the receptor. It is believed that these phosphorylated residues are recognition sites for the binding of other cytoplasmic signaling proteins which link the initial receptor activation at the cell surface
30 to the activation of a specific intracellular target molecule. These signaling proteins

contain an src homology 2 (SH2) domain that is a recognition and binding site for the phosphotyrosine residue. SH2 domains are found in a variety of signaling molecules and oncogenic proteins, such as phospholipase C- γ , Ras GTP-ase activating protein, and GRB2 (Lowenstein, E.J. et al. (1992) Cell 70:431-442).

5 While much is known about key events in the activation of signaling pathways, less is known about how they are switched off. Several SH2-containing proteins have been identified that are induced in murine lymphoid cells by various cytokines, including IL-2, IL-3, IL-6, Interferon- γ , and EPO (Yoshimura, A. et al. (1995) EMBO Journal 14:2816-2826; Starr, R. et al. (1997) Nature 387: 917-921; and Naka, T. et al. (1997)
10 Nature 387: 924-929). A common property of these proteins is the ability to suppress growth and differentiation in murine cells. The induction of these SH2-containing proteins in cytokine stimulated cells suggests that they may function as negative regulators of cytokine signaling. Transcription of the genes encoding four of these proteins, CIS (cytokine-inducible SH2-containing protein), and SOCS-1, -2, and -3 (suppressor of
15 cytokine signaling), is induced by IL-6 both *in vitro* and *in vivo* (Starr et al., *supra*).

The four proteins share little sequence homology in their N-terminal regions, but all contain a central SH2 domain and a conserved C-terminal region designated the "SOCS box". The function of the SOCS box is unknown. However, a conserved core triplet sequence (K/R) (D/E) (Y/F) within the SOCS box is similar to the tyrosine
20 phosphorylation site recognized by the JAK kinase family. This similarity suggests that the SOCS box may provide a site for interaction with, and inhibition of, JAK kinases. The finding that SOCS-1 interacts with the catalytic region of JAK kinases supports this hypothesis (Endo, T. A. et al. (1997) Nature 387: 921-24). Constitutive expression of SOCS-1 in M1 murine lymphoid cells also inhibits the phosphorylation of certain cell
25 signaling components (gp130 and Stat3) in response to IL-6 (Starr et al., *supra*). CIS binds to tyrosine-phosphorylated residues in the beta-chain of the IL-3 and EPO receptors and provides another possible mechanism for suppressing cell signaling by preventing the binding of other signaling proteins (Yoshimura et al., *supra*).

Recently, sixteen additional proteins have been identified containing the SOCS box
30 domain (Hilton, D.J. et al. (1998) Proc. Natl. Acad. Sci. USA 95:114-119). Like the SH2-containing proteins described above, each of the proteins contains a C-terminal SOCS box

and a distinctive motif N-terminal of the SOCS box. In addition to four new SOCS proteins containing the SH2 domain, three additional classes of SOCS proteins were found containing WD-40 repeats (WSB-1 and -2), SPRY domains (SSB-1 to -3), or ankyrin repeats (ASB-1 to -3). A class of small GTPases (Rar proteins) that contain the SOCS box were also identified. The function of WSB, SSB, and ASB proteins are as yet unknown. However, like SH2 domains, WD-40 repeats, ankyrin repeats, and SPRY domains have been implicated in protein-protein interactions (Hilton et al. supra).

Defects or alterations in the activity of signaling proteins such as CIS may play a role in the development of various proliferative disorders and diseases such as cancer. Loss or rearrangement of the putative human gene encoding CIS is associated with the development of renal cell carcinomas and lung cancer (Yoshimura et al., supra). This association suggests that CIS may function as a tumor suppressor gene.

The discovery of new human SOCS proteins and the polynucleotides encoding them satisfies a need in the art by providing new compositions which are useful in the diagnosis, prevention, and treatment of cancer, immune and neurological disorders, and infectious diseases.

SUMMARY OF THE INVENTION

The invention features substantially purified polypeptides, human SOCS proteins, referred to collectively as "HSCOP" and individually as "HSOCP-1", "HSOCP-2", and "HSOCP-3", HSOCP-4", HSOCP-5", HSOCP-6", HSOCP-7", HSOCP-8", and HSOCP-9". In one aspect, the invention provides a substantially purified polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-9, and fragments thereof.

The invention further provides a substantially purified variant having at least 90% amino acid identity to at least one of the amino acid sequences selected from the group consisting of SEQ ID NO:1-9, and fragments thereof. The invention also provides an isolated and purified polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-9, and fragments thereof. The invention also includes an isolated and purified polynucleotide variant having at least

90% polynucleotide sequence identity to the polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-9, and fragments thereof.

Additionally, the invention provides an isolated and purified polynucleotide which
5 hybridizes under stringent conditions to the polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-9, and fragments thereof. The invention also provides an isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide encoding the polypeptide comprising the amino acid sequence selected from the group consisting of
10 SEQ ID NO:1-9, and fragments thereof.

The invention also provides a method for detecting a polynucleotide in a sample containing nucleic acids, the method comprising the steps of (a) hybridizing the complement of the polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-9, and fragments thereof to
15 at least one of the polynucleotides of the sample, thereby forming a hybridization complex; and (b) detecting the hybridization complex, wherein the presence of the hybridization complex correlates with the presence of a polynucleotide in the sample. In one aspect, the method further comprises amplifying the polynucleotide prior to hybridization.

20 The invention also provides an isolated and purified polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:10-18, and fragments thereof. The invention further provides an isolated and purified polynucleotide variant having at least 90% polynucleotide sequence identity to the polynucleotide sequence selected from the group consisting of SEQ ID NO:10-18, and fragments thereof.
25 The invention also provides an isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:10-18, and fragments thereof.

The invention further provides an expression vector containing at least a fragment of the polynucleotide encoding the polypeptide comprising an amino acid sequence
30 selected from the group consisting of SEQ ID NO:1-9, and fragments thereof. In another aspect, the expression vector is contained within a host cell.

The invention also provides a method for producing a polypeptide, the method comprising the steps of: (a) culturing the host cell containing an expression vector containing at least a fragment of a polynucleotide under conditions suitable for the expression of the polypeptide; and (b) recovering the polypeptide from the host cell
5 culture.

The invention also provides a pharmaceutical composition comprising a substantially purified polypeptide having the amino acid sequence selected from the group consisting of SEQ ID NO:1-9, and fragments thereof, in conjunction with a suitable pharmaceutical carrier.

10 The invention further includes a purified antibody which binds to a polypeptide selected from the group consisting of SEQ ID NO:1-9, and fragments thereof. The invention also provides a purified agonist and a purified antagonist to the polypeptide.

The invention also provides a method for treating or preventing a disorder associated with decreased expression or activity of HSCOP, the method comprising
15 administering to a subject in need of such treatment an effective amount of a pharmaceutical composition comprising a substantially purified polypeptide having the amino acid sequence selected from the group consisting of SEQ ID NO:1-9, and fragments thereof, in conjunction with a suitable pharmaceutical carrier.

The invention also provides a method for treating or preventing a disorder
20 associated with increased expression or activity of HSCOP, the method comprising administering to a subject in need of such treatment an effective amount of an antagonist of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-9, and fragments thereof.

25

BRIEF DESCRIPTION OF THE TABLES

Table 1 shows polypeptide and nucleotide sequence identification numbers (SEQ ID NOs), clone identification numbers (clone IDs), cDNA libraries, and cDNA fragments used to assemble full-length sequences encoding HSCOP.

Table 2 shows features of each polypeptide sequence, including potential motifs,
30 homologous sequences, and methods and algorithms used for identification of HSCOP.

Table 3 shows the tissue-specific expression patterns of each nucleic acid sequence

as determined by northern analysis, diseases, disorders, or conditions associated with these tissues, and the vector into which each cDNA was cloned.

Table 4 describes the tissues used to construct the cDNA libraries from which cDNA clones encoding HSCOP were isolated.

5 Table 5 shows the tools, programs, and algorithms used to analyze HSCOP, along with applicable descriptions, references, and threshold parameters.

DESCRIPTION OF THE INVENTION

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

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

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

DEFINITIONS

30 "HSCOP" refers to the amino acid sequences of substantially purified HSCOP obtained from any species, particularly a mammalian species, including bovine, ovine,

porcine, murine, equine, and preferably the human species, from any source, whether natural, synthetic, semi-synthetic, or recombinant.

The term "agonist" refers to a molecule which, when bound to HSCOP, increases or prolongs the duration of the effect of HSCOP. Agonists may include proteins, nucleic acids, carbohydrates, or any other molecules which bind to and modulate the effect of HSCOP.

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

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

The terms "amino acid" or "amino acid sequence" refer to an oligopeptide, peptide, polypeptide, or protein sequence, or a fragment of any of these, and to naturally occurring or synthetic molecules. In this context, "fragments," "immunogenic fragments," or "antigenic fragments" refer to fragments of HSCOP which are preferably at least 5 to
5 about 15 amino acids in length, most preferably at least 14 amino acids, and which retain some biological activity or immunological activity of HSCOP. Where "amino acid sequence" is recited to refer to an amino acid sequence of a naturally occurring protein molecule, "amino acid sequence" and like terms are not meant to limit the amino acid sequence to the complete native amino acid sequence associated with the recited protein
10 molecule.

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

The term "antagonist" refers to a molecule which, when bound to HSCOP,
15 decreases the amount or the duration of the effect of the biological or immunological activity of HSCOP. Antagonists may include proteins, nucleic acids, carbohydrates, antibodies, or any other molecules which decrease the effect of HSCOP.

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

The term "antigenic determinant" refers to that fragment of a molecule (i.e., an epitope) that makes contact with a particular antibody. When a protein or a fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the
30 production of antibodies which bind specifically to antigenic determinants (given regions or three-dimensional structures on the protein). An antigenic determinant may compete

with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

The term “antisense” refers to any composition containing a nucleic acid sequence which is complementary to the “sense” strand of a specific nucleic acid sequence.

5 Antisense molecules may be produced by any method including synthesis or transcription. Once introduced into a cell, the complementary nucleotides combine with natural sequences produced by the cell to form duplexes and to block either transcription or translation. The designation “negative” can refer to the antisense strand, and the designation “positive” can refer to the sense strand.

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

15 The terms “complementary” or “complementarity” refer to the natural binding of polynucleotides by base pairing. For example, the sequence “5' A-G-T 3'” bonds to the complementary sequence “3' T-C-A 5'.” Complementarity between two single-stranded molecules may be “partial,” such that only some of the nucleic acids bind, or it may be “complete,” such that total complementarity exists between the single stranded molecules.

20 The degree of complementarity between nucleic acid strands has significant effects on the efficiency and strength of the hybridization between the nucleic acid strands. This is of particular importance in amplification reactions, which depend upon binding between nucleic acids strands, and in the design and use of peptide nucleic acid (PNA) molecules.

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

other components (e.g., Denhardt's solution, dry milk, salmon sperm DNA, etc.).

“Consensus sequence” refers to a nucleic acid sequence which has been resequenced to resolve uncalled bases, extended using XL-PCR kit (Perkin-Elmer, Norwalk CT) in the 5' and/or the 3' direction, and resequenced, or which has been
5 assembled from the overlapping sequences of more than one Incyte Clone using a computer program for fragment assembly, such as the GELVIEW Fragment Assembly system (GCG, Madison WI). Some sequences have been both extended and assembled to produce the consensus sequence.

The term “correlates with expression of a polynucleotide” indicates that the
10 detection of the presence of nucleic acids, the same or related to a nucleic acid sequence encoding HSCOP, by northern analysis is indicative of the presence of nucleic acids encoding HSCOP in a sample, and thereby correlates with expression of the transcript from the polynucleotide encoding HSCOP.

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

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

The term “similarity” refers to a degree of complementarity. There may be partial similarity or complete similarity. The word “identity” may substitute for the word
25 “similarity.” A partially complementary sequence that at least partially inhibits an identical sequence from hybridizing to a target nucleic acid is referred to as “substantially similar.” The inhibition of hybridization of the completely complementary sequence to the target sequence may be examined using a hybridization assay (Southern or northern blot, solution hybridization, and the like) under conditions of reduced stringency. A
30 substantially similar sequence or hybridization probe will compete for and inhibit the binding of a completely similar (identical) sequence to the target sequence under

conditions of reduced stringency. This is not to say that conditions of reduced stringency are such that non-specific binding is permitted, as reduced stringency conditions require that the binding of two sequences to one another be a specific (i.e., a selective) interaction. The absence of non-specific binding may be tested by the use of a second target sequence
5 which lacks even a partial degree of complementarity (e.g., less than about 30% similarity or identity). In the absence of non-specific binding, the substantially similar sequence or probe will not hybridize to the second non-complementary target sequence.

The phrases "percent identity" and "% identity" refer to the percentage of sequence similarity found in a comparison of two or more amino acid or nucleic acid sequences.
10 Percent identity can be determined electronically, e.g., by using the MEGALIGN program (DNASTAR, Madison WI) which creates alignments between two or more sequences according to methods selected by the user, e.g., the clustal method. (See, e.g., Higgins, D.G. and P.M. Sharp (1988) *Gene* 73:237-244.) The clustal algorithm groups sequences into clusters by examining the distances between all pairs. The clusters are aligned
15 pairwise and then in groups. The percentage similarity between two amino acid sequences, e.g., sequence A and sequence B, is calculated by dividing the length of sequence A, minus the number of gap residues in sequence A, minus the number of gap residues in sequence B, into the sum of the residue matches between sequence A and sequence B, times one hundred. Gaps of low or of no similarity between the two amino
20 acid sequences are not included in determining percentage similarity. Percent identity between nucleic acid sequences can also be counted or calculated by other methods known in the art, e.g., the Jotun Hein method. (See, e.g., Hein, J. (1990) *Methods Enzymol.* 183:626-645.) Identity between sequences can also be determined by other methods known in the art, e.g., by varying hybridization conditions.

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

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

"Hybridization" refers to any process by which a strand of nucleic acid binds with

a complementary strand through base pairing.

The term “hybridization complex” refers to a complex formed between two nucleic acid sequences by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g., C_{0t} or R_{0t} analysis) or
5 formed between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been fixed).

The words “insertion” or “addition” refer to changes in an amino acid or nucleotide
10 sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively, to the sequence found in the naturally occurring molecule.

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

The term “microarray” refers to an arrangement of distinct polynucleotides on a substrate.

The terms “element” or “array element” in a microarray context, refer to hybridizable polynucleotides arranged on the surface of a substrate.

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

The phrases “nucleic acid” or “nucleic acid sequence” refer to a nucleotide, oligonucleotide, polynucleotide, or any fragment thereof. These phrases also refer to DNA
25 or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material. In this context, “fragments” refers to those nucleic acid sequences which, when translated, would produce polypeptides retaining some functional characteristic, e.g., antigenicity, or structural domain characteristic, e.g., ATP-
30 binding site, of the full-length polypeptide.

The terms “operably associated” or “operably linked” refer to functionally related

nucleic acid sequences. A promoter is operably associated or operably linked with a coding sequence if the promoter controls the translation of the encoded polypeptide. While operably associated or operably linked nucleic acid sequences can be contiguous and in the same reading frame, certain genetic elements, e.g., repressor genes, are not
5 contiguously linked to the sequence encoding the polypeptide but still bind to operator sequences that control expression of the polypeptide.

The term "oligonucleotide" refers to a nucleic acid sequence of at least about 6 nucleotides to 60 nucleotides, preferably about 15 to 30 nucleotides, and most preferably about 20 to 25 nucleotides, which can be used in PCR amplification or in a hybridization
10 assay or microarray. "Oligonucleotide" is substantially equivalent to the terms "amplimer," "primer," "oligomer," and "probe," as these terms are commonly defined in the art.

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

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

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

The term "stringent conditions" refers to conditions which permit hybridization

between polynucleotides and the claimed polynucleotides. Stringent conditions can be defined by salt concentration, the concentration of organic solvent, e.g., formamide, temperature, and other conditions well known in the art. In particular, stringency can be increased by reducing the concentration of salt, increasing the concentration of formamide,
5 or raising the hybridization temperature.

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

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

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

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

A "variant" of HSCOP polypeptides refers to an amino acid sequence that is altered by one or more amino acid residues. The variant may have "conservative" changes, wherein a substituted amino acid has similar structural or chemical properties
30 (e.g., replacement of leucine with isoleucine). More rarely, a variant may have "nonconservative" changes (e.g., replacement of glycine with tryptophan). Analogous

minor variations may also include amino acid deletions or insertions, or both. Guidance in determining which amino acid residues may be substituted, inserted, or deleted without abolishing biological or immunological activity may be found using computer programs well known in the art, for example, LASERGENE software (DNASTAR).

5 The term “variant,” when used in the context of a polynucleotide sequence, may encompass a polynucleotide sequence related to HSCOP. This definition may also include, for example, “allelic” (as defined above), “splice,” “species,” or “polymorphic” variants. A splice variant may have significant identity to a reference molecule, but will generally have a greater or lesser number of polynucleotides due to alternate splicing of
10 exons during mRNA processing. The corresponding polypeptide may possess additional functional domains or an absence of domains. Species variants are polynucleotide sequences that vary from one species to another. The resulting polypeptides generally will have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a
15 given species. Polymorphic variants also may encompass “single nucleotide polymorphisms” (SNPs) in which the polynucleotide sequence varies by one base. The presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

20 **THE INVENTION**

The invention is based on the discovery of new human SOCS proteins (HSCOP), the polynucleotides encoding HSCOP, and the use of these compositions for the diagnosis, treatment, or prevention of cancer, immune and neurological disorders, and infectious diseases.

25 Table 1 lists the Incyte clones used to assemble full length nucleotide sequences encoding HSCOP. Columns 1 and 2 show the sequence identification numbers (SEQ ID NOs) of the polypeptide and nucleotide sequences, respectively. Column 3 shows the clone IDs of the Incyte clones in which nucleic acids encoding each HSCOP were identified, and column 4 shows the cDNA libraries from which these clones were isolated.
30 Column 5 shows Incyte clones and their corresponding cDNA libraries. Clones for which cDNA libraries are not indicated were derived from pooled cDNA libraries. The clones in

column 5 were used to assemble the consensus nucleotide sequence of each HSCOP and are useful as fragments in hybridization technologies.

The columns of Table 2 show various properties of each of the polypeptides of the invention: column 1 references the SEQ ID NO; column 2 shows the number of amino acid residues in each polypeptide; column 3, potential phosphorylation sites; column 4, potential glycosylation sites; column 5, the amino acid residues comprising signature sequences and motifs; column 6, homologous sequences; and column 7, analytical methods used to identify each protein through sequence homology and protein motifs.

The columns of Table 3 show the tissue-specificity and diseases, disorders, or conditions associated with nucleotide sequences encoding HSCOP. The first column of Table 3 lists the nucleotide SEQ ID NOs. Column 2 lists tissue categories which express HSCOP as a fraction of total tissue categories expressing HSCOP. Column 3 lists diseases, disorders, or conditions associated with those tissues expressing HSCOP. Column 4 lists the vectors used to subclone the cDNA library.

The columns of Table 4 show descriptions of the tissues used to construct the cDNA libraries from which cDNA clones encoding HSCOP were isolated. Column 1 references the nucleotide SEQ ID NOs, column 2 shows the cDNA libraries from which these clones were isolated, and column 3 shows the tissue origins and other descriptive information relevant to the cDNA libraries in column 2.

The following fragments of the nucleotide sequences encoding HSCOP are useful, for example, in hybridization or amplification technologies to identify SEQ ID NO:10-18 and to distinguish between SEQ ID NO:10-18 and related polynucleotide sequences. Useful fragments include the fragment of SEQ ID NO:10 from about nucleotide 243 to about nucleotide 317; the fragment of SEQ ID NO:11 from about nucleotide 834 to about nucleotide 911; the fragment of SEQ ID NO:12 from about nucleotide 233 to about nucleotide 298; the fragment of SEQ ID NO:13 from about nucleotide 242 to about nucleotide 307; the fragment of SEQ ID NO:14 from about nucleotide 447 to about nucleotide 524; the fragment of SEQ ID NO:15 from about nucleotide 287 to about nucleotide 364; the fragment of SEQ ID NO:16 from about nucleotide 100 to about nucleotide 138; the fragments of SEQ ID NO:17 from about nucleotide 433 to about nucleotide 477 and from about nucleotide 1189 to about nucleotide 1233; and the fragment

of SEQ ID NO:18 from about nucleotide 973 to about nucleotide 1017. Polypeptides encoded by these fragments are useful, for example, as immunogenic peptides.

The invention also encompasses HSCOP variants. A preferred HSCOP variant is one which has at least about 80%, more preferably at least about 90%, and most preferably
5 at least about 95% amino acid sequence identity to the HSCOP amino acid sequence, and which contains at least one functional or structural characteristic of HSCOP.

The invention also encompasses polynucleotides which encode HSCOP. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:10-18, which encodes
10 HSCOP.

The invention also encompasses a variant of a polynucleotide sequence encoding HSCOP. In particular, such a variant polynucleotide sequence will have at least about 70%, more preferably at least about 85%, and most preferably at least about 95% polynucleotide sequence identity to the polynucleotide sequence encoding HSCOP. A
15 particular aspect of the invention encompasses a variant of a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:10-18 which has at least about 70%, more preferably at least about 85%, and most preferably at least about 95% polynucleotide sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NO:10-18. Any one of the polynucleotide variants described above
20 can encode an amino acid sequence which contains at least one functional or structural characteristic of HSCOP.

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

30 Although nucleotide sequences which encode HSCOP and its variants are preferably capable of hybridizing to the nucleotide sequence of the naturally occurring

HSCOP under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding HSCOP or its derivatives possessing a substantially different codon usage, e.g., inclusion of non-naturally occurring codons. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding HSCOP and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

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

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed polynucleotide sequences, and, in particular, to those shown in SEQ ID NO:10-18 and fragments thereof under various conditions of stringency. (See, e.g., Wahl, G.M. and S.L. Berger (1987) *Methods Enzymol.* 152:399-407; Kimmel, A.R. (1987) *Methods Enzymol.* 152:507-511.) For example, stringent salt concentration will ordinarily be less than about 750 mM NaCl and 75 mM trisodium citrate, preferably less than about 500 mM NaCl and 50 mM trisodium citrate, and most preferably less than about 250 mM NaCl and 25 mM trisodium citrate. Low stringency hybridization can be obtained in the absence of organic solvent, e.g., formamide, while high stringency hybridization can be obtained in the presence of at least about 35% formamide, and most preferably at least about 50% formamide. Stringent temperature conditions will ordinarily include temperatures of at least about 30°C, more preferably of at least about 37°C, and most preferably of at least about 42°C. Varying additional parameters, such as hybridization time, the concentration of detergent, e.g., sodium dodecyl sulfate (SDS), and the inclusion or exclusion of carrier DNA, are well known to those skilled in the art.

Various levels of stringency are accomplished by combining these various conditions as needed. In a preferred embodiment, hybridization will occur at 30°C in 750 mM NaCl, 75 mM trisodium citrate, and 1% SDS. In a more preferred embodiment, hybridization will occur at 37°C in 500 mM NaCl, 50 mM trisodium citrate, 1% SDS, 35% formamide, and
5 100 µg/ml denatured salmon sperm DNA (ssDNA). In a most preferred embodiment, hybridization will occur at 42°C in 250 mM NaCl, 25 mM trisodium citrate, 1% SDS, 50 % formamide, and 200 µg/ml ssDNA. Useful variations on these conditions will be readily apparent to those skilled in the art.

The washing steps which follow hybridization can also vary in stringency. Wash
10 stringency conditions can be defined by salt concentration and by temperature. As above, wash stringency can be increased by decreasing salt concentration or by increasing temperature. For example, stringent salt concentration for the wash steps will preferably be less than about 30 mM NaCl and 3 mM trisodium citrate, and most preferably less than about 15 mM NaCl and 1.5 mM trisodium citrate. Stringent temperature conditions for
15 the wash steps will ordinarily include temperature of at least about 25°C, more preferably of at least about 42°C, and most preferably of at least about 68°C. In a preferred embodiment, wash steps will occur at 25°C in 30 mM NaCl, 3 mM trisodium citrate, and 0.1% SDS. In a more preferred embodiment, wash steps will occur at 42°C in 15 mM NaCl, 1.5 mM trisodium citrate, and 0.1% SDS. In a most preferred embodiment, wash
20 steps will occur at 68°C in 15 mM NaCl, 1.5 mM trisodium citrate, and 0.1% SDS. Additional variations on these conditions will be readily apparent to those skilled in the art.

Methods for DNA sequencing are well known in the art and may be used to practice any of the embodiments of the invention. The methods may employ such
25 enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE (US Biochemical, Cleveland OH), Taq polymerase (Perkin-Elmer), thermostable T7 polymerase (Amersham Pharmacia Biotech, Piscataway NJ), or combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system (Life Technologies, Gaithersburg MD). Preferably, sequence preparation is automated with
30 machines such as the Hamilton MICROLAB 2200 (Hamilton, Reno NV), Peltier Thermal Cycler 200 (PTC200; MJ Research, Watertown MA) and the ABI CATALYST 800

(Perkin-Elmer). Sequencing is then carried out using the ABI 373 or 377 DNA sequencing systems (Perkin-Elmer), the MEGABACE 1000 DNA sequencing system (Molecular Dynamics, Sunnyvale CA), or other systems known in the art. The resulting sequences are analyzed using a variety of algorithms which are well known in the art.

5 (See, e.g., Ausubel, F.M. (1997) Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7; Meyers, R.A. (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856-853.)

The nucleic acid sequences encoding HSCOP may be extended utilizing a partial nucleotide sequence and employing various PCR-based methods known in the art to detect

10 upstream sequences, such as promoters and regulatory elements. For example, one method which may be employed, restriction-site PCR, uses universal and nested primers to amplify unknown sequence from genomic DNA within a cloning vector. (See, e.g., Sarkar, G. (1993) *PCR Methods Applic.* 2:318-322.) Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown sequence from a

15 circularized template. The template is derived from restriction fragments comprising a known genomic locus and surrounding sequences. (See, e.g., Triglia, T. et al. (1988) *Nucleic Acids Res.* 16:8186.) A third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA. (See, e.g., Lagerstrom, M. et al. (1991) *PCR Methods Applic.* 1:111-119.) In this

20 method, multiple restriction enzyme digestions and ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art. (See, e.g., Parker, J.D. et al. (1991) *Nucleic Acids Res.* 19:3055-306). Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries

25 (Clontech, Palo Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO 4.06 Primer Analysis software (National Biosciences, Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or

30 more, and to anneal to the template at temperatures of about 68°C to 72°C.

When screening for full-length cDNAs, it is preferable to use libraries that have

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

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

15 In another embodiment of the invention, polynucleotide sequences or fragments thereof which encode HSCOP may be cloned in recombinant DNA molecules that direct expression of HSCOP, or fragments or functional equivalents thereof, in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be
20 produced and used to express HSCOP.

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

 In another embodiment, sequences encoding HSCOP may be synthesized, in whole
30 or in part, using chemical methods well known in the art. (See, e.g., Caruthers, M.H. et al. (1980) Nucl. Acids Res. Symp. Ser. 215-223, and Horn, T. et al. (1980) Nucl. Acids Res.

Symp. Ser. 225-232.) Alternatively, HSCOP itself or a fragment thereof may be synthesized using chemical methods. For example, peptide synthesis can be performed using various solid-phase techniques. (See, e.g., Roberge, J.Y. et al. (1995) *Science* 269:202-204.) Automated synthesis may be achieved using the ABI 431A Peptide
5 Synthesizer (Perkin-Elmer). Additionally, the amino acid sequence of HSCOP, or any part thereof, may be altered during direct synthesis and/or combined with sequences from other proteins, or any part thereof, to produce a variant polypeptide.

The peptide may be substantially purified by preparative high performance liquid chromatography. (See, e.g., Chiez, R.M. and F.Z. Regnier (1990) *Methods Enzymol.*
10 182:392-421.) The composition of the synthetic peptides may be confirmed by amino acid analysis or by sequencing. (See, e.g., Creighton, T. (1984) Proteins. Structures and Molecular Properties, WH Freeman, New York NY.)

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

Methods which are well known to those skilled in the art may be used to construct

expression vectors containing sequences encoding HSCOP and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. (See, e.g., Sambrook, J. et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview NY, ch. 4, 8, and 16-17; Ausubel, F.M. et al. (1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, ch. 9, 13, and 16.)

A variety of expression vector/host systems may be utilized to contain and express sequences encoding HSCOP. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus); plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems. The invention is not limited by the host cell employed.

In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotide sequences encoding HSCOP. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding HSCOP can be achieved using a multifunctional E. coli vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or pSPORT1 plasmid (Life Technologies). Ligation of sequences encoding HSCOP into the vector's multiple cloning site disrupts the *lacZ* gene, allowing a colorimetric screening procedure for identification of transformed bacteria containing recombinant molecules. In addition, these vectors may be useful for in vitro transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence. (See, e.g., Van Heeke, G. and S.M. Schuster (1989) *J. Biol. Chem.* 264:5503-5509.) When large quantities of HSCOP are needed, e.g. for the production of antibodies, vectors which direct high level expression of HSCOP may be used. For example, vectors containing the strong, inducible T5 or T7 bacteriophage promoter may be used.

Yeast expression systems may be used for production of HSCOP. A number of vectors containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH, may be used in the yeast Saccharomyces cerevisiae or Pichia pastoris.

In addition, such vectors direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign sequences into the host genome for stable propagation. (See, e.g., Ausubel, 1995, *supra*; Grant et al. (1987) *Methods Enzymol.* 153:516-54; and Scorer, C. A. et al. (1994) *Bio/Technology* 12:181-184.)

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

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

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

For long term production of recombinant proteins in mammalian systems, stable
30 expression of HSCOP in cell lines is preferred. For example, sequences encoding HSCOP can be transformed into cell lines using expression vectors which may contain viral origins

of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for about 1 to 2 days in enriched media before being switched to selective media. The purpose of the selectable marker is to confer resistance to a selective agent, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk* or *apr* cells, respectively. (See, e.g., Wigler, M. et al. (1977) Cell 11:223-232; Lowy, I. et al. (1980) Cell 22:817-823.) Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate; *neo* confers resistance to the aminoglycosides, neomycin and G-418; and *als* or *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively. (See, e.g., Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. 77:3567-3570; Colbere-Garapin, F. et al. (1981) J. Mol. Biol. 150:1-14.) Additional selectable genes have been described, e.g., *trpB* and *hisD*, which alter cellular requirements for metabolites. (See, e.g., Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. 85:8047-8051.) Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP; Clontech), β glucuronidase and its substrate β -glucuronide, or luciferase and its substrate luciferin may be used. These markers can be used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system. (See, e.g., Rhodes, C.A. (1995) Methods Mol. Biol. 55:121-131.)

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

tandem gene as well.

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

Immunological methods for detecting and measuring the expression of HSCOP using either specific polyclonal or monoclonal antibodies are known in the art. Examples
10 of such techniques include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on HSCOP is preferred, but a competitive binding assay may be employed. These and other assays are well known in the art. (See, e.g., Hampton, R. et al.
15 (1990) Serological Methods, a Laboratory Manual, APS Press, St Paul MN, Sect. IV; Coligan, J. E. et al. (1997) Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York NY; and Pound, J.D. (1998) Immunochemical Protocols, Humana Press, Totowa NJ).

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

inhibitors, magnetic particles, and the like.

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

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

In another embodiment of the invention, natural, modified, or recombinant nucleic
20 acid sequences encoding HSCOP may be ligated to a heterologous sequence resulting in translation of a fusion protein in any of the aforementioned host systems. For example, a chimeric HSCOP protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of HSCOP activity. Heterologous protein and peptide moieties may also
25 facilitate purification of fusion proteins using commercially available affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST), maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose,
30 phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoaffinity purification of fusion proteins using

commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the HSCOP encoding sequence and the heterologous protein sequence, so that HSCOP may be cleaved away from the heterologous moiety following
5 purification. Methods for fusion protein expression and purification are discussed in Ausubel (1995, supra, ch 10). A variety of commercially available kits may also be used to facilitate expression and purification of fusion proteins.

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

Fragments of HSCOP may be produced not only by recombinant production, but also by direct peptide synthesis using solid-phase techniques. (See, e.g., Creighton, supra,
15 pp. 55-60.) Protein synthesis may be performed by manual techniques or by automation. Automated synthesis may be achieved, for example, using the ABI 431A Peptide Synthesizer (Perkin-Elmer). Various fragments of HSCOP may be synthesized separately and then combined to produce the full length molecule.

THERAPEUTICS

20 Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between regions of HSCOP and human SOCS proteins. In addition, the expression of HSCOP is closely associated with cancer, inflammation and the immune response, and neurological tissues. Therefore, HSCOP appears to play a role in cancer, immune and neurological disorders, and infectious diseases. In the treatment of diseases or disorders
25 associated with increased HSCOP expression or activity, it is desirable to decrease the expression or activity of HSCOP. In the treatment of the above conditions associated with decreased HSCOP expression or activity, it is desirable to increase the expression or activity of HSCOP.

Therefore, in one embodiment, HSCOP or a fragment or derivative thereof may be
30 administered to a subject to treat or prevent a disorder associated with decreased expression or activity of HSCOP. Examples of such disorders include, but are not limited

to, a cancer such as adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; an immune disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; a neurological disorder such as akathisia, Alzheimer's disease, amnesia, amyotrophic lateral sclerosis, bipolar disorder, catatonia, cerebral neoplasms, dementia, depression, diabetic neuropathy, Down's syndrome, tardive dyskinesia, dystonias, epilepsy, Huntington's disease, peripheral neuropathy, multiple sclerosis, neurofibromatosis, Parkinson's disease, paranoid psychoses, postherpetic neuralgia, schizophrenia, and Tourette's disorder; and an infectious disease such a viral infection, e.g., those caused by adenoviruses (acute respiratory disease, pneumonia), arenaviruses (lymphocytic choriomeningitis), bunyaviruses (Hantavirus), coronaviruses (pneumonia, chronic bronchitis), hepadnaviruses (hepatitis), herpesviruses (herpes simplex virus, varicella-zoster virus, Epstein-Barr virus, cytomegalovirus), flaviviruses (yellow fever), orthomyxoviruses (influenza), papillomaviruses (cancer), paramyxoviruses (measles, mumps), picornoviruses (rhinovirus, poliovirus, coxsackie-virus), polyomaviruses (BK virus, JC virus), poxviruses (smallpox), reovirus (Colorado tick fever), retroviruses

(human immunodeficiency virus, human T lymphotropic virus), rhabdoviruses (rabies), rotaviruses (gastroenteritis), and togaviruses (encephalitis, rubella), and bacterial, fungal, parasitic, protozoal, and helminthic infections.

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

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

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

In a further embodiment, an antagonist of HSCOP may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of HSCOP. Examples of such disorders include, but are not limited to, those described above. In one aspect, an antibody which specifically binds HSCOP may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissue which express HSCOP.

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

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

dosages of each agent, thus reducing the potential for adverse side effects.

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

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

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to HSCOP have an amino acid sequence consisting of at least about 5 amino
20 acids, and, more preferably, of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or fragments are identical to a portion of the amino acid sequence of the natural protein and contain the entire amino acid sequence of a small, naturally occurring molecule. Short stretches of HSCOP amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be
25 produced.

Monoclonal antibodies to HSCOP may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique. (See, e.g., Kohler, G. et al.
30 (1975) Nature 256:495-497; Kozbor, D. et al. (1985) J. Immunol. Methods 81:31-42; Cote, R.J. et al. (1983) Proc. Natl. Acad. Sci. 80:2026-2030; and Cole, S.P. et al. (1984)

Mol. Cell Biol. 62:109-120.)

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

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

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

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

Various methods such as Scatchard analysis in conjunction with

radioimmunoassay techniques may be used to assess the affinity of antibodies for HSCOP. Affinity is expressed as an association constant, K_a , which is defined as the molar concentration of HSOCH-antibody complex divided by the molar concentrations of free antigen and free antibody under equilibrium conditions. The K_a determined for a
5 preparation of polyclonal antibodies, which are heterogeneous in their affinities for multiple HSCOP epitopes, represents the average affinity, or avidity, of the antibodies for HSCOP. The K_a determined for a preparation of monoclonal antibodies, which are monospecific for a particular HSCOP epitope, represents a true measure of affinity. High-affinity antibody preparations with K_a ranging from about 10^9 to 10^{12} L/mole are preferred
10 for use in immunoassays in which the HSOCH-antibody complex must withstand rigorous manipulations. Low-affinity antibody preparations with K_a ranging from about 10^6 to 10^7 L/mole are preferred for use in immunopurification and similar procedures which ultimately require dissociation of HSCOP, preferably in active form, from the antibody (Catty, D. (1988) Antibodies. Volume I: A Practical Approach, IRL Press, Washington,
15 DC; Liddell, J. E. and Cryer, A. (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY).

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

25 In another embodiment of the invention, the polynucleotides encoding HSCOP, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, the complement of the polynucleotide encoding HSCOP may be used in situations in which it would be desirable to block the transcription of the mRNA. In particular, cells may be transformed with sequences complementary to polynucleotides encoding HSCOP.
30 Thus, complementary molecules or fragments may be used to modulate HSCOP activity, or to achieve regulation of gene function. Such technology is now well known in the art,

and sense or antisense oligonucleotides or larger fragments can be designed from various locations along the coding or control regions of sequences encoding HSCOP.

Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide
5 sequences to the targeted organ, tissue, or cell population. Methods which are well known to those skilled in the art can be used to construct vectors to express nucleic acid sequences complementary to the polynucleotides encoding HSCOP. (See, e.g., Sambrook, supra; Ausubel, 1995, supra.)

Genes encoding HSCOP can be turned off by transforming a cell or tissue with
10 expression vectors which express high levels of a polynucleotide, or fragment thereof, encoding HSCOP. Such constructs may be used to introduce untranslatable sense or antisense sequences into a cell. Even in the absence of integration into the DNA, such vectors may continue to transcribe RNA molecules until they are disabled by endogenous nucleases. Transient expression may last for a month or more with a non-replicating
15 vector, and may last even longer if appropriate replication elements are part of the vector system.

As mentioned above, modifications of gene expression can be obtained by designing complementary sequences or antisense molecules (DNA, RNA, or PNA) to the control, 5', or regulatory regions of the gene encoding HSCOP. Oligonucleotides derived
20 from the transcription initiation site, e.g., between about positions -10 and +10 from the start site, are preferred. Similarly, inhibition can be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex
25 DNA have been described in the literature. (See, e.g., Gee, J.E. et al. (1994) in Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177.) A complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

30 Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific

hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. For example, engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding HSCOP.

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

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

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

30 Many methods for introducing vectors into cells or tissues are available and equally suitable for use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors may be

introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection, by liposome injections, or by polycationic amino polymers may be achieved using methods which are well known in the art. (See, e.g., Goldman, C.K. et al. (1997) Nature Biotechnology 15:462-466.)

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

An additional embodiment of the invention relates to the administration of a pharmaceutical or sterile composition, in conjunction with a pharmaceutically acceptable carrier, for any of the therapeutic effects discussed above. Such pharmaceutical
10 compositions may consist of HSCOP, antibodies to HSCOP, and mimetics, agonists, antagonists, or inhibitors of HSCOP. The compositions may be administered alone or in combination with at least one other agent, such as a stabilizing compound, which may be administered in any sterile, biocompatible pharmaceutical carrier including, but not limited
15 to, saline, buffered saline, dextrose, and water. The compositions may be administered to a patient alone, or in combination with other agents, drugs, or hormones.

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

In addition to the active ingredients, these pharmaceutical compositions may contain suitable pharmaceutically-acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. Further details on techniques for formulation and administration may
25 be found in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing, Easton PA).

Pharmaceutical compositions for oral administration can be formulated using pharmaceutically acceptable carriers well known in the art in dosages suitable for oral administration. Such carriers enable the pharmaceutical compositions to be formulated as
30 tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions, and the like, for ingestion by the patient.

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

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

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

Pharmaceutical formulations suitable for parenteral administration may be
25 formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks' solution, Ringer's solution, or physiologically buffered saline. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable
30 lipophilic solvents or vehicles include fatty oils, such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate, triglycerides, or liposomes. Non-lipid polycationic amino

polymers may also be used for delivery. Optionally, the suspension may also contain suitable stabilizers or agents to increase the solubility of the compounds and allow for the preparation of highly concentrated solutions.

For topical or nasal administration, penetrants appropriate to the particular barrier
5 to be permeated are used in the formulation. Such penetrants are generally known in the art.

The pharmaceutical compositions of the present invention may be manufactured in a manner that is known in the art, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping, or
10 lyophilizing processes.

The pharmaceutical composition may be provided as a salt and can be formed with many acids, including but not limited to, hydrochloric, sulfuric, acetic, lactic, tartaric, malic, and succinic acid. Salts tend to be more soluble in aqueous or other protonic solvents than are the corresponding free base forms. In other cases, the preferred
15 preparation may be a lyophilized powder which may contain any or all of the following: 1 mM to 50 mM histidine, 0.1% to 2% sucrose, and 2% to 7% mannitol, at a pH range of 4.5 to 5.5, that is combined with buffer prior to use.

After pharmaceutical compositions have been prepared, they can be placed in an appropriate container and labeled for treatment of an indicated condition. For
20 administration of HSCOP, such labeling would include amount, frequency, and method of administration.

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

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

A therapeutically effective dose refers to that amount of active ingredient, for

example HSCOP or fragments thereof, antibodies of HSCOP, and agonists, antagonists or inhibitors of HSCOP, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the ED₅₀ (the dose therapeutically
5 effective in 50% of the population) or LD₅₀ (the dose lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, and it can be expressed as the LD₅₀/ED₅₀ ratio. Pharmaceutical compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal
10 studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the ED₅₀ with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

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

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

DIAGNOSTICS

In another embodiment, antibodies which specifically bind HSCOP may be used
30 for the diagnosis of disorders characterized by expression of HSCOP, or in assays to monitor patients being treated with HSCOP or agonists, antagonists, or inhibitors of

HSCOP. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for HSCOP include methods which utilize the antibody and a label to detect HSCOP in human body fluids or in extracts of cells or tissues. The antibodies may be used with or without modification, and may be
5 labeled by covalent or non-covalent attachment of a reporter molecule. A wide variety of reporter molecules, several of which are described above, are known in the art and may be used.

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

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

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

Probes may also be used for the detection of related sequences, and should preferably have at least 50% sequence identity to any of the HSCOP encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:10-18 or from genomic sequences including
5 promoters, enhancers, and introns of the HSCOP gene.

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

Polynucleotide sequences encoding HSCOP may be used for the diagnosis of
15 disorders associated with expression of HSCOP. Examples of such disorders include, but are not limited to, a cancer such as adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate,
20 salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; an immune disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis,
25 diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis,
30 polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis,

thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; a neurological disorder such as akathisia, Alzheimer's disease, amnesia, amyotrophic lateral sclerosis, bipolar disorder, 5 catatonia, cerebral neoplasms, dementia, depression, diabetic neuropathy, Down's syndrome, tardive dyskinesia, dystonias, epilepsy, Huntington's disease, peripheral neuropathy, multiple sclerosis, neurofibromatosis, Parkinson's disease, paranoid psychoses, postherpetic neuralgia, schizophrenia, and Tourette's disorder; and an infectious disease such a viral infection, e.g., those caused by adenoviruses (acute 10 respiratory disease, pneumonia), arenaviruses (lymphocytic choriomeningitis), bunyaviruses (Hantavirus), coronaviruses (pneumonia, chronic bronchitis), hepadnaviruses (hepatitis), herpesviruses (herpes simplex virus, varicella-zoster virus, Epstein-Barr virus, cytomegalovirus), flaviviruses (yellow fever), orthomyxoviruses (influenza), papillomaviruses (cancer), paramyxoviruses (measles, mumps), picornoviruses 15 (rhinovirus, poliovirus, coxsackie-virus), polyomaviruses (BK virus, JC virus), poxviruses (smallpox), reovirus (Colorado tick fever), retroviruses (human immunodeficiency virus, human T lymphotropic virus), rhabdoviruses (rabies), rotaviruses (gastroenteritis), and togaviruses (encephalitis, rubella), and bacterial, fungal, parasitic, protozoal, and helminthic infections. The polynucleotide sequences encoding HSCOP may be used in 20 Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; in dipstick, pin, and multiformat ELISA-like assays; and in microarrays utilizing fluids or tissues from patients to detect altered HSCOP expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding HSCOP may be useful in 25 assays that detect the presence of associated disorders, particularly those mentioned above. The nucleotide sequences encoding HSCOP may be labeled by standard methods and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantitated and compared with a standard value. If the amount of signal in 30 the patient sample is significantly altered in comparison to a control sample then the presence of altered levels of nucleotide sequences encoding HSCOP in the sample

indicates the presence of the associated disorder. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the treatment of an individual patient.

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

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

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

Additional diagnostic uses for oligonucleotides designed from the sequences encoding HSCOP may involve the use of PCR. These oligomers may be chemically synthesized, generated enzymatically, or produced *in vitro*. Oligomers will preferably contain a fragment of a polynucleotide encoding HSCOP, or a fragment of a polynucleotide complementary to the polynucleotide encoding HSCOP, and will be employed under optimized conditions for identification of a specific gene or condition.

Oligomers may also be employed under less stringent conditions for detection or quantitation of closely related DNA or RNA sequences.

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

10 In further embodiments, oligonucleotides or longer fragments derived from any of the polynucleotide sequences described herein may be used as targets in a microarray. The microarray can be used to monitor the expression level of large numbers of genes simultaneously and to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a
15 disorder, to diagnose a disorder, and to develop and monitor the activities of therapeutic agents.

Microarrays may be prepared, used, and analyzed using methods known in the art. (See, e.g., Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) Proc. Natl. Acad. Sci. 93:10614-10619; Baldeschweiler et al. (1995) PCT application
20 WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) Proc. Natl. Acad. Sci. 94:2150-2155; and Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662.)

In another embodiment of the invention, nucleic acid sequences encoding HSCOP may be used to generate hybridization probes useful in mapping the naturally occurring
25 genomic sequence. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries. (See, e.g., Harrington, J.J. et al. (1997) Nat Genet. 15:345-355; Price, C.M. (1993) Blood
30 Rev. 7:127-134; and Trask, B.J. (1991) Trends Genet. 7:149-154.)

Fluorescent in situ hybridization (FISH) may be correlated with other physical

chromosome mapping techniques and genetic map data. (See, e.g., Heinz-Ulrich, et al. (1995) in Meyers, supra, pp. 965-968.) Examples of genetic map data can be found in various scientific journals or at the Online Mendelian Inheritance in Man (OMIM) site. Correlation between the location of the gene encoding HSCOP on a physical chromosomal
5 map and a specific disorder, or a predisposition to a specific disorder, may help define the region of DNA associated with that disorder. The nucleotide sequences of the invention may be used to detect differences in gene sequences among normal, carrier, and affected individuals.

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

In another embodiment of the invention, HSCOP, its catalytic or immunogenic fragments, or oligopeptides thereof can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening
25 may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes between HSCOP and the agent being tested may be measured.

Another technique for drug screening provides for high throughput screening of compounds having suitable binding affinity to the protein of interest. (See, e.g., Geysen,
30 et al. (1984) PCT application WO84/03564.) In this method, large numbers of different small test compounds are synthesized on a solid substrate. The test compounds are reacted

with HSCOP, or fragments thereof, and washed. Bound HSCOP is then detected by methods well known in the art. Purified HSCOP can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

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

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

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

The entire disclosure of all applications, patents, and publications, cited above and
below, and of US provisional applications 60/087,104 (filed May 28, 1998), and
20 09/216,006 (filed December 17, 1998) are hereby incorporated by reference.

EXAMPLES

I. Construction of cDNA Libraries

RNA was purchased from Clontech or isolated from tissues described in Table 4. Some tissues were homogenized and lysed in guanidinium isothiocyanate, while others
25 were homogenized and lysed in phenol or in a suitable mixture of denaturants, such as TRIZOL (Life Technologies), a monophasic solution of phenol and guanidine isothiocyanate. The resulting lysates were centrifuged over CsCl cushions or extracted with chloroform. RNA was precipitated from the lysates with either isopropanol or sodium acetate and ethanol, or by other routine methods.

30 Phenol extraction and precipitation of RNA were repeated as necessary to increase RNA purity. In some cases, RNA was treated with DNase. For most libraries, poly(A+)

RNA was isolated using oligo d(T)-coupled paramagnetic particles (Promega), OLIGOTEX latex particles (QIAGEN, Chatsworth CA), or an OLIGOTEX mRNA purification kit (QIAGEN). Alternatively, RNA was isolated directly from tissue lysates using other RNA isolation kits, e.g., the POLY(A)PURE mRNA purification kit (Ambion,
5 Austin TX).

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP vector system (Stratagene) or SUPERSRIPT plasmid system (Life Technologies), using the recommended procedures or similar methods known
10 in the art. (See, e.g., Ausubel, 1997, supra, units 5.1-6.6). Reverse transcription was initiated using oligo d(T) or random primers. Synthetic oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-1000 bp) using SEPHACRYL S1000, SEPHAROSE CL2B, or SEPHAROSE CL4B column
15 chromatography (Amersham Pharmacia Biotech) or preparative agarose gel electrophoresis. cDNAs were ligated into compatible restriction enzyme sites of the polylinker of a suitable plasmid, e.g., PBLUESCRIPT plasmid (Stratagene), pSPORT1 plasmid (Life Technologies), or pINCY (Incyte Pharmaceuticals, Palo Alto CA). Recombinant plasmids were transformed into competent E. coli cells including XL1-Blue,
20 XL1-BlueMRF, or SOLR from Stratagene or DH5 α , DH10B, or ElectroMAX DH10B from Life Technologies.

II. Isolation of cDNA Clones

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

30 Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR in a high-throughput format (Rao, V.B. (1994) Anal. Biochem. 216:1-14). Host cell

lysis and thermal cycling steps were carried out in a single reaction mixture. Samples were processed and stored in 384-well plates, and the concentration of amplified plasmid DNA was quantified fluorometrically using PICOGREEN dye (Molecular Probes, Eugene OR) and a Fluoroskan II fluorescence scanner (Labsystems Oy, Helsinki, Finland).

5 **III. Sequencing and Analysis**

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

20 The polynucleotide sequences derived from cDNA sequencing were assembled and analyzed using a combination of software programs which utilize algorithms well known to those skilled in the art. Table 5 summarizes the tools, programs, and algorithms used and provides applicable descriptions, references, and threshold parameters. The first column of Table 5 shows the tools, programs, and algorithms used, the second column
25 provides brief descriptions thereof, the third column presents appropriate references, all of which are incorporated by reference herein in their entirety, and the fourth column presents, where applicable, the scores, probability values, and other parameters used to evaluate the strength of a match between two sequences (the higher the score, the greater the homology between two sequences). Sequences were analyzed using MACDNASIS
30 PRO software (Hitachi Software Engineering, South San Francisco CA) and LASERGENE software (DNASTAR).

The polynucleotide sequences were validated by removing vector, linker, and polyA sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programming, and dinucleotide nearest neighbor analysis. The sequences were then queried against a selection of public databases such as GenBank
5 primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS to acquire annotation, using programs based on BLAST, FASTA, and BLIMPS. The sequences were assembled into full length polynucleotide sequences using programs based on Phred, Phrap, and Consed, and were screened for open reading frames using programs based on GeneMark, BLAST, and FASTA. The full length polynucleotide sequences were
10 translated to derive the corresponding full length amino acid sequences, and these full length sequences were subsequently analyzed by querying against databases such as the GenBank databases (described above), SwissProt, BLOCKS, PRINTS, Prosite, and Hidden Markov Model (HMM)-based protein family databases such as PFAM. HMM is a probabilistic approach which analyzes consensus primary structures of gene families.
15 (See, e.g., Eddy, S.R. (1996) *Curr. Opin. Str. Biol.* 6:361-365.)

The programs described above for the assembly and analysis of full length polynucleotide and amino acid sequences were also used to identify polynucleotide sequence fragments from SEQ ID NO:10-18. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies were
20 described in The Invention section above.

IV. Northern Analysis

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

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

$$\frac{\% \text{ sequence identity} \times \% \text{ maximum BLAST score}}{100}$$

100

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. For example, with a product score of 40, the match
5 will be exact within a 1% to 2% error, and, with a product score of 70, the match will be exact. Similar molecules are usually identified by selecting those which show product scores between 15 and 40, although lower scores may identify related molecules.

The results of northern analyses are reported as a percentage distribution of libraries in which the transcript encoding HSCOP occurred. Analysis involved the
10 categorization of cDNA libraries by organ/tissue and disease, disorder, or condition. The organ/tissue categories included cardiovascular, dermatologic, developmental, endocrine, gastrointestinal, hematopoietic/immune, musculoskeletal, nervous, reproductive, and urologic. The disease/disorder/condition categories included cancer, inflammation/trauma, cell proliferation, neurological, and pooled. For each category, the number of libraries
15 expressing the sequence of interest was counted and divided by the total number of libraries across all categories. Percentage values of tissue-specific and disease-, disorder-, or condition-specific expression are reported in Table 3.

V. Extension of HSCOP Encoding Polynucleotides

20 The full length nucleic acid sequences of SEQ ID NO:10-18 were produced by extension of an appropriate fragment of the full length molecule using oligonucleotide primers designed from this fragment. One primer was synthesized to initiate 5' extension of the known fragment, and the other primer, to initiate 3' extension of the known fragment. The initial primers were designed using OLIGO 4.06 software (National
25 Biosciences), or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations was avoided.

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

High fidelity amplification was obtained by PCR using methods well known in the

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

The concentration of DNA in each well was determined by dispensing 100 μ l PICOGREEN quantitation reagent (0.25% (v/v) PICOGREEN; Molecular Probes, Eugene OR) dissolved in 1X TE and 0.5 μ l of undiluted PCR product into each well of an opaque
15 fluorimeter plate (Corning Costar, Acton MA), allowing the DNA to bind to the reagent. The plate was scanned in a Fluoroskan II (Labsystems Oy, Helsinki, Finland) to measure the fluorescence of the sample and to quantify the concentration of DNA. A 5 μ l to 10 μ l aliquot of the reaction mixture was analyzed by electrophoresis on a 1 % agarose mini-gel to determine which reactions were successful in extending the sequence.

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

The cells were lysed, and DNA was amplified by PCR using Taq DNA

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

10 In like manner, the nucleotide sequences of SEQ ID NO:10-18 are used to obtain 5' regulatory sequences using the procedure above, oligonucleotides designed for such extension, and an appropriate genomic library.

VI. Labeling and Use of Individual Hybridization Probes

Hybridization probes derived from SEQ ID NO:10-18 are employed to screen
15 cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250 μ Ci of [γ -³²P] adenosine triphosphate
20 (Amersham Pharmacia Biotech), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech). An aliquot containing 10⁷ counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following
25 endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba I, or Pvu II (DuPont NEN).

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

VII. Microarrays

A chemical coupling procedure and an ink jet device can be used to synthesize array elements on the surface of a substrate. (See, e.g., Baldeschweiler, *supra*.) An array analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced by hand or using available methods and machines and contain any appropriate number of elements. After hybridization, nonhybridized probes are removed and a scanner used to determine the levels and patterns of fluorescence. The degree of complementarity and the relative abundance of each probe which hybridizes to an element on the microarray may be assessed through analysis of the scanned images.

Full-length cDNAs, Expressed Sequence Tags (ESTs), or fragments thereof may comprise the elements of the microarray. Fragments suitable for hybridization can be selected using software well known in the art such as LASERGENE software (DNASTAR). Full-length cDNAs, ESTs, or fragments thereof corresponding to one of the nucleotide sequences of the present invention, or selected at random from a cDNA library relevant to the present invention, are arranged on an appropriate substrate, e.g., a glass slide. The cDNA is fixed to the slide using, e.g., UV cross-linking followed by thermal and chemical treatments and subsequent drying. (See, e.g., Schena, M. et al. (1995) *Science* 270:467-470; Shalon, D. et al. (1996) *Genome Res.* 6:639-645.) Fluorescent probes are prepared and used for hybridization to the elements on the substrate. The substrate is analyzed by procedures described above.

VIII. Complementary Polynucleotides

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

transcript.

IX. Expression of HSCOP

Expression and purification of HSCOP is achieved using bacterial or virus-based expression systems. For expression of HSCOP in bacteria, cDNA is subcloned into an appropriate vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the *trp-lac (tac)* hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express HSCOP upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of HSCOP in eukaryotic cells is achieved by infecting insect or mammalian cell lines with recombinant Autographica californica nuclear polyhedrosis virus (AcMNPV), commonly known as baculovirus. The nonessential polyhedrin gene of baculovirus is replaced with cDNA encoding HSCOP by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription. Recombinant baculovirus is used to infect Spodoptera frugiperda (Sf9) insect cells in most cases, or human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus. (See Engelhard, E. K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945.)

In most expression systems, HSCOP is synthesized as a fusion protein with, e.g., glutathione S-transferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid, single-step, affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-kilodalton enzyme from Schistosoma japonicum, enables the purification of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Amersham Pharmacia Biotech). Following purification, the GST moiety can be proteolytically cleaved from HSCOP at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak). 6-His, a stretch of six consecutive histidine residues, enables purification on

metal-chelate resins (QIAGEN). Methods for protein expression and purification are discussed in Ausubel (1995, supra, ch 10 and 16). Purified HSCOP obtained by these methods can be used directly in the following activity assay.

X. Demonstration of HSCOP Activity

5 HSCOP activity is demonstrated by the inhibition of differentiation in murine M1 cells transfected with the gene expressing HSCOP and induced to differentiate by treatment with IL-6 (Starr et al., supra). Differentiation is measured in the parent M1 cell line and in M1 cells transfected with HSCOP by the appearance of differentiated colonies arising from cells grown in semi-soft agar culture. The percent inhibition of
10 differentiation in M1 transfected cells compared to the parent M1 cell line is proportional to the activity of HSCOP in the former cells.

XI. Functional Assays

HSCOP function is assessed by expressing the sequences encoding HSCOP at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned
15 into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include pCMV SPORT (Life Technologies) and pCR3.1 (Invitrogen, Carlsbad CA), both of which contain the cytomegalovirus promoter. 5-10 μ g of recombinant vector are transiently transfected into a human cell line, preferably of endothelial or hematopoietic origin, using either liposome formulations or
20 electroporation. 1-2 μ g of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP; Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM),
25 an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP and to evaluate properties, for example, their apoptotic state. FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity
30 as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in

expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M. G. (1994) Flow Cytometry, Oxford, New York
5 NY.

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

15 **XII. Production of HSCOP Specific Antibodies**

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

20 Alternatively, the HSCOP amino acid sequence is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity, and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are well described in the art. (See, e.g., Ausubel,
25 1995, supra, ch. 11.)

Typically, oligopeptides 15 residues in length are synthesized using an ABI 431A Peptide Synthesizer (Perkin-Elmer) using fmoc-chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity. (See, e.g., Ausubel, 1995, supra.) Rabbits are
30 immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for antipeptide activity by, for example, binding the peptide to plastic,

blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radioiodinated goat anti-rabbit IgG.

XIII. Purification of Naturally Occurring HSCOP Using Specific Antibodies

Naturally occurring or recombinant HSCOP is substantially purified by
5 immunoaffinity chromatography using antibodies specific for HSCOP. An
immunoaffinity column is constructed by covalently coupling anti-HSOCH antibody to an
activated chromatographic resin, such as CNBr-activated SEPHAROSE (Amersham
Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the
manufacturer's instructions.

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

XIV. Identification of Molecules Which Interact with HSCOP

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

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

Table I

Protein SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
1	10	1758450	PITUNOT03	1758450H1 (PITUNOT03), 2522888H1 (BRAITUT21), 2638457T6 (BONTNOT01), 4177041H1 (BRAINOT22), 983327R6 (TONGTUT01)
2	11	1834242	BRAINON01	1834242H1 (BRAINON01), 2729019F6 (OVRTUT05), 1834242X14R1 (BRAINON01), 1834242X13R1 (BRAINON01), 782734R1 (MYOMNOT01), 1421816F1 (KIDNNOT09), 1266788T1 (BRAINOT09), 1266937F1 (BRAINOT09), 1757409H1 (PITUNOT03)
3	12	1849725	LUNGFET03	1849725H1 (LUNGFET03), 158419R1 (ADENINB01), 1611584F6 (COLNTUT06), SNBA01842F1
4	13	2547840	LUNGTUT06	2547840H1 (LUNGTUT06), 2109531H1 (BRAITUT03), 1282090F6 (COLNNOT16), 2953401H1 (KIDNFET01), 1378436F1 (LUNGNOT10), 1282090T6 (COLNNOT16), 2014223H1 (TESTNOT03)
5	14	3071986	UTRSNOR01	3071986H1 (UTRSNOR01), 2860082H1 (SININOT03), 265704H1 (HNT2AGT01), 1476725F1 (CORPNOT02), 1605966F6 (LUNGNOT15), 1512163F1 (LUNGNOT14), 1235056T1 (LUNGFET03), 1502645F1 (BRAITUT07)
6	15	3484619	KIDNNOT31	3484619H1 (KIDNNOT31), 2631528H1 (COLNTUT15), 1319593F1 (BLADNOT04), 2849278F6 (BRSTTUT13), 4024760H1 (BRAXNOT02), 2738625F6 (OVARNOT09)

Table 1 cont.

Protein SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
7	16	1275743	TESTTUT02	841412R1 (PROSTUT05), 1275743H1 (TESTTUT02), 2571403R6 (HIPOAZT01), 3319512F6 (PROSBPT03), 3964671X314D1 (PROSNOT14)
8	17	1722533	BLADNOT06	585432X15 (PROSNOT02), 1271329T6 (TESTTUT02), 1579164F1 (DUODNOT01), 1722533F6 and 1722533H1 (BLADNOT06), 2635492F6 (BONTNOT01), 2880628H1 and 2882159F6 (UTRSTUT05), 3203865H1 (PENCNOT02), 4852494H1 (TESTNOT10)
9	18	1759763	PITUNOT03	034803H1 (THPINOB01), 161745R6 (ADENINB01), 595902H1 (BRAVUNT02), 626174R6 (PGANNOT01), 953742R1 (SCORNON01), 1759763H1 (PITUNOT03), 2138314F6 (ENDCNOT01), 2532454T6 (GBLANOT02), 3053743H1 (LNODNOT08)

Table 2

Protein Seq ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential glycosylation sites	Signature Sequence	Homologous Sequence	Analytical Methods
1	288	S17 S64 T201 S206 S279 S19 T53 T249	N48 N86 N223	Ankyrin Repeats: T53-N80 N86-N112 R118-N145 N151-R178 SOCS box: P247-E288	Ankyrin protein	BLAST HMM PFAM
2	423	S22 S44 S92 S101 T299 T329 T4 S15 S91 S101 S126 T196 S244 S314 S387 T419 Y420	N94 N167	WD-40 repeats: L168-D201 L212-S244 L254-D286 SOCS box: L385-F423	WSB-2	BLAST HMM PFAM PRINTS MOTIFS
3	349	S34 S57 S80 S125 T138 S21		SOCS box: V307-F347	SOCS-containing protein	HMM
4	355	S126 S44 S46 S48 S69 S171 S340 S211 S264 S322 S340 Y102	N111	SOCS box: A274-V316	SOCS-containing protein	HMM
5	421	T18 S118 S328 T57 S83 T159 T194 T380 S394 S418 Y419	N37 N77 N80	WD-40 repeats: L166-D199 L210-N242 L252-D284 SOCS box: V384-I421	WSB-1	BLAST HMM PFAM PRINTS MOTIFS

Table 2 cont.

Protein Seq ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential glycosylation sites	Signature Sequence	Homologous Sequence	Analytical Methods
6	278	S219 T254 Y277		Ankyrin repeats: S52-D79 G85-N112 T117-E144 F150-N177 H182-Y209 SOCS box: P239-N278	Ankyrin protein	BLAST HMM PFAM
7	281	S33 T151 S166 S191 T97 T217 S256 S262	N272	Ras family domain: K16-C278 prenyl group binding site: C278-S281 SOCS box: V189-H237	Rar protein or Ras-like GTPase	BLAST, BLOCKS PRINTS, MOTIFS, PFAM, HMM
8	635	S413 S24 S26 S38 T47 T169 S247 T333 S545 S579 S6 T57 T180 S201 S345 T360 S390 S522 Y210	N202 N304 N331	Ankyrin repeats: E137-R364 S368-A400 R410-A472 SOCS box: P593-Q635	Ankyrin protein	BLAST PFAM HMM
9	518	T108 T109 S154 S234 S412 T270 S443 S473 S476	N106 N139	Ankyrin repeats: D9-M74 E78-S143 C145-N243 N279-C311 GTP-binding: L300-S317 SOCS box: V459-M506	Ankyrin protein	BLAST PFAM HMM BLOCKS

Table 3

Polynucleotide SEQ ID NO:	Tissue Expression (Fraction of Total)	Disease, Disorder or Condition (Fraction of Total)	Vector
10	Reproductive (0.269) Cardiovascular (0.154) Developmental (0.115)	Cancer (0.346) Inflammation (0.269) Cell proliferation (0.192)	pSPORT1
11	Nervous (0.244) Reproductive (0.233) Hematopoietic/Immune (0.144)	Cancer (0.389) Cell proliferation (0.256) Inflammation (0.256)	pSPORT1
12	Gastrointestinal (0.250) Hematopoietic/Immune (0.250) Reproductive (0.250)	Inflammation (0.417) Cancer (0.250) Cell proliferation (0.250)	pINCY
13	Reproductive (0.218) Hematopoietic/Immune (0.179) Nervous (0.141)	Cancer (0.449) Inflammation (0.269) Cell proliferation (0.205)	pINCY
14	Reproductive (0.222) Cardiovascular (0.167) Hematopoietic/Immune (0.148)	Cancer (0.454) Inflammation (0.324) Cell proliferation (0.185)	pINCY
15	Reproductive (0.400) Gastrointestinal (0.200) Nervous (0.200)	Cancer (0.667) Inflammation (0.133) Neurological (0.067)	pINCY

Table 3 cont.

Polynucleotide SEQ ID NO:	Tissue Expression (Fraction of Total)	Disease, Disorder or Condition (Fraction of Total)	Vector
16	Reproductive (0.400) Nervous (0.360) Cardiovascular (0.080)	Cancer (0.680) Inflammation (0.120) Neurological (0.120)	pINCY
17	Reproductive (0.438) Gastrointestinal (0.156) Cardiovascular (0.094)	Cancer (0.438) Inflammation (0.250) Cell proliferation (0.156)	pINCY
18	Nervous (0.286) Cardiovascular (0.163) Hematopoietic/Immune (0.143)	Cancer (0.327) Cell proliferation (0.265) Inflammation (0.224)	pSPORT1

Table 4

Polynucleotide SEQ ID NO:	Library	Library Description
10	PITUNOT03	The library was constructed using RNA isolated from pituitary tissue of a 46-year-old Caucasian male who died from colon cancer. Patient history included arthritis and peptic ulcer disease.
11	BRAINON01	This normalized brain library was constructed from 4.88 million independent clones from the BRAINOT03 library. Starting RNA was made from nontumorous brain tissue removed from a 26-year-old Caucasian male during cranioplasty and excision of a cerebral meningeal lesion. Pathology for the associated tumor tissue indicated a grade 4 oligoastrocytoma. The patient presented with epilepsy, ptosis of the eyelid, hemiplegia and migraine. Patient history included radiation therapy, and hypercholesterolemia. The normalization and hybridization conditions were adapted from Soares et al., PNAS (1994) 91:9928, using a significantly longer (48-hour) reannealing hybridization period.
12	LUNGFET03	The library was constructed using RNA isolated from lung tissue removed from a Caucasian female fetus who died at 20 weeks' gestation. Family history included bronchitis in the mother during the first trimester.
13	LUNGTUT06	The library was constructed using RNA isolated from apical lung tumor tissue removed from an 80-year-old Caucasian female during a segmental lung resection. Pathology indicated a metastatic granulosa cell tumor. Patient history included benign hypertension, nonspecific reaction to a tuberculin skin test, pelvic soft tissue tumor, and acquired antibody E from a previous transfusion. Family history included tuberculosis, benign hypertension, lung cancer, and atherosclerotic coronary artery disease.

Table 4 cont.

Polynucleotide SEQ ID NO:	Library	Library Description
14	UTRSNOR01	The library was constructed using RNA isolated from nontumorous uterine endometrium tissue removed from a 29-year-old Caucasian female during a vaginal hysterectomy and cystocele repair. Pathology indicated the endometrium was secretory, and the cervix showed mild chronic cervicitis with focal squamous metaplasia. Pathology for the associated tumor tissue indicated intramural uterine leiomyoma. Patient history included hypothyroidism and pelvic floor relaxation. Family history included benign hypertension, Type II diabetes, and hyperlipidemia.
15	KIDNNOT31	The library was constructed using RNA isolated from tissue that had kidney markers.
16	TESTTUT02	The library was constructed using RNA isolated from testicular tumor removed from a 31-year-old Caucasian male during unilateral orchiectomy. Pathology indicated embryonal carcinoma.
17	BLADNOT06	The library was constructed using RNA isolated from the posterior wall bladder tissue removed from a 66-year-old Caucasian male during a radical prostatectomy, radical cystectomy and urinary diversion. Pathology for the associated tumor tissue indicated grade 3 transitional cell carcinoma on the anterior wall of the bladder and urothelium. Patient history included lung neoplasm, and tobacco abuse in remission. Family history included a malignant breast neoplasm, tuberculosis, cerebrovascular disease, atherosclerotic coronary artery disease, and lung cancer.
18	PITUNOT03	The library was constructed using RNA isolated from pituitary tissue of a 46-year-old Caucasian male who died from colon cancer.

Table 5

Program	Description	Reference	Parameter Threshold
ABI FACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) <i>J. Mol. Biol.</i> 215:403-410; Altschul, S.F. et al. (1997) <i>Nucleic Acids Res.</i> 25: 3389-3402.	ESTs: Probability value= 1.0E-8 or less Full Length sequences: Probability value= 1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises at least five functions: fasta, tfasta, fastx, tfastx, and ssearch.	Pearson, W.R. and D.J. Lipman (1988) <i>Proc. Natl. Acad. Sci.</i> 85:2444-2448; Pearson, W.R. (1990) <i>Methods Enzymol.</i> 183: 63-98; and Smith, T.F. and M. S. Waterman (1981) <i>Adv. Appl. Math.</i> 2:482-489.	ESTs: fasta E value=1.06E-6 Assembled ESTs: fasta Identity= 95% or greater and Match length=200 bases or greater; fastx E value=1.0E-8 or less Full Length sequences: fastx score=100 or greater
BLIMPS	A BLocks IMProved Searcher that matches a sequence against those in BLOCKS, PRINTS, DOMO, PRODOM, and PFAM databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S and J.G. Henikoff, <i>Nucl. Acid Res.</i> , 19:6565-72, 1991. J.G. Henikoff and S. Henikoff (1996) <i>Methods Enzymol.</i> 266:88-105; and Attwood, T.K. et al. (1997) <i>J. Chem. Inf. Comput. Sci.</i> 37: 417-424.	Score=1000 or greater; Ratio of Score/Strength = 0.75 or larger; and, if applicable, Probability value= 1.0E-3 or less
HMMER	An algorithm for searching a query sequence against hidden Markov model (HMM)-based databases of protein family consensus sequences, such as PFAM.	Krogh, A. et al. (1994) <i>J. Mol. Biol.</i> , 235:1501-1531; Sonnhammer, E.L.L. et al. (1988) <i>Nucleic Acids Res.</i> 26:320-322.	Score=10-50 bits for PFAM hits, depending on individual protein families

Table 5 (cont.)

Program	Description	Reference	Parameter Threshold
ProfileScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Gribskov, M. et al. (1988) CABIOS 4:61-66; Gribskov, et al. (1989) Methods Enzymol. 183:146-159; Bairoch, A. et al. (1997) Nucleic Acids Res. 25: 217-221.	Normalized quality score ≥ GCG-specified "HIGH" value for that particular Prosite motif. Generally, score = 1.4-2.1.
Phred	A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Phis Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M. S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	Score = 120 or greater; Match length = 56 or greater
Consed	A graphical tool for viewing and editing Phrap assemblies	Gordon, D. et al. (1998) Genome Res. 8:195-202.	
SPScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielson, H. et al. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12: 431-439.	Score = 3.5 or greater
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Bairoch et al. <i>supra</i> ; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

What is claimed is:

1. A substantially purified polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-9, and fragments thereof.
- 5 2. A substantially purified variant having at least 90% amino acid sequence identity to the amino acid sequence of claim 1.
3. An isolated and purified polynucleotide encoding the polypeptide of claim 1.
4. An isolated and purified polynucleotide variant having at least 90%
10 polynucleotide sequence identity to the polynucleotide of claim 3.
5. An isolated and purified polynucleotide which hybridizes under stringent conditions to the polynucleotide of claim 3.
6. An isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide of claim 3.
- 15 7. A method for detecting a polynucleotide, the method comprising the steps of:
 - (a) hybridizing the polynucleotide of claim 6 to at least one nucleic acid in a sample, thereby forming a hybridization complex; and
 - (b) detecting the hybridization complex, wherein the presence of the
20 hybridization complex correlates with the presence of the polynucleotide in the sample.
8. The method of claim 7 further comprising amplifying the polynucleotide prior to hybridization.
9. An isolated and purified polynucleotide comprising a polynucleotide
25 sequence selected from the group consisting of SEQ ID NO:10-18, and fragments thereof.
10. An isolated and purified polynucleotide variant having at least 90% polynucleotide sequence identity to the polynucleotide of claim 9.
11. An isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide of claim 9.
- 30 12. An expression vector comprising at least a fragment of the polynucleotide of claim 3.

13. A host cell comprising the expression vector of claim 12.
14. A method for producing a polypeptide, the method comprising the steps of:
 - a) culturing the host cell of claim 13 under conditions suitable for the expression of the polypeptide; and
 - 5 b) recovering the polypeptide from the host cell culture.
15. A pharmaceutical composition comprising the polypeptide of claim 1 in conjunction with a suitable pharmaceutical carrier.
16. A purified antibody which specifically binds to the polypeptide of claim 1.
17. A purified agonist of the polypeptide of claim 1.
- 10 18. A purified antagonist of the polypeptide of claim 1.
19. A method for treating or preventing a disorder associated with decreased expression or activity of HSCOP, the method comprising administering to a subject in need of such treatment an effective amount of the pharmaceutical composition of claim 15.
- 15 20. A method for treating or preventing a disorder associated with increased expression or activity of HSCOP, the method comprising administering to a subject in need of such treatment an effective amount of the antagonist of claim 18.

SEQUENCE LISTING

<110> INCYTE PHARMACEUTICALS, INC.

LAL, Preeti
 HILLMAN, Jennifer L.
 GORGONE, Gina
 CORLEY, Neil C.
 PATTERSON, Chandra
 YUE, Henry
 TANG, Y. Tom
 AZIMZAI, Yalda

<120> HUMAN SOCS PROTEINS

<130> PF-0525 PCT

<140> To Be Assigned

<141> Herewith

<150> 60/087,104; 09/216,006

<151> 1998-05-28; 1998-12-17

<160> 18

<170> PERL Program

<210> 1

<211> 288

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte clone 1758450

<400> 1

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

```

Asn Asp Thr Pro Leu Ser Trp Ala Ala Met Lys Gly Asn Leu Glu
      155                               160           165
Ser Val Ser Ile Leu Leu Asp Tyr Gly Ala Glu Val Arg Val Ile
      170                               175           180
Asn Leu Ile Gly Gln Thr Pro Ile Ser Arg Leu Val Ala Leu Leu
      185                               190           195
Val Arg Gly Leu Gly Thr Glu Lys Glu Asp Ser Cys Phe Glu Leu
      200                               205           210
Leu His Arg Ala Val Gly His Phe Glu Leu Arg Lys Asn Gly Thr
      215                               220           225
Met Pro Arg Glu Val Ala Arg Asp Pro Gln Leu Cys Glu Lys Leu
      230                               235           240
Thr Val Leu Cys Ser Ala Pro Gly Thr Leu Lys Thr Leu Ala Arg
      245                               250           255
Tyr Ala Val Arg Arg Ser Leu Gly Leu Gln Tyr Leu Pro Asp Ala
      260                               265           270
Val Lys Gly Leu Pro Leu Pro Ala Ser Leu Lys Glu Tyr Leu Leu
      275                               280           285
Leu Leu Glu
    
```

```

<210> 2
<211> 423
<212> PRT
<213> Homo sapiens
    
```

```

<220>
<221> misc_feature
<223> Incyte clone 1834242
    
```

```

<400> 2
Met Lys Leu Thr Pro Arg Thr Ala Gly Arg Ala Trp Ala Gln Ser
  1                               5           10           15
Arg Lys Gly Lys Arg Ser Ser Trp Gly Gly Thr Ala Ala Val Ala
      20                               25           30
Glu Leu Lys Pro Gly Arg Pro His Gln Phe Asp Trp Lys Ser Ser
      35                               40           45
Cys Glu Thr Trp Ser Val Ala Phe Ser Pro Asp Gly Ser Trp Phe
      50                               55           60
Ala Trp Ser Gln Gly His Cys Ile Val Lys Leu Ile Pro Trp Pro
      65                               70           75
Leu Glu Glu Gln Phe Ile Pro Lys Gly Phe Glu Ala Lys Ser Arg
      80                               85           90
Ser Ser Lys Asn Glu Thr Lys Gly Arg Gly Ser Pro Lys Glu Lys
      95                               100          105
Thr Leu Asp Cys Gly Gln Ile Val Trp Gly Leu Ala Phe Ser Pro
      110                              115          120
Trp Pro Ser Pro Pro Ser Arg Lys Leu Trp Ala Arg His His Pro
      125                              130          135
Gln Val Pro Asp Val Ser Cys Leu Val Leu Ala Thr Gly Leu Asn
      140                              145          150
Asp Gly Gln Ile Lys Ile Trp Glu Val Gln Thr Gly Leu Leu Leu
      155                              160          165
Leu Asn Leu Ser Gly His Gln Asp Val Val Arg Asp Leu Ser Phe
      170                              175          180
    
```

Thr Pro Ser Gly Ser Leu Ile Leu Val Ser Ala Ser Arg Asp Lys
 185 190 195
 Thr Leu Arg Ile Trp Asp Leu Asn Lys His Gly Lys Gln Ile Gln
 200 205 210
 Val Leu Ser Gly His Leu Gln Trp Val Tyr Cys Cys Ser Ile Ser
 215 220 225
 Pro Asp Cys Ser Met Leu Cys Ser Ala Ala Gly Glu Lys Ser Val
 230 235 240
 Phe Leu Trp Ser Met Arg Ser Tyr Thr Leu Ile Arg Lys Leu Glu
 245 250 255
 Gly His Gln Ser Ser Val Val Ser Cys Asp Phe Ser Pro Asp Ser
 260 265 270
 Ala Leu Leu Val Thr Ala Ser Tyr Asp Thr Asn Val Ile Met Trp
 275 280 285
 Asp Pro Tyr Thr Gly Glu Arg Leu Arg Ser Leu His His Thr Gln
 290 295 300
 Val Asp Pro Ala Met Asp Asp Ser Asp Val His Ile Ser Ser Leu
 305 310 315
 Arg Ser Val Cys Phe Ser Pro Glu Gly Leu Tyr Leu Ala Thr Val
 320 325 330
 Ala Asp Asp Arg Leu Leu Arg Ile Trp Ala Leu Glu Leu Lys Thr
 335 340 345
 Pro Ile Ala Phe Ala Pro Met Thr Asn Gly Leu Cys Cys Thr Phe
 350 355 360
 Phe Pro His Gly Gly Val Ile Ala Thr Gly Thr Arg Asp Gly His
 365 370 375
 Val Gln Phe Trp Thr Ala Pro Arg Val Leu Ser Ser Leu Lys His
 380 385 390
 Leu Cys Arg Lys Ala Leu Arg Ser Phe Leu Thr Thr Tyr Gln Val
 395 400 405
 Leu Ala Leu Pro Ile Pro Lys Lys Met Lys Glu Phe Leu Thr Tyr
 410 415 420
 Arg Thr Phe

<210> 3
 <211> 349
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte clone 1849725

<400> 3
 Met Glu Asp Pro Gln Ser Lys Glu Pro Ala Gly Glu Ala Val Ala
 1 5 10 15
 Pro Ala Leu Leu Glu Ser Pro Arg Pro Glu Gly Gly Glu Glu Pro
 20 25 30
 Pro Arg Pro Ser Pro Glu Glu Thr Gln Gln Cys Lys Phe Asp Gly
 35 40 45
 Gln Glu Thr Lys Gly Ser Lys Phe Ile Thr Ser Ser Ala Ser Asp
 50 55 60
 Phe Ser Asp Pro Val Tyr Lys Glu Ile Ala Ile Thr Asn Gly Cys
 65 70 75
 Ile Asn Arg Met Ser Lys Glu Glu Leu Arg Ala Lys Leu Ser Glu

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

<210> 4

<211> 355

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte clone 2547840

<400> 4

Met Ala Arg Arg Pro Arg Asn Ser Arg Ala Trp His Phe Val Leu					
1	5		10		15
Ser Ala Ala Arg Arg Asp Ala Asp Ala Arg Ala Val Ala Leu Ala					
	20		25		30
Gly Ser Thr Asn Trp Gly Tyr Asp Ser Asp Gly Gln His Ser Asp					
	35		40		45
Ser Asp Ser Asp Pro Glu Tyr Ser Thr Leu Pro Pro Ser Ile Pro					

	50		55		60
Ser Ala Val Pro Val Thr Gly Glu Ser Phe Cys Asp Cys Ala Gly					
	65		70		75
Gln Ser Glu Ala Ser Phe Cys Ser Ser Leu His Ser Ala His Arg					
	80		85		90
Gly Arg Asp Cys Arg Cys Gly Glu Glu Asp Glu Tyr Phe Asp Trp					
	95		100		105
Val Trp Asp Asp Leu Asn Lys Ser Ser Ala Thr Leu Leu Ser Cys					
	110		115		120
Asp Asn Arg Lys Val Ser Phe His Met Glu Tyr Ser Cys Gly Thr					
	125		130		135
Ala Ala Ile Arg Gly Thr Lys Glu Leu Gly Glu Gly Gln His Phe					
	140		145		150
Trp Glu Ile Lys Met Thr Ser Pro Val Tyr Gly Thr Asp Met Met					
	155		160		165
Val Gly Ile Gly Thr Ser Asp Val Asp Leu Asp Lys Tyr Arg His					
	170		175		180
Thr Phe Cys Ser Leu Leu Gly Arg Asp Glu Asp Ser Trp Gly Leu					
	185		190		195
Ser Tyr Thr Gly Leu Leu His His Lys Gly Asp Lys Thr Ser Phe					
	200		205		210
Ser Ser Arg Phe Gly Gln Gly Ser Ile Ile Gly Val His Leu Asp					
	215		220		225
Thr Trp His Gly Thr Leu Thr Phe Phe Lys Asn Arg Lys Cys Ile					
	230		235		240
Gly Val Ala Ala Thr Lys Leu Gln Asn Lys Arg Phe Tyr Pro Met					
	245		250		255
Val Cys Ser Thr Ala Ala Arg Ser Ser Met Lys Val Thr Arg Ser					
	260		265		270
Cys Ala Ser Ala Thr Ser Leu Gln Tyr Leu Cys Cys His Arg Leu					
	275		280		285
Arg Gln Leu Arg Pro Asp Ser Gly Asp Thr Leu Glu Gly Leu Pro					
	290		295		300
Leu Pro Pro Gly Leu Lys Gln Val Leu His Asn Lys Leu Gly Trp					
	305		310		315
Val Leu Ser Met Ser Cys Ser Arg Arg Lys Ala Pro Val Ser Asp					
	320		325		330
Pro Gln Ala Ala Thr Ser Ala His Pro Ser Ser Arg Glu Pro Arg					
	335		340		345
Pro Cys Gln Arg Lys Arg Cys Arg Arg Thr					
	350		355		

<210> 5
 <211> 421
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte clone 3071986

<400> 5
 Met Ala Ser Phe Pro Pro Arg Val Asn Glu Lys Glu Ile Val Arg
 1 5 10 15

Leu	Arg	Thr	Ile	Gly	Glu	Leu	Leu	Ala	Pro	Ala	Ala	Pro	Phe	Asp
				20					25					30
Lys	Lys	Cys	Gly	Arg	Glu	Asn	Trp	Thr	Val	Ala	Phe	Ala	Pro	Asp
				35					40					45
Gly	Ser	Tyr	Phe	Ala	Trp	Ser	Gln	Gly	His	Arg	Thr	Val	Lys	Leu
				50					55					60
Val	Pro	Trp	Ser	Gln	Cys	Leu	Gln	Asn	Phe	Leu	Leu	His	Gly	Thr
				65					70					75
Lys	Asn	Val	Thr	Asn	Ser	Ser	Ser	Leu	Arg	Leu	Pro	Arg	Gln	Asn
				80					85					90
Ser	Asp	Gly	Gly	Gln	Lys	Asn	Lys	Pro	Arg	Glu	His	Ile	Ile	Asp
				95					100					105
Cys	Gly	Asp	Ile	Val	Trp	Ser	Leu	Ala	Phe	Gly	Ser	Ser	Val	Pro
				110					115					120
Glu	Lys	Gln	Ser	Arg	Cys	Val	Asn	Ile	Glu	Trp	His	Arg	Phe	Arg
				125					130					135
Phe	Gly	Gln	Asp	Gln	Leu	Leu	Leu	Ala	Thr	Gly	Leu	Asn	Asn	Gly
				140					145					150
Arg	Ile	Lys	Ile	Trp	Asp	Val	Tyr	Thr	Gly	Lys	Leu	Leu	Leu	Asn
				155					160					165
Leu	Val	Asp	His	Thr	Glu	Val	Val	Arg	Asp	Leu	Thr	Phe	Ala	Pro
				170					175					180
Asp	Gly	Ser	Leu	Ile	Leu	Val	Ser	Ala	Ser	Arg	Asp	Lys	Thr	Leu
				185					190					195
Arg	Val	Trp	Asp	Leu	Lys	Asp	Asp	Gly	Asn	Met	Met	Lys	Val	Leu
				200					205					210
Arg	Gly	His	Gln	Asn	Trp	Val	Tyr	Ser	Cys	Ala	Phe	Ser	Pro	Asp
				215					220					225
Ser	Ser	Met	Leu	Cys	Ser	Val	Gly	Ala	Ser	Lys	Ala	Val	Phe	Leu
				230					235					240
Trp	Asn	Met	Asp	Lys	Tyr	Thr	Met	Ile	Arg	Lys	Leu	Glu	Gly	His
				245					250					255
His	His	Asp	Val	Val	Ala	Cys	Asp	Phe	Ser	Pro	Asp	Gly	Ala	Leu
				260					265					270
Leu	Ala	Thr	Ala	Ser	Tyr	Asp	Thr	Arg	Val	Tyr	Ile	Trp	Asp	Pro
				275					280					285
His	Asn	Gly	Asp	Ile	Leu	Met	Glu	Phe	Gly	His	Leu	Phe	Pro	Pro
				290					295					300
Pro	Thr	Pro	Ile	Phe	Ala	Gly	Gly	Ala	Asn	Asp	Arg	Trp	Val	Arg
				305					310					315
Ser	Val	Ser	Phe	Ser	His	Asp	Gly	Leu	His	Val	Ala	Ser	Leu	Ala
				320					325					330
Asp	Asp	Lys	Met	Val	Arg	Phe	Trp	Arg	Ile	Asp	Glu	Asp	Tyr	Pro
				335					340					345
Val	Gln	Val	Ala	Pro	Leu	Ser	Asn	Gly	Leu	Cys	Cys	Ala	Phe	Ser
				350					355					360
Thr	Asp	Gly	Ser	Val	Leu	Ala	Ala	Gly	Thr	His	Asp	Gly	Ser	Val
				365					370					375
Tyr	Phe	Trp	Ala	Thr	Pro	Arg	Gln	Val	Pro	Ser	Leu	Gln	His	Leu
				380					385					390
Cys	Arg	Met	Ser	Ile	Arg	Arg	Val	Met	Pro	Thr	Gln	Glu	Val	Gln
				395					400					405
Glu	Leu	Pro	Ile	Pro	Ser	Lys	Leu	Leu	Glu	Phe	Leu	Ser	Tyr	Arg
				410					415					420
Ile														

<210> 6
 <211> 278
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte clone 3484619

<400> 6
 Met Glu Pro Arg Ala Ala Asp Gly Cys Phe Leu Gly Asp Val Gly
 1 5 10 15
 Phe Trp Val Glu Arg Thr Pro Val His Glu Ala Ala Gln Arg Gly
 20 25 30
 Glu Ser Leu Gln Leu Gln Gln Leu Ile Glu Ser Gly Ala Cys Val
 35 40 45
 Asn Gln Val Thr Val Asp Ser Ile Thr Pro Leu His Ala Ala Ser
 50 55 60
 Leu Gln Gly Gln Ala Arg Cys Val Gln Leu Leu Leu Ala Ala Gly
 65 70 75
 Ala Gln Val Asp Ala Arg Asn Ile Asp Gly Ser Thr Pro Leu Cys
 80 85 90
 Asp Ala Cys Ala Ser Gly Ser Ile Glu Cys Val Lys Leu Leu Leu
 95 100 105
 Ser Tyr Gly Ala Lys Val Asn Pro Pro Leu Tyr Thr Ala Ser Pro
 110 115 120
 Leu His Glu Ala Cys Met Ser Gly Ser Ser Glu Cys Val Arg Leu
 125 130 135
 Leu Ile Asp Val Gly Ala Asn Leu Glu Ala His Asp Cys His Phe
 140 145 150
 Gly Thr Pro Leu His Val Ala Cys Ala Arg Glu His Leu Asp Cys
 155 160 165
 Val Lys Val Leu Leu Asn Ala Gly Ala Asn Val Asn Ala Ala Lys
 170 175 180
 Leu His Glu Thr Ala Leu His His Ala Ala Lys Val Lys Asn Val
 185 190 195
 Asp Leu Ile Glu Met Leu Ile Glu Phe Gly Gly Asn Ile Tyr Ala
 200 205 210
 Arg Asp Asn Arg Gly Lys Lys Pro Ser Asp Tyr Thr Trp Ser Ser
 215 220 225
 Ser Ala Pro Ala Lys Cys Phe Glu Tyr Tyr Glu Lys Thr Pro Leu
 230 235 240
 Thr Leu Ser Gln Leu Cys Arg Val Asn Leu Arg Lys Ala Thr Gly
 245 250 255
 Val Arg Gly Leu Glu Lys Ile Ala Lys Leu Asn Ile Pro Pro Arg
 260 265 270
 Leu Ile Asp Tyr Leu Ser Tyr Asn
 275

<210> 7
 <211> 281
 <212> PRT
 <213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte clone 1275743

<400> 7

```

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

<210> 8

<211> 635

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte clone 1722533

<400> 8

```

Met Ala Thr Gln Ile Ser Thr Arg Gly Ser Gln Cys Thr Ile Gly
 1          5          10          15
    
```

Gln Glu Glu Tyr Ser Leu Tyr Ser Ser Leu Ser Glu Asp Glu Leu
 20 25 30
 Val Gln Met Ala Ile Glu Gln Ser Leu Ala Asp Lys Thr Arg Gly
 35 40 45
 Pro Thr Thr Ala Glu Ala Thr Ala Ser Ala Cys Thr Asn Arg Gln
 50 55 60
 Pro Ala His Phe Tyr Pro Trp Thr Arg Ser Thr Ala Pro Pro Glu
 65 70 75
 Ser Ser Pro Ala Arg Ala Pro Met Gly Leu Phe Gln Gly Val Met
 80 85 90
 Gln Lys Tyr Ser Ser Ser Leu Phe Lys Thr Ser Gln Leu Ala Pro
 95 100 105
 Ala Asp Pro Leu Ile Lys Ala Ile Lys Asp Gly Asp Glu Glu Ala
 110 115 120
 Leu Lys Thr Met Ile Lys Glu Gly Lys Asn Leu Ala Glu Pro Asn
 125 130 135
 Lys Glu Gly Trp Leu Pro Leu His Glu Ala Ala Tyr Tyr Gly Gln
 140 145 150
 Val Gly Cys Leu Lys Val Leu Gln Arg Ala Tyr Pro Gly Thr Ile
 155 160 165
 Asp Gln Arg Thr Leu Gln Glu Glu Thr Ala Val Tyr Leu Ala Thr
 170 175 180
 Cys Arg Gly His Leu Asp Cys Leu Leu Ser Leu Leu Gln Ala Gly
 185 190 195
 Ala Glu Pro Asp Ile Ser Asn Lys Ser Arg Glu Thr Pro Leu Tyr
 200 205 210
 Lys Ala Cys Glu Arg Lys Asn Ala Glu Ala Val Lys Ile Leu Val
 215 220 225
 Gln His Asn Ala Asp Thr Asn His Arg Cys Asn Arg Gly Trp Thr
 230 235 240
 Ala Leu His Glu Ser Val Ser Arg Asn Asp Leu Glu Val Met Gln
 245 250 255
 Ile Leu Val Ser Gly Gly Ala Lys Val Glu Ser Lys Asn Ala Tyr
 260 265 270
 Gly Ile Thr Pro Leu Phe Val Ala Ala Gln Ser Gly Gln Leu Glu
 275 280 285
 Ala Leu Arg Phe Leu Ala Lys Tyr Gly Ala Asp Ile Asn Thr Gln
 290 295 300
 Ala Ser Asp Asn Ala Ser Ala Leu Tyr Glu Ala Cys Lys Asn Glu
 305 310 315
 His Glu Glu Val Val Glu Phe Leu Leu Ser Gln Gly Ala Asp Ala
 320 325 330
 Asn Lys Thr Asn Lys Asp Gly Leu Leu Pro Leu His Ile Ala Ser
 335 340 345
 Lys Lys Gly Asn Tyr Arg Ile Val Gln Met Leu Leu Pro Val Thr
 350 355 360
 Ser Arg Thr Arg Ile Arg Arg Ser Gly Val Ser Pro Leu His Leu
 365 370 375
 Ala Ala Glu Arg Asn His Asp Glu Val Leu Glu Ala Leu Leu Ser
 380 385 390
 Ala Arg Phe Asp Val Asn Thr Pro Leu Ala Pro Glu Arg Ala Arg
 395 400 405
 Leu Tyr Glu Asp Arg Arg Thr Ser Ala Leu Tyr Phe Ala Val Val
 410 415 420
 Asn Asn Asn Val Tyr Ala Thr Glu Leu Leu Leu Gln His Gly Ala
 425 430 435
 Asp Pro Asn Arg Asp Val Ile Ser Pro Leu Leu Val Ala Ile Arg

	440		445		450
His Gly Cys Leu Arg Thr Met Gln Leu Leu Leu Asp His Gly Ala					
	455		460		465
Asn Ile Asp Ala Tyr Ile Ala Thr His Pro Thr Ala Phe Pro Ala					
	470		475		480
Thr Ile Met Phe Ala Met Lys Cys Leu Ser Leu Leu Lys Phe Leu					
	485		490		495
Met Asp Leu Gly Cys Asp Gly Glu Pro Cys Phe Ser Cys Leu Tyr					
	500		505		510
Gly Asn Gly Pro His Pro Pro Ala Pro Gln Pro Ser Ser Arg Phe					
	515		520		525
Asn Asp Ala Pro Ala Ala Asp Lys Glu Pro Ser Val Val Gln Phe					
	530		535		540
Cys Glu Phe Val Ser Ala Pro Glu Val Ser Arg Trp Ala Gly Pro					
	545		550		555
Ile Ile Asp Val Leu Leu Asp Tyr Val Gly Asn Val Gln Leu Cys					
	560		565		570
Ser Arg Leu Lys Glu His Ile Asp Ser Phe Glu Asp Trp Ala Val					
	575		580		585
Ile Lys Glu Lys Ala Glu Pro Pro Arg Pro Leu Ala His Leu Cys					
	590		595		600
Arg Leu Arg Val Arg Lys Ala Ile Gly Lys Tyr Arg Ile Lys Leu					
	605		610		615
Leu Asp Thr Leu Pro Leu Pro Gly Arg Leu Ile Arg Tyr Leu Lys					
	620		625		630
Tyr Glu Asn Thr Gln					
	635				

<210> 9
 <211> 518
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte clone 1759763

<400> 9

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

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

<210> 10
 <211> 1117
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte clone 1758450

<400> 10
 cacgccttga cagcggcttt caacccccac ctccagcccag caattcggca gtttgaggca 60
 tgtgaacacc ttgagccttg atgagttcca gtatgtggta tattatgcag agcattcaga 120
 gcaaatactc tctctccgag cgcttaatcc gaacaattgc tgccatccgt tccttcccac 180
 atgataatgt agaggacctc atcagagggg gagcagatgt gaactgcact catggcacac 240
 tgaagccctt gcaactgtgc tgtatgggtg cagatgctga ctgtgtggag ttacttctgg 300
 aaaaaggagc cgaggtgaat gccctggatg ggtataaccg aacagccctc cactatgcag 360
 cagagaaaga tgaggcttgt gtggagggtc tattggagta tggtgcaaac cccaatgctt 420
 tggatggcaa cagagatacc ccacttccct gggcagcctt taagaacaat gctgagtgtg 480
 tgcgggctct cctagagagc ggggcctctg tcaatgccct ggattacaac aatgatacac 540
 cgctcagctg ggctgccatg aagggaaatc ttgagagtgt cagcatcctt ctggattatg 600
 ggcagaggtt cagagtcac c aacctaatag gccagacacc catctcccgc ctggtggctc 660
 tgctagtcag gggacttggg acagagaaag aggactcttg ctttgagctc ctccacagag 720
 ctggttgaca ctttgaattg aggaaaaatg gcaccatgcc acgagaggtg gccagagacc 780
 cgcagctatg tgaaaaactg actgttctgt gctcagctcc aggaactcta aaaacactcg 840
 ctccgctatg cgtgcccgt agcctgggac tccagtatct ccccgatgca gtgaagggcc 900
 ttccactgcc agcttctttg aaggaatacc tgttactttt agaatagccg gagaagatgt 960
 ttgcaccatc gtgcaggcag ctctgggtga ggttgcctc gcagtactcc ttgtcacaga 1020
 aacagaaaa acagttgttt cctgatgtgt gggttataga tttogaagca acatgtcaca 1080
 acaataacct gcatagcaac tcccctttcc aaacaaa 1117

<210> 11
 <211> 2589
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte clone 1834242

<400> 11
 cttgaatgaa gctgacacca agaaccgcgg gaagagcttg ggcccaaagc aggaaagggg 60
 agcgtctgag ttggggagga accgctgctg tggccgaact caagcccggg cgccccacc 120
 agtttgattg gaagtccagc tgtgaaacct ggagcgtcgc cttctcccca gatggctcct 180
 gttttgcttg gtctcaagga cactgcatcg tcaaactgat cccctggccg ttggaggagc 240
 agttcatccc taaagggttt gaagccaaaa gccgaagtag caaaaatgag acgaaagggc 300
 ggggcagccc aaaagagaag acgctggact gtggtcagat tgtctggggg ctggccttca 360
 gccctggccc ttccccacc agcaggaagc tctgggcacg ccaccacccc caagtggccg 420
 atgtctcttg octggttctt gctacgggac tcaacgatgg gcagatcaag atctgggagg 480
 tgcagacagg gctcctgctt ttgaatcttt ccggccacca agatgtcgtg agagatctga 540
 gcttcacacc cagtggcagt ttgatcttgg tctccgcgtc acgggataag actcttcgca 600
 tctgggacct gaataaacac ggtaaacaga ttcaagtgtt atcgggccac ctgcagtggg 660
 tttactgctg ttccatctcc ccagactgca gcattgctgt ctctgcagct ggagagaagt 720
 cggctcttct atggagcatg aggtcctaca cgtaattcg gaagctagag ggccatcaaa 780
 gcagtgttgt ctcttctgac ttctcccccg actctgccct gcttgtcacg gcttcttacg 840
 ataccaatgt gattatgtgg gacccttaca ccggcgaaag gctgaggtca ctccaccaca 900
 cccaggttga ccccgccatg gatgacagt acgtccacat tagctcactg agatctgtgt 960

```

gcttctctcc ggaaggcttg taccttgcca cggtggcaga tgacagactc ctcaggatct 1020
gggccctgga actgaaaact cccattgcat ttgctcctat gaccaatggg ctttgcctgca 1080
cattttttcc acatggtgga gtcattgcca cagggacaag agatggccac gtccagttct 1140
ggacagctcc tagggtcctg tctcactga agcacttatg ccggaaagcc cttcgaagtt 1200
tcctaacaac ttaccaagtc ctagcactgc caatcccaa gaaaatgaaa gagttcctca 1260
catacaggac tttttaagca acaccacatc ttgtgcttct ttgtagcagg gtaaactcgtc 1320
ctgtcaaagg gagttgctgg aataatgggc caaacatctg gtcttgcat gaaatagcat 1380
ttctttggga ttgtgaatag aatgtagcaa aaccagattc cagtgtacta gtcattggatc 1440
tttctctccc tggcatgtga aagtcagtct tagaggaaga gattccactt gcacggcaac 1500
agagccttac gttaaatctt cagtccagtt atgaacagca agtggtgaac tctttctgct 1560
tgttttgatt caaagtgcag ttactgatgt tgttttgatt atgcaactaa gtaggcctcc 1620
agagcctctc tagtggcaga gcagctcaca ctccctccgc tgggaacgat ggcttctgcc 1680
tagtacctat cettgtgttt ctgatgcagt ggtagcattg gttcaagttc tctcctgctg 1740
tggtcagagt tgcttcgatg ttggccaagt gcttttcttc ttgggctccc ttctgacctg 1800
caggacagtt ttctggagc catttggtat gaggtattaa tttagcttaa ctaaattaca 1860
ggggactcag aggccgtgct cctgaccgat ccagacacta ttactggctt tttttttttt 1920
tttttaacaa tgggtgtgat gtgcaggaaa tgacaaatgt gtatgtcaga ttatacaagg 1980
atgtattctt aaaccgcatg actattcaga tggctactga gttatcagtg gccatttatt 2040
agcatcatat ttatttgtat tttctcaaca gatgttaagg tacaactgtg tttttctcga 2100
ttatctaaaa accatagtac ttaaattgaa cagttgcaa gatgtcttaa ttgtgtaaag 2160
aattgggtga gtcatgact tagctgatac tcttatgtac gagatctgct tctgtggtt 2220
aacttcattg gattaatcag ctggtttcaa ctctactgctg aacacaaaat agctccttaa 2280
aagtactgtt ctcttctcagt ggcattgtat tatctaatca agacacctca ttcaaacaaa 2340
acctgcctta ggaaaattta atatatattta aattatttta aaagaaatac aacatcttat 2400
tcttttagctt tcttaatcgg tgctttatgg aggccagtgt aacgttacat gactcgttga 2460
gaaagttgag gaatttcctc taccaccttt gttgcttgaa gaaaaacatg tcttttcaaa 2520
atgagaggct ttcattgaag aaaagaaaaa aacaacagtt aaaagctaaa aaaaaaaaaa 2580
aaaaaaaaa 2589

```

```

<210> 12
<211> 2038
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte clone 1849725

```

```

<400> 12
cgggaacgcg agcccggtaa tttttcaacg gagaaaggcg aggctttcgg gctctgcaga 60
gtgagagtta gcaagtgtcc ggctccagca actctcctct ggcgtgacag ccggcatgga 120
ggatccacag agtaaagagc ctgcccgcga ggcctggct cccgcgctgc tggagtgcgc 180
gcgcccgag ggcggggag agccgcgcgcg tcccagtccc gaggaaactc aacagtgtaa 240
atthgatggc caggagacaa aaggatccaa gttcattacc tccagtgcga gtgacttcag 300
tgaccgggtt tacaagaga ttgccattac gaatggctgt attaatagaa tgagtaagga 360
agaactcaga gctaagcttt cagaattcaa gcttgaaact agaggagtaa aggatgttct 420
aaagaagaga ctgaaaaact attataagaa gcagaagctg atgctgaaag agagcaattt 480
tgctgacagt tattatgact acatttgtat tattgacttt gaagccactt gtgaagaagg 540
aaaccacct gagtttgtac atgaaataat tgaatttccg gttgttttac tgaatacgca 600
tactttagaa atagaagaca cgtttcagca gtatgtaaga ccagagatta acacacagct 660
gtctgatttc tgcacagtc taactggaat tactcaggat caggtagaca gagctgatac 720
cttccctcag gtactaaaaa aagtaattga ctggatgaaa ttgaaggaat taggaacaaa 780
gtataaatac tcacttttaa cagatggttc ttgggatatg agtaagttct tgaacattca 840
gtgtcaactc agcaggctca aataccctcc ttttgcgaaa aagtggatca atattcggaa 900
gtcatatgga aatttttaca aggttcctag aagccaaacc aaactgacaa taatgcttga 960
aaaattagga atggattatg atgggcggcc tcaactgtggt cttgatgact ctaagaatat 1020

```

```

cgcccgaata gcagttcgaa tgcttcagga tgggtgtgaa ctccgaatca acgagaaaat 1080
gcatgcagga cagctaataga gtgtgtcctc ttccttacca atagagggca ctccaccacc 1140
acaaatgccca catttttagaa agtaacagtt ttgtgtgtgg atcattccaa ttgaagttgc 1200
tatgaagagg tagcagatga atctcattga attagtcctg tagtgcaaac tttaagcacc 1260
ttaaacaatt taaaatctta ttacaggtga tagagataga tacatgtatg tgaacagatt 1320
ttgtaggaag gcatactgaa ttctttgtca ccaagcactt ttgataatgg acaggaatcc 1380
ggtaacctag ataaccaagg tcttgggtcta acacaatggg atattttaat aattttaaag 1440
aggggggttcc acaggttata aattcccttt ttttgggtgt ttaaaaaaat ggcccaaaaa 1500
tctcctaaat atggggcctt ggtgtctctc ggtttgaaa atggggccaac aaatcccttt 1560
aactcggggg tggtggttta acaataaat gggtagaat ggggtggggg tttcccttt 1620
taaatttaaa accatttcca ccttaagga ttggtaaaca cccccctaa atccccttta 1680
aaaaaattgg tcccgggaaa aattgggatt tgggggcaaa agggtaagga attcctgtaa 1740
tccctaaagg cctctctttg ggggaatfff tccccagggg gaatatatcc ccttaaggtg 1800
ccccctttt gtggaatfff tttcccaaaa aggggtttat aataaatgtt gggaaaagt 1860
ttccaccccc aaggggaaat ggggtggggg gggaaaattt tccggtaaaa gaggtgacac 1920
tttggggtag atgaccata aatacttgcg cctcaagggg gtttgcccct attttcaaaa 1980
aactccccta aaaatttggg gggaggagaa ttttatttgg attagggggg tttatata 2038

```

<210> 13

<211> 1537

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte clone 2547840

<400> 13

```

tgggtcggac gcgtgggagg aaagtgggtc agggccgggc cggcggagcg cgcagcgggg 60
cagccagatt ctttccacca tggccagacg cccccggaac agcagggcct ggcacttctg 120
cctgagtgca gcccggcag agcagatgc cggggccgtg gctctagcag gctccactaa 180
ctggggctac gactctgatg ggcagcacag cgactcggac tccgaccccg agtactccac 240
gctgccgcca tccatcccc gtgcgggtgcc cgtgaccggc gactccttct gtgactgtgc 300
tgggcagagc gaggcctcct tctgtagcag cctgcactcg gccacccggg gcagggactg 360
ccgtgcgga gaggaagacg agtatcttga ctgggtctgg gatgacttaa ataagtcatc 420
agccaccctg ctgagctgtg acaaccgtaa ggtcagcttc cacatggagt acagctgcgg 480
cacagcggcc atccggggca ccaaggagct gggggagggc cagcacttct gggagatcaa 540
gatgacctct cccgtctacg gcaccgacat gatgggtggc atcgggacgt cggatgtgga 600
cctggacaaa taccgccaca cgttctgcag cctgtcgggc agggatgagg acagctgggg 660
cctctcctac acgggcctcc tccaccacaa gggcgacaag accagcttct cgtcgcgggt 720
cggccagggc tccatcattg gcgtgcacct ggacacctgg cacgycacac tcacctttt 780
caagaacagg aagtgtatag gtgtggcagc caccaagctg cagaacaaga gattctacce 840
gatggtgtgc tccacggcgg cccggagcag catgaaggtc acccgctcct gtgccagcgc 900
cacttccctc cagtacctgt gctgccaccg cctgcgccag ctgcggccag actcgggaga 960
cacgctggag ggtctgccgc tgcccgggg cctcaagcag gtgctacaca acaagctggg 1020
ctgggtcctg agcatgagtt gcagccgccc caaggctcca gtgtccgatc cccaggcagc 1080
gacctccgcc cccccagca gtcgcgagcc tggccctgct cagaggaagc gctgccgccc 1140
gacctgactg acttcccagt ggaactgcct tcttgggctg ggacagcccc tttcctctgt 1200
cccttctttc tctgtccctt ccttccagcc aactccagg gcggagtgg atgaggcccc 1260
tccggaggga gccatctctt gctcccagg ctgggacagt cctttctgtg ggggtcttag 1320
ggcccctctg ctgctgtgct ggggtgggaa gggctgccc tgagccccag gtcttgtggg 1380
aggctgcgag gacgagagcc tggctggagc ccgcttggct gttcccacag ggccctcgggt 1440
tttccaaact tgctctgcat gctgtcagcg gctgccccgc cgtcatagac ttaaaggact 1500
gcaataaatg tagagttgat gtctaacaaa aaaaaaa 1537

```

<210> 14
 <211> 2203
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte clone 3071986

<400> 14
 ctgtcttctc cgcagcgcg aggctgggta cagggctctat tgtctgtggt tgactccgta 60
 ctttggctctg aggccttcgg gagctttccc gaggcagtta gcagaagccg cagcggccgc 120
 ccccgcccgt ttcctctgtc cctgggcccg ggagggacca acttggcgtc acgcccctca 180
 geggtcgcca ctctcttctc tgttgttggg tccgcatcgt attcccggaa tcagacgggtg 240
 ccccatagat ggccagcttt cccccgaggg tcaacgagaa agagatcgtg agattacgta 300
 ctatagtgga acttttagct cctgcagctc cttttgacaa gaaatgtggt cgtgaaaatt 360
 ggactgttgc ttttgctcca gatggttcat actttgcttg gtcacaagga catcgcacag 420
 taaacttgt tccgtgggcc cagtgccttc agaactttct cttgcatggc accaagaatg 480
 ttaccaattc aagcagttta agattgcaa gacaaaatag tgatggtggt cagaaaaata 540
 agcctcgtga acatattata gactgtggag atatagtctg gagtcttctt tttgggtcat 600
 cagttccaga aaaacagagt cgctgtgtaa atatagaatg gcatcgcttc agatttggac 660
 aagatcagct acttcttgct acaggggtga acaatgggcg tatcaaaata tgggatgtat 720
 atacaggaaa actcctcctt aacttggtag atcatactga agtggtcaga gatttaactt 780
 ttgctccaga tggaaagcttg atcctgggtg cagcttcaag agacaaaact ctcagagtat 840
 gggacctgaa agatgatgga aacatgatga aagtattgag ggggcatcag aattgggtgt 900
 acagctgtgc attctctcct gactcttcta tgctgtgttc agtcggagcc agtaaagcag 960
 ttttctttg gaatatggat aaatacacca tgatacggaa actagaagga catcaccatg 1020
 atgtggtagc ttgtgacttt tctcctgatg gagcattact ggctactgca tcttatgata 1080
 ctcgagtata tatctgggat ccacataatg gagacattct gatggaattt gggcacctgt 1140
 ttccccacc tactccaata tttgctggag gagcaaatga cgggtgggta cgatctgtat 1200
 cttttagcca tgatggactg catgttgcaa gccttgctga tgataaaatg gtgaggttct 1260
 ggagaattga tgaggattat ccagtgcaag ttgcaccttt gagcaatggt ctttgcctgtg 1320
 ccttctctac tgatggcagt gttttagctg ctgggacaca tgacggaagt gtgtatTTTT 1380
 gggccactcc acggcaggtc cctagcctgc aacatttatg tcgcatgtca atccgaagag 1440
 tgatgcccac ccaagaagtt caggagctgc cgattccttc caagcttttg gagtttctct 1500
 cgtatcgtat ttagaagatt ctgccttccc tagtagtagg gactgacaga atacacttaa 1560
 cacaaacctc aagctttact gacttcaatt atctgttttt aaagacgtag aagatttatt 1620
 taatttgata tgttcttgta ctgcattttg atcagttgag cttttaaaat attatttata 1680
 gacaatagaa gtatttctga acatatcaaa tataaatttt tttaaagatc taactgtgaa 1740
 aacatacata cctgtacata tttagatata agctgctata tgttgaatgg acccttttgc 1800
 ttttctgatt tttagtctg acatgtatat attgcttcag tagagccaca atatgtatct 1860
 ttgctgtaaa gtgcaaggaa attttaaatt ctgggacact gagttagatg gtaaatactg 1920
 acttacgaaa gttgaattgg gtgaggcggg caaatcacct gaggtcagca gtttgagact 1980
 agcctggcaa acatgatgaa accctgtctc tactaaaaat acaaaaaaaaa aaaaaattag 2040
 ccaggcgtgg tgggtcacac ctgtagtctc agctacttgg gaggctgagg caggagaatt 2100
 gcttgaacct aggaggtgga ggttgacgta agccaagatc acaccactgc actccaacct 2160
 ggacaacaga gcgagactcc atctcaaaaa aaaaataaaa agg 2203

<210> 15
 <211> 1622
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte clone 3484619

<400> 15

```

ccgcatggag ccccgggagg cggacggctg ctctcctgggc gacgtgggtt tctgggtgga 60
gcggaacccct gtgcacgagg cagcccagcg gggtgagagc ctgcagctgc aacagctgat 120
cgagagcggc gcctgctgta accaggtcac cgtggactcc atcacgcccc tgcacgcagc 180
cagtctgcag ggccaggcgc ggtgtgtgca gctgctgctg gcggctgggg cccagggtgga 240
tgctcgcaac atcgacggca gcaccccgct ctgcatgccc tgcgcctcgg gcagcatcga 300
gtgtgtgaag ctcttgctgt cctacggggc caaggtcaac cctcccctgt acacagcgtc 360
ccccctgcac gaggcctgca tgagcgggag ttccgaatgt gtgaggcttc ttattgacgt 420
cggggccaat ctggaagcgc acgattgcca ttttgggacc cctctgcacg ttgcctgtgc 480
ccgggagcat ctggactgtg tcaaagtgtc gctcaatgca ggggccaacg tgaatgcggc 540
aaagcttcat gagactgccc ttcaccacgc ggccaaggct aagaatgtg acctcatcga 600
gatgcttacc gagtttggcg gcaacatcta cgcccgggac aaccgcggga agaagccgtc 660
tgactacacg tggagcagca gcgctcccgc caagtgtctc gagtactacg aaaagacacc 720
tctgactctg tcacagctct gcagggtgaa cttaggaaag gccactggcg tccgagggct 780
ggagaagatt gccaagttaa acatcccgcc ccggctcatt gattacctct cctacaactg 840
aattgcaggt ggggtccgga ccgtgactgc ccccgttgtg cccagcattg cccgggtgag 900
ggctctgcct gttcctctga agcagcgtga ttgctgtaga tagaacaacg ctccctcgag 960
tcccttctct cgatcctgtt taggcttctc tctgggatcc tggataatgt ttccagggtg 1020
ttgggaaggc ctgctgtctc ggtcacagtt gtgggtgtgg ccctgcgctg ttctacagaa 1080
cctaccctct caatgggcat gggcccaacc atccagtttt cctcttttac ggaccatcct 1140
caaaggcact ctccaggacag acggcgtggg gagcacagag gaggctggga gagctgggga 1200
ctgagggcat tgttctgat tctcactcac cggggcagcc tgccgcagat gcacagggcc 1260
caggtgcagg ccaccacctc cgggtcggca ccaggactgc cctcgggtgt cataggggat 1320
ggctgggccc acggaaggtc ggcctgggat gtggcctggg actgctgctc tgctggctgc 1380
tgtgtggatg cttttcctgg agcactttcc aaggcatccc ccagcccaaa gcctgcgcgc 1440
atctgtcact cagggacttt ctatgggtct ttgtggggga aggcctggc tttgtattcc 1500
cacaagtagc actgagtttc ttaggaaatt tgtcttcagt attaagtctc caactcttgt 1560
aaaaagttaa tatgtaggat aaaaaccttt tagaggacac gtagggcgga ccactaagggt 1620
tt
    
```

<210> 16

<211> 1385

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte clone 1275743

<400> 16

```

cgcaacgggc gcaggtgcgg ggcgaggct ctctcacgcc ggggcctcac ccggcgggtgc 60
ttcggcaggg ggcggcgcg gggcgaggc ggcgcgcca tgggctcgca gggcagtcgg 120
gtgaagagct acgactacct gctcaagttc ctgctgggtg gcgacagcga cgtgggcaag 180
ggcgagatcc tggagagcct gcaggacggc gggcgagagt ccccgtagc ctacagtaac 240
gggatcgact acaagaccac caccatcctg ctggacggcc ggcgctgaa gctggagctc 300
tgggacacgt cgggccaggg ccggttctgc accatcttca ggtcctactc caggggctgc 360
caggggatcc tcttgggtga tgacatcacc aaccgctggg cctttgacgg catcgaccgc 420
tggatcaagg agatcgatga gcatgcaccc ggagtcccc ggatcttggg tggaaaccgg 480
ctgcacctgg ctttcaagcg gcaggtccc acggagcagg cccgcgcgta cgcagagaag 540
aactgcatga ctttcttga ggtcagcccc ctgtgcaact tcaacgtcat cgagtccttc 600
acggagctat cccgcatcgt gctcatgcgg cacggcatgg agaagatctg gaggcccaac 660
cgagtgttca gcctgcagga cctctgctgc cgggccatcg tctcctgcac ccccgtagc 720
ctcatcgaca agcttccact gcccgccacc atcaagagcc acctcaagtc cttctcgatg 780
gccaacggca tgaacgcgg catgatgcac ggccgttcc actccctggc cagcggggcc 840
ggggcgggcg gcagcaagg caacagcctc aagaggtcca agtccatccg tccaccccag 900
agccccccc agaactgctc gcggagtaac tgcaagatct cctagcgggg atgggggggg 960
    
```

```

ccgcctgtgc agatgccagg agggctcgag ctggacactc ctggctggac gccaggccag 1020
tgccgcctac gtggagactg tccacacagc tgcctcagaa gcgccgggct ttcctcacac 1080
ctgagccggg tgcgaggagg agcatgcacg gaccaagcgc ggcaggcggg aggagggggc 1140
gcggctgggc tgctggtgct tccgggaatc ttggtcgaa acaagccggg cctccccagc 1200
tgcttgggct tgaccggcgg ggagcctggt tggcctttct tatttatata gagaacactt 1260
cacttttttg tacattttta aggggccttc aggggaagcct ggggtgtggcc cgggtggtggt 1320
gcactggtga cttcatggcc acgccagctg cggggacgca cttgggactc ctcgagaggg 1380
gactc 1385

```

```

<210> 17
<211> 2790
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte clone 1722533

```

```

<400> 17
cccggaggaa cggaaggcag gattgcagct tcctcagtgc aacctccaaa caggaaatct 60
gagatgtata acatcttata gtttggcttg tcaacgttgg tcatgcggtg gccccaaaat 120
aaactccctg cttcaaagga cagcgtttca gaactgcctg gcagagcagc cagaagcttg 180
gggccagggc agaaggaaaa ctcggggagc atgttctgaa ttaagacact ttcaagaaaa 240
tcctttgtat taccctgaa ttgtaccctt gtttcagagc ctaacagggt tttctgattt 300
gctgttcct cctccccact ggggtgtctgt tctggaggcc aggggtgagag gtggaggagg 360
atggccacgc agatcagcac tcggggcagc cagtgtacca ttgggcagga ggagtacagc 420
ctgtacagca gcctgagcga ggatgaactg gtgcagatgg ccatcgagca gaggcctagc 480
gacaagacaa ggggccaac cactgctgag gccaccgct ctgcatgtac caaccgcca 540
cctgccatt tctaccatg gaccaggtcc actgcacctc ctgagagttc gccggcccgg 600
gccccaatgg gcttgttcca aggggtcatg cagaaataca gcagcagctt gttcaagacc 660
tcccagctgg cgcctgcgga ccccttgata aaggccatca aggatggcga tgaagaggcc 720
ttgaagacca tgatcaagga agggaagaat ctgcgagagc ccaacaagga gggctggctg 780
ccgctgcacg aggccgcata ctatggccag gtgggctgcc tgaagtccet gcagcgagcg 840
taccagggga ccatcgacca gcgcacctg caggaggaaa cagccgttta cttggcaacg 900
tgcaggggccc acctggactg tctcctgtca ctgctccaag caggggcaga gccggacatc 960
tccaacaaat cccgagagac accgctctac aaagcctgtg agcgcagaagaa cgcggaggcc 1020
gtgaagattc tgggtgcagca caacgcagac accaaccacc gctgcaaccg cggctggacc 1080
gctctgcacg agtctgtgtc tcgcaatgac ctggaggtca tgcagatcct ggtgagcggg 1140
ggagccaagg tggaaatcaa gaacgcctac ggcatcacc ccttgttctg ggccgcccag 1200
agtggacagt tggaggcctt gaggttctta gccaaagtac gtgctgacat caacacgcag 1260
gccagcgaca acgctgtctgc cctctacgag gcctgcaaga atgagcatga ggaggtggtg 1320
gagtttctgc tgtcacaggg tgccgagcgc aacaagacca acaaggacgg cttgctccc 1380
ctgcacateg cctccaagaa gggcaactac aggatcgtgc agatgctgct gccggtgacc 1440
agccgcacgc gcatacgccg tagcggcgtc agtccgctgc acctggcggc cgagcgcaac 1500
cacgacgagg tgctggaggc gctgctgagc gcgcgcttgc acgtgaacac gccgctggcg 1560
cccgaacgcg cgcgcctcta cgaagaccgg cgcacgtcgc cgctgtactt cgcggtggtc 1620
aacaacaacg tgtacgccac cgagctgctg ctgcaacacg gcgccgacc caaccgagc 1680
gtcatcagcc ccttgtctgt ggccatccgc cacggctgcc tgcgcacaat gcagctgctg 1740
ctggaccacg gcgcgaacat cgacgcctat atcgccaacg accccaccgc cttccccgcc 1800
accatcatgt tcgcatgaa gtgcctgtcg ctgctcaagt tcctcatgga cctgggctgc 1860
gacggcgagc cctgcttctc atgcctctac ggcaacggcc cgcacccgcc ggccccgag 1920
ccctccagca ggttcaacga cgcgcccgcg gccgacaagg agcccagcgt ggtgcaagtc 1980
tgtgagttcg tatctgcccc agaggtgagc cgtggggcgg ggcccatcat cgatgtcctc 2040
ctggactacg tgggcaacgt gcagctctgc tcgcggctga aggaacacat cgacagcttt 2100

```

```

gaggactggg ccgtcatcaa ggagaaggca gaacctccaa gacctctggc tcacctttgc 2160
cgactgcggg ttcgaaaggc cattgggaaa taccgataaa aactcctaga caccttgccg 2220
ctoccaggca ggctgattag atacctgaaa tacgagaaca cccagtaact ggggccacgg 2280
ggagagagga gtagccccctc agactcttct tactaagtct caggacgtcg gtgttcccaa 2340
ctccaagggg acctgggtgac agacgaggct gcaggctgcc tccctctcag cctggacagc 2400
taccaggatc tcaactgggtc tcagggccca gagctttggc cagagcagag aacagaatgt 2460
gtcaaggaga agaatcattt gtttacaac tgatgagcag atcccagacc ttctctacct 2520
tcaggaatgg cagaaacctc tattcctggg gccagggcag agcttgaggt gttctgggga 2580
aggtgggtgct cagagccttc cctgtgcccc tccacttggt ctggaaaact caccactga 2640
cttcagagct ttctctccaa agactaagat gaagacgtgg cccaaggtag ggggtagggg 2700
gagcctgggt cttggagggc tttgttaagt attaataaa taaatgttac acatgtgaca 2760
cctgcccagt gaaaaactaa aaaaaaaaaa 2790

```

<210> 18

<211> 2263

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte clone 1759763

<400> 18

```

cggaccgtcc gcctacgggg gccgggacgt cgctgcgcg tctctcgttt tcggacggct 60
gcagcatcgc ggtggggatc gaaagcgggg gcttctggga cgcagctctg gagacgcggc 120
ctcggaccag ccatttcggt gtagaagtgg cagcacggca gactggtcaa acaaatggat 180
tttacagagg cttacgcgga cacgtgctct acagttggac ttgctgccag ggaaggcaat 240
gttaaagtct taaggaaact gtcaaaaaag ggccgaagtg tcgatgttgc tgataacagg 300
ggatggatgc caattcatga agcagcttat cacaactctg tagaatgttt gcaaatgtta 360
attaatgcag attcatctga aactacatt aagatgaaga cttttgaagg tttctgtgct 420
ttgcatctcg ctgcaagtca aggacattgg aaaatcgtac agattctttt agaagctggg 480
gcagatccta atgcaactac tttagaagaa acgacaccat tgtttttagc tgttgaaaat 540
ggacagatag atgtgttaag gctgttgctt caacacggag caaatgttaa tggatcccat 600
tctatgtgtg gatggaactc cttgcaccag gcttcttttc aggaaaatgc tgagatcata 660
aaattgcttc ttagaaaagg agcaacaag gaatgccagg atgactttgg aatcacacct 720
ttatltgtgg ctgctcagta tggcaagcta gaaagcttga gcatacttat ttcatcaggt 780
gcaaatgtca attgtcaagc cttggacaaa tccagtcggg cagatcctga tctttactgt 840
aatgagaca gttggcagtt acctattcat gcagctgcac aaatgggcca tacaanaatc 960
ttggacttgt taataccact tactaacggg gcctgtgaca ctgggctaaa caaagtaagc 1020
cctgtttact cagcagtggt tgggggacat gaagattgcc tagaaatatt actccggaat 1080
ggctacagcc cagacgcccc ggcgtgcctt gtttttgat tcagttctcc tgtgtgcatg 1140
gctttccaaa aggactgtga gttctttgga attgtgaaca ttcttttgaa atatggagcc 1200
cagataaatg aacttcattt ggcatactgc ctgaagtacg agaagtttcc gatatttcgc 1260
tactttttga ggaaagggtg ctcatgggga ccatggaacc atatatatga atttgtaaat 1320
catgcaatta aagcacaagc aaaatataag gagtgggttg cacatcttct ggttgctgga 1380
tttgaccac tgattctact gtgcaattct tggattgact cagtcagcat tgacaccctt 1440
atcttcactt tggagtttac taattggaag acacttgcac cagctgttga aaggatgctc 1500
tctgctcgtg cctcaaacgc ttggattcta cagcaacata ttgccactgt tccatccctg 1560
accatcttt gtcgtttgga aattcggctc agtctaaaat cagaacgtct acggctctgac 1620
agttatatta gtcagctgcc acttcccaga agcctacata attattgct ctatgaagac 1680
gttctgagga tgtatgaagt tccagaactg gcagctattc aagatggata aatcagtgaa 1740
actacttaac acagctaatt ttttctctg aaaaatcacc gagacaaaag agccacagag 1800
tacaagtttt tatgatttta tagtcaaaag atgattattg attgtcagat aggttagggt 1860
ttggggggcc agtagttcag tgagaatgtt tatgtttaca actagccttc ccagtaaaaa 1920
aaaaaaaaaa aaaaaaaaaa aattgtaaac atcacttata ttactttatt gcagcttcat 1980

```

```
caccagtaca ttatatgttg taatatttat ttacctgac attttgatca tttctgctt 2040
tattttgcta ataaactgtg atgttacttc tagtgctaaa catggcatat ttccacctat 2100
gattcgtggt tacctggtat taggagctca gaatggaatg cataaagctt cactggaagt 2160
gtatacaact gtggtgtaga atctgttatt attatcatta ttattttatt tagacttgac 2220
tatctcttat gtttattaaa gaacatgttt tcctaaaaaa aaa 2263
```