

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
27 December 2001 (27.12.2001)

PCT

(10) International Publication Number
WO 01/98332 A2

(51) International Patent Classification⁷: **C07K 14/00**

YUE, Henry [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US).

(21) International Application Number: PCT/US01/19930

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

(22) International Filing Date: 22 June 2001 (22.06.2001)

(25) Filing Language: English

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

(26) Publication Language: English

(30) Priority Data:
60/213,957 22 June 2000 (22.06.2000) US

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

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

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): **LAL, Preeti** [IN/US]; P.O. Box 5142, Santa Clara, CA 95056 (US). **GANDHI, Ameena, R.** [US/US]; 837 Roble Avenue, #1, Menlo Park, CA 94025 (US). **YAO, Monique, G.** [US/US]; 111 Frederick Court, Mountain View, CA 94023 (US). **NGUYEN, Dannel, B.** [US/US]; 1403 Ridgewood Drive, San Jose, CA 94043 (US). **TANG, Y., Tom** [US/US]; 4230 Ranwick Court, San Jose, CA 95118 (US).

Published:

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

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

(54) Title: SECRETED REDOX PROTEINS

(57) Abstract: The invention provides human secreted redox proteins (SRP) and polynucleotides which identify and encode SRP. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating, or preventing disorders associated with aberrant expression of SRP.



WO 01/98332 A2

SECRETED REDOX PROTEINS

TECHNICAL FIELD

This invention relates to nucleic acid and amino acid sequences of secreted redox proteins and to the use of these sequences in the diagnosis, treatment, and prevention of cell proliferative, neurological, and metabolic disorders, and in the assessment of the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of secreted redox proteins.

BACKGROUND OF THE INVENTION

10 Overview of Antioxidant Defenses

Eukaryotic cells extract energy and synthesize macromolecules by a complex series of oxidation-reduction reactions collectively referred to as aerobic metabolism. One consequence of aerobic metabolism is the production of free radicals in the form of superoxides ($O_2^{\cdot-}$) and hydroxyl ions ($OH\cdot$). Superoxides are produced within cells by mitochondria and the endoplasmic reticulum as a consequence of "leakage" of electrons onto O_2 from their correct paths in electron transfer chains. Hydroxyl ions are produced by ionizing radiation and by the reaction of $O_2^{\cdot-}$ with hydrogen peroxide (H_2O_2) at iron- or copper-containing sites. Free radicals, especially hydroxyl ions, are extremely reactive and can interact with almost all molecules, including proteins, carbohydrates, DNA, and lipids. These interactions can lead to the formation of nonradical hydroperoxides, such as phospholipid hydroperoxides. Interaction of hydroxyl ions with DNA may be a significant contributor to the age-dependent development of cancer. Cells also use free radicals and their derivatives in beneficial ways, such as cytochrome P450-mediated oxidations, regulation of smooth muscle tone, and killing of microorganisms by macrophages and granulocytes (Bast, A. et al. (1991) Am. J. Med. 91(3C):2S-13S).

Most of the superoxides generated undergo dismutation by superoxide dismutase (SOD) generating hydrogen peroxide. H_2O_2 can also be produced by the action of other oxidase enzymes. Hydrogen peroxide resembles water in its molecular structure and is very diffusible within and between cells. Hydrogen peroxide can be removed by catalase, but the most important enzyme in human cells for removing hydrogen peroxide and hydroperoxides is glutathione peroxidase (GSHPx). GSHPx removes peroxides by using them to oxidize reduced glutathione (GSH) to oxidized glutathione (GSSG). These enzymes are discussed below.

Antioxidant defenses exist as a balanced, coordinated system. Disruption in this balance leads to clinical manifestations. For example, an excess of SOD in relation to GSHPx may be relevant to the clinical condition known as Down's Syndrome. Low antioxidant levels or high $O_2^{\cdot-}$ and H_2O_2 levels produce oxidative stress. Oxidative stress induced by phagocytes at sites of chronic

inflammation lead to rheumatoid arthritis in the joints and inflammatory bowel diseases in the intestine. Asthma is also a manifestation of an inflammatory reaction in the lung and is related to oxygen free radical formation (Sies, H. (1991) Am. J. Med. 91 (3C):31S-38S).

Glutathione Peroxidases

5 The family of glutathione peroxidases encompass three tetrameric glutathione peroxidases (GPx1-3) and the monomeric phospholipid hydroperoxide glutathione peroxidase (PHGPx/GPx4). Although the overall homology between tetrameric enzymes and GPx4 is less than 30%, a pronounced similarity has been detected in clusters involved in the active site and a common catalytic triad has been defined by structural and kinetic data. (Epp, O. et al. (1983) Eur. J. Biochem. 133:51-10 69). The family members show different tissue distributions. GPx1 is ubiquitously expressed in cells, whereas GPx2 is present in the liver and colon, and GPx3 is present in plasma. GPx4 is found at low levels in all tissues but is expressed at high level in the testis. These tissue localization patterns may be important for regulating the level and targets of glutathione peroxidase activity (Ursini, F. et al (1995) Methods Enzymol. 252:38-53).

15 GPx4 is unique in both its structure and activity. GPx4 is the only monomeric glutathione peroxidase found in mammals. It is also the only mammalian glutathione peroxidase to show high affinity for and reactivity with phospholipid hydroperoxides, and to be membrane associated. The inhibition of lipid peroxidation by GPx4 requires glutathione and physiological levels of vitamin E, suggesting a tandem mechanism for the antioxidant activities of GPx4 and vitamin E. GPx4 also has 20 alternative transcription and translation start sites which determine its subcellular localization (Esworthy, R.S. et al. (1994) Gene 144:317-318; and Maiorino, M. et al. (1990) Methods Enzymol. 186:448-450).

Glutathione S-Transferases (GST)

25 The glutathione S-transferases (GST) are a ubiquitous family of enzymes with dual substrate specificities that perform important biochemical functions of xenobiotic biotransformation and detoxification, drug metabolism, and protection of tissues against peroxidative damage. The basic reaction catalyzed by these enzymes is the conjugation of an electrophile with reduced glutathione (GSH) and results in either activation or deactivation/detoxification of the chemical. The absolute requirement for binding reduced GSH to a wide variety of chemicals necessitates a diversity in GST 30 structures in various organisms and cell types.

 GSTs are homodimeric or heterodimeric proteins localized in the cell cytosol. The major isozymes share common structural and catalytic properties and in many have been classified into four major classes, Alpha, Mu, Pi, and Theta. The two largest classes, Alpha and Mu, are identified by their respective isoelectric points; pI ~ 7.5-9.0 (Alpha), and pI ~ 6.6 (Mu). Each GST possesses a 35 common binding site for GSH and a variable hydrophobic binding site. The hydrophobic binding site

in each isozyme is specific for particular electrophilic substrates. Specific amino acid residues within GSTs have been identified as important for these binding sites and for catalytic activity. Residues Q67, T68, D101, E104, and R131 are important for the binding of GSH (Lee, H.-C. et al. (1995) J. Biol. Chem. 270:99-109). Residues R13, R20, and R69 are important for the catalytic activity of GST (Stenberg G et al. (1991) Biochem. J. 274:549-555).

While, GSTs normally perform the essential function of deactivation and detoxification of potentially mutagenic and carcinogenic chemicals, dysfunction or inappropriate expression of GSTs are detrimental. Some forms of rat and human GSTs are reliable preneoplastic markers of carcinogenesis. Expression of human GSTs in bacterial strains, such as Salmonella typhimurium, used in the well known Ames test for mutagenicity, has helped to establish the role of these enzymes in mutagenesis. Dihalomethanes, which produce liver tumors in mice, are believed to be activated by GST. This view is supported by the finding that dihalomethanes are more mutagenic in transformed bacterial cells expressing human GST than in non-transformed cells (Thier, R. et al. (1993) Proc. Natl. Acad. Sci. USA 90:8567-8580). The mutagenicity of ethylene dibromide and ethylene dichloride is increased in bacterial cells expressing the human Alpha GST, A1-1, while the mutagenicity of aflatoxin B1 is substantially reduced by enhancing the expression of GST (Simula, T.P. et al. (1993) Carcinogenesis 14:1371-1376). Thus, control of GST activity may be useful in the control of mutagenesis and carcinogenesis.

GST has been implicated in the acquired resistance of many cancers to drug treatment, the phenomenon known as multi-drug resistance (MDR). MDR occurs when a cancer patient is treated with a cytotoxic drug such as cyclophosphamide and subsequently becomes resistant to this drug and to a variety of other cytotoxic agents as well. Increased GST levels are associated with some of these drug resistant cancers, and it is believed that this increase occurs in response to the drug agent which is then deactivated by the GST catalyzed GSH conjugation reaction. The increased GST levels then protect the cancer cells from other cytotoxic agents for which GST has affinity. Increased levels of A1-1 in tumors has been linked to drug resistance induced by cyclophosphamide treatment (Dirven H.A. et al. (1994) Cancer Res. 54:6215-6220). Thus control of GST activity in cancerous tissues may be useful in treating MDR in cancer patients.

Glutaredoxin

The reduction of ribonucleotides to the corresponding deoxyribonucleotides, needed for DNA synthesis during cell proliferation, is catalyzed by the enzyme ribonucleotide diphosphate reductase. Glutaredoxin is a glutathione (GSH)-dependent hydrogen donor for ribonucleotide diphosphate reductase and contains the active site consensus sequence -C-P-Y-C-. This sequence is conserved in glutaredoxin from such different organisms as E. coli, vaccinia virus, yeast, plants, and mammalian cells. Glutaredoxin has inherent GSH-disulphide oxidoreductase (thioltransferase) activity in a

coupled system with GSH, NADPH, and GSH-reductase, catalyzing the reduction of low molecular weight disulphides as well as proteins. Glutaredoxin has been proposed to exert a general thiol redox control of protein activity by acting both as an effective protein disulphide reductase, similar to thioredoxin, and as a specific GSH-mixed disulphide reductase (Padilla, C.A. et al. (1996) FEBS Lett. 378:69-73).

In addition to their important role in DNA synthesis and cell division, glutaredoxin and other thioproteins provide effective antioxidant defense against oxygen radicals and hydrogen peroxide (Schallreuter, K.U. and J.M. Wood (1991) Melanoma Res. 1:159-167). Glutaredoxin is the principal agent responsible for protein dethiolation in vivo and reduces dehydroascorbic acid in normal human neutrophils (Jung, C.H. and J.A. Thomas (1996) Arch. Biochem. Biophys. 335:61-72; Park, J.B. and M. Levine (1996) Biochem. J. 315:931-938).

Peroxisomal β -oxidation Enzymes

Another example of the importance of redox reactions in cell metabolism is the degradation of saturated and unsaturated fatty acids by mitochondrial and peroxisomal beta-oxidation enzymes which sequentially remove two-carbon units from Coenzyme A (CoA)-activated fatty acids. The main beta-oxidation pathway degrades both saturated and unsaturated fatty acids while the auxiliary pathway performs additional steps required for the degradation of unsaturated fatty acids.

The pathways of mitochondrial and peroxisomal beta-oxidation use similar enzymes, but have different substrate specificities and functions. Mitochondria oxidize short-, medium-, and long-chain fatty acids to produce energy for cells. Mitochondrial beta-oxidation is a major energy source for cardiac and skeletal muscle. In liver, it provides ketone bodies to the peripheral circulation when glucose levels are low as in starvation, endurance exercise, and diabetes (Eaton, S. et al. (1996) Biochem. J. 320:345-357). Peroxisomes oxidize medium-, long-, and very-long-chain fatty acids, dicarboxylic fatty acids, branched fatty acids, prostaglandins, xenobiotics, and bile acid intermediates. The chief roles of peroxisomal beta-oxidation are to shorten toxic lipophilic carboxylic acids to facilitate their excretion and to shorten very-long-chain fatty acids prior to mitochondrial beta-oxidation (Mannaerts, G.P. and P.P. Van Veldhoven (1993) Biochimie 75:147-158).

The auxiliary beta-oxidation enzyme 2,4-dienoyl-CoA reductase catalyzes the following reaction:

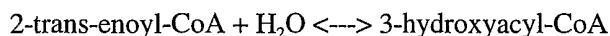


This reaction removes even-numbered double bonds from unsaturated fatty acids prior to their entry into the main beta-oxidation pathway (Koivuranta, K.T. et al. (1994) Biochem. J. 304:787-792). The enzyme may also remove odd-numbered double bonds from unsaturated fatty acids (Smeland, T.E. et

al. (1992) Proc. Natl. Acad. Sci. USA 89:6673-6677).

Rat 2,4-dienoyl-CoA reductase is located in both mitochondria and peroxisomes (Dommes, V. et al. (1981) J. Biol. Chem. 256:8259-8262). Two immunologically different forms of rat mitochondrial enzyme exist with molecular masses of 60 kDa and 120 kDa (Hakkola, E.H. and Hiltunen, J.K. (1993) Eur. J. Biochem. 215:199-204). The 120 kDa mitochondrial rat enzyme is synthesized as a 335 amino acid precursor with a 29 amino acid N-terminal leader peptide which is cleaved to form the mature enzyme (Hirose, A. et al. (1990) Biochim. Biophys. Acta 1049:346-349). A human mitochondrial enzyme 83% similar to rat enzyme is synthesized as a 335 amino acid residue precursor with a 19 amino acid N-terminal leader peptide (Koivuranta, supra). These cloned human and rat mitochondrial enzymes function as homotetramers (Koivuranta, supra). A Saccharomyces cerevisiae peroxisomal 2,4-dienoyl-CoA reductase is 295 amino acids long, contains a C-terminal peroxisomal targeting signal, and functions as a homodimer (Coe, J.G.S. et al. (1994) Mol. Gen. Genet. 244:661-672; and Gurvitz, A. et al. (1997) J. Biol. Chem. 272:22140-22147). All 2,4-dienoyl-CoA reductases have a fairly well conserved NADPH binding site motif of sequence (Koivuranta, supra).

The main pathway beta-oxidation enzyme enoyl-CoA hydratase catalyzes the following reaction:



This reaction hydrates the double bond between C-2 and C-3 of 2-trans-enoyl-CoA, which is generated from saturated and unsaturated fatty acids (Engel, C.K. et al. (1996) EMBO J. 15:5135-5145). This step is downstream from the step catalyzed by 2,4-dienoyl-reductase. Different enoyl-CoA hydratases act on short-, medium-, and long-chain fatty acids (Eaton, supra). Mitochondrial and peroxisomal enoyl-CoA hydratases occur as both mono-functional enzymes and as part of multi-functional enzyme complexes. Human liver mitochondrial short-chain enoyl-CoA hydratase is synthesized as a 290 amino acid precursor with a 29 amino acid N-terminal leader peptide (Kanazawa, M. et al. (1993) Enzyme Protein 47:9-13; and Janssen, U. et al. (1997) Genomics 40:470-475). Rat short-chain enoyl-CoA hydratase is 87% identical to the human sequence in the mature region of the protein and functions as a homohexamer (Kanazawa, supra; and Engel, supra). A mitochondrial trifunctional protein exists that has long-chain enoyl-CoA hydratase, 3-hydroxyacyl-CoA dehydrogenase, and long-chain 3-oxothiolase activities (Eaton, supra). In human peroxisomes, enoyl-CoA hydratase activity is found in both a 327 amino acid residue mono-functional enzyme and as part of a multi-functional enzyme, also known as bifunctional enzyme, which possesses enoyl-CoA hydratase, enoyl-CoA isomerase, and 3-hydroxyacyl-CoA hydroxymethylglutaryl-CoA lyase activities (FitzPatrick, D.R. et al. (1995) Genomics 27:457-466; and Hoefler, G. et al. (1994) Genomics 19:60-67). A 339 amino

acid residue human protein with short-chain enoyl-CoA hydratase activity also acts as an AU-specific RNA binding protein (Nakagawa, J. et al. (1995) Proc. Natl. Acad. Sci. USA 92:2051-2055). All enoyl-CoA hydratases share homology near two active site glutamic acid residues, with 17 amino acid residues highly conserved (Wu, W.-J. et al. (1997) Biochemistry 36:2211-2220).

5 Inherited deficiencies in mitochondrial and peroxisomal beta-oxidation enzymes are associated with severe diseases, some of which manifest themselves soon after birth and lead to death within a few years. Mitochondrial beta-oxidation associated deficiencies include, e.g., carnitine palmitoyl transferase and carnitine deficiency, very-long-chain acyl-CoA dehydrogenase deficiency, medium-chain acyl-CoA dehydrogenase deficiency, short-chain acyl-CoA dehydrogenase deficiency,
10 electron transport flavoprotein and electron transport flavoprotein:ubiquinone oxidoreductase deficiency, trifunctional protein deficiency, and short-chain 3-hydroxyacyl-CoA dehydrogenase deficiency (Eaton, supra). Mitochondrial trifunctional protein (including enoyl-CoA hydratase) deficient patients have reduced long-chain enoyl-CoA hydratase activities and suffer from non-ketotic hypoglycemia, sudden infant death syndrome, cardiomyopathy, hepatic dysfunction, and muscle
15 weakness, and may die at an early age (Eaton, supra). A patient with a deficiency in mitochondrial 2,4-dienoyl-CoA reductase was hypotonic soon after birth, had feeding difficulties, and died at four months from respiratory acidosis (Roe, C.R. et al. (1990) J. Clin. Invest. 85:1703-1707).

Defects in mitochondrial beta-oxidation are associated with Reye's syndrome, a disease characterized by hepatic dysfunction and encephalopathy that sometimes follows viral infection in
20 children. Reye's syndrome patients may have elevated serum levels of free fatty acids (Cotran, R.S. et al. (1994) Robbins Pathologic Basis of Disease, W.B. Saunders Co., Philadelphia, PA, p.866). Patients with mitochondrial short-chain 3-hydroxyacyl-CoA dehydrogenase deficiency and medium-chain 3-hydroxyacyl-CoA dehydrogenase deficiency also exhibit Reye-like illnesses (Eaton, supra; and Egidio, R.J. et al. (1989) Am. Fam. Physician 39:221-226).

25 Inherited conditions associated with peroxisomal beta-oxidation include Zellweger syndrome, neonatal adrenoleukodystrophy, infantile Refsum's disease, acyl-CoA oxidase deficiency, peroxisomal thiolase deficiency, and bifunctional protein deficiency (Suzuki, Y. et al. (1994) Am. J. Hum. Genet. 54:36-43; Hoefler, supra). Patients with peroxisomal bifunctional enzyme, including enoyl-CoA hydratase, deficiency suffer from hypotonia, seizures, psychomotor defects, and defective
30 neuronal migration; accumulate very-long-chain fatty acids; and typically die within a few years of birth (Watkins, P.A. et al. (1989) J. Clin. Invest. 83:771-777).

Peroxisomal beta-oxidation is impaired in cancerous tissue. Although neoplastic human breast epithelial cells have the same number of peroxisomes as do normal cells, fatty acyl-CoA oxidase activity is lower than in control tissue (el Bouhtoury, F., et al. (1992) J. Pathol. 166:27-35).
35 Human colon carcinomas have fewer peroxisomes than normal colon tissue and have lower fatty-

acyl-CoA oxidase and bifunctional enzyme (including enoyl-CoA hydratase) activities than normal tissue (Cable, S., et al. (1992) Virchows Arch. B Cell Pathol. Incl. Mol. Pathol. 62:221-226).

Protein Disulfide Isomerases, Thioredoxins, and Glutaredoxins

Cells contain a number of specialized molecules that assist in the formation of protein secondary and tertiary structure by orchestrating the formation of disulfide bonds. Although incubation of reduced, unfolded proteins in buffers with defined ratios of oxidized and reduced thiols can lead to native conformation, the rate of folding is slow and the attainment of native conformation decreases proportionately to the size and number of cysteines in the protein. Certain cellular compartments such as the endoplasmic reticulum of eukaryotes and the periplasmic space of prokaryotes are maintained in a more oxidized state than the surrounding cytosol. Correct disulfide formation can occur in these compartments but at a rate that is insufficient for normal cell processes and not adequate for synthesizing secreted proteins. The protein disulfide isomerases, thioredoxins and glutaredoxins are able to catalyze the formation of disulfide bonds and regulate the redox environment in cells to enable the necessary thiol:disulfide exchanges (Loferer, H. (1995) J. Biol Chem. 270:26178-26183).

Each of these proteins have somewhat different functions but all belong to a group of disulfide-containing redox proteins that contain a conserved active-site sequence and are ubiquitously distributed in eukaryotes and prokaryotes. Protein disulfide isomerases are found in the endoplasmic reticulum of eukaryotes and in the periplasmic space of prokaryotes. They function by exchanging their own disulfide for a thiol in a folding peptide chain. In contrast, the reduced thioredoxins and glutaredoxins are generally found in the cytoplasm and function by directly reducing disulfides in the substrate proteins.

These catalytic molecules not only facilitate disulfide formation but also regulate and participate in a wide variety of physiological processes. The thioredoxin system serves, for example, as a hydrogen donor for ribonucleotide reductase and as a regulator of enzymes by redox control. It also modulates the activity of transcription factors such as NF- κ B, AP-1, and steroid receptors.

Secreted Redox Proteins

Redox polypeptides are also released into the extracellular environment and may have similar or distinct functions compared to their intracellular homologs. Several cytokines or secreted cytokine-like factors such as adult T-cell leukemia-derived factor, 3B6-interleukin-1, T-hybridoma-derived (MP-6) B cell stimulatory factor, and early pregnancy factor have been reported to be identical to thioredoxin (Holmgren, A. (1985) Annu. Rev. Biochem. 54:237-271; Abate, C. et al. (1990) Science 249:1157-1161; Tagaya, Y. et al. (1989) EMBO J. 8:757-764; Wakasugi, H. (1987) Proc. Natl. Acad. Sci. USA 84:804-808; Rosen, A. et al. (1995) Int. Immunol. 7:625-633). Thus thioredoxin secreted by stimulated lymphocytes (Yodoi, J. and Tursz, T. (1991) Adv. Cancer Res.

57:381-411; Tagaya, N. et al. (1990) Proc. Natl. Acad. Sci. USA 87:8282-8286) has extracellular activities including a role as a regulator of cell growth and a mediator in the immune system (Miranda-Vizuet, A. et al. (1996) J. Biol. Chem. 271:19099-19103; Yamauchi, A. et al. (1992) Mol. Immunol. 29:263-270).

5 The selenoprotein thioredoxin reductase is secreted by both normal and neoplastic cells and has been implicated as both a growth factor and as a polypeptide involved in apoptosis (Soderberg, A. et al. (2000) Cancer Res. 60:2281-2289). An extracellular plasmin reductase secreted by hamster ovary cells (HT-1080) has been show to participate in the generation of angiostatin from plasmin. In this case, the reduction of the plasmin disulfide bonds triggers the proteolytic cleavage of plasmin
10 which yields the angiogenesis inhibitor, angiostatin (Stathakis, P. et al. (1997) J. Biol. Chem. 272:20641-20645). Low levels of reduced sulphhydryl groups in plasma has been associated with rheumatoid arthritis. The failure of these sulphhydryl groups to scavenge active oxygen species (e.g., hydrogen peroxide produced by activated neutrophils) results in oxidative damage to surrounding tissues and the resulting inflammation (Hall, N.D. et al. (1994) Rheumatol. Int. 4:35-38).

15 The discovery of new human glutathione peroxidases, thioredoxins, beta-oxidation enzymes, glutaredoxin β , and other molecules involved in maintaining the correct oxidation potential and directing critical redox reactions in cells or in extracellular spaces, 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 cell proliferative, neurological, and metabolic disorders, and in the
20 assessment of the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of secreted redox proteins.

SUMMARY OF THE INVENTION

 The invention features purified polypeptides, secreted redox proteins, referred to collectively
25 as "SRP" and individually as "SRP-1," "SRP-2," and "SRP-3." In one aspect, the invention provides an isolated polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, c) a biologically active fragment of a polypeptide
30 having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3. In one alternative, the invention provides an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:1-3.

 The invention further provides an isolated polynucleotide encoding a polypeptide selected
35 from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the

group consisting of SEQ ID NO:1-3, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, and d) an immunogenic fragment of a polypeptide
5 having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3. In one alternative, the polynucleotide encodes a polypeptide selected from the group consisting of SEQ ID NO:1-3. In another alternative, the polynucleotide is selected from the group consisting of SEQ ID NO:4-6.

Additionally, the invention provides a recombinant polynucleotide comprising a promoter
10 sequence operably linked to a polynucleotide encoding a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group
15 consisting of SEQ ID NO:1-3, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3. In one alternative, the invention provides a cell transformed with the recombinant polynucleotide. In another alternative, the invention provides a transgenic organism comprising the recombinant polynucleotide.

The invention also provides a method for producing a polypeptide selected from the group
20 consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, and d) an immunogenic fragment of a polypeptide having an amino
25 acid sequence selected from the group consisting of SEQ ID NO:1-3. The method comprises a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding the polypeptide, and b) recovering the polypeptide so expressed.

Additionally, the invention provides an isolated antibody which specifically binds to a
30 polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, and d) an
35 immunogenic fragment of a polypeptide having an amino acid sequence selected from the group

consisting of SEQ ID NO:1-3.

The invention further provides an isolated polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:4-6, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at
5 least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:4-6, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). In one alternative, the polynucleotide comprises at least 60 contiguous nucleotides.

Additionally, the invention provides a method for detecting a target polynucleotide in a
10 sample, said target polynucleotide having a sequence of a polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:4-6, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:4-6, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide
15 complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). The method comprises a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and b)
20 detecting the presence or absence of said hybridization complex, and optionally, if present, the amount thereof. In one alternative, the probe comprises at least 60 contiguous nucleotides.

The invention further provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of
25 SEQ ID NO:4-6, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:4-6, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). The method comprises a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and b)
30 detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.

The invention further provides a composition comprising an effective amount of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, b) a polypeptide comprising a
35 naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected

from the group consisting of SEQ ID NO:1-3, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, and a pharmaceutically acceptable excipient. In one embodiment, the composition comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-3. The invention additionally provides a method of treating a disease or condition associated with decreased expression of functional SRP, comprising administering to a patient in need of such treatment the composition.

The invention also provides a method for screening a compound for effectiveness as an agonist of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting agonist activity in the sample. In one alternative, the invention provides a composition comprising an agonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with decreased expression of functional SRP, comprising administering to a patient in need of such treatment the composition.

Additionally, the invention provides a method for screening a compound for effectiveness as an antagonist of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting antagonist activity in the sample. In one alternative, the invention provides a composition comprising an antagonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with overexpression of functional SRP, comprising administering to a patient in need of such treatment the composition.

The invention further provides a method of screening for a compound that specifically binds

to a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3. The method comprises a) combining the polypeptide with at least one test compound under suitable conditions, and b) detecting binding of the polypeptide to the test compound, thereby identifying a compound that specifically binds to the polypeptide.

The invention further provides a method of screening for a compound that modulates the activity of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3. The method comprises a) combining the polypeptide with at least one test compound under conditions permissive for the activity of the polypeptide, b) assessing the activity of the polypeptide in the presence of the test compound, and c) comparing the activity of the polypeptide in the presence of the test compound with the activity of the polypeptide in the absence of the test compound, wherein a change in the activity of the polypeptide in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide.

The invention further provides a method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence selected from the group consisting of SEQ ID NO:4-6, the method comprising a) exposing a sample comprising the target polynucleotide to a compound, and b) detecting altered expression of the target polynucleotide.

The invention further provides a method for assessing toxicity of a test compound, said method comprising a) treating a biological sample containing nucleic acids with the test compound; b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide selected from the group consisting of i) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:4-6, ii) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:4-6, iii) a polynucleotide having a sequence complementary to i), iv) a polynucleotide complementary to the polynucleotide of

ii), and v) an RNA equivalent of i)-iv). Hybridization occurs under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide selected from the group consisting of i) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:4-6, ii) a
5 polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:4-6, iii) a polynucleotide complementary to the polynucleotide of i), iv) a polynucleotide complementary to the polynucleotide of ii), and v) an RNA equivalent of i)-iv). Alternatively, the target polynucleotide comprises a fragment of a polynucleotide sequence selected from the group consisting of i)-v) above; c)
10 quantifying the amount of hybridization complex; and d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.

15 BRIEF DESCRIPTION OF THE TABLES

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

Table 2 shows the GenBank identification number and annotation of the nearest GenBank homolog for polypeptides of the invention. The probability score for the match between each
20 polypeptide and its GenBank homolog is also shown.

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

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

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

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

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

DESCRIPTION OF THE INVENTION

35 Before the present proteins, nucleotide sequences, and methods are described, it is understood

that this invention is not limited to the particular machines, materials and methods described, as these may vary. It is also to be understood that the terminology used herein is for the 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.

5 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.

10 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 are cited for the purpose of describing and disclosing
15 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

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

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

An “allelic variant” is an alternative form of the gene encoding SRP. Allelic variants may result from at least one mutation in the nucleic acid sequence and may result in altered mRNAs or in polypeptides whose structure or function may or may not be altered. A gene may have none, one, or many allelic variants of its naturally occurring form. Common mutational changes which give rise to
30 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 SRP include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polypeptide the same as SRP or a
35 polypeptide with at least one functional characteristic of SRP. Included within this definition are

polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding SRP, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding SRP.

The encoded protein may also be "altered," and may contain deletions, insertions, or substitutions of

5 amino acid residues which produce a silent change and result in a functionally equivalent SRP.

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 SRP is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid, and positively charged amino acids may include lysine and arginine. Amino acids with uncharged polar side chains having similar hydrophilicity values may include: asparagine and glutamine; and serine and threonine. Amino acids with uncharged side chains having similar hydrophilicity values may include: leucine, isoleucine, and valine; glycine and alanine; and phenylalanine and tyrosine.

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

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

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

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

35 The term "antigenic determinant" refers to that region of a molecule (i.e., an epitope) that

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

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

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

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

25 A "composition comprising a given polynucleotide sequence" and a "composition 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 SRP or fragments of SRP may be employed as hybridization probes. The probes may be stored in freeze-dried form and may be
30 associated with a stabilizing agent such as a carbohydrate. In hybridizations, the probe may be deployed in an aqueous solution containing salts (e.g., NaCl), detergents (e.g., sodium dodecyl sulfate; SDS), and other components (e.g., Denhardt's solution, dry milk, salmon sperm DNA, etc.).

"Consensus sequence" refers to a nucleic acid sequence which has been subjected to repeated DNA sequence analysis to resolve uncalled bases, extended using the XL-PCR kit (Applied
35 Biosystems, Foster City CA) in the 5' and/or the 3' direction, and resequenced, or which has been

assembled from one or more overlapping cDNA, EST, or genomic DNA fragments using a computer program for fragment assembly, such as the GELVIEW fragment assembly system (GCG, Madison WI) or Phrap (University of Washington, Seattle WA). Some sequences have been both extended and assembled to produce the consensus sequence.

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

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

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

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

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

one biological or immunological function of the polypeptide from which it was derived.

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

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

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

A fragment of SEQ ID NO:4-6 comprises a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:4-6, for example, as distinct from any other sequence in the genome from which the fragment was obtained. A fragment of SEQ ID NO:4-6 is useful, for example, in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:4-6 from related polynucleotide sequences. The precise length of a fragment of SEQ ID NO:4-6 and the region of SEQ ID NO:4-6 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A fragment of SEQ ID NO:1-3 is encoded by a fragment of SEQ ID NO:4-6. A fragment of SEQ ID NO:1-3 comprises a region of unique amino acid sequence that specifically identifies SEQ ID NO:1-3. For example, a fragment of SEQ ID NO:1-3 is useful as an immunogenic peptide for the development of antibodies that specifically recognize SEQ ID NO:1-3. The precise length of a fragment of SEQ ID NO:1-3 and the region of SEQ ID NO:1-3 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

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

“full length” polynucleotide sequence encodes a “full length” polypeptide sequence.

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

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

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

Alternatively, a suite of commonly used and freely available sequence comparison algorithms is provided by the National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST) (Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410), which is available from several sources, including the NCBI, Bethesda, MD, and on the Internet at <http://www.ncbi.nlm.nih.gov/BLAST/>. The BLAST software suite includes various sequence analysis programs including “blastn,” that is used to align a known polynucleotide sequence with other polynucleotide sequences from a variety of databases. Also available is a tool called “BLAST 2 Sequences” that is used for direct pairwise comparison of two nucleotide sequences. “BLAST 2 Sequences” can be accessed and used interactively at <http://www.ncbi.nlm.nih.gov/gorf/bl2.html>. The “BLAST 2 Sequences” tool can be used for both blastn and blastp (discussed below). BLAST programs are commonly used with gap and other parameters set to default settings. For example, to compare two nucleotide sequences, one may use blastn with the “BLAST 2 Sequences” tool Version 2.0.12 (April-21-2000) set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62

Reward for match: 1

Penalty for mismatch: -2

Open Gap: 5 and Extension Gap: 2 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 11

Filter: on

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

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

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

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

Alternatively the NCBI BLAST software suite may be used. For example, for a pairwise
comparison of two polypeptide sequences, one may use the “BLAST 2 Sequences” tool Version
30 2.0.12 (April-21-2000) with blastp set at default parameters. Such default parameters may be, for
example:

Matrix: BLOSUM62

Open Gap: 11 and Extension Gap: 1 penalties

Gap x drop-off: 50

35 *Expect: 10*

Word Size: 3

Filter: on

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

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

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

“Hybridization” refers to the process by which a polynucleotide strand anneals with a complementary strand through base pairing under defined hybridization conditions. Specific hybridization is an indication that two nucleic acid sequences share a high degree of complementarity. Specific hybridization complexes form under permissive annealing conditions and remain hybridized after the “washing” step(s). The washing step(s) is particularly important in determining the stringency of the hybridization process, with more stringent conditions allowing less non-specific binding, i.e., binding between pairs of nucleic acid strands that are not perfectly matched. Permissive conditions for annealing of nucleic acid sequences are routinely determinable by one of ordinary skill in the art and may be consistent among hybridization experiments, whereas wash conditions may be varied among experiments to achieve the desired stringency, and therefore hybridization specificity. Permissive annealing conditions occur, for example, at 68°C in the presence of about 6 x SSC, about 1% (w/v) SDS, and about 100 µg/ml sheared, denatured salmon sperm DNA.

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

High stringency conditions for hybridization between polynucleotides of the present invention include wash conditions of 68°C in the presence of about 0.2 x SSC and about 0.1% SDS, for 1 hour. Alternatively, temperatures of about 65°C, 60°C, 55°C, or 42°C may be used. SSC concentration may be varied from about 0.1 to 2 x SSC, with SDS being present at about 0.1%.

5 Typically, blocking reagents are used to block non-specific hybridization. Such blocking reagents include, for instance, sheared and denatured salmon sperm DNA at about 100-200 µg/ml. Organic solvent, such as formamide at a concentration of about 35-50% v/v, may also be used under particular circumstances, such as for RNA:DNA hybridizations. Useful variations on these wash conditions will be readily apparent to those of ordinary skill in the art. Hybridization, particularly under high
10 stringency conditions, may be suggestive of evolutionary similarity between the nucleotides. Such similarity is strongly indicative of a similar role for the nucleotides and their encoded polypeptides.

The term "hybridization complex" refers to a complex formed between two nucleic acid sequences by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g., C₀t or R₀t analysis) or formed between one
15 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" and "addition" refer to changes in an amino acid or nucleotide sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively.

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

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

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

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

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

The phrases "nucleic acid" and "nucleic acid sequence" refer to a nucleotide, oligonucleotide, polynucleotide, or any fragment thereof. These phrases also refer to DNA or RNA 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.

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

10 "Peptide nucleic acid" (PNA) refers to an antisense molecule or anti-gene agent which comprises an oligonucleotide of at least about 5 nucleotides in length linked to a peptide backbone 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.

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

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

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

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

Methods for preparing and using probes and primers are described in the references, for example Sambrook, J. et al. (1989) Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold
35 Spring Harbor Press, Plainview NY; Ausubel, F.M. et al. (1987) Current Protocols in Molecular

Biology, Greene Publ. Assoc. & Wiley-Intersciences, New York NY; Innis, M. et al. (1990) PCR Protocols, A Guide to Methods and Applications, Academic Press, San Diego CA. PCR primer pairs can be derived from a known sequence, for example, by using computer programs intended for that purpose such as Primer (Version 0.5, 1991, Whitehead Institute for Biomedical Research, Cambridge MA).

Oligonucleotides for use as primers are selected using software known in the art for such purpose. For example, OLIGO 4.06 software is useful for the selection of PCR primer pairs of up to 100 nucleotides each, and for the analysis of oligonucleotides and larger polynucleotides of up to 5,000 nucleotides from an input polynucleotide sequence of up to 32 kilobases. Similar primer selection programs have incorporated additional features for expanded capabilities. For example, the PrimOU primer selection program (available to the public from the Genome Center at University of Texas South West Medical Center, Dallas TX) is capable of choosing specific primers from megabase sequences and is thus useful for designing primers on a genome-wide scope. The Primer3 primer selection program (available to the public from the Whitehead Institute/MIT Center for Genome Research, Cambridge MA) allows the user to input a "mispriming library," in which sequences to avoid as primer binding sites are user-specified. Primer3 is useful, in particular, for the selection of oligonucleotides for microarrays. (The source code for the latter two primer selection programs may also be obtained from their respective sources and modified to meet the user's specific needs.) The PrimeGen program (available to the public from the UK Human Genome Mapping Project Resource Centre, Cambridge UK) designs primers based on multiple sequence alignments, thereby allowing selection of primers that hybridize to either the most conserved or least conserved regions of aligned nucleic acid sequences. Hence, this program is useful for identification of both unique and conserved oligonucleotides and polynucleotide fragments. The oligonucleotides and polynucleotide fragments identified by any of the above selection methods are useful in hybridization technologies, for example, as PCR or sequencing primers, microarray elements, or specific probes to identify fully or partially complementary polynucleotides in a sample of nucleic acids. Methods of oligonucleotide selection are not limited to those described above.

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

transform a cell.

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

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

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

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

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

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

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

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

“Substrate” refers to any suitable rigid or semi-rigid support including membranes, filters, chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, tubing, plates, polymers,

microparticles and capillaries. The substrate can have a variety of surface forms, such as wells, trenches, pins, channels and pores, to which polynucleotides or polypeptides are bound.

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

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

A "transgenic organism," as used herein, is any organism, including but not limited to
15 animals and plants, in which one or more of the cells of the organism contains heterologous nucleic acid introduced by way of human intervention, such as by transgenic techniques well known in the art. The nucleic acid is introduced into the cell, directly or indirectly by introduction into a precursor of the cell, by way of deliberate genetic manipulation, such as by microinjection or by infection with a recombinant virus. The term genetic manipulation does not include classical cross-breeding, or in
20 vitro fertilization, but rather is directed to the introduction of a recombinant DNA molecule. The transgenic organisms contemplated in accordance with the present invention include bacteria, cyanobacteria, fungi, plants and animals. The isolated DNA of the present invention can be introduced into the host by methods known in the art, for example infection, transfection, transformation or transconjugation. Techniques for transferring the DNA of the present invention
25 into such organisms are widely known and provided in references such as Sambrook et al. (1989), supra.

A "variant" of a particular nucleic acid sequence is defined as a nucleic acid sequence having at least 40% sequence identity to the particular nucleic acid sequence over a certain length of one of the nucleic acid sequences using blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-
30 1999) set at default parameters. Such a pair of nucleic acids may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% or greater sequence identity over a certain defined length. A variant may be described as, for example, an "allelic" (as defined above), "splice," "species," or "polymorphic" variant. A splice variant may have
35 significant identity to a reference molecule, but will generally have a greater or lesser number of

polynucleotides due to alternative splicing of exons during mRNA processing. The corresponding polypeptide may possess additional functional domains or lack domains that are present in the reference molecule. Species variants are polynucleotide sequences that vary from one species to another. The resulting polypeptides will generally have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. Polymorphic variants also may encompass "single nucleotide polymorphisms" (SNPs) in which the polynucleotide sequence varies by one nucleotide base. The presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

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

THE INVENTION

The invention is based on the discovery of new human secreted redox proteins (SRP), the polynucleotides encoding SRP, and the use of these compositions for the diagnosis, treatment, or prevention of cell proliferative, neurological, and metabolic disorders.

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

Table 2 shows sequences with homology to the polypeptides of the invention as identified by BLAST analysis against the GenBank protein (genpept) database. Columns 1 and 2 show the polypeptide sequence identification number (Polypeptide SEQ ID NO:) and the corresponding Incyte polypeptide sequence number (Incyte Polypeptide ID) for polypeptides of the invention. Column 3 shows the GenBank identification number (Genbank ID NO:) of the nearest GenBank homolog. Column 4 shows the probability score for the match between each polypeptide and its GenBank homolog. Column 5 shows the annotation of the GenBank homolog along with relevant citations

where applicable, all of which are expressly incorporated by reference herein.

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

Together, Tables 2 and 3 summarize the properties of polypeptides of the invention, and these properties establish that the claimed polypeptides are secreted redox proteins. For example, SEQ ID NO:2 is 54% identical to Caenorhabditis elegans glycerol-3-phosphate dehydrogenase (GenBank ID g7331888) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 1.5e-123, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. Data from BLIMPS analyses provide further corroborative evidence that SEQ ID NO:2 is a glycerol-3-phosphate dehydrogenase. SEQ ID NO:1 and SEQ ID NO:3 were analyzed and annotated in a similar manner. SEQ ID NO:1-3 all contain signal peptide motifs based on SPScan analysis. The algorithms and parameters for the analysis of SEQ ID NO:1-3 are described in Table 7.

As shown in Table 4, the full length polynucleotide sequences of the present invention were assembled using cDNA sequences or coding (exon) sequences derived from genomic DNA, or any combination of these two types of sequences. Columns 1 and 2 list the polynucleotide sequence identification number (Polynucleotide SEQ ID NO:) and the corresponding Incyte polynucleotide consensus sequence number (Incyte Polynucleotide ID) for each polynucleotide of the invention. Column 3 shows the length of each polynucleotide sequence in basepairs. Column 4 lists fragments of the polynucleotide sequences which are useful, for example, in hybridization or amplification technologies that identify SEQ ID NO:4-6 or that distinguish between SEQ ID NO:4-6 and related polynucleotide sequences. Column 5 shows identification numbers corresponding to cDNA sequences, coding sequences (exons) predicted from genomic DNA, and/or sequence assemblages comprised of both cDNA and genomic DNA. These sequences were used to assemble the full length polynucleotide sequences of the invention. Columns 6 and 7 of Table 4 show the nucleotide start (5') and stop (3') positions of the cDNA and/or genomic sequences in column 5 relative to their respective full length sequences.

The identification numbers in Column 5 of Table 4 may refer specifically, for example, to

Incyte cDNAs along with their corresponding cDNA libraries. For example, 6709408F8 is the identification number of an Incyte cDNA sequence, and HEAADIR01 is the cDNA library from which it is derived. Incyte cDNAs for which cDNA libraries are not indicated were derived from pooled cDNA libraries (e.g., 70790623V1). Alternatively, the identification numbers in column 5 may refer to GenBank cDNAs or ESTs which contributed to the assembly of the full length polynucleotide sequences.

In some cases, Incyte cDNA coverage redundant with the sequence coverage shown in column 5 was obtained to confirm the final consensus polynucleotide sequence, but the relevant Incyte cDNA identification numbers are not shown.

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

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

The invention also encompasses polynucleotides which encode SRP. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:4-6, which encodes SRP. The polynucleotide sequences of SEQ ID NO:4-6, as presented in the Sequence Listing, embrace the equivalent RNA sequences, wherein occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

The invention also encompasses a variant of a polynucleotide sequence encoding SRP. In particular, such a variant polynucleotide sequence will have at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to the polynucleotide sequence encoding SRP. A particular aspect of the invention encompasses a variant of a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:4-6 which has at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NO:4-6. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of SRP.

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 SRP, some bearing minimal

similarity to the polynucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of polynucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the
5 polynucleotide sequence of naturally occurring SRP, and all such variations are to be considered as being specifically disclosed.

Although nucleotide sequences which encode SRP and its variants are generally capable of hybridizing to the nucleotide sequence of the naturally occurring SRP under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding SRP or
10 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 SRP and its derivatives without altering the encoded amino acid sequences
15 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 SRP and SRP 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
20 reagents well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding SRP 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:4-6 and fragments thereof under various conditions of stringency. (See, e.g., Wahl, G.M. and
25 S.L. Berger (1987) *Methods Enzymol.* 152:399-407; Kimmel, A.R. (1987) *Methods Enzymol.* 152:507-511.) Hybridization conditions, including annealing and wash conditions, are described in "Definitions."

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

(Applied Biosystems). Sequencing is then carried out using either the ABI 373 or 377 DNA sequencing system (Applied Biosystems), the MEGABACE 1000 DNA sequencing system (Molecular Dynamics, Sunnyvale CA), or other systems known in the art. The resulting sequences are analyzed using a variety of algorithms which are well known in the art. (See, e.g., Ausubel, F.M. (1997) Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7; Meyers, R.A. (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856-853.)

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

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

Capillary electrophoresis systems which are commercially available may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different nucleotide-

specific, laser-stimulated fluorescent dyes, and a charge coupled device camera for detection of the emitted wavelengths. Output/light intensity may be converted to electrical signal using appropriate software (e.g., GENOTYPER and SEQUENCE NAVIGATOR, Applied Biosystems), and the entire process from loading of samples to computer analysis and electronic data display may be computer
5 controlled. Capillary electrophoresis is especially preferable for sequencing small DNA fragments which may be present in limited amounts in a particular sample.

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

The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter SRP-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
15 shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, oligonucleotide-mediated site-directed mutagenesis may be used to introduce mutations that create new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

The nucleotides of the present invention may be subjected to DNA shuffling techniques such
20 as MOLECULARBREEDING (Maxygen Inc., Santa Clara CA; described in U.S. Patent Number 5,837,458; Chang, C.-C. et al. (1999) Nat. Biotechnol. 17:793-797; Christians, F.C. et al. (1999) Nat. Biotechnol. 17:259-264; and Cramer, A. et al. (1996) Nat. Biotechnol. 14:315-319) to alter or improve the biological properties of SRP, such as its biological or enzymatic activity or its ability to bind to other molecules or compounds. DNA shuffling is a process by which a library of gene
25 variants is produced using PCR-mediated recombination of gene fragments. The library is then subjected to selection or screening procedures that identify those gene variants with the desired properties. These preferred variants may then be pooled and further subjected to recursive rounds of DNA shuffling and selection/screening. Thus, genetic diversity is created through "artificial" breeding and rapid molecular evolution. For example, fragments of a single gene containing random
30 point mutations may be recombined, screened, and then reshuffled until the desired properties are optimized. Alternatively, fragments of a given gene may be recombined with fragments of homologous genes in the same gene family, either from the same or different species, thereby maximizing the genetic diversity of multiple naturally occurring genes in a directed and controllable manner.

35 In another embodiment, sequences encoding SRP may be synthesized, in whole or in part,

using chemical methods well known in the art. (See, e.g., Caruthers, M.H. et al. (1980) Nucleic Acids Symp. Ser. 7:215-223; and Horn, T. et al. (1980) Nucleic Acids Symp. Ser. 7:225-232.)

Alternatively, SRP itself or a fragment thereof may be synthesized using chemical methods. For example, peptide synthesis can be performed using various solution-phase or solid-phase techniques.

- 5 (See, e.g., Creighton, T. (1984) Proteins, Structures and Molecular Properties, WH Freeman, New York NY, pp. 55-60; and Roberge, J.Y. et al. (1995) Science 269:202-204.) Automated synthesis may be achieved using the ABI 431A peptide synthesizer (Applied Biosystems). Additionally, the amino acid sequence of SRP, or any part thereof, may be altered during direct synthesis and/or combined with sequences from other proteins, or any part thereof, to produce a variant polypeptide or
10 a polypeptide having a sequence of a naturally occurring 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. 182:392-421.) The composition of the synthetic peptides may be confirmed by amino acid analysis or by sequencing. (See, e.g., Creighton, supra, pp. 28-53.)

- 15 In order to express a biologically active SRP, the nucleotide sequences encoding SRP or derivatives thereof may be inserted into an appropriate expression 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
20 encoding SRP. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of sequences encoding SRP. Such signals include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where sequences encoding SRP and its initiation codon and upstream regulatory sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be
25 needed. However, in 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, e.g., Scharf, D. et al. (1994) Results Probl.
30 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 SRP 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
35 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 SRP. 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. (See, e.g., Sambrook, supra; Ausubel, supra; Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509; Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945; Takamatsu, N. (1987) EMBO J. 6:307-311; The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196; Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659; and Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355.) Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted organ, tissue, or cell population. (See, e.g., Di Nicola, M. et al. (1998) Cancer Gen. Ther. 5(6):350-356; Yu, M. et al. (1993) Proc. Natl. Acad. Sci. USA 90(13):6340-6344; Buller, R.M. et al. (1985) Nature 317(6040):813-815; McGregor, D.P. et al. (1994) Mol. Immunol. 31(3):219-226; and Verma, I.M. and N. Somia (1997) Nature 389:239-242.) The invention is not limited by the host cell employed.

In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotide sequences encoding SRP. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding SRP 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 SRP 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 SRP are needed, e.g. for the production of antibodies, vectors which direct high level expression of SRP may be used. For example, vectors containing the strong, inducible SP6 or T7 bacteriophage promoter may be used.

Yeast expression systems may be used for production of SRP. A number of vectors containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH promoters, may be used in the yeast Saccharomyces cerevisiae or Pichia pastoris. In addition, such vectors direct either the secretion or intracellular retention of expressed proteins and enable

integration of foreign sequences into the host genome for stable propagation. (See, e.g., Ausubel, 1995, supra; Bitter, G.A. et al. (1987) *Methods Enzymol.* 153:516-544; and Scorer, C.A. et al. (1994) *Bio/Technology* 12:181-184.)

Plant systems may also be used for expression of SRP. Transcription of sequences encoding
5 SRP may be driven by viral promoters, e.g., the 35S and 19S promoters of CaMV used alone or in
combination with the omega leader sequence from TMV (Takamatsu, N. (1987) *EMBO J.*
6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock
promoters may be used. (See, e.g., Coruzzi, G. et al. (1984) *EMBO J.* 3:1671-1680; Broglie, R. et al.
(1984) *Science* 224:838-843; and Winter, J. et al. (1991) *Results Probl. Cell Differ.* 17:85-105.)
10 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.)

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 SRP may be ligated into an
15 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 SRP in host cells. (See, e.g., Logan, J. and T. Shenk (1984) *Proc.*
Natl. Acad. Sci. USA 81:3655-3659.) In addition, transcription enhancers, such as the Rous sarcoma
virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-
20 based vectors may also be used for high-level protein expression.

Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of
DNA than can be contained in and expressed from a plasmid. HACs of about 6 kb to 10 Mb are
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.*
25 15:345-355.)

For long term production of recombinant proteins in mammalian systems, stable expression
of SRP in cell lines is preferred. For example, sequences encoding SRP 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
30 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.

35 Any number of selection systems may be used to recover transformed cell lines. These

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

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, the presence and expression of the gene may need to be confirmed. For example, if the sequence encoding SRP is inserted within a marker gene sequence, transformed cells containing sequences encoding SRP can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding SRP 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 SRP and that express SRP may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations, PCR amplification, and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein sequences.

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

Totowa NJ.)

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

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

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

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

maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoaffinity
5 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 SRP encoding sequence and the heterologous protein sequence, so that SRP may be cleaved away from the heterologous moiety following purification. Methods for fusion protein expression and purification are discussed in Ausubel (1995, supra, ch. 10).
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 SRP may be achieved in vitro using the TNT rabbit reticulocyte lysate or wheat germ extract system (Promega). These systems couple transcription and translation of protein-coding sequences operably associated with the
15 T7, T3, or SP6 promoters. Translation takes place in the presence of a radiolabeled amino acid precursor, for example, ³⁵S-methionine.

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

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

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

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

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

In another embodiment, polynucleotides encoding SRP or their mammalian homologs may be “knocked out” in an animal model system using homologous recombination in embryonic stem (ES) cells. Such techniques are well known in the art and are useful for the generation of animal models of human disease. (See, e.g., U.S. Patent Number 5,175,383 and U.S. Patent Number 5,767,337.) For
20 example, mouse ES cells, such as the mouse 129/SvJ cell line, are derived from the early mouse embryo and grown in culture. The ES cells are transformed with a vector containing the gene of interest disrupted by a marker gene, e.g., the neomycin phosphotransferase gene (neo; Capecchi, M.R. (1989) Science 244:1288-1292). The vector integrates into the corresponding region of the host genome by homologous recombination. Alternatively, homologous recombination takes place using
25 the Cre-loxP system to knockout a gene of interest in a tissue- or developmental stage-specific manner (Marth, J.D. (1996) Clin. Invest. 97:1999-2002; Wagner, K.U. et al. (1997) Nucleic Acids Res. 25:4323-4330). Transformed ES cells are identified and microinjected into mouse cell blastocysts such as those from the C57BL/6 mouse strain. The blastocysts are surgically transferred to pseudopregnant dams, and the resulting chimeric progeny are genotyped and bred to produce
30 heterozygous or homozygous strains. Transgenic animals thus generated may be tested with potential therapeutic or toxic agents.

Polynucleotides encoding SRP may also be manipulated in vitro in ES cells derived from human blastocysts. Human ES cells have the potential to differentiate into at least eight separate cell lineages including endoderm, mesoderm, and ectodermal cell types. These cell lineages differentiate
35 into, for example, neural cells, hematopoietic lineages, and cardiomyocytes (Thomson, J.A. et al.

(1998) Science 282:1145-1147).

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

10 THERAPEUTICS

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between regions of SRP and secreted redox proteins. In addition, the expression of SRP is closely associated with diseased tissues. Therefore, SRP appears to play a role in cell proliferative, neurological, and metabolic disorders. In the treatment of disorders associated with increased SRP
15 expression or activity, it is desirable to decrease the expression or activity of SRP. In the treatment of disorders associated with decreased SRP expression or activity, it is desirable to increase the expression or activity of SRP.

Therefore, in one embodiment, SRP 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 SRP.

20 Examples of such disorders include, but are not limited to, a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone,
25 bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer’s disease, Pick’s disease, Huntington’s disease, dementia, Parkinson’s disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and
30 other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease; prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome; fatal familial insomnia, nutritional and
35 metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal

hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis;

5 inherited, metabolic, endocrine, and toxic myopathies; myasthenia gravis, periodic paralysis; mental disorders including mood, anxiety, and schizophrenic disorders; seasonal affective disorder (SAD); akathisia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, Tourette's disorder, progressive supranuclear palsy, corticobasal degeneration, and familial frontotemporal dementia; a genetic disorder such as

10 adrenoleukodystrophy, Alport's syndrome, choroideremia, Duchenne and Becker muscular dystrophy, Down's syndrome, cystic fibrosis, chronic granulomatous disease, Gaucher's disease, Huntington's chorea, Marfan's syndrome, muscular dystrophy, myotonic dystrophy, pycnodysostosis, Refsum's syndrome, retinoblastoma, sickle cell anemia, thalassemia, Werner syndrome, von Willebrand's disease, Wilms' tumor, Zellweger syndrome, peroxisomal acyl-CoA oxidase deficiency,

15 peroxisomal thiolase deficiency, peroxisomal bifunctional protein deficiency, mitochondrial carnitine palmitoyl transferase and carnitine deficiency, mitochondrial very-long-chain acyl-CoA dehydrogenase deficiency, mitochondrial medium-chain acyl-CoA dehydrogenase deficiency, mitochondrial short-chain acyl-CoA dehydrogenase deficiency, mitochondrial electron transport flavoprotein and electron transport flavoprotein:ubiquinone oxidoreductase deficiency, mitochondrial

20 trifunctional protein deficiency, and mitochondrial short-chain 3-hydroxyacyl-CoA dehydrogenase deficiency; and a smooth muscle disorder such as angina, anaphylactic shock, arrhythmias, asthma, cardiovascular shock, Cushing's syndrome, hypertension, hypoglycemia, myocardial infarction, migraine, and pheochromocytoma, and myopathies including cardiomyopathy, encephalopathy, epilepsy, Kearns-Sayre syndrome, lactic acidosis, myoclonic disorder, and ophthalmoplegia, and

25 other metabolic disorders, such as hyperammonemia, trimethylaminuria (fish-odor syndrome), 3-hydroxydicarboxylic aciduria, dicarboxylic aciduria, xanthinuria, congenital lipid adrenal hyperplasia (CLAH), albinism type III, hyperinsulinism-hyperammonemia syndrome (HHS), glutaric acidemia type I (GA-I), familial recurrent myoglobinuria, insulin resistance, hereditary thymine-uraciluria (familial pyrimidinemia), idiopathic sideroblastic anemia (AISA), neonatal

30 adrenoleukodystrophy, hypoxia, increased damage to tissues caused by trauma, radiation and ultraviolet exposure, liver dysfunction, marked obesity, methemoglobinemia (HM1, HM2, and HM3), hypertrophic hirsutism with amenorrhea, and Hermansky-Pudlack syndrome.

In another embodiment, a vector capable of expressing SRP 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 SRP including, but not limited to, those described above.

In a further embodiment, a composition comprising a substantially purified SRP 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 SRP including, but not limited to, those
5 provided above.

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

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

15 In an additional embodiment, a vector expressing the complement of the polynucleotide encoding SRP may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of SRP 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
20 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.

25 An antagonist of SRP may be produced using methods which are generally known in the art. In particular, purified SRP may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind SRP. Antibodies to SRP 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
30 produced by a Fab expression library. Neutralizing antibodies (i.e., those which inhibit dimer formation) are generally preferred for therapeutic use.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others may be immunized by injection with SRP or with any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to
35 increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral

gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are especially preferable.

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

Monoclonal antibodies to SRP 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. (1975) Nature 256:495-497; Kozbor, D. et al. (1985) J. Immunol. Methods 81:31-42; Cote, R.J. et al. (1983) Proc. Natl. Acad. Sci. USA 80:2026-2030; and Cole, S.P. et al. (1984) Mol. Cell Biol. 62:109-120.)

In addition, techniques developed for the production of "chimeric antibodies," such as the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used. (See, e.g., Morrison, S.L. et al. (1984) Proc. Natl. Acad. Sci. USA 81:6851-6855; Neuberger, M.S. et al. (1984) Nature 312:604-608; and Takeda, S. et al. (1985) Nature 314:452-454.) Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce SRP-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. USA 88:10134-10137.)

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

Antibody fragments which contain specific binding sites for SRP 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 SRP and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering SRP epitopes is generally used, 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 SRP. Affinity is expressed as an association constant, K_a , which is defined as the molar concentration of SRP-antibody complex divided by the molar concentrations of free antigen and free antibody under equilibrium conditions. The K_a determined for a preparation of polyclonal antibodies, which are heterogeneous in their affinities for multiple SRP epitopes, represents the average affinity, or avidity, of the antibodies for SRP. The K_a determined for a preparation of monoclonal antibodies, which are monospecific for a particular SRP epitope, represents a true measure of affinity. High-affinity antibody preparations with K_a ranging from about 10^9 to 10^{12} L/mole are preferred for use in immunoassays in which the SRP-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 SRP, preferably in active form, from the antibody (Catty, D. (1988) Antibodies, Volume I: A Practical Approach, IRL Press, Washington DC; Liddell, J.E. and A. Cryer (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY).

The titer and avidity of polyclonal antibody preparations may be further evaluated to determine the quality and suitability of such preparations for certain downstream applications. For example, a polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml, preferably 5-10 mg specific antibody/ml, is generally employed in procedures requiring precipitation of SRP-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.)

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

In therapeutic use, any gene delivery system suitable for introduction of the antisense sequences into appropriate target cells can be used. Antisense sequences can be delivered intracellularly in the form of an expression plasmid which, upon transcription, produces a sequence complementary to at least a portion of the cellular sequence encoding the target protein. (See, e.g., Slater, J.E. et al. (1998) *J. Allergy Clin. Immunol.* 102(3):469-475; and Scanlon, K.J. et al. (1995) 5 9(13):1288-1296.) Antisense sequences can also be introduced intracellularly through the use of viral vectors, such as retrovirus and adeno-associated virus vectors. (See, e.g., Miller, A.D. (1990) *Blood* 76:271; Ausubel, *supra*; Uckert, W. and W. Walther (1994) *Pharmacol. Ther.* 63(3):323-347.) Other gene delivery mechanisms include liposome-derived systems, artificial viral envelopes, and other 10 systems known in the art. (See, e.g., Rossi, J.J. (1995) *Br. Med. Bull.* 51(1):217-225; Boado, R.J. et al. (1998) *J. Pharm. Sci.* 87(11):1308-1315; and Morris, M.C. et al. (1997) *Nucleic Acids Res.* 25(14):2730-2736.)

In another embodiment of the invention, polynucleotides encoding SRP may be used for somatic or germline gene therapy. Gene therapy may be performed to (i) correct a genetic deficiency 15 (e.g., in the cases of severe combined immunodeficiency (SCID)-X1 disease characterized by X-linked inheritance (Cavazzana-Calvo, M. et al. (2000) *Science* 288:669-672), severe combined immunodeficiency syndrome associated with an inherited adenosine deaminase (ADA) deficiency (Blaese, R.M. et al. (1995) *Science* 270:475-480; Bordignon, C. et al. (1995) *Science* 270:470-475), cystic fibrosis (Zabner, J. et al. (1993) *Cell* 75:207-216; Crystal, R.G. et al. (1995) *Hum. Gene* 20 *Therapy* 6:643-666; Crystal, R.G. et al. (1995) *Hum. Gene Therapy* 6:667-703), thalassemias, familial hypercholesterolemia, and hemophilia resulting from Factor VIII or Factor IX deficiencies (Crystal, R.G. (1995) *Science* 270:404-410; Verma, I.M. and N. Somia (1997) *Nature* 389:239-242)), (ii) express a conditionally lethal gene product (e.g., in the case of cancers which result from unregulated cell proliferation), or (iii) express a protein which affords protection against intracellular parasites 25 (e.g., against human retroviruses, such as human immunodeficiency virus (HIV) (Baltimore, D. (1988) *Nature* 335:395-396; Poeschla, E. et al. (1996) *Proc. Natl. Acad. Sci. USA.* 93:11395-11399), hepatitis B or C virus (HBV, HCV); fungal parasites, such as Candida albicans and Paracoccidioides brasiliensis; and protozoan parasites such as Plasmodium falciparum and Trypanosoma cruzi). In the case where a genetic deficiency in SRP expression or regulation causes disease, the expression of 30 SRP from an appropriate population of transduced cells may alleviate the clinical manifestations caused by the genetic deficiency.

In a further embodiment of the invention, diseases or disorders caused by deficiencies in SRP are treated by constructing mammalian expression vectors encoding SRP and introducing these vectors by mechanical means into SRP-deficient cells. Mechanical transfer technologies for use with 35 cells in vivo or ex vitro include (i) direct DNA microinjection into individual cells, (ii) ballistic gold

particle delivery, (iii) liposome-mediated transfection, (iv) receptor-mediated gene transfer, and (v) the use of DNA transposons (Morgan, R.A. and W.F. Anderson (1993) *Annu. Rev. Biochem.* 62:191-217; Ivics, Z. (1997) *Cell* 91:501-510; Boulay, J.-L. and H. R  capon (1998) *Curr. Opin. Biotechnol.* 9:445-450).

5 Expression vectors that may be effective for the expression of SRP include, but are not limited to, the PCDNA 3.1, EPITAG, PRCCMV2, PREP, PVAX vectors (Invitrogen, Carlsbad CA), PCMV-SCRIPT, PCMV-TAG, PEGSH/PERV (Stratagene, La Jolla CA), and PTET-OFF, PTET-ON, PTRE2, PTRE2-LUC, PTK-HYG (Clontech, Palo Alto CA). SRP may be expressed using (i) a constitutively active promoter, (e.g., from cytomegalovirus (CMV), Rous sarcoma virus (RSV), SV40
10 virus, thymidine kinase (TK), or β -actin genes), (ii) an inducible promoter (e.g., the tetracycline-regulated promoter (Gossen, M. and H. Bujard (1992) *Proc. Natl. Acad. Sci. USA* 89:5547-5551; Gossen, M. et al. (1995) *Science* 268:1766-1769; Rossi, F.M.V. and H.M. Blau (1998) *Curr. Opin. Biotechnol.* 9:451-456), commercially available in the T-REX plasmid (Invitrogen)); the ecdysone-inducible promoter (available in the plasmids PVGRXR and PIND; Invitrogen); the
15 FK506/rapamycin inducible promoter; or the RU486/mifepristone inducible promoter (Rossi, F.M.V. and Blau, H.M. *supra*)), or (iii) a tissue-specific promoter or the native promoter of the endogenous gene encoding SRP from a normal individual.

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

25 In another embodiment of the invention, diseases or disorders caused by genetic defects with respect to SRP expression are treated by constructing a retrovirus vector consisting of (i) the polynucleotide encoding SRP under the control of an independent promoter or the retrovirus long terminal repeat (LTR) promoter, (ii) appropriate RNA packaging signals, and (iii) a Rev-responsive element (RRE) along with additional retrovirus *cis*-acting RNA sequences and coding sequences
30 required for efficient vector propagation. Retrovirus vectors (e.g., PFB and PFBNEO) are commercially available (Stratagene) and are based on published data (Riviere, I. et al. (1995) *Proc. Natl. Acad. Sci. USA* 92:6733-6737), incorporated by reference herein. The vector is propagated in an appropriate vector producing cell line (VPCL) that expresses an envelope gene with a tropism for receptors on the target cells or a promiscuous envelope protein such as VSVg (Armentano, D. et al.
35 (1987) *J. Virol.* 61:1647-1650; Bender, M.A. et al. (1987) *J. Virol.* 61:1639-1646; Adam, M.A. and

A.D. Miller (1988) J. Virol. 62:3802-3806; Dull, T. et al. (1998) J. Virol. 72:8463-8471; Zufferey, R. et al. (1998) J. Virol. 72:9873-9880). U.S. Patent Number 5,910,434 to Rigg ("Method for obtaining retrovirus packaging cell lines producing high transducing efficiency retroviral supernatant") discloses a method for obtaining retrovirus packaging cell lines and is hereby incorporated by
5 reference. Propagation of retrovirus vectors, transduction of a population of cells (e.g., CD4⁺ T-cells), and the return of transduced cells to a patient are procedures well known to persons skilled in the art of gene therapy and have been well documented (Ranga, U. et al. (1997) J. Virol. 71:7020-7029; Bauer, G. et al. (1997) Blood 89:2259-2267; Bonyhadi, M.L. (1997) J. Virol. 71:4707-4716; Ranga, U. et al. (1998) Proc. Natl. Acad. Sci. USA 95:1201-1206; Su, L. (1997) Blood 89:2283-
10 2290).

In the alternative, an adenovirus-based gene therapy delivery system is used to deliver polynucleotides encoding SRP to cells which have one or more genetic abnormalities with respect to the expression of SRP. The construction and packaging of adenovirus-based vectors are well known to those with ordinary skill in the art. Replication defective adenovirus vectors have proven to be
15 versatile for importing genes encoding immunoregulatory proteins into intact islets in the pancreas (Csete, M.E. et al. (1995) Transplantation 27:263-268). Potentially useful adenoviral vectors are described in U.S. Patent Number 5,707,618 to Armentano ("Adenovirus vectors for gene therapy"), hereby incorporated by reference. For adenoviral vectors, see also Antinozzi, P.A. et al. (1999) Annu. Rev. Nutr. 19:511-544 and Verma, I.M. and N. Somia (1997) Nature 18:389:239-242, both
20 incorporated by reference herein.

In another alternative, a herpes-based, gene therapy delivery system is used to deliver polynucleotides encoding SRP to target cells which have one or more genetic abnormalities with respect to the expression of SRP. The use of herpes simplex virus (HSV)-based vectors may be especially valuable for introducing SRP to cells of the central nervous system, for which HSV has a
25 tropism. The construction and packaging of herpes-based vectors are well known to those with ordinary skill in the art. A replication-competent herpes simplex virus (HSV) type 1-based vector has been used to deliver a reporter gene to the eyes of primates (Liu, X. et al. (1999) Exp. Eye Res. 169:385-395). The construction of a HSV-1 virus vector has also been disclosed in detail in U.S. Patent Number 5,804,413 to DeLuca ("Herpes simplex virus strains for gene transfer"), which is
30 hereby incorporated by reference. U.S. Patent Number 5,804,413 teaches the use of recombinant HSV d92 which consists of a genome containing at least one exogenous gene to be transferred to a cell under the control of the appropriate promoter for purposes including human gene therapy. Also taught by this patent are the construction and use of recombinant HSV strains deleted for ICP4, ICP27 and ICP22. For HSV vectors, see also Goins, W.F. et al. (1999) J. Virol. 73:519-532 and Xu, H. et al.
35 (1994) Dev. Biol. 163:152-161, hereby incorporated by reference. The manipulation of cloned

herpesvirus sequences, the generation of recombinant virus following the transfection of multiple plasmids containing different segments of the large herpesvirus genomes, the growth and propagation of herpesvirus, and the infection of cells with herpesvirus are techniques well known to those of ordinary skill in the art.

5 In another alternative, an alphavirus (positive, single-stranded RNA virus) vector is used to deliver polynucleotides encoding SRP to target cells. The biology of the prototypic alphavirus, Semliki Forest Virus (SFV), has been studied extensively and gene transfer vectors have been based on the SFV genome (Garoff, H. and K.-J. Li (1998) Curr. Opin. Biotechnol. 9:464-469). During alphavirus RNA replication, a subgenomic RNA is generated that normally encodes the viral capsid
10 proteins. This subgenomic RNA replicates to higher levels than the full length genomic RNA, resulting in the overproduction of capsid proteins relative to the viral proteins with enzymatic activity (e.g., protease and polymerase). Similarly, inserting the coding sequence for SRP into the alphavirus genome in place of the capsid-coding region results in the production of a large number of SRP-coding RNAs and the synthesis of high levels of SRP in vector transduced cells. While alphavirus
15 infection is typically associated with cell lysis within a few days, the ability to establish a persistent infection in hamster normal kidney cells (BHK-21) with a variant of Sindbis virus (SIN) indicates that the lytic replication of alphaviruses can be altered to suit the needs of the gene therapy application (Dryga, S.A. et al. (1997) Virology 228:74-83). The wide host range of alphaviruses will allow the introduction of SRP into a variety of cell types. The specific transduction of a subset of cells in a
20 population may require the sorting of cells prior to transduction. The methods of manipulating infectious cDNA clones of alphaviruses, performing alphavirus cDNA and RNA transfections, and performing alphavirus infections, are well known to those with ordinary skill in the art.

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

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

endonucleolytic cleavage of sequences encoding SRP.

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 render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

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

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

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

polynucleotide encoding SRP may be therapeutically useful.

At least one, and up to a plurality, of test compounds may be screened for effectiveness in altering expression of a specific polynucleotide. A test compound may be obtained by any method commonly known in the art, including chemical modification of a compound known to be effective in
5 altering polynucleotide expression; selection from an existing, commercially-available or proprietary library of naturally-occurring or non-natural chemical compounds; rational design of a compound based on chemical and/or structural properties of the target polynucleotide; and selection from a library of chemical compounds created combinatorially or randomly. A sample comprising a polynucleotide encoding SRP is exposed to at least one test compound thus obtained. The sample
10 may comprise, for example, an intact or permeabilized cell, or an *in vitro* cell-free or reconstituted biochemical system. Alterations in the expression of a polynucleotide encoding SRP are assayed by any method commonly known in the art. Typically, the expression of a specific nucleotide is detected by hybridization with a probe having a nucleotide sequence complementary to the sequence of the polynucleotide encoding SRP. The amount of hybridization may be quantified, thus forming
15 the basis for a comparison of the expression of the polynucleotide both with and without exposure to one or more test compounds. Detection of a change in the expression of a polynucleotide exposed to a test compound indicates that the test compound is effective in altering the expression of the polynucleotide. A screen for a compound effective in altering expression of a specific polynucleotide can be carried out, for example, using a *Schizosaccharomyces pombe* gene expression system
20 (Atkins, D. et al. (1999) U.S. Patent No. 5,932,435; Arndt, G.M. et al. (2000) Nucleic Acids Res. 28:E15) or a human cell line such as HeLa cell (Clarke, M.L. et al. (2000) Biochem. Biophys. Res. Commun. 268:8-13). A particular embodiment of the present invention involves screening a combinatorial library of oligonucleotides (such as deoxyribonucleotides, ribonucleotides, peptide nucleic acids, and modified oligonucleotides) for antisense activity against a specific polynucleotide
25 sequence (Bruice, T.W. et al. (1997) U.S. Patent No. 5,686,242; Bruice, T.W. et al. (2000) U.S. Patent No. 6,022,691).

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

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

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

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

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

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

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

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

A therapeutically effective dose refers to that amount of active ingredient, for example SRP
35 or fragments thereof, antibodies of SRP, and agonists, antagonists or inhibitors of SRP, 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 ED_{50} (the dose therapeutically effective in 50% of the population) or LD_{50} (the dose lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, which can be expressed as the LD_{50}/ED_{50} ratio. Compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the ED_{50} 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 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 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 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 SRP may be used for the diagnosis of disorders characterized by expression of SRP, or in assays to monitor patients being treated with SRP or agonists, antagonists, or inhibitors of SRP. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for SRP include methods which utilize the antibody and a label to detect SRP in human body fluids or in extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by covalent or non-covalent attachment of a reporter molecule. A wide variety of reporter molecules, several of which are described above, are known in the art and may be used.

A variety of protocols for measuring SRP, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of SRP expression. Normal or standard values for SRP expression are established by combining body fluids or cell extracts taken

from normal mammalian subjects, for example, human subjects, with antibodies to SRP under conditions suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, such as photometric means. Quantities of SRP expressed in subject, control, and disease samples from biopsied tissues are compared with the standard values. Deviation
5 between standard and subject values establishes the parameters for diagnosing disease.

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

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

Probes may also be used for the detection of related sequences, and may have at least 50%
20 sequence identity to any of the SRP 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:4-6 or from genomic sequences including promoters, enhancers, and introns of the SRP gene.

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

30 Polynucleotide sequences encoding SRP may be used for the diagnosis of disorders associated with expression of SRP. Examples of such disorders include, but are not limited to, a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including
35 adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in

particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms,

5 Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous

10 system disease; prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome; fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial

15 nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis; inherited, metabolic, endocrine, and toxic myopathies; myasthenia gravis, periodic paralysis; mental disorders including mood, anxiety, and schizophrenic disorders; seasonal affective disorder (SAD); akathisia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic

20 neuralgia, Tourette's disorder, progressive supranuclear palsy, corticobasal degeneration, and familial frontotemporal dementia; a genetic disorder such as adrenoleukodystrophy, Alport's syndrome, choroideremia, Duchenne and Becker muscular dystrophy, Down's syndrome, cystic fibrosis, chronic granulomatous disease, Gaucher's disease, Huntington's chorea, Marfan's syndrome, muscular dystrophy, myotonic dystrophy, pycnodysostosis, Refsum's syndrome, retinoblastoma, sickle cell

25 anemia, thalassemia, Werner syndrome, von Willebrand's disease, Wilms' tumor, Zellweger syndrome, peroxisomal acyl-CoA oxidase deficiency, peroxisomal thiolase deficiency, peroxisomal bifunctional protein deficiency, mitochondrial carnitine palmitoyl transferase and carnitine deficiency, mitochondrial very-long-chain acyl-CoA dehydrogenase deficiency, mitochondrial medium-chain acyl-CoA dehydrogenase deficiency, mitochondrial short-chain acyl-CoA

30 dehydrogenase deficiency, mitochondrial electron transport flavoprotein and electron transport flavoprotein:ubiquinone oxidoreductase deficiency, mitochondrial trifunctional protein deficiency, and mitochondrial short-chain 3-hydroxyacyl-CoA dehydrogenase deficiency; and a smooth muscle disorder such as angina, anaphylactic shock, arrhythmias, asthma, cardiovascular shock, Cushing's syndrome, hypertension, hypoglycemia, myocardial infarction, migraine, and pheochromocytoma,

and myopathies including cardiomyopathy, encephalopathy, epilepsy, Kearns-Sayre syndrome, lactic acidosis, myoclonic disorder, and ophthalmoplegia, and other metabolic disorders, such as hyperammonemia, trimethylaminuria (fish-odor syndrome), 3-hydroxydicarboxylic aciduria, dicarboxylic aciduria, xanthinuria, congenital lipoid adrenal hyperplasia (CLAH), albinism type III, 5 hyperinsulinism-hyperammonemia syndrome (HHS), glutaric acidemia type I (GA-I), familial recurrent myoglobinuria, insulin resistance, hereditary thymine-uraciluria (familial pyrimidinemia), idiopathic sideroblastic anemia (AISA), neonatal adrenoleukodystrophy, hypoxia, increased damage to tissues caused by trauma, radiation and ultraviolet exposure, liver dysfunction, marked obesity, methemoglobinemia (HM1, HM2, and HM3), hypertrophic hirsutism with amenorrhea, and 10 Hermansky-Pudlack syndrome. The polynucleotide sequences encoding SRP may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; in dipstick, pin, and multiformat ELISA-like assays; and in microarrays utilizing fluids or tissues from patients to detect altered SRP expression. Such qualitative or quantitative methods are well known in the art.

15 In a particular aspect, the nucleotide sequences encoding SRP may be useful in assays that detect the presence of associated disorders, particularly those mentioned above. The nucleotide sequences encoding SRP may be labeled by standard methods and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantified and compared with a 20 standard value. If the amount of signal in the patient sample is significantly altered in comparison to a control sample then the presence of altered levels of nucleotide sequences encoding SRP 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.

25 In order to provide a basis for the diagnosis of a disorder associated with expression of SRP, 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 SRP, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects 30 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, 35 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 SRP 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 SRP, or a fragment of a polynucleotide complementary to the polynucleotide encoding SRP, and will be employed under optimized conditions for identification of a specific gene or condition. Oligomers may also be employed under less stringent conditions for detection or quantification of closely related DNA or RNA sequences.

In a particular aspect, oligonucleotide primers derived from the polynucleotide sequences encoding SRP may be used to detect single nucleotide polymorphisms (SNPs). SNPs are substitutions, insertions and deletions that are a frequent cause of inherited or acquired genetic disease in humans. Methods of SNP detection include, but are not limited to, single-stranded conformation polymorphism (SSCP) and fluorescent SSCP (fSSCP) methods. In SSCP, oligonucleotide primers derived from the polynucleotide sequences encoding SRP are used to amplify DNA using the polymerase chain reaction (PCR). The DNA may be derived, for example, from diseased or normal tissue, biopsy samples, bodily fluids, and the like. SNPs in the DNA cause differences in the secondary and tertiary structures of PCR products in single-stranded form, and these differences are detectable using gel electrophoresis in non-denaturing gels. In fSSCP, the oligonucleotide primers are fluorescently labeled, which allows detection of the amplimers in high-throughput equipment such as DNA sequencing machines. Additionally, sequence database analysis methods, termed *in silico* SNP (isSNP), are capable of identifying polymorphisms by comparing the sequence of individual overlapping DNA fragments which assemble into a common consensus sequence. These computer-based methods filter out sequence variations due to laboratory preparation of DNA and sequencing errors using statistical models and automated analyses of DNA sequence chromatograms. In the alternative, SNPs may be detected and characterized by mass spectrometry using, for example, the high throughput MASSARRAY system (Sequenom, Inc., San Diego CA).

Methods which may also be used to quantify the expression of SRP include radiolabeling or

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

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

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

A particular embodiment relates to the use of the polynucleotides of the present invention to generate a transcript image of a tissue or cell type. A transcript image represents the global pattern of gene expression by a particular tissue or cell type. Global gene expression patterns are analyzed by quantifying the number of expressed genes and their relative abundance under given conditions and at a given time. (See Seilhamer et al., "Comparative Gene Transcript Analysis," U.S. Patent Number 5,840,484, expressly incorporated by reference herein.) Thus a transcript image may be generated by hybridizing the polynucleotides of the present invention or their complements to the totality of transcripts or reverse transcripts of a particular tissue or cell type. In one embodiment, the hybridization takes place in high-throughput format, wherein the polynucleotides of the present invention or their complements comprise a subset of a plurality of elements on a microarray. The resultant transcript image would provide a profile of gene activity.

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

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

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

Another particular embodiment relates to the use of the polypeptide sequences of the present invention to analyze the proteome of a tissue or cell type. The term proteome refers to the global
30 pattern of protein expression in a particular tissue or cell type. Each protein component of a proteome can be subjected individually to further analysis. Proteome expression patterns, or profiles, are analyzed by quantifying the number of expressed proteins and their relative abundance under given conditions and at a given time. A profile of a cell's proteome may thus be generated by separating and analyzing the polypeptides of a particular tissue or cell type. In one embodiment, the
35 separation is achieved using two-dimensional gel electrophoresis, in which proteins from a sample are

separated by isoelectric focusing in the first dimension, and then according to molecular weight by sodium dodecyl sulfate slab gel electrophoresis in the second dimension (Steiner and Anderson, supra). The proteins are visualized in the gel as discrete and uniquely positioned spots, typically by staining the gel with an agent such as Coomassie Blue or silver or fluorescent stains. The optical density of each protein spot is generally proportional to the level of the protein in the sample. The optical densities of equivalently positioned protein spots from different samples, for example, from biological samples either treated or untreated with a test compound or therapeutic agent, are compared to identify any changes in protein spot density related to the treatment. The proteins in the spots are partially sequenced using, for example, standard methods employing chemical or enzymatic cleavage followed by mass spectrometry. The identity of the protein in a spot may be determined by comparing its partial sequence, preferably of at least 5 contiguous amino acid residues, to the polypeptide sequences of the present invention. In some cases, further sequence data may be obtained for definitive protein identification.

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

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

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

the amino acid residues of the individual proteins and comparing these partial sequences to the polypeptides of the present invention.

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

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

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

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

region of DNA associated with that disorder and thus may further positional cloning efforts.

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

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

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

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

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

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

in any way whatsoever.

The disclosures of all patents, applications and publications, mentioned above and below, including U.S. Ser. No. 60/213,957, are expressly incorporated by reference herein.

5

EXAMPLES

I. Construction of cDNA Libraries

Incyte cDNAs were derived from cDNA libraries described in the LIFESEQ GOLD database (Incyte Genomics, Palo Alto CA) and shown in Table 4, column 5. Some tissues were homogenized and lysed in guanidinium isothiocyanate, while others were homogenized and lysed in phenol or in a
10 suitable mixture of denaturants, such as TRIZOL (Life Technologies), a monophasic solution of phenol and guanidine isothiocyanate. The resulting lysates were centrifuged over CsCl cushions or extracted with chloroform. RNA was precipitated from the lysates with either isopropanol or sodium acetate and ethanol, or by other routine methods.

Phenol extraction and precipitation of RNA were repeated as necessary to increase RNA
15 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, Austin TX).

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA
20 libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP vector system (Stratagene) or SUPERScript plasmid system (Life Technologies), using the recommended procedures or similar methods known in the art. (See, e.g., Ausubel, 1997, supra, units 5.1-6.6.) Reverse transcription was initiated using oligo d(T) or random primers. Synthetic
25 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 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.,
30 PBLUESCRIPT plasmid (Stratagene), PSPORT1 plasmid (Life Technologies), PCDNA2.1 plasmid (Invitrogen, Carlsbad CA), PBK-CMV plasmid (Stratagene), or pINCY (Incyte Genomics, Palo Alto CA), or derivatives thereof. Recombinant plasmids were transformed into competent *E. coli* cells including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene or DH5 α , DH10B, or ElectroMAX DH10B from Life Technologies.

35 II. Isolation of cDNA Clones

Plasmids obtained as described in Example I were recovered from host cells by in vivo excision using the UNIZAP vector system (Stratagene) or by cell lysis. Plasmids were purified using at least one of the following: a Magic or WIZARD Minipreps DNA purification system (Promega); an AGTC Miniprep purification kit (Edge Biosystems, Gaithersburg MD); and QIAWELL 8 Plasmid, 5 QIAWELL 8 Plus Plasmid, QIAWELL 8 Ultra Plasmid purification systems or the R.E.A.L. PREP 96 plasmid purification kit from QIAGEN. Following precipitation, plasmids were resuspended in 0.1 ml of distilled water and stored, with or without lyophilization, at 4°C.

Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR in a high-throughput format (Rao, V.B. (1994) Anal. Biochem. 216:1-14). Host cell lysis and thermal 10 cycling steps were carried out in a single reaction mixture. Samples were processed and stored in 384-well plates, and the concentration of amplified plasmid DNA was quantified fluorometrically using PICOGREEN dye (Molecular Probes, Eugene OR) and a FLUOROSKAN II fluorescence scanner (Labsystems Oy, Helsinki, Finland).

III. Sequencing and Analysis

15 Incyte cDNA recovered in plasmids as described in Example II were sequenced as follows. Sequencing reactions were processed using standard methods or high-throughput instrumentation such as the ABI CATALYST 800 (Applied Biosystems) thermal cycler or the PTC-200 thermal cycler (MJ Research) in conjunction with the HYDRA microdispenser (Robbins Scientific) or the MICROLAB 2200 (Hamilton) liquid transfer system. cDNA sequencing reactions were prepared 20 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 (Applied Biosystems). Electrophoretic separation of cDNA sequencing reactions and detection of labeled polynucleotides were carried out using the MEGABACE 1000 DNA sequencing system (Molecular Dynamics); the ABI PRISM 373 or 377 sequencing system (Applied Biosystems) in conjunction with standard ABI 25 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 VIII.

The polynucleotide sequences derived from Incyte cDNAs were validated by removing 30 vector, linker, and poly(A) sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programming, and dinucleotide nearest neighbor analysis. The Incyte cDNA sequences or translations thereof were then queried against a selection of public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS, PRINTS, DOMO, PRODOM, and hidden Markov model (HMM)-based protein family 35 databases such as PFAM. (HMM is a probabilistic approach which analyzes consensus primary

structures of gene families. See, for example, Eddy, S.R. (1996) *Curr. Opin. Struct. Biol.* 6:361-365.) The queries were performed using programs based on BLAST, FASTA, BLIMPS, and HMMER. The Incyte cDNA sequences were assembled to produce full length polynucleotide sequences.

Alternatively, GenBank cDNAs, GenBank ESTs, stitched sequences, stretched sequences, or

5 Genscan-predicted coding sequences (see Examples IV and V) were used to extend Incyte cDNA assemblages to full length. Assembly was performed using programs based on Phred, Phrap, and Consed, and cDNA assemblages were screened for open reading frames using programs based on GeneMark, BLAST, and FASTA. The full length polynucleotide sequences were translated to derive the corresponding full length polypeptide sequences. Alternatively, a polypeptide of the invention
10 may begin at any of the methionine residues of the full length translated polypeptide. Full length polypeptide sequences were subsequently analyzed by querying against databases such as the GenBank protein databases (genpept), SwissProt, BLOCKS, PRINTS, DOMO, PRODOM, Prosite, and hidden Markov model (HMM)-based protein family databases such as PFAM. Full length polynucleotide sequences are also analyzed using MACDNASIS PRO software (Hitachi Software
15 Engineering, South San Francisco CA) and LASERGENE software (DNASTAR). Polynucleotide and polypeptide sequence alignments are generated using default parameters specified by the CLUSTAL algorithm as incorporated into the MEGALIGN multisequence alignment program (DNASTAR), which also calculates the percent identity between aligned sequences.

Table 7 summarizes the tools, programs, and algorithms used for the analysis and assembly of
20 Incyte cDNA and full length sequences and provides applicable descriptions, references, and threshold parameters. The first column of Table 7 shows the tools, programs, and algorithms used, the second column provides brief descriptions thereof, the third column presents appropriate 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
25 strength of a match between two sequences (the higher the score or the lower the probability value, the greater the identity between two sequences).

The programs described above for the assembly and analysis of full length polynucleotide and polypeptide sequences were also used to identify polynucleotide sequence fragments from SEQ
ID NO:4-6. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and
30 amplification technologies are described in Table 4, column 4.

IV. Identification and Editing of Coding Sequences from Genomic DNA

Putative secreted redox proteins were initially identified by running the Genscan gene identification program against public genomic sequence databases (e.g., gbpr1 and gbhtg). Genscan is a general-purpose gene identification program which analyzes genomic DNA sequences from a
35 variety of organisms (See Burge, C. and S. Karlin (1997) *J. Mol. Biol.* 268:78-94, and Burge, C. and

S. Karlin (1998) Curr. Opin. Struct. Biol. 8:346-354). The program concatenates predicted exons to form an assembled cDNA sequence extending from a methionine to a stop codon. The output of Genscan is a FASTA database of polynucleotide and polypeptide sequences. The maximum range of sequence for Genscan to analyze at once was set to 30 kb. To determine which of these Genscan
5 predicted cDNA sequences encode secreted redox proteins, the encoded polypeptides were analyzed by querying against PFAM models for secreted redox proteins. Potential secreted redox proteins were also identified by homology to Incyte cDNA sequences that had been annotated as secreted redox proteins. These selected Genscan-predicted sequences were then compared by BLAST analysis to the genpept and gbprl public databases. Where necessary, the Genscan-predicted sequences were
10 then edited by comparison to the top BLAST hit from genpept to correct errors in the sequence predicted by Genscan, such as extra or omitted exons. BLAST analysis was also used to find any Incyte cDNA or public cDNA coverage of the Genscan-predicted sequences, thus providing evidence for transcription. When Incyte cDNA coverage was available, this information was used to correct or confirm the Genscan predicted sequence. Full length polynucleotide sequences were obtained by
15 assembling Genscan-predicted coding sequences with Incyte cDNA sequences and/or public cDNA sequences using the assembly process described in Example III. Alternatively, full length polynucleotide sequences were derived entirely from edited or unedited Genscan-predicted coding sequences.

V. Assembly of Genomic Sequence Data with cDNA Sequence Data

20 "Stitched" Sequences

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

type (cDNA to genomic sequence). The resultant stitched sequences were translated and compared by BLAST analysis to the genpept and gbpr public databases. Incorrect exons predicted by Genscan were corrected by comparison to the top BLAST hit from genpept. Sequences were further extended with additional cDNA sequences, or by inspection of genomic DNA, when necessary.

5 **“Stretched” Sequences**

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

VI. Chromosomal Mapping of SRP Encoding Polynucleotides

The sequences which were used to assemble SEQ ID NO:4-6 were compared with sequences
20 from the Incyte LIFESEQ database and public domain databases using BLAST and other implementations of the Smith-Waterman algorithm. Sequences from these databases that matched SEQ ID NO:4-6 were assembled into clusters of contiguous and overlapping sequences using assembly algorithms such as Phrap (Table 7). Radiation hybrid and genetic mapping data available from public resources such as the Stanford Human Genome Center (SHGC), Whitehead Institute for
25 Genome Research (WIGR), and Généthon were used to determine if any of the clustered sequences had been previously mapped. Inclusion of a mapped sequence in a cluster resulted in the assignment of all sequences of that cluster, including its particular SEQ ID NO:, to that map location.

Map locations are represented by ranges, or intervals, of human chromosomes. The map position of an interval, in centiMorgans, is measured relative to the terminus of the chromosome's p-
30 arm. (The centiMorgan (cM) is a unit of measurement based on recombination frequencies between chromosomal markers. On average, 1 cM is roughly equivalent to 1 megabase (Mb) of DNA in humans, although this can vary widely due to hot and cold spots of recombination.) The cM distances are based on genetic markers mapped by Généthon which provide boundaries for radiation hybrid markers whose sequences were included in each of the clusters. Human genome maps and
35 other resources available to the public, such as the NCBI "GeneMap'99" World Wide Web site

(<http://www.ncbi.nlm.nih.gov/genemap/>), can be employed to determine if previously identified disease genes map within or in proximity to the intervals indicated above.

VII. Analysis of Polynucleotide Expression

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

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

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

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

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

germ cells; hemic and immune system; liver; musculoskeletal system; nervous system; pancreas; respiratory system; sense organs; skin; stomatognathic system; unclassified/mixed; or urinary tract.

The number of libraries in each category is counted and divided by the total number of libraries across all categories. Similarly, each human tissue is classified into one of the following

5 disease/condition categories: cancer, cell line, developmental, inflammation, neurological, trauma, cardiovascular, pooled, and other, and the number of libraries in each category is counted and divided by the total number of libraries across all categories. The resulting percentages reflect the tissue- and disease-specific expression of cDNA encoding SRP. cDNA sequences and cDNA library/tissue information are found in the LIFESEQ GOLD database (Incyte Genomics, Palo Alto CA).

10 **VIII. Extension of SRP Encoding Polynucleotides**

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

Selected human cDNA libraries were used to extend the sequence. If more than one
20 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 2-mercaptoethanol, Taq DNA polymerase (Amersham Pharmacia Biotech), ELONGASE enzyme
25 (Life Technologies), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair PCI A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the alternative, the parameters for primer pair T7 and SK+ were as follows: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times;
30 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 fluorimeter plate (Corning Costar, Acton MA), allowing the DNA to bind to the reagent. The plate was scanned in a Fluoroskan II
35 (Labsystems Oy, Helsinki, Finland) to measure the fluorescence of the sample and to quantify the

concentration of DNA. A 5 μ l to 10 μ l aliquot of the reaction mixture was analyzed by electrophoresis on a 1 % agarose gel to determine which reactions were successful in extending the sequence.

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

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

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

IX. Labeling and Use of Individual Hybridization Probes

Hybridization probes derived from SEQ ID NO:4-6 are employed to screen 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 (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 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 conditions of up to, for example, 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate. Hybridization patterns are visualized using autoradiography or an alternative imaging means and compared.

X. Microarrays

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

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

Tissue or Cell Sample Preparation

Total RNA is isolated from tissue samples using the guanidinium thiocyanate method and poly(A)⁺ RNA is purified using the oligo-(dT) cellulose method. Each poly(A)⁺ RNA sample is reverse transcribed using MMLV reverse-transcriptase, 0.05 pg/μl oligo-(dT) primer (21mer), 1X

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

Microarray Preparation

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

20 Purified array elements are immobilized on polymer-coated glass slides. Glass microscope slides (Corning) are cleaned by ultrasound in 0.1% SDS and acetone, with extensive distilled water washes between and after treatments. Glass slides are etched in 4% hydrofluoric acid (VWR Scientific Products Corporation (VWR), West Chester PA), washed extensively in distilled water, and coated with 0.05% aminopropyl silane (Sigma) in 95% ethanol. Coated slides are cured in a
25 110°C oven.

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

30 Microarrays are UV-crosslinked using a STRATALINKER UV-crosslinker (Stratagene). Microarrays are washed at room temperature once in 0.2% SDS and three times in distilled water. Non-specific binding sites are blocked by incubation of microarrays in 0.2% casein in phosphate buffered saline (PBS) (Tropix, Inc., Bedford MA) for 30 minutes at 60°C followed by washes in 0.2% SDS and distilled water as before.

Hybridization

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

Detection

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

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

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

The output of the photomultiplier tube is digitized using a 12-bit RTI-835H analog-to-digital

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

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

XI. Complementary Polynucleotides

Sequences complementary to the SRP-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring SRP. 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 SRP. 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 SRP-encoding transcript.

XII. Expression of SRP

Expression and purification of SRP is achieved using bacterial or virus-based expression systems. For expression of SRP 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 SRP upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of SRP 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 SRP 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, SRP 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-
 5 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 SRP at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity
 10 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 SRP obtained by these methods can be used directly in the assays shown in Examples XVI, XVII, and XVIII, where applicable.

XIII. Functional Assays

15 SRP function is assessed by expressing the sequences encoding SRP at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include PCMV SPORT (Life Technologies) and PCR3.1 (Invitrogen, Carlsbad CA), both of which contain the cytomegalovirus promoter. 5-10 μ g of recombinant vector are transiently transfected into
 20 a human cell line, for example, an endothelial or hematopoietic cell line, using either liposome formulations or electroporation. 1-2 μ g of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP;
 25 Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP and to evaluate the apoptotic state of the cells and other cellular properties. FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide;
 30 changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are
 35 discussed in Ormerod, M.G. (1994) Flow Cytometry, Oxford, New York NY.

The influence of SRP on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding SRP 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 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 SRP and other genes of interest can be analyzed by northern analysis or microarray techniques.

XIV. Production of SRP Specific Antibodies

SRP 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.

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

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

XV. Purification of Naturally Occurring SRP Using Specific Antibodies

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

Media containing SRP are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of SRP (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/SRP 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 SRP is collected.

XVI. Identification of Molecules Which Interact with SRP

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

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

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

XVII. Demonstration of SRP Activity

SRP activity is determined by measuring the enzyme catalyzed conjugation of GSH with 1-chloro-2,4-dinitrobenzene (CDNB), a common substrate for most GSTs. SRP is incubated with 1 mM CDNB and 2.5 mM GSH together in 0.1M potassium phosphate buffer, pH 6.5, at 25 °C. The
20 conjugation reaction is measured by the change in absorbance at 340 nm using an ultraviolet spectrophotometer. Enzyme activity is proportional to the change in absorbance at 340 nm.

In the alternative, SRP activity is measured based upon the increase in extinction coefficient of NADH coenzyme at 340 nm for the measurement of oxidation activity, or the decrease in extinction coefficient of NADH coenzyme at 340 nm for the measurement of reduction activity
25 (Dalziel, K. (1963) *J. Biol. Chem.* 238:2850-2858). One of the following substrates may be used: 3-phosphoglycerate, hydroxypyruvate, ethanol, acetaldehyde, glycine, or hypotaurine. The reaction mixture contains 1-2 mg/ml (OR), 15 mM substrate, 2.4 mM NAD^+ or 2.0 mM NADH, in 0.1 M phosphate buffer, pH 7.1 (NAD^+ , oxidation reaction) or 0.1 M Na_2HPO_4 buffer, pH 7.4 (NADH, reduction reaction) in a total volume of 0.1 ml. Changes in absorbance at 340 nm (A_{340}) are measured
30 at ambient temperature using a recording spectrophotometer (Shimadzu Scientific Instruments, Inc., Pleasanton CA). The amount of NADH is stoichiometrically equivalent to the amount of substrate initially present, and the change in A_{340} is a direct measure of the amount of NADH produced; $\Delta A_{340} = 6620[\text{NADH}]$.

XVIII. Plasmapheresis of a Patient's Blood to Remove Toxins

SRP is used to purify a patient's blood after exposure to a chemical toxin. In this case, rather than administering a pharmaceutical composition of SRP to the patient, the toxin is removed from the blood by plasmapheresis in a manner similar to that used for kidney dialysis. SRP is stabilized in a mini-filtration column and the patient's plasma is passed through the column where the toxin is
5 removed by binding to SRP before the plasma is returned to the patient.

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 certain embodiments, it
10 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 1

Incyte Project ID	Polypeptide SEQ ID NO:	Incyte Polypeptide ID	Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID
2457018	1	2457018CD1	4	2457018CB1
4907793	2	4907793CD1	5	4907793CB1
888992	3	888992CD1	6	888992CB1

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO:	Probability Score	GenBank Homolog
2	4907793CD1	g7331888	1.50E-123	Contains similarity to Mycoplasma genitalium glycerol-3-phosphate dehydrogenase (SW:P47285) [Caenorhabditis elegans] The C. elegans Sequencing Consortium (1998) Science 282:2012-2018
3	888992CD1	g6018392	5.40E-74	Similar to Thioredoxins; cDNA EST yk491h3.5 comes from this gene [Caenorhabditis elegans] The C. elegans Sequencing Consortium (1998) Science 282:2012-2018
		g12654285	0	Similar to peptide N-glycanase homolog (S.cerevisiae) [Homo sapiens]

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
1	2457018CD1	71	T39 S43		Histidinol dehydrogenase signature: A14-F70 Signal cleavage: M1-G37	ProfileScan
2	4907793CD1	463	S31 S47 S63 S107 S130 T215 Y238 T245 S263 T316 S344 T381 T393	N364	FAD-DEPENDENT GLYCEROL-3-PHOSPHATE DEHYDROGENASE DM08855 P13032 6-270: T46-L288 CONSERVED GLYCEROL3PHOSPHATE DEHYDROGENASE GLPA ALPAGABP TRANSMEMBRANE PD011272: F49-V161 Similarity to ganglion mother cells (GMC) oxidoreductases: BL00623A: D50-I68; Q253-T273; V303-L325 UbiH/COQ6 monooxygenase protein: BL01304A: D50-S63 Flavin-containing rat monoamine oxidase: PR00757A: D50-L69, V96-Q115 DIMETHYLANILINE MONOOXYGENASE [N-OXIDE FORMING] 3 (HEPATIC FLAVIN-CONTAINING MONO-OXYGENASE 3; FMO 3) (DIMETHYL-ANILINE OXIDASE 3): PR00370A: D50-R65 FAD-dependent glycerol-3-phosphate dehydrogenase: PR01001A: F49-L61, K81-N93 Peptide reductase protein: PD01888A: A375-P390 RGD (cell surface receptor domain): R282-D284 Signal cleavage: M1-A64	SPScan BLAST-DOMO BLAST-PRODOR BLIMPS-BLOCKS BLIMPS-PRINTS BLIMPS-PRINTS BLIMPS-PRINTS BLIMPS-PRINTS BLIMPS-PRODOR MOTIFS SPScan

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
3	888992CD1	654	T27 T37 Y51 S53 S62 S94 S110 S111 S116 S119 S224 S267 Y282 S289 S380 S515 S602 S619 S636 S207 S374 T388 S392 S446 T449 T479 S552 S568	N117 N222 N511	Transglutaminase-like superfamily: T304-C355 RGD (cell surface receptor domain): R620-D622 Signal cleavage: M1-A13	HMMER-PFAM MOTIFS SPScan

Table 4

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragments	Sequence Fragments	5', Position	3', Position
4	2457018CB1	2305	1-131, 2170-2305, 1082-1143	6709408F8 (HEAADIR01)	1533	2286
				7260286H1 (UTRETM01)	1	494
				1315086F6 (BLADTUT02)	878	1473
				5742422R9 (LUNGNON03)	1094	1617
				2457018T6 (ENDANOT01)	449	1066
				6312338F6 (NERDTDN03)	1654	2305
				2330733T6 (COLNNOT11)	488	1075
5	4907793CB1	1742	872-1018	70790623V1	1125	1629
				70787836V1	1163	1742
				70789052V1	556	1068
				70789169V1	952	1570
				70789897V1	412	1049
				4216786F6 (ADRENOT15)	1	488
6	888992CB1	2187	921-942	70715608V1	701	1239
				6597709H1 (ARTANOT06)	1300	2025
				426440R6 (BLADNOT01)	1210	1758
				6592338H1 (UTREDIT14)	1581	2187
				2604204F6 (LUNGTUT07)	1	539
				4761979F6 (PLACNOT05)	513	1109

Table 5

Polynucleotide SEQ ID NO:	Incyte Project ID	Representative Library
4	2457018CB1	COLNNOT16
5	4907793CB1	ADRENOT15
6	888992CB1	BRAITUT08

Table 6

Library	Vector	Library Description
ADRENOT15	pINCY	Library was constructed using RNA isolated from adrenal tissue removed from a Caucasian female fetus, who died from anencephalus after 16-weeks' gestation.
BRAITUT08	pINCY	Library was constructed using RNA isolated from brain tumor tissue removed from the left frontal lobe of a 47-year-old Caucasian male during excision of cerebral meningeal tissue. Pathology indicated grade 4 fibrillary astrocytoma with focal tumoral radionecrosis. Patient history included cerebrovascular disease, deficiency anemia, hyperlipidemia, epilepsy, and tobacco use. Family history included cerebrovascular disease and a malignant prostate neoplasm.
COLNNOT16	pINCY	Library was constructed using RNA isolated from sigmoid colon tissue removed from a 62-year-old Caucasian male during a sigmoidectomy and permanent colostomy.

Table 7

Program	Description	Reference	Parameter Threshold
ABIFACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25:3389-3402.	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) Proc. Natl. Acad. Sci. USA 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183:63-98; and Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489.	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 (1991) Nucleic Acids Res. 19:6565-6572; Henikoff, J.G. and S. Henikoff (1996) Methods Enzymol. 266:88-105; and Attwood, T.K. et al. (1997) J. Chem. Inf. Comput. Sci. 37:417-424.	Probability value= 1.0E-3 or less
HMMER	An algorithm for searching a query sequence against hidden Markov model (HMM)-based databases of protein family consensus sequences, such as PFAM.	Krogh, A. et al. (1994) J. Mol. Biol. 235:1501-1531; Sonnhammer, E.L.L. et al. (1988) Nucleic Acids Res. 26:320-322; Durbin, R. et al. (1998) Our World View, in a Nutshell, Cambridge Univ. Press, pp. 1-350.	PFAM hits: Probability value= 1.0E-3 or less Signal peptide hits: Score= 0 or greater

Table 7 (cont.)

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

What is claimed is:

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

10. An isolated antibody which specifically binds to a polypeptide of claim 1.
11. An isolated polynucleotide selected from the group consisting of:
- a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting
 - 5 of SEQ ID NO:4-6,
 - b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:4-6,
 - c) a polynucleotide complementary to a polynucleotide of a),
 - d) a polynucleotide complementary to a polynucleotide of b), and
 - 10 e) an RNA equivalent of a)-d).
12. An isolated polynucleotide comprising at least 60 contiguous nucleotides of a polynucleotide of claim 11.
13. A method for detecting a target polynucleotide in a sample, said target polynucleotide
- 15 having a sequence of a polynucleotide of claim 11, the method comprising:
- a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization
 - 20 complex is formed between said probe and said target polynucleotide or fragments thereof, and
 - b) detecting the presence or absence of said hybridization complex, and, optionally, if present, the amount thereof.
14. A method of claim 13, wherein the probe comprises at least 60 contiguous nucleotides.
- 25
15. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 11, the method comprising:
- a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and
 - 30 b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.
16. A composition comprising a polypeptide of claim 1 and a pharmaceutically acceptable excipient.

17. A composition of claim 16, wherein the polypeptide has an amino acid sequence selected from the group consisting of SEQ ID NO:1-3.

18. A method for treating a disease or condition associated with decreased expression of functional SRP, comprising administering to a patient in need of such treatment the composition of claim 16.

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

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

20. A composition comprising an agonist compound identified by a method of claim 19 and a pharmaceutically acceptable excipient.

21. A method for treating a disease or condition associated with decreased expression of functional SRP, comprising administering to a patient in need of such treatment a composition of claim 20.

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

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

23. A composition comprising an antagonist compound identified by a method of claim 22 and a pharmaceutically acceptable excipient.

24. A method for treating a disease or condition associated with overexpression of functional SRP, comprising administering to a patient in need of such treatment a composition of claim 23.

25. A method of screening for a compound that specifically binds to the polypeptide of claim 1, said method comprising the steps of:

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

compound that specifically binds to the polypeptide of claim 1.

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

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

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

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

28. A method for assessing toxicity of a test compound, said method comprising:

- a) treating a biological sample containing nucleic acids with the test compound;
- 25 b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide of claim 11 under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide comprising a polynucleotide sequence of a polynucleotide of claim 11 or fragment thereof;
- 30 c) quantifying the amount of hybridization complex; and
- d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.

29. A diagnostic test for a condition or disease associated with the expression of SRP in a biological sample comprising the steps of:

- a) combining the biological sample with an antibody of claim 10, under conditions suitable for the antibody to bind the polypeptide and form an antibody:polypeptide complex; and
- b) detecting the complex, wherein the presence of the complex correlates with the presence of the polypeptide in the biological sample.

30. The antibody of claim 10, wherein the antibody is:

- a) a chimeric antibody,
- b) a single chain antibody,
- c) a Fab fragment,
- d) a $F(ab')_2$ fragment, or
- e) a humanized antibody.

31. A composition comprising an antibody of claim 10 and an acceptable excipient.

32. A method of diagnosing a condition or disease associated with the expression of SRP in a subject, comprising administering to said subject an effective amount of the composition of claim 31.

33. A composition of claim 31, wherein the antibody is labeled.

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

35. A method of preparing a polyclonal antibody with the specificity of the antibody of claim 10 comprising:

- a) immunizing an animal with a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, or an immunogenic fragment thereof, under conditions to elicit an antibody response;
- b) isolating antibodies from said animal; and
- c) screening the isolated antibodies with the polypeptide, thereby identifying a polyclonal antibody which binds specifically to a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3.

36. An antibody produced by a method of claim 35.

37. A composition comprising the antibody of claim 36 and a suitable carrier.

5 38. A method of making a monoclonal antibody with the specificity of the antibody of claim
10 comprising:

a) immunizing an animal with a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, or an immunogenic fragment thereof, under conditions to elicit an antibody response;

10 b) isolating antibody producing cells from the animal;

c) fusing the antibody producing cells with immortalized cells to form monoclonal antibody-producing hybridoma cells;

d) culturing the hybridoma cells; and

e) isolating from the culture monoclonal antibody which binds specifically to a polypeptide
15 having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3.

39. A monoclonal antibody produced by a method of claim 38.

40. A composition comprising the antibody of claim 39 and a suitable carrier.

20

41. The antibody of claim 10, wherein the antibody is produced by screening a Fab expression library.

42. The antibody of claim 10, wherein the antibody is produced by screening a recombinant
25 immunoglobulin library.

43. A method for detecting a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3 in a sample, comprising the steps of:

a) incubating the antibody of claim 10 with a sample under conditions to allow specific
30 binding of the antibody and the polypeptide; and

b) detecting specific binding, wherein specific binding indicates the presence of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3 in the sample.

44. A method of purifying a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3 from a sample, the method comprising:

a) incubating the antibody of claim 10 with a sample under conditions to allow specific binding of the antibody and the polypeptide; and

5 b) separating the antibody from the sample and obtaining the purified polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3.

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

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

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

48. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:4.

15

49. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:5.

50. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:6.

<110> INCYTE GENOMICS, INC.
 LAL, Preeti
 GANDHI, Ameena R.
 YAO, Monique G.
 NGUYEN, Danniel B.
 TANG, Y. Tom
 YUE, Henry

<120> SECRETED REDOX PROTEINS

<130> PF-0796 PCT

<140> To Be Assigned

<141> Herewith

<150> 60/213,957

<151> 2000-06-22

<160> 6

<170> PERL Program

<210> 1

<211> 71

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2457018CD1

<400> 1

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

<210> 2

<211> 463

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 4907793CD1

<400> 2

Met	Val	Pro	Ala	Leu	Arg	Tyr	Leu	Val	Gly	Ala	Cys	Gly	Arg	Ala
1				5					10					15
Arg	Gly	Leu	Phe	Ala	Gly	Gly	Ser	Pro	Gly	Ala	Cys	Gly	Phe	Ala
				20					25					30
Ser	Gly	Arg	Pro	Arg	Pro	Leu	Cys	Gly	Gly	Ser	Arg	Ser	Ala	Ser
				35					40					45
Thr	Ser	Ser	Phe	Asp	Ile	Val	Ile	Val	Gly	Gly	Gly	Ile	Val	Gly
				50					55					60
Leu	Ala	Ser	Ala	Arg	Ala	Leu	Ile	Leu	Arg	His	Pro	Ser	Leu	Ser
				65					70					75

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

<210> 3

<211> 654

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 888992CD1

<400> 3

Met	Ala	Ala	Ala	Ala	Leu	Gly	Ser	Ser	Ser	Gly	Ser	Ala	Ser	Pro
1				5					10					15
Ala	Val	Ala	Glu	Leu	Cys	Gln	Asn	Thr	Pro	Glu	Thr	Phe	Leu	Glu
				20					25					30
Ala	Ser	Lys	Leu	Leu	Leu	Thr	Tyr	Ala	Asp	Asn	Ile	Leu	Arg	Asn
				35					40					45
Pro	Asn	Asp	Glu	Lys	Tyr	Arg	Ser	Ile	Arg	Ile	Gly	Asn	Thr	Ala
				50					55					60
Phe	Ser	Thr	Arg	Leu	Leu	Pro	Val	Arg	Gly	Ala	Val	Glu	Cys	Leu
				65					70					75
Phe	Glu	Met	Gly	Phe	Glu	Glu	Gly	Glu	Thr	His	Leu	Ile	Phe	Pro
				80					85					90
Lys	Lys	Ala	Ser	Val	Glu	Gln	Leu	Gln	Lys	Ile	Arg	Asp	Leu	Ile
				95					100					105
Ala	Ile	Glu	Arg	Ser	Ser	Arg	Leu	Asp	Gly	Ser	Asn	Lys	Ser	His
				110					115					120
Lys	Val	Lys	Ser	Ser	Gln	Gln	Pro	Ala	Ala	Ser	Thr	Gln	Leu	Pro
				125					130					135
Thr	Thr	Pro	Ser	Ser	Asn	Pro	Ser	Gly	Leu	Asn	Gln	His	Thr	Arg
				140					145					150
Asn	Arg	Gln	Gly	Gln	Ser	Ser	Asp	Pro	Pro	Ser	Ala	Ser	Thr	Val
				155					160					165
Ala	Ala	Asp	Ser	Ala	Ile	Leu	Glu	Val	Leu	Gln	Ser	Asn	Ile	Gln
				170					175					180
His	Val	Leu	Val	Tyr	Glu	Asn	Pro	Ala	Leu	Gln	Glu	Lys	Ala	Leu
				185					190					195
Ala	Cys	Ile	Pro	Val	Gln	Glu	Leu	Lys	Arg	Lys	Ser	Gln	Glu	Lys
				200					205					210
Leu	Ser	Arg	Ala	Arg	Lys	Leu	Asp	Lys	Gly	Ile	Asn	Ile	Ser	Asp
				215					220					225
Glu	Asp	Phe	Leu	Leu	Leu	Glu	Leu	Leu	His	Trp	Phe	Lys	Glu	Glu
				230					235					240
Phe	Phe	His	Trp	Val	Asn	Asn	Val	Leu	Cys	Ser	Lys	Cys	Gly	Gly
				245					250					255
Gln	Thr	Arg	Ser	Arg	Asp	Arg	Ser	Leu	Leu	Pro	Ser	Asp	Asp	Glu
				260					265					270
Leu	Lys	Trp	Gly	Ala	Lys	Glu	Val	Glu	Asp	His	Tyr	Cys	Asp	Ala
				275					280					285
Cys	Gln	Phe	Ser	Asn	Arg	Phe	Pro	Arg	Tyr	Asn	Asn	Pro	Glu	Lys
				290					295					300
Leu	Leu	Glu	Thr	Arg	Cys	Gly	Arg	Cys	Gly	Glu	Trp	Ala	Asn	Cys
				305					310					315
Phe	Thr	Leu	Cys	Cys	Arg	Ala	Val	Gly	Phe	Glu	Ala	Arg	Tyr	Val
				320					325					330
Trp	Asp	Tyr	Thr	Asp	His	Val	Trp	Thr	Glu	Val	Tyr	Ser	Pro	Ser
				335					340					345
Gln	Gln	Arg	Trp	Leu	His	Cys	Asp	Ala	Cys	Glu	Asp	Val	Cys	Asp
				350					355					360
Lys	Pro	Leu	Leu	Tyr	Glu	Ile	Gly	Trp	Gly	Lys	Lys	Leu	Ser	Tyr
				365					370					375
Val	Ile	Ala	Phe	Ser	Lys	Asp	Glu	Val	Val	Asp	Val	Thr	Trp	Arg
				380					385					390
Tyr	Ser	Cys	Lys	His	Glu	Glu	Val	Ile	Ala	Arg	Arg	Thr	Lys	Val
				395					400					405
Lys	Glu	Ala	Leu	Leu	Arg	Asp	Thr	Ile	Asn	Gly	Leu	Asn	Lys	Gln
				410					415					420
Arg	Gln	Leu	Phe	Leu	Ser	Glu	Asn	Arg	Arg	Lys	Glu	Leu	Leu	Gln
				425					430					435
Arg	Ile	Ile	Val	Glu	Leu	Val	Glu	Phe	Ile	Ser	Pro	Lys	Thr	Pro
				440					445					450
Lys	Pro	Gly	Glu	Leu	Gly	Gly	Arg	Ile	Ser	Gly	Ser	Val	Ala	Trp
				455					460					465
Arg	Val	Ala	Arg	Gly	Glu	Met	Gly	Leu	Gln	Arg	Lys	Glu	Thr	Leu

	470		475		480
Phe Ile Pro Cys	Glu Asn Glu Lys Ile	Ser Lys Gln Leu His	Leu		
	485		490		495
Cys Tyr Asn Ile	Val Lys Asp Arg Tyr	Val Arg Val Ser Asn	Asn		
	500		505		510
Asn Gln Thr Ile	Ser Gly Trp Glu Asn	Gly Val Trp Lys Met	Glu		
	515		520		525
Ser Ile Phe Arg	Lys Val Glu Thr Asp	Trp His Met Val Tyr	Leu		
	530		535		540
Ala Arg Lys Glu	Gly Ser Ser Phe Ala	Tyr Ile Ser Trp Lys	Phe		
	545		550		555
Glu Cys Gly Ser	Val Gly Leu Lys Val	Asp Ser Ile Ser Ile	Arg		
	560		565		570
Thr Ser Ser Gln	Thr Phe Gln Thr Gly	Thr Val Glu Trp Lys	Leu		
	575		580		585
Arg Ser Asp Thr	Ala Gln Val Glu Leu	Thr Gly Asp Asn Ser	Leu		
	590		595		600
His Ser Tyr Ala	Asp Phe Ser Gly Ala	Thr Glu Val Ile Leu	Glu		
	605		610		615
Ala Glu Leu Ser	Arg Gly Asp Gly Asp	Val Ala Trp Gln His	Thr		
	620		625		630
Gln Leu Phe Arg	Gln Ser Leu Asn Asp	His Glu Glu Asn Cys	Leu		
	635		640		645
Glu Ile Ile Ile	Lys Phe Ser Asp Leu				
	650				

<210> 4

<211> 2305

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2457018CB1

<400> 4

```

ctctgcgccc cgcgtagttc cggtagggcgac tgcggcgcat ggcggtgagc ggtgtggaga 60
agacgcgcgg gtggctgggc cttgcattgt tggcggctct cggtgaccct cgtgctaccc 120
ccctttctcc aggccttggg acatcttgct ggggaggcag cggcggcccc aggcccggt 180
actccctgcg cgtcccgcgg agcccggctt cccggcccag tttccagcgc ccggaatcct 240
tccactgtct gtctctgccc agagcaacct acgtgcagta acgctgactc cagagcgcac 300
ccgttggggc atgaaggcgg cacagcgtcg aaaaaacaaa agaataagaa gaaaacgcgg 360
aacagggcct ctgtggcaaa tggaggcgag aaggcctcag agaaactcgc cccagaagaa 420
gttcccctaa gcgctgaggg ccaggcacia cagttggccc aggaattggc ttggtgtgtg 480
gagcaactgg agctgggcct caagaggcag aaaccacccc cgaaacagaa agagcaggct 540
attggagcaa tccgaaccct gcgcagcaaa agaacgccct tgccccggaa gaggcagctg 600
atgcactcct tgtttggaga ctatagggct cagatggaag ccgaatggcg tgaggccctg 660
cgggctctca gagctgctgc ttattcagcc caggtgcaac ctgtagatgg agccaccaga 720
aagaagagcc aaagggtctg caggcctcgc tctatatgga gagccaaagc cactctggac 780
atgcctgatg aagagtttag gttcaatttc ttttagcgtc tccccgaacc tgaaacaatc 840
cccctccctt ggggtggtgt aggggtttgt tttgagtgc gagcctttcc aggacttctg 900
ttgtcagaga accctggagt tggctgttcc ctggctggtc caaggatttg tagctgttgt 960
gaagggtgtg gaccatcaga taggcaaaag accccgttcg ttttctgatg aaatgttctc 1020
tctttcagaa gagagagaga ggtgcattta gaaaatatgc aataaattga agtgagtgtt 1080
caaagtattg tagaaggaat attgtactca gtcttttagga ttagattaag tggctgttgg 1140
taacaaagat tagtggagaa gctgtataat cgtacactgg ttttcacttt tgaaaggaat 1200
ccctgtcaaa tgggttagtg cttaatgctg ttagtgcata ttgccctaatt ctctattttt 1260
gataaaattg gataaggagt gaaagagtat cgtgaccacc tatgttagag gaagtacaga 1320
agatgcaggg gtgtggtatc cctgggtcca gtccctcacc tggtagcctt gtgcatgttg 1380
ccttcattcc tgagcaggta tcatcctcag ggaaccagca tggcacctac caggccaggc 1440
tctgttctta ggagcaagga gcttcttgcg ctaacagttc tggcctgaga cctggattga 1500
gccttggcag acttcttgtc taaatgttgg ccattcagtc tcaggccctc tggtccatgg 1560
aattgggaat ctccagggtg cctaatacctc attggtggct tgatgtttgc tggatatctt 1620

```

```

caaactcagt tcccagacta gattgatacc tggagcccag ctgcctactc agcattttcca 1680
cttgggtgct tcataggcat ttcaaacctg atgtgtttta aacacttgat taggctccgg 1740
ttttcctttg gcttctgctt ttcagtgaat ggcatgactg cctatgtggg tggcaagcca 1800
cccaggtgcc gaggaagag actgagggca cgagctgttc cagtataata aaatatataa 1860
aataagaaga gttatactag atctagatca tagacatgat tataatgtgag tatcattaat 1920
cattagttaa tagcaattac tctttattcc aatattataa taatcctcac tctacaatca 1980
taacctagga aaaaccaggc catacagaga taggagccga ggggacatag tgcgaagtgg 2040
ccagaagaca agagtgtgag ccttctctta tgcctggaca gggccaccag agggcttggg 2100
ctagcagtaa caccagtgtc tgggaagatg cctgttgcaa agtggaccat ggtctagcag 2160
tagcatcagt gtcaaggaaa aacaccact acttagcaga ctgggaaaag gacctccct 2220
ttccccgggg ggaagttag agaagactac tctccactc ttgtgggagg gcctgacat 2280
caggtcaggc cgccgcagtt atcaa 2305

```

<210> 5

<211> 1742

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 4907793CB1

<400> 5

```

ggccgtggag ggtcaagtgg cttcttctga gcgctgaggg aggggagcgt gcgtagggga 60
tggtgccagc gctgcgttat ttggttgggt cctgcccagc ggcccgcggg cttttcgccg 120
gtggctcccc tggggcgtgc ggggttcgct ctgggagggc aagaccgctg tgtggaggta 180
ggcgacgcgc cagcaccagc tcatttgata tagtcattgt tgggtggcga attgtggggc 240
ttgcctctgc cagagcactc atcctgcgac atccatcact ttctattggt gttctggaaa 300
aggagaaaaga tttagctggt caccagactg gacataacag tgggtgtcata catagtggaa 360
tttattataa acctgagtct ctgaaagcca aattatgtgt acaaggtgca gccctcctct 420
atgagtactg tcagcaaaaag ggaatttcct acaagcagtg tggcaagcct atagtagctg 480
ttgaacaaga agaaattccc agacttcagg ccctatatga gaaaggcctc cagaatgggt 540
tcccgggcct gaggtgatc cagcaggagg atataaaaaa gaaggagcca tattgtaggg 600
gtctaattgc tattgattgt ccacatactg gcattgtgga ctatcgccag gtggctttgt 660
catttgccca ggatttccaa gaagcagggt gctctgtctt gaccaatttt gaagtaaaag 720
gtattgaaat ggctaaagaa agtccttcaa gaagtataga tggaaatgcaa tatccaattg 780
ttataaagaa taaaaaggga gaggaatttc gatgtcagta tgttgtgaca tgtgcaggac 840
tttactcaga ccgtatttca gagttgagtg gctgcactcc tgatcctcga attgtaccat 900
tccgggggaga ttacctgctt ttgaagccag aaaaatgtta tcttgtaaaa ggaaatattt 960
atccggtccc agatagccgg tttcctttcc taggagttca cttcacacca aggatggatg 1020
gcagtatttg gctagggcct aatgcagttc ttgcctttta acgagagggg tacagaccct 1080
ttgacttcag tgccacagat gttatggata taattatcaa tagtggcctg attaaactgg 1140
catcccagaa tttttcctat ggagttactg aaatgtataa agcatgtttt cttggtgcaa 1200
cagtgaagta tcttcaaaaa ttcattccctg aaattactat cagtgatata cttagggggc 1260
cagctggagt aagagcccag gccctggata gagatggaaa tctggtagaa gattttgtat 1320
ttgatgcagg agttggggat attggaaatc gcattcttca tgtgagaaat gcaccttctc 1380
ctgctgctac ttcttccatt gcaatttctg gaatgattgc agatgaagta caacaaagat 1440
ttgaattata aataatgaaa ggagctaggt atgcactgta atcttcatgt cagcaacaag 1500
aatgtactaa ttgcattctt taatttaaaa aatggttttt aaactttcgt ttttaggtca 1560
aagataacca ctgcattgtt aaaatgatgt aacatagaaa taataatttt ctacaaaaat 1620
tagccaggcg tgggtggcgca tgcctgtaat ctcagctact caggaggctg aggcaggaga 1680
atcaattgaa ccgagcagac ggaggtgagc tgagagctga gattgcacca ctgcactcca 1740
gc 1742

```

<210> 6

<211> 2187

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 888992CB1

<400> 6

```

gggcggcggc ggcgcccgcct ggcgctcaag catggcgggcg gcggcattgg gcagctcctc 60
aggctcggcg tccccggccg tggctgagct ctgccagaac accccggaga cctttttgga 120
ggcctccaag ctgctgctca cctatgctga caacatcctc agaaacccta atgatgaaaa 180
atatagatcc atccggattg gaaacacagc cttttctact agactcttgc ctgtcagagg 240
agctgttgaa tgtttatattg aaatgggctt tgaagaggga gaaacacatc tcatctttcc 300
taaaaaagct tcagtggagc agctgcaaaa aattcgtgac ctgattgcca tagagagaag 360
tagcagactg gatggctcaa ataagagcca caaagtaaag tcactctcagc aacctgcagc 420
cagtacccag cttcctacaa caccatcttc aaatcccagt ggggttaaacc agcacacaag 480
gaaccgtcaa gggcagtcac cagatccacc atctgcttca acggttgctg ctgactcagc 540
cattctagaa gttcttcagt ccaacattca gcatgtgctg gtctatgaaa atcctgctct 600
tcaggagaaa gcgttggtt gtattccggt ccaagaacta aaaaggaaat cacaagaaaa 660
gttatcgaga gctagaaaaat tggataaagg tatcaatata agt gatgagg attttctttt 720
gctggagctt ttgcaactgg ttaaggaaga attttttcac tgggtgaata acgttttgtg 780
cagcaaatgt ggtggacaga ctaggtctag agatagatca ttactgcca gtgatgatga 840
gctgaagtgg ggtgcaaagg aagtgaaga tcattactgt gatgcctgcc agttcagcaa 900
tcgattccca agatataata accctgagaa acttttggaa acaagatgtg gacgggtgtg 960
cgagtgggcc aattgtttta cactgtgctg ccgagctgta ggggttgaag ctcgctatgt 1020
ttgggattac acagaccatg tctggacaga agtctattct cttctcagc agcgggtggt 1080
gcaactgtat gcatgtgaag atgtctgtga caagccactc ctttatgaaa taggatgggg 1140
caagaagctt tcctatgtca tagcattttc aaaagatgag gtatgtgatg tcacttggcg 1200
atattcctgc aaacatgaag aggtgattgc cagaagaact aagggttaaag aagcattact 1260
tcgagacact attaatgggc ttaataagca gaggcaactg tttttgtcag aaaacagaag 1320
gaaagaactt ctccagagga taattgtgga gcttgttgaa tttatatctc caaaacccc 1380
taaacctgga gaacttgggg gaagaatatc tgggtcagtg gcttggagag tagcccgagg 1440
tgaaatgggt ctacagagaa aagaaacctt gtttattccc tgtgaaaatg agaagatttc 1500
taaacagctc cacctttgtt acaatattgt gaaagatcgt tatgttcgag tttcaaataa 1560
caatcaaacc atttctggat gggagaaatgg cgtgtggaaa atggaatcta tattcagaaa 1620
agttgaaaca gactggcaca tgggtatattt ggcccgaag gaaggatcat cttttgctta 1680
tatttcctgg aagtttgagt gtgggtcagt tggcctaaaa gtagatagca tttctattag 1740
aacaagtagt caaacttttc agactggaac agtagaatgg aaattgcat ctgatacagc 1800
acaagtagaa ctgacaggcg ataacagtct tcaactctat gctgattttt ctggtgccac 1860
tgaagttatt ttggaagcag aattaagcag aggagatggt gatgtcgctt ggcaacacac 1920
ccagctgttt agacaaagct taaatgacca tgaagaaaat tgtttggaga taattataaa 1980
attcagtgac ctttgagaac ctgaacatta tagaaaagct ggcaataatc aaggacttac 2040
tgaagtagtc tgttggttca gtgcatgctt agttggcagt taccaccctg tgctagcata 2100
tttcttttgc tagctatcca tcatgtaacc ctcatgaaaa ttatctttat acgtggacta 2160
taataaaaata ttgaattaaa aaaaaaa 2187

```