

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
20 November 2003 (20.11.2003)

PCT

(10) International Publication Number
WO 03/094862 A2

- (51) International Patent Classification⁷: **A61K**
- (21) International Application Number: PCT/US03/15357
- (22) International Filing Date: 13 May 2003 (13.05.2003)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data: 60/380,566 13 May 2002 (13.05.2002) US
- (71) Applicant (for all designated States except US): **RIGEL PHARMACEUTICALS, INC.** [US/US]; 1180 Veterans Blvd., South San Francisco, CA 94080 (US).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): **LORENS, James** [US/US]; 7 Berenda Way, Portola Valley, CA 94028 (US). **XU, Weiduan** [CN/US]; 1800 22nd Ave., #203, San Francisco, CA 94122 (US).
- (74) Agents: **PARENT, Annette, S.** et al.; Townsend and Townsend and Crew LLP, Two Embarcadero Center, 8th Floor, San Francisco, CA 94111-3834 (US).
- (81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.
- (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

[Continued on next page]

(54) Title: tRNA SYNTHASE: MODULATORS OF ANGIOGENESIS

```

FEATURES             Location/Qualifiers
     source            1..1942
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /map="1p13.3-p13.1"
     gene              1..1942
                     /gene="SARS"
                     /note="synonym: SERP, SERP2"
                     /db_xref="locusID:6101"
                     /db_xref="MIM:607529"
     CDS                16..1482
                     /gene="SARS"
                     /SC_number="6.1.3.11"
                     /note="serine tRNA ligase"
                     /codon_start=1
                     /product="seryl-tRNA synthetase"
                     /protein_id="NP_059294.2"
                     /db_xref="GI:1630958"
                     /db_xref="locusID:6101"
                     /db_xref="MIM:607529"
                     /translation="MGLDLDPRVDEKGGDPAIRDTQEKRFQDPLVDGLVVDSEKR
RCRFADNMLKLNLCSTKGGKMKKEPVGDEESVPEVLSFDLADALAMAVSQ
KKVVELLDRAILKCDNERIKLBARFEMLRRTGNLHSPVFQNDQVDFVSRVWQ
DCTPFRVYRVLVYKDTGTREYVYVAGRCRFLKLVFLDQALSTVALKTLQSS
QYPIVYPPFRKRVQVLAQLSDFDEBLKVKICKSEKSDCEYDEEYLIATSEQPI
NAIHRDHWLRBDEPDKVAGLSTCRQGVGGRDRGCFVHQFKIEQFVSSPHD
NKKWREEMPTTTRTYGDLIPVIVTVDGSLDASMLDAMFQSGARFKLI
VSCSNDYQARRLRIRYQZTKXMDKVEFVMLNATCATRTICALENYQTEKI
TPSKLKEFPNPLQLBLLPFFKFAFTEQBFERKQKHGEGSKKAARDVTLNLLQN"
     NEFTD*            82..1482
                     /gene="SARS"
                     /note="Region: Seryl-tRNA synthetase (translation,
ribosomal structure and biogenesis)"
                     /db_xref="CDD:C000112"
     polyA_signal       1831..1836
                     /gene="SARS"
     polyA_site         1852
                     /gene="SARS"
                     /evidence=experimental
     polyA_site         1895
                     /gene="SARS"
                     /evidence=experimental

```

```

BASE COUNT          530 a 460 c 538 g 414 t
ORIGIN
1  gcaatgagcgg ggtcaacagg tgaatgctgc ggcgcgaccc ttgcttccct gacgtgtggc
61  ccggaggggaa agaagatgat gccgcatctg gatttgttcc ggttggatca agagggggac
121  ccagccctcca tccgagagcc gaggagagag cgttctcaag accgggagct agcgagccag
181  ctggtgaagg cagacagcga gtcgagcaga tttgatcttc ggcgcagaca cttgaacaag
241  ctgaagacc  tatcgagca gacatcaga gagaacaagc agaaanaagc gccagtgagg
301  gatgagagc ctgaccaga gaatctgctg agttctgacc accttactgc agagcttca
361  gtaactctga aagtctaca aatcaaaaaa gccccactcc tcaatgatga agccatctg
421  agttgctgag cgaagcggat aaagctgaaa ccagagcggg ttgagaaacc ccagagatc
481  agttctctcc tgcctctct agacatctc agaaagagag aggaagtcga caaaaagaa
541  gggagcattt ggggctggtg taacgtcagg aagaagtae ccaatctgga cctggtgctg
601  atgtgagtg gcttgaagg cyaaaagggg gccctggagg ctggagagcc aggtatctc
661  cttaaggggg tctgtgctt ccggagcagc gctctcacc agcaagcctc tggatcttg
721  ggaagtcagg gtaaatctc cactatcacc ccccttcca tgaagagga gtcctagc
781  gaggtagaac agctagcca gttgatgaa gaacttata agtgatctg caaagcagc
841  gaagagctg acacacccc ctacgtcag agtacctg tgcacacc ccagcacc
901  attcctgcc tgcctcagg tgaatgaca cgtctcagc agatctcc cagatctg
961  ggcctctaca cctgcttcg tcaagaggtg gctccatg gccgagcaac ccggtgctc
1021  ttcagagcc acagcttg gaagctgaa cagtttgtt cctcaccacc ccagagcaac
1081  aactcaaggg agatctga agatctcatt acccagag agagatctc ccagctcag
1141  gggatctctt acccattgt gaattctgc tcaagttct tgaatcatg tgcagtaag
1201  agcctgacc tgaagagct gctctcagg tcaagagctc tccgtgact ggtctctgt
1261  tctaatgaa cgtatccc ggtctcagg cctcaaac gtaatggga acccaagaa
1321  atgattgaca agtggagtt tgcctatg ctcaatgta ccatgtgag cactaccgt
1381  accatctgac cctctcagg gaactaccag acagagagc gaatcactg gctggaaaa
1441  tgaagagct tcatctccc agatctgaa gactgacc cctctgta gctctgctc
1501  atcgagcag agcatcaaa gacgagag agcaacagc agggcagca aagaagaca
1561  gacgcaagag agctaccct agaaacagc cctctgaca tgaagagcc ccagctgta
1621  accctctccc cctcaatt gacagattt cactctgag tggtaaac tcaagctg
1681  ccaaacagca ggaagccaa gaccccttc atcccctg cccctctga cctgtagct
1741  gaggagggaa cagtgacag taccacagc atgtctctg cctccctgac gggatagag
1801  acccaactc gatgctgac gaacctatg aataagcat cctctggagc ggtctagaa
1861  tctctctcag tctctctc aggtctgca cccgagacc aaaaaaaa aaaaaaaa
1921  aaaaaaaa aaaaaaaa aa

```

(57) Abstract: The present invention relates to regulation of angiogenesis. More particularly, the present invention is directed to nucleic acids encoding seryl-tRNA synthetase protein, which is involved in modulation of angiogenesis. The invention further relates to methods for identifying and using agents, including small organic molecules, antibodies, peptides, cyclic peptides, nucleic acids, antisense nucleic acids, RNAi, and ribozymes, that modulate angiogenesis via modulation of seryl-tRNA synthetase protein; as well as to the use of expression profiles and compositions in diagnosis and therapy of diseases related to angiogenesis.



WO 03/094862 A2



Declaration under Rule 4.17:

— *of inventorship (Rule 4.17(iv)) for US only*

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.

tRNA SYNTHASE: Modulators of Angiogenesis

CROSS-REFERENCES TO RELATED APPLICATIONS

This application claims the benefit of priority of US provisional application
5 60/380,566, filed May 13, 2002, which is herein incorporated by reference for all purposes.

BACKGROUND OF THE INVENTION

The present invention relates to regulation of angiogenesis. More particularly,
the present invention is directed to nucleic acids encoding "seryl-tRNA synthetase protein",
10 which is involved in modulation of angiogenesis. The invention further relates to methods
for identifying and using agents, including small organic molecules, antibodies, peptides,
cyclic peptides, nucleic acids, antisense nucleic acids, RNAi, and ribozymes, that modulate
angiogenesis via modulation of seryl-tRNA synthetase protein; as well as to the use of
expression profiles and compositions in diagnosis and therapy of diseases related to
15 angiogenesis.

BRIEF SUMMARY OF THE INVENTION

Angiogenesis is typically limited in a normal adult to the placenta, ovary,
endometrium and sites of wound healing. However, angiogenesis, or its absence, plays an
20 important role in the maintenance of a variety of pathological states. Some of these states are
characterized by neovascularization, e.g., cancer, diabetic retinopathy, glaucoma, and age
related macular degeneration. Others, e.g., stroke, infertility, heart disease, ulcers, and
scleroderma, are diseases of angiogenic insufficiency. Therefore, there is a need to identify
nucleic acids encoding proteins involved in the regulation of angiogenesis, to identify, e.g.,
25 modulators of angiogenesis, as well as new therapeutic and diagnostic applications.

In one aspect, the present invention provides a method for identifying a
compound that modulates angiogenesis, the method comprising the steps of: (i) contacting
the compound with a seryl-tRNA synthetase polypeptide, the polypeptide encoded by a
nucleic acid that hybridizes under stringent conditions to a nucleic acid encoding a
30 polypeptide comprising an amino acid sequence of SEQ ID NO:2; and (ii) determining the
functional effect of the compound upon the seryl-tRNA synthetase polypeptide.

In one embodiment, the functional effect is determined *in vitro*. In another embodiment, the functional effect is a physical effect. In another embodiment, the functional effect is determined by measuring ligand or substrate binding to the polypeptide. In another embodiment, the functional effect is a chemical effect. In another embodiment, the functional effect is determined by measuring aminoacyl-tRNA synthetase activity of the polypeptide.

In one embodiment, the polypeptide is expressed in a eukaryotic host cell. In another embodiment, the functional effect is a physical effect. In another embodiment, the functional effect is determined by measuring ligand or substrate binding to the polypeptide. In another embodiment, the functional effect is a chemical or phenotypic effect. In another embodiment, the host cell is an endothelial cell. In another embodiment, the functional effect is determined by measuring $\alpha v\beta 3$ expression, haptotaxis, or chemotaxis.

In one embodiment, modulation is inhibition of angiogenesis.

In one embodiment, the polypeptide is recombinant. In another embodiment, the polypeptide comprises a sequence of SEQ ID NO:2.

In one embodiment, the compound is an antibody, antisense molecule, RNAi molecule, peptide, cyclic peptide, ribozyme, or a small organic molecule.

In another aspect, the present invention provides a method of modulating angiogenesis in a subject, the method comprising the step of administering to the subject a therapeutically effective amount of a compound identified using the methods described herein.

In one embodiment, the compound is an antibody, antisense molecule, RNAi molecule, peptide, cyclic peptide, ribozyme, or a small organic molecule.

In one embodiment, the subject is a human.

In one embodiment, the compound inhibits angiogenesis.

In another aspect, the present invention provides a method of modulating angiogenesis in a subject, the method comprising the step of administering to the subject a therapeutically effective amount of a seryl-tRNA synthetase polypeptide, the polypeptide encoded by a nucleic acid that hybridizes under stringent conditions to a nucleic acid encoding a polypeptide comprising an amino acid sequence of SEQ ID NO:2.

In another aspect, the present invention provides a method of modulating angiogenesis in a subject, the method comprising the step of administering to the subject a therapeutically effective amount of a nucleic acid encoding a seryl-tRNA synthetase

polypeptide, wherein the nucleic acid hybridizes under stringent conditions to a nucleic acid encoding a polypeptide comprising an amino acid sequence of SEQ ID NO:2.

BRIEF DESCRIPTION OF THE DRAWINGS

5 Figure 1 provides a nucleotide and amino acid sequence of seryl tRNA synthetase.

Figure 2 provides the sequence of the seryl tRNA synthetase cDNA isolated in the assay for angiogenesis modulators.

10 DETAILED DESCRIPTION OF THE INVENTION

Introduction

For the first time, a "seryl-tRNA synthetase" has been identified as a protein involved in regulating angiogenesis. The human seryl-tRNA synthetase gene (Accession number NM_006513) of the present invention encodes the full-length seryl tRNA synthetase protein (Accession number NP_006504.2). The human seryl-tRNA synthetase gene was first
15 cloned from a brain cDNA library (Vincent *et al.*, *Eur. J. Biochem.* 250:77-84 (1997)).

The seryl-tRNA synthetase protein is a member of a family of aminoacyl-tRNA synthetases that attach the correct amino acid molecule to a tRNA for protein translation. Aminoacyl-tRNA synthetases catalyze a two step reaction, first activating an
20 amino acid by forming an aminoacyl-adenylate and then transferring the aminoacyl-adenylate to the correct tRNA molecule at either the 2' or 3' terminal OH group.

Aminoacyl-tRNA synthetases belong to one of two classes. Class I enzymes are generally monomeric and attach the target amino acid to the 2' terminal OH of the tRNA molecule. The catalytic site of Class I enzymes all share a structure called the Rossmann fold.
25 Class II molecules are generally dimeric or tetrameric and usually attach the amino acid to the 3' OH of the correct tRNA molecule. Class II enzymes appear to have a novel antiparallel fold at the catalytic site (Cusack, *Cur. Op. Struct. Biol.* 7:881-889 (1997)).

Some class I aminoacyl-tRNA synthetases have been shown to function as cytokine molecules. The first, tyrosyl-tRNA synthetase is cleaved into two domains with
30 cytokine activity. The carboxy terminal domain is not found in other aminoacyl-tRNA synthetases. The NH3 domain includes the active site of the enzyme and is structurally similar to other Class I aminoacyl-tRNA synthetases. The NH3 domain induces migration of polymorphonuclear leukocytes. The full-length tyrosyl-tRNA synthetase is not active as a cytokine (Wakasugi *et al.*, *Science* 284:147-151 (1999))

A second Class I aminoacyl-tRNA synthetase also has been shown to have cytokine activity. Fragments including the carboxy terminus of the human tryptophanyl-tRNA synthetase were able to inhibit angiogenesis in a variety of assay. All of the truncations included the active site of the enzyme. However, full-length tryptophanyl-tRNA synthetase had no activity in the angiogenesis assays (Wakasugi *et al.*, *Proc. Nat'l Acad. Sci. USA* 99:173-177 (2002); Otani *et al.*, *Proc. Nat'l Acad. Sci. USA* 99:178-183 (2002)).

The present invention identifies a full-length human Class II aminoacyl-tRNA synthetase (seryl-tRNA synthetase) involved in modulation of angiogenesis. Without wishing to be bound by theory, this full-length aminoacyl-tRNA synthetase may also act as a cytokine. Angiogenesis assays described herein reveal for the first time that expression of a cDNA encoding seryl-tRNA synthetase exerted a negative effect on $\alpha v\beta 3$ surface expression. In addition, endothelial cells expressing the partial sequence were strongly inhibited in their haptotactic response to vitronectin, which is an indicator of an anti-angiogenic phenotype. Without wishing to be bound by theory, it appears that a truncated seryl-tRNA synthetase sequence could act as a negative transdominant mutant by providing an anti-angiogenic phenotype.

The seryl-tRNA synthetase protein and other members of the angiogenesis pathway therefore represent targets for the development of angiogenic drugs, preferably anti-angiogenic drugs, e.g., anti-angiogenic drugs for treatment of neovascularization, e.g., cancer, diabetic retinopathy, glaucoma, and age related macular degeneration, or angiogenic drugs for treatment of angiogenic insufficiency, e.g., stroke, infertility, heart disease, ulcers, and scleroderma, are diseases of angiogenic insufficiency. Modulators include small organic molecules, nucleic acids, peptides, cyclic peptides, antibodies, antisense molecules, RNAi molecules, and ribozymes. The nucleic acids and proteins of the invention are also useful for diagnostic applications, using, e.g., nucleic acid probes, oligonucleotides, and antibodies.

Definitions

By "disorder associated with angiogenesis" or "disease associated with angiogenesis" herein is meant a disease state which is marked by either an excess or a deficit of vessel development. Angiogenesis disorders associated with increased angiogenesis include, but are not limited to, cancer and proliferative diabetic retinopathy. Pathological states for which it may be desirable to increase angiogenesis include stroke, heart disease, infertility, ulcers, and scleredema. An increase in angiogenesis may also be desirable in transplantation or for artificial or *in vitro* growth of organs.

The terms “seryl-tRNA synthetase” or a nucleic acid encoding “seryl-tRNA synthetase” refer to nucleic acid and polypeptide polymorphic variants, alleles, mutants, and interspecies homologs that: (1) have an amino acid sequence that has greater than about 60% amino acid sequence identity, 65%, 70%, 75%, 80%, 85%, 90%, preferably 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% or greater amino acid sequence identity, preferably over a region of over a region of at least about 25, 50, 100, 200, 500, 1000, or more amino acids, to a polypeptide encoded by a nucleic acid of SEQ ID NO:1 or an amino acid sequence of SEQ ID NO:2; (2) specifically bind to antibodies, e.g., polyclonal antibodies, raised against an immunogen comprising an amino acid sequence of SEQ ID NO:2, immunogenic fragments thereof, and conservatively modified variants thereof; (3) specifically hybridize under stringent hybridization conditions to a nucleic acid encoding SEQ ID NO:2, e.g., a nucleic acid sequence of SEQ IN NO:1 or its complement, and conservatively modified variants thereof; (4) have a nucleic acid sequence that has greater than about 95%, preferably greater than about 96%, 97%, 98%, 99%, or higher nucleotide sequence identity, preferably over a region of at least about 25, 50, 100, 200, 500, 1000, or more nucleotides, to SEQ ID NO:1 or its complement. A polynucleotide or polypeptide sequence is typically from a mammal including, but not limited to, primate, e.g., human; rodent, e.g., rat, mouse, hamster; cow, pig, horse, sheep, or any mammal. The nucleic acids and proteins of the invention include both naturally occurring or recombinant molecules. A nucleotide and amino acid sequence of seryl-tRNA synthetase is found in Figure 2. In addition, accession numbers for the human seryl-tRNA synthetase gene (NM_006513), and the full-length seryl tRNA synthetase protein (NP_006504.2) are provided.

The phrase “functional effects” in the context of assays for testing compounds that modulate activity of a seryl-tRNA synthetase protein includes the determination of a parameter that is indirectly or directly under the influence of a seryl-tRNA synthetase polypeptide, e.g., a chemical or phenotypic effect such as loss-of angiogenesis phenotype represented by a change in expression of a cell surface marker, such as $\alpha v\beta 3$ integrin, or changes in cellular proliferation, especially endothelial cell proliferation; or enzymatic activity, or, e.g., a physical effect such as ligand binding or inhibition of ligand binding. A functional effect therefore includes ligand binding activity, the ability of cells to proliferate, expression in cells undergoing angiogenesis, and other characteristics of angiogenic cells. “Functional effects” include *in vitro*, *in vivo*, and *ex vivo* activities. Angiogenesis assays are described, e.g., in *Angiogenesis Protocols* (Murray, ed., 2001).

By “determining the functional effect” is meant assaying for a compound that increases or decreases a parameter that is indirectly or directly under the influence of a seryl-tRNA synthetase protein, e.g., measuring physical and chemical or phenotypic effects. Such functional effects can be measured by any means known to those skilled in the art, e.g., changes in spectroscopic characteristics (e.g., fluorescence, absorbance, refractive index); hydrodynamic (e.g., shape), chromatographic; or solubility properties for the protein; ligand binding assays, e.g., binding to antibodies; measuring inducible markers or transcriptional activation of the angiogenesis protein; measuring changes in enzymatic activity, e.g., phosphatase activity; measuring changes in cell surface markers, e.g., $\alpha\beta3$ integrin; and measuring cellular proliferation, particularly endothelial cell proliferation. Determination of the functional effect of a compound on angiogenesis can also be performed using angiogenesis assays known to those of skill in the art such as endothelial cell tube formation assays; chemo taxis assays, haptotaxis assays; differentiation assays using matrigel or co-culture with smooth muscle cells, the chick CAM assay; the mouse corneal assay; and assays that assess vascularization of an implanted tumor. The functional effects can be evaluated by many means known to those skilled in the art, e.g., microscopy for quantitative or qualitative measures of alterations in morphological features, e.g., tube or blood vessel formation, measurement of changes in RNA or protein levels for angiogenesis-associated sequences, measurement of RNA stability, identification of downstream or reporter gene expression (CAT, luciferase, β -gal, GFP and the like), e.g., via chemiluminescence, fluorescence, colorimetric reactions, antibody binding, inducible markers, etc.

“Ligand” refers to a molecule that is specifically bound by a protein. An antibody is one example of a ligand.

“Substrate” refers to a molecule that binds to an enzyme and is part of a specific chemical reaction catalyzed by the enzyme. Substrates of seryl-tRNA synthetase include serine and serine-accepting tRNA.

“Inhibitors,” “activators,” and “modulators” of seryl-tRNA synthetase polynucleotide and polypeptide sequences are used to refer to activating, inhibitory, or modulating molecules identified using *in vitro* and *in vivo* assays of seryl-tRNA synthetase polynucleotide and polypeptide sequences. Inhibitors are compounds that, e.g., bind to, partially or totally block activity, decrease, prevent, delay activation, inactivate, desensitize, or down regulate the activity or expression of seryl-tRNA synthetase proteins, e.g., antagonists. “Activators” are compounds that increase, open, activate, facilitate, enhance activation, sensitize, agonize, or up regulate seryl-tRNA synthetase protein activity, e.g.,

agonists. Inhibitors, activators, or modulators also include genetically modified versions of seryl-tRNA synthetase proteins, e.g., versions with altered activity, as well as naturally occurring and synthetic ligands, antagonists, agonists, antibodies, peptides, cyclic peptides, nucleic acids, antisense molecules, ribozymes, RNAi, small organic molecules and the like.

5 Such assays for inhibitors and activators include, e.g., expressing seryl-tRNA synthetase protein *in vitro*, in cells, or cell extracts, applying putative modulator compounds, and then determining the functional effects on activity, as described above.

Samples or assays comprising seryl-tRNA synthetase proteins that are treated with a potential activator, inhibitor, or modulator are compared to control samples without
10 the inhibitor, activator, or modulator to examine the extent of inhibition. Control samples (untreated with inhibitors) are assigned a relative protein activity value of 100%. Inhibition of seryl-tRNA synthetase is achieved when the activity value relative to the control is about 80%, preferably 50%, more preferably 25-0%. Activation of seryl-tRNA synthetase is achieved when the activity value relative to the control (untreated with activators) is 110%,
15 more preferably 150%, more preferably 200-500% (i.e., two to five fold higher relative to the control), more preferably 1000-3000% higher.

The term "test compound" or "drug candidate" or "modulator" or grammatical equivalents as used herein describes any molecule, either naturally occurring or synthetic, e.g., protein, oligopeptide (e.g., from about 5 to about 25 amino acids in length, preferably
20 from about 10 to 20 or 12 to 18 amino acids in length, preferably 12, 15, or 18 amino acids in length), small organic molecule, polysaccharide, lipid, fatty acid, polynucleotide, oligonucleotide, etc., to be tested for the capacity to directly or indirectly modulation lymphocyte activation. The test compound can be in the form of a library of test compounds, such as a combinatorial or randomized library that provides a sufficient range of diversity.
25 Test compounds are optionally linked to a fusion partner, e.g., targeting compounds, rescue compounds, dimerization compounds, stabilizing compounds, addressable compounds, and other functional moieties. Conventionally, new chemical entities with useful properties are generated by identifying a test compound (called a "lead compound") with some desirable property or activity, e.g., inhibiting activity, creating variants of the lead compound, and
30 evaluating the property and activity of those variant compounds. Often, high throughput screening (HTS) methods are employed for such an analysis.

A "small organic molecule" refers to an organic molecule, either naturally occurring or synthetic, that has a molecular weight of more than about 50 daltons and less

than about 2500 daltons, preferably less than about 2000 daltons, preferably between about 100 to about 1000 daltons, more preferably between about 200 to about 500 daltons.

“RNAi molecule” or an “siRNA” refers to a nucleic acid that forms a double stranded RNA, which double stranded RNA has the ability to reduce or inhibit expression of a gene or target gene when the siRNA expressed in the same cell as the gene or target gene. “siRNA” thus refers to the double stranded RNA formed by the complementary strands. The complementary portions of the siRNA that hybridize to form the double stranded molecule typically have substantial or complete identity. In one embodiment, an siRNA refers to a nucleic acid that has substantial or complete identity to a target gene and forms a double stranded siRNA. The sequence of the siRNA can correspond to the full length target gene, or a subsequence thereof. Typically, the siRNA is at least about 15-50 nucleotides in length (e.g., each complementary sequence of the double stranded siRNA is 15-50 nucleotides in length, and the double stranded siRNA is about 15-50 base pairs in length, preferably about 20-30 base nucleotides, preferably about 20-25 nucleotides in length, e.g., 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, or 30 nucleotides in length.

“Biological sample” include sections of tissues such as biopsy and autopsy samples, and frozen sections taken for histologic purposes. Such samples include blood, sputum, tissue, cultured cells, e.g., primary cultures, explants, and transformed cells, stool, urine, etc. A biological sample is typically obtained from a eukaryotic organism, most preferably a mammal such as a primate, e.g., chimpanzee or human; cow; dog; cat; a rodent, e.g., guinea pig, rat, mouse; rabbit; or a bird; reptile; or fish.

The terms “identical” or percent “identity,” in the context of two or more nucleic acids or polypeptide sequences, refer to two or more sequences or subsequences that are the same or have a specified percentage of amino acid residues or nucleotides that are the same (i.e., about 70% identity, preferably 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or higher identity over a specified region (e.g., SEQ ID NO:1 or 2), when compared and aligned for maximum correspondence over a comparison window or designated region) as measured using a BLAST or BLAST 2.0 sequence comparison algorithms with default parameters described below, or by manual alignment and visual inspection (*see, e.g.*, NCBI web site or the like). Such sequences are then said to be “substantially identical.” This definition also refers to, or may be applied to, the complement of a test sequence. The definition also includes sequences that have deletions and/or additions, as well as those that have substitutions. As described below, the preferred algorithms can account for gaps and the like. Preferably, identity exists over a region that is

at least about 25 amino acids or nucleotides in length, or more preferably over a region that is 50-100 amino acids or nucleotides in length.

For sequence comparison, typically one sequence acts as a reference sequence, to which test sequences are compared. When using a sequence comparison algorithm, test and reference sequences are entered into a computer, subsequence coordinates are designated, if necessary, and sequence algorithm program parameters are designated. Preferably, default program parameters can be used, or alternative parameters can be designated. The sequence comparison algorithm then calculates the percent sequence identities for the test sequences relative to the reference sequence, based on the program parameters.

A "comparison window", as used herein, includes reference to a segment of any one of the number of contiguous positions selected from the group consisting of from 20 to 600, usually about 50 to about 200, more usually about 100 to about 150 in which a sequence may be compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned. Methods of alignment of sequences for comparison are well-known in the art. Optimal alignment of sequences for comparison can be conducted, e.g., by the local homology algorithm of Smith & Waterman, *Adv. Appl. Math.* 2:482 (1981), by the homology alignment algorithm of Needleman & Wunsch, *J. Mol. Biol.* 48:443 (1970), by the search for similarity method of Pearson & Lipman, *Proc. Nat'l. Acad. Sci. USA* 85:2444 (1988), by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI), or by manual alignment and visual inspection (*see, e.g., Current Protocols in Molecular Biology* (Ausubel *et al.*, eds. 1995 supplement)).

A preferred example of algorithm that is suitable for determining percent sequence identity and sequence similarity are the BLAST and BLAST 2.0 algorithms, which are described in Altschul *et al.*, *Nuc. Acids Res.* 25:3389-3402 (1997) and Altschul *et al.*, *J. Mol. Biol.* 215:403-410 (1990), respectively. BLAST and BLAST 2.0 are used, with the parameters described herein, to determine percent sequence identity for the nucleic acids and proteins of the invention. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/>). This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence, which either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold (Altschul *et al.*, *supra*).

These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Cumulative scores are calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always > 0) and N (penalty score for mismatching residues; always < 0). For amino acid sequences, a scoring matrix is used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T, and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, an expectation (E) of 10, M=5, N=-4 and a comparison of both strands. For amino acid sequences, the BLASTP program uses as defaults a wordlength of 3, and expectation (E) of 10, and the BLOSUM62 scoring matrix (see Henikoff & Henikoff, *Proc. Natl. Acad. Sci. USA* 89:10915 (1989)) alignments (B) of 50, expectation (E) of 10, M=5, N=-4, and a comparison of both strands.

“Nucleic acid” refers to deoxyribonucleotides or ribonucleotides and polymers thereof in single- or double-stranded form. The term encompasses nucleic acids containing known nucleotide analogs or modified backbone residues or linkages, which are synthetic, naturally occurring, and non-naturally occurring, which have similar binding properties as the reference nucleic acid, and which are metabolized in a manner similar to the reference nucleotides. Examples of such analogs include, without limitation, phosphorothioates, phosphoramidates, methyl phosphonates, chiral-methyl phosphonates, 2-O-methyl ribonucleotides, peptide-nucleic acids (PNAs).

Unless otherwise indicated, a particular nucleic acid sequence also implicitly encompasses conservatively modified variants thereof (e.g., degenerate codon substitutions) and complementary sequences, as well as the sequence explicitly indicated. Specifically, degenerate codon substitutions may be achieved by generating sequences in which the third position of one or more selected (or all) codons is substituted with mixed-base and/or deoxyinosine residues (Batzer *et al.*, *Nucleic Acid Res.* 19:5081 (1991); Ohtsuka *et al.*, *J. Biol. Chem.* 260:2605-2608 (1985); Rossolini *et al.*, *Mol. Cell. Probes* 8:91-98 (1994)). The term nucleic acid is used interchangeably with gene, cDNA, mRNA, oligonucleotide, and polynucleotide.

A particular nucleic acid sequence also implicitly encompasses “splice variants.” Similarly, a particular protein encoded by a nucleic acid implicitly encompasses any protein encoded by a splice variant of that nucleic acid. “Splice variants,” as the name suggests, are products of alternative splicing of a gene. After transcription, an initial nucleic acid transcript may be spliced such that different (alternate) nucleic acid splice products encode different polypeptides. Mechanisms for the production of splice variants vary, but include alternate splicing of exons. Alternate polypeptides derived from the same nucleic acid by read-through transcription are also encompassed by this definition. Any products of a splicing reaction, including recombinant forms of the splice products, are included in this definition.

The terms “polypeptide,” “peptide” and “protein” are used interchangeably herein to refer to a polymer of amino acid residues. The terms apply to amino acid polymers in which one or more amino acid residue is an artificial chemical mimetic of a corresponding naturally occurring amino acid, as well as to naturally occurring amino acid polymers and non-naturally occurring amino acid polymer.

The term “amino acid” refers to naturally occurring and synthetic amino acids, as well as amino acid analogs and amino acid mimetics that function in a manner similar to the naturally occurring amino acids. Naturally occurring amino acids are those encoded by the genetic code, as well as those amino acids that are later modified, e.g., hydroxyproline, γ -carboxyglutamate, and O-phosphoserine. Amino acid analogs refers to compounds that have the same basic chemical structure as a naturally occurring amino acid, i.e., an α carbon that is bound to a hydrogen, a carboxyl group, an amino group, and an R group, e.g., homoserine, norleucine, methionine sulfoxide, methionine methyl sulfonium. Such analogs have modified R groups (e.g., norleucine) or modified peptide backbones, but retain the same basic chemical structure as a naturally occurring amino acid. Amino acid mimetics refers to chemical compounds that have a structure that is different from the general chemical structure of an amino acid, but that functions in a manner similar to a naturally occurring amino acid.

Amino acids may be referred to herein by either their commonly known three letter symbols or by the one-letter symbols recommended by the IUPAC-IUB Biochemical Nomenclature Commission. Nucleotides, likewise, may be referred to by their commonly accepted single-letter codes.

“Conservatively modified variants” applies to both amino acid and nucleic acid sequences. With respect to particular nucleic acid sequences, conservatively modified variants refers to those nucleic acids which encode identical or essentially identical amino

acid sequences, or where the nucleic acid does not encode an amino acid sequence, to essentially identical sequences. Because of the degeneracy of the genetic code, a large number of functionally identical nucleic acids encode any given protein. For instance, the codons GCA, GCC, GCG and GCU all encode the amino acid alanine. Thus, at every
5 position where an alanine is specified by a codon, the codon can be altered to any of the corresponding codons described without altering the encoded polypeptide. Such nucleic acid variations are “silent variations,” which are one species of conservatively modified variations. Every nucleic acid sequence herein which encodes a polypeptide also describes every possible silent variation of the nucleic acid. One of skill will recognize that each codon
10 in a nucleic acid (except AUG, which is ordinarily the only codon for methionine, and TGG, which is ordinarily the only codon for tryptophan) can be modified to yield a functionally identical molecule. Accordingly, each silent variation of a nucleic acid which encodes a polypeptide is implicit in each described sequence with respect to the expression product, but not with respect to actual probe sequences.

15 As to amino acid sequences, one of skill will recognize that individual substitutions, deletions or additions to a nucleic acid, peptide, polypeptide, or protein sequence which alters, adds or deletes a single amino acid or a small percentage of amino acids in the encoded sequence is a “conservatively modified variant” where the alteration results in the substitution of an amino acid with a chemically similar amino acid.
20 Conservative substitution tables providing functionally similar amino acids are well known in the art. Such conservatively modified variants are in addition to and do not exclude polymorphic variants, interspecies homologs, and alleles of the invention.

The following eight groups each contain amino acids that are conservative substitutions for one another: 1) Alanine (A), Glycine (G); 2) Aspartic acid (D), Glutamic acid (E); 3) Asparagine (N), Glutamine (Q); 4) Arginine (R), Lysine (K); 5) Isoleucine (I),
25 Leucine (L), Methionine (M), Valine (V); 6) Phenylalanine (F), Tyrosine (Y), Tryptophan (W); 7) Serine (S), Threonine (T); and 8) Cysteine (C), Methionine (M) (*see, e.g., Creighton, Proteins* (1984)).

30 Macromolecular structures such as polypeptide structures can be described in terms of various levels of organization. For a general discussion of this organization, *see, e.g., Alberts et al., Molecular Biology of the Cell* (3rd ed., 1994) and Cantor and Schimmel, *Biophysical Chemistry Part I: The Conformation of Biological Macromolecules* (1980). “Primary structure” refers to the amino acid sequence of a particular peptide. “Secondary structure” refers to locally ordered, three dimensional structures within a polypeptide. These

structures are commonly known as domains, e.g., transmembrane domains, pore domains, and cytoplasmic tail domains. Domains are portions of a polypeptide that form a compact unit of the polypeptide and are typically 15 to 350 amino acids long. Exemplary domains include domains with enzymatic activity, e.g., phosphatase domains, ligand binding domains, etc. Typical domains are made up of sections of lesser organization such as stretches of β -sheet and α -helices. "Tertiary structure" refers to the complete three dimensional structure of a polypeptide monomer. "Quaternary structure" refers to the three dimensional structure formed by the noncovalent association of independent tertiary units. Anisotropic terms are also known as energy terms.

A "label" or a "detectable moiety" is a composition detectable by spectroscopic, photochemical, biochemical, immunochemical, chemical, or other physical means. For example, useful labels include ^{32}P , fluorescent dyes, electron-dense reagents, enzymes (e.g., as commonly used in an ELISA), biotin, digoxigenin, or haptens and proteins which can be made detectable, e.g., by incorporating a radiolabel into the peptide or used to detect antibodies specifically reactive with the peptide.

The term "recombinant" when used with reference, e.g., to a cell, or nucleic acid, protein, or vector, indicates that the cell, nucleic acid, protein or vector, has been modified by the introduction of a heterologous nucleic acid or protein or the alteration of a native nucleic acid or protein, or that the cell is derived from a cell so modified. Thus, for example, recombinant cells express genes that are not found within the native (non-recombinant) form of the cell or express native genes that are otherwise abnormally expressed, under expressed or not expressed at all.

The term "heterologous" when used with reference to portions of a nucleic acid indicates that the nucleic acid comprises two or more subsequences that are not found in the same relationship to each other in nature. For instance, the nucleic acid is typically recombinantly produced, having two or more sequences from unrelated genes arranged to make a new functional nucleic acid, e.g., a promoter from one source and a coding region from another source. Similarly, a heterologous protein indicates that the protein comprises two or more subsequences that are not found in the same relationship to each other in nature (e.g., a fusion protein).

The phrase "stringent hybridization conditions" refers to conditions under which a probe will hybridize to its target subsequence, typically in a complex mixture of nucleic acids, but to no other sequences. Stringent conditions are sequence-dependent and will be different in different circumstances. Longer sequences hybridize specifically at

higher temperatures. An extensive guide to the hybridization of nucleic acids is found in Tijssen, *Techniques in Biochemistry and Molecular Biology--Hybridization with Nucleic Probes*, "Overview of principles of hybridization and the strategy of nucleic acid assays" (1993). Generally, stringent conditions are selected to be about 5-10°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength pH. The T_m is the temperature (under defined ionic strength, pH, and nucleic concentration) at which 50% of the probes complementary to the target hybridize to the target sequence at equilibrium (as the target sequences are present in excess, at T_m , 50% of the probes are occupied at equilibrium). Stringent conditions may also be achieved with the addition of destabilizing agents such as formamide. For selective or specific hybridization, a positive signal is at least two times background, preferably 10 times background hybridization. Exemplary stringent hybridization conditions can be as following: 50% formamide, 5x SSC, and 1% SDS, incubating at 42°C, or, 5x SSC, 1% SDS, incubating at 65°C, with wash in 0.2x SSC, and 0.1% SDS at 65°C.

Nucleic acids that do not hybridize to each other under stringent conditions are still substantially identical if the polypeptides which they encode are substantially identical. This occurs, for example, when a copy of a nucleic acid is created using the maximum codon degeneracy permitted by the genetic code. In such cases, the nucleic acids typically hybridize under moderately stringent hybridization conditions. Exemplary "moderately stringent hybridization conditions" include a hybridization in a buffer of 40% formamide, 1 M NaCl, 1% SDS at 37°C, and a wash in 1X SSC at 45°C. A positive hybridization is at least twice background. Those of ordinary skill will readily recognize that alternative hybridization and wash conditions can be utilized to provide conditions of similar stringency. Additional guidelines for determining hybridization parameters are provided in numerous reference, e.g., and *Current Protocols in Molecular Biology*, ed. Ausubel, *et al.*

For PCR, a temperature of about 36°C is typical for low stringency amplification, although annealing temperatures may vary between about 32°C and 48°C depending on primer length. For high stringency PCR amplification, a temperature of about 62°C is typical, although high stringency annealing temperatures can range from about 50°C to about 65°C, depending on the primer length and specificity. Typical cycle conditions for both high and low stringency amplifications include a denaturation phase of 90°C - 95°C for 30 sec - 2 min., an annealing phase lasting 30 sec. - 2 min., and an extension phase of about 72°C for 1 - 2 min. Protocols and guidelines for low and high stringency amplification

reactions are provided, e.g., in Innis *et al.* (1990) *PCR Protocols, A Guide to Methods and Applications*, Academic Press, Inc. N.Y.).

“Antibody” refers to a polypeptide comprising a framework region from an immunoglobulin gene or fragments thereof that specifically binds and recognizes an antigen.

5 The recognized immunoglobulin genes include the kappa, lambda, alpha, gamma, delta, epsilon, and mu constant region genes, as well as the myriad immunoglobulin variable region genes. Light chains are classified as either kappa or lambda. Heavy chains are classified as gamma, mu, alpha, delta, or epsilon, which in turn define the immunoglobulin classes, IgG, IgM, IgA, IgD and IgE, respectively. Typically, the antigen-binding region of an antibody
10 will be most critical in specificity and affinity of binding.

An exemplary immunoglobulin (antibody) structural unit comprises a tetramer. Each tetramer is composed of two identical pairs of polypeptide chains, each pair having one “light” (about 25 kD) and one “heavy” chain (about 50-70 kD). The N-terminus of each chain defines a variable region of about 100 to 110 or more amino acids primarily
15 responsible for antigen recognition. The terms variable light chain (V_L) and variable heavy chain (V_H) refer to these light and heavy chains respectively.

Antibodies exist, e.g., as intact immunoglobulins or as a number of well-characterized fragments produced by digestion with various peptidases. Thus, for example, pepsin digests an antibody below the disulfide linkages in the hinge region to produce $F(ab)'_2$,
20 a dimer of Fab which itself is a light chain joined to V_H - C_{H1} by a disulfide bond. The $F(ab)'_2$ may be reduced under mild conditions to break the disulfide linkage in the hinge region, thereby converting the $F(ab)'_2$ dimer into an Fab' monomer. The Fab' monomer is essentially Fab with part of the hinge region (*see Fundamental Immunology* (Paul ed., 3d ed. 1993)). While various antibody fragments are defined in terms of the digestion of an intact
25 antibody, one of skill will appreciate that such fragments may be synthesized *de novo* either chemically or by using recombinant DNA methodology. Thus, the term antibody, as used herein, also includes antibody fragments either produced by the modification of whole antibodies, or those synthesized *de novo* using recombinant DNA methodologies (e.g., single chain Fv) or those identified using phage display libraries (*see, e.g., McCafferty et al., Nature*
30 348:552-554 (1990))

For preparation of antibodies, e.g., recombinant, monoclonal, or polyclonal antibodies, many techniques known in the art can be used (*see, e.g., Kohler & Milstein, Nature* 256:495-497 (1975); Kozbor *et al., Immunology Today* 4: 72 (1983); Cole *et al.*, pp. 77-96 in *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc. (1985); Coligan,

Current Protocols in Immunology (1991); Harlow & Lane, *Antibodies, A Laboratory Manual* (1988); and Goding, *Monoclonal Antibodies: Principles and Practice* (2d ed. 1986)). The genes encoding the heavy and light chains of an antibody of interest can be cloned from a cell, e.g., the genes encoding a monoclonal antibody can be cloned from a hybridoma and used to produce a recombinant monoclonal antibody. Gene libraries encoding heavy and light chains of monoclonal antibodies can also be made from hybridoma or plasma cells. Random combinations of the heavy and light chain gene products generate a large pool of antibodies with different antigenic specificity (see, e.g., Kuby, *Immunology* (3rd ed. 1997)). Techniques for the production of single chain antibodies or recombinant antibodies (U.S. Patent 4,946,778, U.S. Patent No. 4,816,567) can be adapted to produce antibodies to polypeptides of this invention. Also, transgenic mice, or other organisms such as other mammals, may be used to express humanized or human antibodies (see, e.g., U.S. Patent Nos. 5,545,807; 5,545,806; 5,569,825; 5,625,126; 5,633,425; 5,661,016, Marks *et al.*, *Bio/Technology* 10:779-783 (1992); Lonberg *et al.*, *Nature* 368:856-859 (1994); Morrison, *Nature* 368:812-13 (1994); Fishwild *et al.*, *Nature Biotechnology* 14:845-51 (1996); Neuberger, *Nature Biotechnology* 14:826 (1996); and Lonberg & Huszar, *Intern. Rev. Immunol.* 13:65-93 (1995)). Alternatively, phage display technology can be used to identify antibodies and heteromeric Fab fragments that specifically bind to selected antigens (see, e.g., McCafferty *et al.*, *Nature* 348:552-554 (1990); Marks *et al.*, *Biotechnology* 10:779-783 (1992)). Antibodies can also be made bispecific, i.e., able to recognize two different antigens (see, e.g., WO 93/08829, Traunecker *et al.*, *EMBO J.* 10:3655-3659 (1991); and Suresh *et al.*, *Methods in Enzymology* 121:210 (1986)). Antibodies can also be heteroconjugates, e.g., two covalently joined antibodies, or immunotoxins (see, e.g., U.S. Patent No. 4,676,980, WO 91/00360; WO 92/200373; and EP 03089).

Methods for humanizing or primatizing non-human antibodies are well known in the art. Generally, a humanized antibody has one or more amino acid residues introduced into it from a source which is non-human. These non-human amino acid residues are often referred to as import residues, which are typically taken from an import variable domain. Humanization can be essentially performed following the method of Winter and co-workers (see, e.g., Jones *et al.*, *Nature* 321:522-525 (1986); Riechmann *et al.*, *Nature* 332:323-327 (1988); Verhoeyen *et al.*, *Science* 239:1534-1536 (1988) and Presta, *Curr. Op. Struct. Biol.* 2:593-596 (1992)), by substituting rodent CDRs or CDR sequences for the corresponding sequences of a human antibody. Accordingly, such humanized antibodies are chimeric antibodies (U.S. Patent No. 4,816,567), wherein substantially less than an intact

human variable domain has been substituted by the corresponding sequence from a non-human species. In practice, humanized antibodies are typically human antibodies in which some CDR residues and possibly some FR residues are substituted by residues from analogous sites in rodent antibodies.

5 A “chimeric antibody” is an antibody molecule in which (a) the constant region, or a portion thereof, is altered, replaced or exchanged so that the antigen binding site (variable region) is linked to a constant region of a different or altered class, effector function and/or species, or an entirely different molecule which confers new properties to the chimeric antibody, e.g., an enzyme, toxin, hormone, growth factor, drug, etc.; or (b) the variable
10 region, or a portion thereof, is altered, replaced or exchanged with a variable region having a different or altered antigen specificity.

In one embodiment, the antibody is conjugated to an “effector” moiety. The effector moiety can be any number of molecules, including labeling moieties such as radioactive labels or fluorescent labels, or can be a therapeutic moiety. In one aspect the
15 antibody modulates the activity of the protein.

The phrase “specifically (or selectively) binds” to an antibody or “specifically (or selectively) immunoreactive with,” when referring to a protein or peptide, refers to a binding reaction that is determinative of the presence of the protein, often in a heterogeneous population of proteins and other biologics. Thus, under designated immunoassay conditions,
20 the specified antibodies bind to a particular protein at least two times the background and more typically more than 10 to 100 times background. Specific binding to an antibody under such conditions requires an antibody that is selected for its specificity for a particular protein. For example, polyclonal antibodies raised to seryl-tRNA synthetase protein as shown in SEQ ID NO:2, polymorphic variants, alleles, orthologs, and conservatively modified variants, or
25 splice variants, or portions thereof, can be selected to obtain only those polyclonal antibodies that are specifically immunoreactive with seryl-tRNA synthetase proteins and not with other proteins. This selection may be achieved by subtracting out antibodies that cross-react with other molecules. A variety of immunoassay formats may be used to select antibodies specifically immunoreactive with a particular protein. For example, solid-phase ELISA
30 immunoassays are routinely used to select antibodies specifically immunoreactive with a protein (*see, e.g., Harlow & Lane, Antibodies, A Laboratory Manual* (1988) for a description of immunoassay formats and conditions that can be used to determine specific immunoreactivity).

Assays for proteins that modulate angiogenesis

High throughput functional genomics assays can be used to identify modulators of angiogenesis. Such assays can monitor changes in cell surface marker expression, $\alpha v\beta 3$ integrin production, proliferation, and differentiation using either cell lines or primary cells. Typically, early passage or primary endothelial cells are contacted with a cDNA or a random peptide library (encoded by nucleic acids). The cDNA library can comprise sense, antisense, full length, and truncated cDNAs. The peptide library is encoded by nucleic acids. The effect of the cDNA or peptide library on the endothelial cells is then monitored, using an assay such as cell surface marker expression (e.g., $\alpha v\beta 3$ integrin) or a phenotypic assay for angiogenesis such as migration towards an ECM (extracellular matrix) component (see, e.g., Klemke *et al.*, *J. Cell Biol.* 4:961-972 (1998)) or endothelial cell tube formation assays, as well as other bioassays such as the chick CAM assay, the mouse corneal assay, haptotaxis assays, and assays measuring the effect of administering potential modulators on implanted tumors. The effect of the cDNA or peptide can be validated and distinguished from somatic mutations, using, e.g., regulatable expression of the nucleic acid such as expression from a tetracycline promoter. cDNAs and nucleic acids encoding peptides can be rescued using techniques known to those of skill in the art, e.g., using a sequence tags.

Proteins interacting with the peptide or with the protein encoded by the cDNA (e.g., seryl-tRNA synthetase) can be isolated using a yeast two-hybrid system, mammalian two hybrid system, or phage display screen, etc. Targets so identified can be further used as bait in these assays to identify additional members of the angiogenesis pathway, which members are also targets for drug development (see, e.g., Fields *et al.*, *Nature* 340:245 (1989); Vasavada *et al.*, *Proc. Nat'l Acad. Sci. USA* 88:10686 (1991); Fearon *et al.*, *Proc. Nat'l Acad. Sci. USA* 89:7958 (1992); Dang *et al.*, *Mol. Cell. Biol.* 11:954 (1991); Chien *et al.*, *Proc. Nat'l Acad. Sci. USA* 9578 (1991); and U.S. Patent Nos. 5,283,173, 5,667,973, 5,468,614, 5,525,490, and 5,637,463).

Suitable endothelial cell lines include human umbilical vein cells (see, e.g., Jaffe *et al.*, *J. Clin. Invest.* 52:2745-2754 (1973)); human adult dermal capillary-derived cells (see, e.g., Davison *et al.*, *In Vitro* 19:937-945 (1983)); human adipose capillary derived cells (see, e.g., Kern *et al.*, *J. Clin. Invest.* 71:1822-1829 (1983); bovine aorta (see, e.g., Booyse *et al.*, *Thromb. Diathes. Ahemorrh.* 34:825-839 (1975); and rat brain capillary derived cells (see, e.g., Bowman *et al.*, *In Vitro* 17:353-362 (1981)). For culture of endothelial cell lines, explants, and primary cells, see Freshney *et al.*, *Culture of Animal Cells* (3rd ed. 1994). Suitable angiogenesis cell surface markers include $\alpha v\beta 3$ integrin (see, e.g., Elicerir &

Cheresh, *Cancer J. Sci. Am.* 6 Supp. 3:S245-249 (2000), Maeshima *et al.*, *J. Biol. Chem.* (June 8, 2001)).

Cell surface markers such as $\alpha\beta3$ can be assayed using fluorescently labeled antibodies and FACS. Cell proliferation can be measured using ^3H -thymidine or dye inclusion. Angiogenesis phenotype is measured by loss of phenotype observation.

cDNA libraries are made from any suitable source, preferably from endothelial cells. Libraries encoding random peptides are made according to techniques well known to those of skill in the art (*see, e.g.*, U.S. Patent No. 6,153,380, 6,114,111, and 6,180,343). Any suitable vector can be used for the cDNA and peptide libraries, including, e.g., retroviral vectors.

In a preferred embodiment, target proteins that modulate angiogenesis are identified using a high throughput cell based assay (using a microtiter plate format) and FACS screening for $\alpha\beta3$ cell surface expression. cDNA libraries are made which include, e.g., sense, antisense, full length, and truncated cDNAs. The cDNAs are cloned into a retroviral vector. Endothelial cells are infected with the library, cultured for a short effector phase and then the cells with reduced $\alpha\beta3$ surface levels are enriched by antibody staining and magnetic cell sorting. The enriched cell population is then sorted into microtiter plates using fluorescent antibodies and FACS. Resultant cell colonies are analyzed by immunofluorescence for reduced $\alpha\beta3$ surface levels. Selected colonies are infected with wild type MMLV virus to rescue the proviral vector. The infectious supernatant is used to infect endothelial cells, which are subsequently analyzed for $\alpha\beta3$ levels by FACS. The cDNA is isolated and sequenced to determine if it represents a wild type or mutated cDNA, e.g., whether the cDNA represents a negative transdominant mutant. Optionally, a marker such as GFP can be used to select for retrovirally infected cells. Using this system, a cDNA encoding seryl-tRNA synthetase was identified as a target for anti-angiogenic drug therapy.

Isolation of nucleic acids encoding seryl-tRNA synthetase

This invention relies on routine techniques in the field of recombinant genetics. Basic texts disclosing the general methods of use in this invention include Sambrook *et al.*, *Molecular Cloning, A Laboratory Manual* (2nd ed. 1989); Kriegler, *Gene Transfer and Expression: A Laboratory Manual* (1990); and *Current Protocols in Molecular Biology* (Ausubel *et al.*, eds., 1994).

Seryl-tRNA synthetase nucleic acids, polymorphic variants, orthologs, and alleles that are substantially identical to the amino acid sequence of SEQ ID NO:2 can be

isolated using seryl-tRNA synthetase nucleic acid probes and oligonucleotides under stringent hybridization conditions, by screening libraries. Alternatively, expression libraries can be used to clone seryl-tRNA synthetase protein, polymorphic variants, orthologs, and alleles by detecting expressed homologs immunologically with antisera or purified antibodies made against human seryl-tRNA synthetase or portions thereof.

To make a cDNA library, one should choose a source that is rich in seryl-tRNA synthetase RNA, e.g., endothelial cells. The mRNA is then made into cDNA using reverse transcriptase, ligated into a recombinant vector, and transfected into a recombinant host for propagation, screening and cloning. Methods for making and screening cDNA libraries are well known (*see, e.g., Gubler & Hoffman, Gene 25:263-269 (1983); Sambrook et al., supra; Ausubel et al., supra*).

For a genomic library, the DNA is extracted from the tissue and either mechanically sheared or enzymatically digested to yield fragments of about 12-20 kb. The fragments are then separated by gradient centrifugation from undesired sizes and are constructed in bacteriophage lambda vectors. These vectors and phage are packaged *in vitro*. Recombinant phage are analyzed by plaque hybridization as described in Benton & Davis, *Science* 196:180-182 (1977). Colony hybridization is carried out as generally described in Grunstein *et al., Proc. Natl. Acad. Sci. USA., 72:3961-3965 (1975)*.

An alternative method of isolating seryl-tRNA synthetase nucleic acid and its orthologs, alleles, mutants, polymorphic variants, and conservatively modified variants combines the use of synthetic oligonucleotide primers and amplification of an RNA or DNA template (*see U.S. Patents 4,683,195 and 4,683,202; PCR Protocols: A Guide to Methods and Applications (Innis et al., eds, 1990)*). Methods such as polymerase chain reaction (PCR) and ligase chain reaction (LCR) can be used to amplify nucleic acid sequences of human seryl-tRNA synthetase directly from mRNA, from cDNA, from genomic libraries or cDNA libraries. Degenerate oligonucleotides can be designed to amplify seryl-tRNA synthetase homologs using the sequences provided herein. Restriction endonuclease sites can be incorporated into the primers. Polymerase chain reaction or other *in vitro* amplification methods may also be useful, for example, to clone nucleic acid sequences that code for proteins to be expressed, to make nucleic acids to use as probes for detecting the presence of seryl-tRNA synthetase encoding mRNA in physiological samples, for nucleic acid sequencing, or for other purposes. Genes amplified by the PCR reaction can be purified from agarose gels and cloned into an appropriate vector.

Gene expression of seryl-tRNA synthetase can also be analyzed by techniques known in the art, e.g., reverse transcription and amplification of mRNA, isolation of total RNA or poly A⁺ RNA, northern blotting, dot blotting, *in situ* hybridization, RNase protection, high density polynucleotide array technology, e.g., and the like.

5 Nucleic acids encoding seryl-tRNA synthetase protein can be used with high density oligonucleotide array technology (e.g., GeneChipTM) to identify seryl-tRNA synthetase protein, orthologs, alleles, conservatively modified variants, and polymorphic variants in this invention. In the case where the homologs being identified are linked to a known disease such as angiogenesis, they can be used with GeneChipTM as a diagnostic tool
10 in detecting the disease in a biological sample, *see, e.g.,* Gunthand *et al.*, *AIDS Res. Hum. Retroviruses* 14: 869-876 (1998); Kozal *et al.*, *Nat. Med.* 2:753-759 (1996); Matson *et al.*, *Anal. Biochem.* 224:110-106 (1995); Lockhart *et al.*, *Nat. Biotechnol.* 14:1675-1680 (1996); Gingeras *et al.*, *Genome Res.* 8:435-448 (1998); Hacia *et al.*, *Nucleic Acids Res.* 26:3865-3866 (1998).

15 The gene for seryl-tRNA synthetase is typically cloned into intermediate vectors before transformation into prokaryotic or eukaryotic cells for replication and/or expression. These intermediate vectors are typically prokaryote vectors, e.g., plasmids, or shuttle vectors.

20 **Expression in prokaryotes and eukaryotes**

To obtain high level expression of a cloned gene, such as those cDNAs encoding seryl-tRNA synthetase, one typically subclones seryl-tRNA synthetase into an expression vector that contains a strong promoter to direct transcription, a transcription/translation terminator, and if for a nucleic acid encoding a protein, a ribosome
25 binding site for translational initiation. Suitable bacterial promoters are well known in the art and described, e.g., in Sambrook *et al.*, and Ausubel *et al.*, *supra*. Bacterial expression systems for expressing the seryl-tRNA synthetase protein are available in, e.g., *E. coli*, *Bacillus sp.*, and *Salmonella* (Palva *et al.*, *Gene* 22:229-235 (1983); Mosbach *et al.*, *Nature* 302:543-545 (1983). Kits for such expression systems are commercially available.

30 Eukaryotic expression systems for mammalian cells, yeast, and insect cells are well known in the art and are also commercially available.

Selection of the promoter used to direct expression of a heterologous nucleic acid depends on the particular application. The promoter is preferably positioned about the same distance from the heterologous transcription start site as it is from the transcription start

site in its natural setting. As is known in the art, however, some variation in this distance can be accommodated without loss of promoter function.

In addition to the promoter, the expression vector typically contains a transcription unit or expression cassette that contains all the additional elements required for the expression of the seryl-tRNA synthetase encoding nucleic acid in host cells. A typical
5 expression cassette thus contains a promoter operably linked to the nucleic acid sequence encoding seryl-tRNA synthetase and signals required for efficient polyadenylation of the transcript, ribosome binding sites, and translation termination. Additional elements of the cassette may include enhancers and, if genomic DNA is used as the structural gene, introns
10 with functional splice donor and acceptor sites.

In addition to a promoter sequence, the expression cassette should also contain a transcription termination region downstream of the structural gene to provide for efficient termination. The termination region may be obtained from the same gene as the promoter sequence or may be obtained from different genes.

The particular expression vector used to transport the genetic information into
15 the cell is not particularly critical. Any of the conventional vectors used for expression in eukaryotic or prokaryotic cells may be used. Standard bacterial expression vectors include plasmids such as pBR322 based plasmids, pSKF, pET23D, and fusion expression systems such as MBP, GST, and LacZ. Epitope tags can also be added to recombinant proteins to
20 provide convenient methods of isolation, e.g., c-myc.

Expression vectors containing regulatory elements from eukaryotic viruses are typically used in eukaryotic expression vectors, e.g., SV40 vectors, papilloma virus vectors, retroviral vectors, and vectors derived from Epstein-Barr virus. Other exemplary eukaryotic vectors include pMSG, pAV009/A⁺, pMTO10/A⁺, pMAMneo-5, baculovirus pDSVE, and
25 any other vector allowing expression of proteins under the direction of the CMV promoter, SV40 early promoter, SV40 later promoter, metallothionein promoter, murine mammary tumor virus promoter, Rous sarcoma virus promoter, polyhedrin promoter, or other promoters shown effective for expression in eukaryotic cells.

Expression of proteins from eukaryotic vectors can be also be regulated using
30 inducible promoters. With inducible promoters, expression levels are tied to the concentration of inducing agents, such as tetracycline or ecdysone, by the incorporation of response elements for these agents into the promoter. Generally, high level expression is obtained from inducible promoters only in the presence of the inducing agent; basal

expression levels are minimal. Inducible expression vectors are often chosen if expression of the protein of interest is detrimental to eukaryotic cells.

In one embodiment, the vectors of the invention have a regulatable promoter, e.g., tet-regulated systems and the RU-486 system (*see, e.g., Gossen & Bujard, Proc. Nat'l Acad. Sci. USA* 89:5547 (1992); Oligino *et al., Gene Ther.* 5:491-496 (1998); Wang *et al., Gene Ther.* 4:432-441 (1997); Neering *et al., Blood* 88:1147-1155 (1996); and Rendahl *et al., Nat. Biotechnol.* 16:757-761 (1998)). These impart small molecule control on the expression of the candidate target nucleic acids. This beneficial feature can be used to determine that a desired phenotype is caused by a transfected cDNA rather than a somatic mutation.

Some expression systems have markers that provide gene amplification such as thymidine kinase and dihydrofolate reductase. Alternatively, high yield expression systems not involving gene amplification are also suitable, such as using a baculovirus vector in insect cells, with a seryl-tRNA synthetase encoding sequence under the direction of the polyhedrin promoter or other strong baculovirus promoters.

The elements that are typically included in expression vectors also include a replicon that functions in *E. coli*, a gene encoding antibiotic resistance to permit selection of bacteria that harbor recombinant plasmids, and unique restriction sites in nonessential regions of the plasmid to allow insertion of eukaryotic sequences. The particular antibiotic resistance gene chosen is not critical, any of the many resistance genes known in the art are suitable. The prokaryotic sequences are preferably chosen such that they do not interfere with the replication of the DNA in eukaryotic cells, if necessary.

Standard transfection methods are used to produce bacterial, mammalian, yeast or insect cell lines that express large quantities of seryl-tRNA synthetase protein, which are then purified using standard techniques (*see, e.g., Colley et al., J. Biol. Chem.* 264:17619-17622 (1989); *Guide to Protein Purification*, in *Methods in Enzymology*, vol. 182 (Deutscher, ed., 1990)). Transformation of eukaryotic and prokaryotic cells are performed according to standard techniques (*see, e.g., Morrison, J. Bact.* 132:349-351 (1977); Clark-Curtiss & Curtiss, *Methods in Enzymology* 101:347-362 (Wu *et al.*, eds, 1983).

Any of the well-known procedures for introducing foreign nucleotide sequences into host cells may be used. These include the use of calcium phosphate transfection, polybrene, protoplast fusion, electroporation, biolistics, liposomes, microinjection, plasma vectors, viral vectors and any of the other well known methods for introducing cloned genomic DNA, cDNA, synthetic DNA or other foreign genetic material into a host cell (*see, e.g., Sambrook et al., supra*). It is only necessary that the particular

genetic engineering procedure used be capable of successfully introducing at least one gene into the host cell capable of expressing seryl-tRNA synthetase.

After the expression vector is introduced into the cells, the transfected cells are cultured under conditions favoring expression of seryl-tRNA synthetase, which is recovered
5 from the culture using standard techniques identified below.

Purification of seryl-tRNA synthetase -angiogenesis polypeptides

Either naturally occurring or recombinant seryl-tRNA synthetase can be purified for use in functional assays. Naturally occurring seryl-tRNA synthetase can be
10 purified, e.g., from human tissue. Recombinant seryl-tRNA synthetase can be purified from any suitable expression system.

The seryl-tRNA synthetase protein may be purified to substantial purity by standard techniques, including selective precipitation with such substances as ammonium sulfate; column chromatography, immunopurification methods, and others (*see, e.g.*, Scopes,
15 *Protein Purification: Principles and Practice* (1982); U.S. Patent No. 4,673,641; Ausubel *et al.*, *supra*; and Sambrook *et al.*, *supra*).

A number of procedures can be employed when recombinant seryl-tRNA synthetase protein is being purified. For example, proteins having established molecular adhesion properties can be reversibly fused to the seryl-tRNA synthetase protein. With the
20 appropriate ligand, seryl-tRNA synthetase protein can be selectively adsorbed to a purification column and then freed from the column in a relatively pure form. The fused protein is then removed by enzymatic activity. Finally, seryl-tRNA synthetase protein could be purified using immunoaffinity columns.

A. Purification of seryl-tRNA synthetase from recombinant bacteria

Recombinant proteins are expressed by transformed bacteria in large amounts, typically after promoter induction; but expression can be constitutive. Promoter induction with IPTG is one example of an inducible promoter system. Bacteria are grown according to standard procedures in the art. Fresh or frozen bacteria cells are used for isolation of protein.
30

Proteins expressed in bacteria may form insoluble aggregates (“inclusion bodies”). Several protocols are suitable for purification of seryl-tRNA synthetase protein inclusion bodies. For example, purification of inclusion bodies typically involves the extraction, separation and/or purification of inclusion bodies by disruption of bacterial cells, e.g., by incubation in a buffer of 50 mM TRIS/HCL pH 7.5, 50 mM NaCl, 5 mM MgCl₂, 1

mM DTT, 0.1 mM ATP, and 1 mM PMSF. The cell suspension can be lysed using 2-3 passages through a French Press, homogenized using a Polytron (Brinkman Instruments) or sonicated on ice. Alternate methods of lysing bacteria are apparent to those of skill in the art (see, e.g., Sambrook *et al.*, *supra*; Ausubel *et al.*, *supra*).

5 If necessary, the inclusion bodies are solubilized, and the lysed cell suspension is typically centrifuged to remove unwanted insoluble matter. Proteins that formed the inclusion bodies may be renatured by dilution or dialysis with a compatible buffer. Suitable solvents include, but are not limited to urea (from about 4 M to about 8 M), formamide (at least about 80%, volume/volume basis), and guanidine hydrochloride (from about 4 M to
10 about 8 M). Some solvents which are capable of solubilizing aggregate-forming proteins, for example SDS (sodium dodecyl sulfate), 70% formic acid, are inappropriate for use in this procedure due to the possibility of irreversible denaturation of the proteins, accompanied by a lack of immunogenicity and/or activity. Although guanidine hydrochloride and similar
15 agents are denaturants, this denaturation is not irreversible and renaturation may occur upon removal (by dialysis, for example) or dilution of the denaturant, allowing re-formation of immunologically and/or biologically active protein. Other suitable buffers are known to those skilled in the art. Human seryl-tRNA synthetase proteins are separated from other bacterial proteins by standard separation techniques, e.g., with Ni-NTA agarose resin.

 Alternatively, it is possible to purify seryl-tRNA synthetase protein from
20 bacteria periplasm. After lysis of the bacteria, when the seryl-tRNA synthetase protein exported into the periplasm of the bacteria, the periplasmic fraction of the bacteria can be isolated by cold osmotic shock in addition to other methods known to skill in the art. To isolate recombinant proteins from the periplasm, the bacterial cells are centrifuged to form a pellet. The pellet is resuspended in a buffer containing 20% sucrose. To lyse the cells, the
25 bacteria are centrifuged and the pellet is resuspended in ice-cold 5 mM MgSO₄ and kept in an ice bath for approximately 10 minutes. The cell suspension is centrifuged and the supernatant decanted and saved. The recombinant proteins present in the supernatant can be separated from the host proteins by standard separation techniques well known to those of
30 skill in the art.

B. Standard protein separation techniques for purifying seryl-tRNA synthetase proteins

Solubility fractionation

Often as an initial step, particularly if the protein mixture is complex, an initial salt fractionation can separate many of the unwanted host cell proteins (or proteins derived from the cell culture media) from the recombinant protein of interest. The preferred salt is ammonium sulfate. Ammonium sulfate precipitates proteins by effectively reducing the amount of water in the protein mixture. Proteins then precipitate on the basis of their solubility. The more hydrophobic a protein is, the more likely it is to precipitate at lower ammonium sulfate concentrations. A typical protocol includes adding saturated ammonium sulfate to a protein solution so that the resultant ammonium sulfate concentration is between 20-30%. This concentration will precipitate the most hydrophobic of proteins. The precipitate is then discarded (unless the protein of interest is hydrophobic) and ammonium sulfate is added to the supernatant to a concentration known to precipitate the protein of interest. The precipitate is then solubilized in buffer and the excess salt removed if necessary, either through dialysis or diafiltration. Other methods that rely on solubility of proteins, such as cold ethanol precipitation, are well known to those of skill in the art and can be used to fractionate complex protein mixtures.

Size differential filtration

The molecular weight of the seryl-tRNA synthetase protein can be used to isolate it from proteins of greater and lesser size using ultrafiltration through membranes of different pore size (for example, Amicon or Millipore membranes). As a first step, the protein mixture is ultrafiltered through a membrane with a pore size that has a lower molecular weight cut-off than the molecular weight of the protein of interest. The retentate of the ultrafiltration is then ultrafiltered against a membrane with a molecular cut off greater than the molecular weight of the protein of interest. The recombinant protein will pass through the membrane into the filtrate. The filtrate can then be chromatographed as described below.

Column chromatography

The seryl-tRNA synthetase proteins can also be separated from other proteins on the basis of its size, net surface charge, hydrophobicity, and affinity for ligands. In addition, antibodies raised against proteins can be conjugated to column matrices and the

proteins immunopurified. All of these methods are well known in the art. It will be apparent to one of skill that chromatographic techniques can be performed at any scale and using equipment from many different manufacturers (e.g., Pharmacia Biotech).

5 Assays for modulators of seryl-tRNA synthetase protein and angiogenesis

A. Assays

Modulation of a seryl-tRNA synthetase protein, and corresponding modulation of angiogenesis, can be assessed using a variety of *in vitro* and *in vivo* assays, including high throughput ligand binding and cell based assays, as described herein. Such assays can be used to test for inhibitors and activators of seryl-tRNA synthetase protein, and, consequently, inhibitors and activators of angiogenesis. Such modulators of seryl-tRNA synthetase 3 protein are useful for treating angiogenesis disorders. Modulators of seryl-tRNA synthetase protein are tested using either recombinant or naturally occurring seryl-tRNA synthetase, preferably human seryl-tRNA synthetase.

15 Preferably, the seryl-tRNA synthetase protein will have the sequence displayed in SEQ ID NO:2 or a conservatively modified variant thereof. Alternatively, the seryl-tRNA synthetase protein of the assay will be derived from a eukaryote and include an amino acid subsequence having substantial amino acid sequence identity to SEQ ID NO:2. Generally, the amino acid sequence identity will be at least 60%, preferably at least 65%, 20 70%, 75%, 80%, 85%, or 90%, most preferably at least 95%.

Measurement of an angiogenic or loss-of-angiogenesis phenotype on seryl-tRNA synthetase protein or cell expressing seryl-tRNA synthetase protein, either recombinant or naturally occurring, can be performed using a variety of assays, *in vitro*, *in vivo*, and *ex vivo*. For example, recombinant or naturally occurring seryl-tRNA synthetase can be used *in vitro*, in a ligand binding or enzymatic function assay. Seryl-tRNA synthetase 25 present in a cellular extract can also be used in *in vitro* assays. Cell- and animal-based *in vivo* assays can also be used to assay for seryl-tRNA synthetase modulators. Any suitable physical, chemical, or phenotypic change that affects activity or binding can be used to assess the influence of a test compound on the polypeptide of this invention. When the functional effects are determined using intact cells or animals, one can also measure a variety of effects 30 such as, in the case of angiogenesis associated with tumors, tumor growth, neovascularization, cell surface markers such as $\alpha v\beta 3$, hormone release, transcriptional changes to both known and uncharacterized genetic markers (e.g., northern blots), changes in cell metabolism such as cell growth or pH changes, and changes in intracellular second

messengers such as cGMP. In one embodiment, measurement of $\alpha v\beta 3$ integrin cell surface expression and FACS sorting is used to identify modulators of angiogenesis.

In vitro assays

5 Assays to identify compounds with seryl-tRNA synthetase modulating activity, e.g., anti-angiogenic activity, can be performed *in vitro*, e.g., binding assays. Such assays can use full length seryl-tRNA synthetase protein or a variant thereof (*see, e.g.*, SEQ ID NO:2), or a fragment of a seryl-tRNA synthetase protein having a desired activity. Purified recombinant or naturally occurring seryl-tRNA synthetase protein can be used in the
10 *in vitro* methods of the invention. In addition to purified seryl-tRNA synthetase protein, the recombinant or naturally occurring seryl-tRNA synthetase protein can be part of a cellular lysate. As described below, the assay can be either solid state or soluble. Preferably, the protein is bound to a solid support, either covalently or non-covalently. Often, the *in vitro* assays of the invention are ligand binding or ligand affinity assays, either non-competitive or
15 competitive. Other *in vitro* assays include measuring changes in spectroscopic (e.g., fluorescence, absorbance, refractive index), hydrodynamic (e.g., shape), chromatographic, or solubility properties for the protein.

In one embodiment, a high throughput binding assay is performed in which the seryl-tRNA synthetase protein or chimera comprising a fragment thereof is contacted with
20 a potential modulator and incubated for a suitable amount of time. In one embodiment, the potential modulator is bound to a solid support, and the seryl-tRNA synthetase protein is added. In another embodiment, the seryl-tRNA synthetase protein is bound to a solid support. A wide variety of modulators can be used, as described below, including small organic molecules, peptides, and antibodies. A wide variety of assays can be used to identify
25 seryl-tRNA synthetase -modulator binding or phosphatase activity, including labeled protein-protein binding assays, electrophoretic mobility shifts, immunoassays, and the like. In some cases, the binding of the candidate modulator is determined through the use of competitive binding assays, where interference with binding of a known ligand is measured in the presence of a potential modulator. Often, either the potential modulator or the known ligand
30 is labeled.

Cell-based in vivo assays

In another embodiment, seryl-tRNA synthetase protein is expressed in a cell, and functional, e.g., physical and chemical or phenotypic, changes are assayed to identify

angiogenesis modulators, preferably anti-angiogenesis compounds. Cells expressing seryl-tRNA synthetase proteins can also be used in binding assays or enzymatic assays. Any suitable functional effect can be measured, as described herein. For example, ligand binding, cell surface marker expression, cellular proliferation, cellular differentiation assays and cell migration assays are all suitable assays to identify potential modulators using a cell based system. Suitable cells for such cell based assays include both primary endothelial cells and cell lines, as described herein. The seryl-tRNA synthetase protein can be naturally occurring or recombinant. Also, as described above, a fragment of seryl-tRNA synthetase protein with a desired activity can be used in cell based assays.

As described above, in one embodiment, loss-of angiogenesis phenotype is measured by contacting endothelial cells comprising an seryl-tRNA synthetase target with a potential modulator. Modulation of angiogenesis is identified by screening for cell surface marker expression, e.g., $\alpha\beta3$ integrin expression levels, using fluorescent antibodies and FACS sorting.

In another embodiment, cellular seryl-tRNA synthetase polypeptide levels are determined by measuring the level of protein or mRNA. The level of seryl-tRNA synthetase protein or proteins are measured using immunoassays such as western blotting, ELISA and the like with an antibody that selectively binds to the seryl-tRNA synthetase polypeptide or a fragment thereof. For measurement of mRNA, amplification, e.g., using PCR, LCR, or hybridization assays, e.g., northern hybridization, RNase protection, dot blotting, are preferred. The level of protein or mRNA is detected using directly or indirectly labeled detection agents, e.g., fluorescently or radioactively labeled nucleic acids, radioactively or enzymatically labeled antibodies, and the like, as described herein.

Alternatively, seryl-tRNA synthetase expression can be measured using a reporter gene system. Such a system can be devised using a seryl-tRNA synthetase protein promoter operably linked to a reporter gene such as chloramphenicol acetyltransferase, firefly luciferase, bacterial luciferase, β -galactosidase and alkaline phosphatase. Furthermore, the protein of interest can be used as an indirect reporter via attachment to a second reporter such as red or green fluorescent protein (*see, e.g., Mistili & Spector, Nature Biotechnology* 15:961-964 (1997)). The reporter construct is typically transfected into a cell. After treatment with a potential modulator, the amount of reporter gene transcription, translation, or activity is measured according to standard techniques known to those of skill in the art.

In another embodiment, seryl-tRNA synthetase phosphatase activity can be measured, using, e.g., labeled substrate proteins, gel electrophoresis, and ELISA assays.

A variety of phenotypic angiogenesis assays are known to those of skill in the art. Various models have been employed to evaluate angiogenesis (e.g., Croix *et al.*, *Science* 289:1197-1202 (2000) and Kahn *et al.*, *Amer. J. Pathol.* 156:1887-1900). Assessment of
5 angiogenesis in the presence of a potential modulator of angiogenesis can be performed using cell-culture-based angiogenesis assays, e.g., endothelial cell tube formation assays, cellular differentiation assays using matrigel matrix or by co-culture with smooth muscle cells, chemotaxis assays using VEGF or FGF, and haptotaxis assays, as well as other animal based
10 bioassays such as the chick CAM assay, the mouse corneal assay, and assays measuring the effect of administering potential modulators on implanted tumors.

Animal models

A number of animal based assays for angiogenesis phenotypes are known to
15 those of skill in the art and can be used to assay for modulators of angiogenesis. For example, the chick CAM assay is described by O'Reilly, *et al.* *Cell* 79: 315-328 (1994). Briefly, 3 day old chicken embryos with intact yolks are separated from the egg and placed in a petri dish. After 3 days of incubation, a methylcellulose disc containing the protein to be tested is applied to the CAM of individual embryos. After about 48 hours of incubation, the
20 embryos and CAMs are observed to determine whether endothelial growth has been inhibited.

The mouse corneal assay involves implanting a growth factor-containing pellet, along with another pellet containing the suspected endothelial growth inhibitor, in the cornea of a mouse and observing the pattern of capillaries that are elaborated in the cornea.

25 Angiogenesis can also be measured by determining the extent of neovascularization of a tumor. For example, carcinoma cells can be subcutaneously inoculated into athymic nude mice and tumor growth then monitored. Immunoassays using endothelial cell-specific antibodies are typically used to stain for vascularization of tumor and the number of vessels in the tumor.

30 As described above, animal models of angiogenesis find use in screening for modulators of angiogenesis. Similarly, transgenic animal technology including gene knockout technology, for example as a result of homologous recombination with an appropriate gene targeting vector, or gene overexpression, will result in the absence or increased expression of the seryl-tRNA synthetase protein. The same technology can also be

applied to make knock-out cells. When desired, tissue-specific expression or knockout of the seryl-tRNA synthetase protein may be necessary. Transgenic animals generated by such methods find use as animal models of angiogenesis and are additionally useful in screening for modulators of angiogenesis.

5 Knock-out cells and transgenic mice can be made by insertion of a marker gene or other heterologous gene into the endogenous seryl-tRNA synthetase gene site in the mouse genome via homologous recombination. Such mice can also be made by substituting the endogenous seryl-tRNA synthetase with a mutated version of seryl-tRNA synthetase, or by mutating the endogenous seryl-tRNA synthetase, e.g., by exposure to carcinogens.

10 A DNA construct is introduced into the nuclei of embryonic stem cells. Cells containing the newly engineered genetic lesion are injected into a host mouse embryo, which is re-implanted into a recipient female. Some of these embryos develop into chimeric mice that possess germ cells partially derived from the mutant cell line. Therefore, by breeding the chimeric mice it is possible to obtain a new line of mice containing the introduced genetic
15 lesion (*see, e.g.,* Capecchi *et al., Science* 244:1288 (1989)). Chimeric targeted mice can be derived according to Hogan *et al., Manipulating the Mouse Embryo: A Laboratory Manual*, Cold Spring Harbor Laboratory (1988) and *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach*, Robertson, ed., IRL Press, Washington, D.C., (1987).

20 B. Modulators

The compounds tested as modulators of seryl-tRNA synthetase protein can be any small organic molecule, or a biological entity, such as a protein, e.g., an antibody or peptide, a sugar, a nucleic acid, e.g., an antisense oligonucleotide, RNAi, or a ribozyme, or a lipid. Alternatively, modulators can be genetically altered versions of a seryl-tRNA
25 synthetase protein. Typically, test compounds will be small organic molecules, peptides, lipids, and lipid analogs.

Essentially any chemical compound can be used as a potential modulator or ligand in the assays of the invention, although most often compounds can be dissolved in aqueous or organic (especially DMSO-based) solutions are used. The assays are designed to
30 screen large chemical libraries by automating the assay steps and providing compounds from any convenient source to assays, which are typically run in parallel (e.g., in microtiter formats on microtiter plates in robotic assays). It will be appreciated that there are many suppliers of chemical compounds, including Sigma (St. Louis, MO), Aldrich (St. Louis, MO), Sigma-

Aldrich (St. Louis, MO), Fluka Chemika-Biochemica Analytika (Buchs Switzerland) and the like.

In one preferred embodiment, high throughput screening methods involve providing a combinatorial small organic molecule or peptide library containing a large number of potential therapeutic compounds (potential modulator or ligand compounds). Such “combinatorial chemical libraries” or “ligand libraries” are then screened in one or more assays, as described herein, to identify those library members (particular chemical species or subclasses) that display a desired characteristic activity. The compounds thus identified can serve as conventional “lead compounds” or can themselves be used as potential or actual therapeutics.

A combinatorial chemical library is a collection of diverse chemical compounds generated by either chemical synthesis or biological synthesis, by combining a number of chemical “building blocks” such as reagents. For example, a linear combinatorial chemical library such as a polypeptide library is formed by combining a set of chemical building blocks (amino acids) in every possible way for a given compound length (i.e., the number of amino acids in a polypeptide compound). Millions of chemical compounds can be synthesized through such combinatorial mixing of chemical building blocks.

Preparation and screening of combinatorial chemical libraries is well known to those of skill in the art. Such combinatorial chemical libraries include, but are not limited to, peptide libraries (*see, e.g.*, U.S. Patent 5,010,175, Furka, *Int. J. Pept. Prot. Res.* 37:487-493 (1991) and Houghton *et al.*, *Nature* 354:84-88 (1991)). Other chemistries for generating chemical diversity libraries can also be used. Such chemistries include, but are not limited to: peptoids (e.g., PCT Publication No. WO 91/19735), encoded peptides (e.g., PCT Publication No. WO 93/20242), random bio-oligomers (e.g., PCT Publication No. WO 92/00091), benzodiazepines (e.g., U.S. Pat. No. 5,288,514), diversomers such as hydantoins, benzodiazepines and dipeptides (Hobbs *et al.*, *Proc. Nat. Acad. Sci. USA* 90:6909-6913 (1993)), vinylogous polypeptides (Hagihara *et al.*, *J. Amer. Chem. Soc.* 114:6568 (1992)), nonpeptidal peptidomimetics with glucose scaffolding (Hirschmann *et al.*, *J. Amer. Chem. Soc.* 114:9217-9218 (1992)), analogous organic syntheses of small compound libraries (Chen *et al.*, *J. Amer. Chem. Soc.* 116:2661 (1994)), oligocarbamates (Cho *et al.*, *Science* 261:1303 (1993)), and/or peptidyl phosphonates (Campbell *et al.*, *J. Org. Chem.* 59:658 (1994)), nucleic acid libraries (*see* Ausubel, Berger and Sambrook, all *supra*), peptide nucleic acid libraries (*see, e.g.*, U.S. Patent 5,539,083), antibody libraries (*see, e.g.*, Vaughn *et al.*, *Nature Biotechnology*, 14(3):309-314 (1996) and PCT/US96/10287), carbohydrate libraries (*see,*

e.g., Liang *et al.*, *Science*, 274:1520-1522 (1996) and U.S. Patent 5,593,853), small organic molecule libraries (*see, e.g.*, benzodiazepines, Baum C&EN, Jan 18, page 33 (1993); isoprenoids, U.S. Patent 5,569,588; thiazolidinones and metathiazanones, U.S. Patent 5,549,974; pyrrolidines, U.S. Patents 5,525,735 and 5,519,134; morpholino compounds, U.S. Patent 5,506,337; benzodiazepines, 5,288,514, and the like).

Devices for the preparation of combinatorial libraries are commercially available (*see, e.g.*, 357 MPS, 390 MPS, Advanced Chem Tech, Louisville KY, Symphony, Rainin, Woburn, MA, 433A Applied Biosystems, Foster City, CA, 9050 Plus, Millipore, Bedford, MA). In addition, numerous combinatorial libraries are themselves commercially available (*see, e.g.*, ComGenex, Princeton, N.J., Asinex, Moscow, Ru, Tripos, Inc., St. Louis, MO, ChemStar, Ltd, Moscow, RU, 3D Pharmaceuticals, Exton, PA, Martek Biosciences, Columbia, MD, etc.).

C. *Solid State and soluble high throughput assays*

In one embodiment the invention provides soluble assays using a seryl-tRNA synthetase protein, or a cell or tissue expressing a seryl-tRNA synthetase protein, either naturally occurring or recombinant. In another embodiment, the invention provides solid phase based *in vitro* assays in a high throughput format, where the seryl-tRNA synthetase protein is attached to a solid phase substrate. Any one of the assays described herein can be adapted for high throughput screening, e.g., ligand binding, cellular proliferation, cell surface marker flux, e.g., $\alpha\beta3$ integrin, phosphatase activity, etc.. In one preferred embodiment, the cell-based system using $\alpha\beta3$ integrin modulation and FACS assays is used in a high throughput format for identifying modulators of seryl-tRNA synthetase proteins, and therefore modulators of T cell angiogenesis.

In the high throughput assays of the invention, either soluble or solid state, it is possible to screen up to several thousand different modulators or ligands in a single day. This methodology can be used for seryl-tRNA synthetase proteins *in vitro*, or for cell-based assays comprising an seryl-tRNA synthetase protein. In particular, each well of a microtiter plate can be used to run a separate assay against a selected potential modulator, or, if concentration or incubation time effects are to be observed, every 5-10 wells can test a single modulator. Thus, a single standard microtiter plate can assay about 100 (e.g., 96) modulators. If 1536 well plates are used, then a single plate can easily assay from about 100- about 1500 different compounds. It is possible to assay many plates per day; assay screens for up to about 6,000,

20,000, 50,000, or more than 100,000 different compounds are possible using the integrated systems of the invention.

For a solid state reaction, the protein of interest or a fragment thereof, e.g., an extracellular domain, or a cell comprising the protein of interest or a fragment thereof as part of a fusion protein can be bound to the solid state component, directly or indirectly, via
5 covalent or non covalent linkage e.g., via a tag. The tag can be any of a variety of components. In general, a molecule which binds the tag (a tag binder) is fixed to a solid support, and the tagged molecule of interest is attached to the solid support by interaction of the tag and the tag binder.

10 A number of tags and tag binders can be used, based upon known molecular interactions well described in the literature. For example, where a tag has a natural binder, for example, biotin, protein A, or protein G, it can be used in conjunction with appropriate tag binders (avidin, streptavidin, neutravidin, the Fc region of an immunoglobulin, etc.) Antibodies to molecules with natural binders such as biotin are also widely available and
15 appropriate tag binders; *see*, SIGMA Immunochemicals 1998 catalogue SIGMA, St. Louis MO).

Similarly, any haptenic or antigenic compound can be used in combination with an appropriate antibody to form a tag/tag binder pair. Thousands of specific antibodies are commercially available and many additional antibodies are described in the literature.
20 For example, in one common configuration, the tag is a first antibody and the tag binder is a second antibody which recognizes the first antibody. In addition to antibody-antigen interactions, receptor-ligand interactions are also appropriate as tag and tag-binder pairs. For example, agonists and antagonists of cell membrane receptors (e.g., cell receptor-ligand interactions such as transferrin, c-kit, viral receptor ligands, cytokine receptors, chemokine
25 receptors, interleukin receptors, immunoglobulin receptors and antibodies, the cadherein family, the integrin family, the selectin family, and the like; *see, e.g.*, Pigott & Power, *The Adhesion Molecule Facts Book I* (1993). Similarly, toxins and venoms, viral epitopes, hormones (e.g., opiates, steroids, etc.), intracellular receptors (e.g. which mediate the effects of various small ligands, including steroids, thyroid hormone, retinoids and vitamin D;
30 peptides), drugs, lectins, sugars, nucleic acids (both linear and cyclic polymer configurations), oligosaccharides, proteins, phospholipids and antibodies can all interact with various cell receptors.

Synthetic polymers, such as polyurethanes, polyesters, polycarbonates, polyureas, polyamides, polyethyleneimines, polyarylene sulfides, polysiloxanes, polyimides,

and polyacetates can also form an appropriate tag or tag binder. Many other tag/tag binder pairs are also useful in assay systems described herein, as would be apparent to one of skill upon review of this disclosure.

Common linkers such as peptides, polyethers, and the like can also serve as tags, and include polypeptide sequences, such as poly gly sequences of between about 5 and 200 amino acids. Such flexible linkers are known to persons of skill in the art. For example, poly(ethylene glycol) linkers are available from Shearwater Polymers, Inc. Huntsville, Alabama. These linkers optionally have amide linkages, sulfhydryl linkages, or heterofunctional linkages.

Tag binders are fixed to solid substrates using any of a variety of methods currently available. Solid substrates are commonly derivatized or functionalized by exposing all or a portion of the substrate to a chemical reagent which fixes a chemical group to the surface which is reactive with a portion of the tag binder. For example, groups which are suitable for attachment to a longer chain portion would include amines, hydroxyl, thiol, and carboxyl groups. Aminoalkylsilanes and hydroxyalkylsilanes can be used to functionalize a variety of surfaces, such as glass surfaces. The construction of such solid phase biopolymer arrays is well described in the literature. *See, e.g., Merrifield, J. Am. Chem. Soc.* 85:2149-2154 (1963) (describing solid phase synthesis of, e.g., peptides); Geysen *et al., J. Immun. Meth.* 102:259-274 (1987) (describing synthesis of solid phase components on pins); Frank & Doring, *Tetrahedron* 44:60316040 (1988) (describing synthesis of various peptide sequences on cellulose disks); Fodor *et al., Science*, 251:767-777 (1991); Sheldon *et al., Clinical Chemistry* 39(4):718-719 (1993); and Kozal *et al., Nature Medicine* 2(7):753759 (1996) (all describing arrays of biopolymers fixed to solid substrates). Non-chemical approaches for fixing tag binders to substrates include other common methods, such as heat, cross-linking by UV radiation, and the like.

Antibodies to seryl-tRNA synthetase -angiogenesis polypeptides

In addition to the detection of seryl-tRNA synthetase gene and gene expression using nucleic acid hybridization technology, one can also use immunoassays to detect seryl-tRNA synthetase proteins of the invention. Such assays are useful for screening for modulators of seryl-tRNA synthetase and angiogenesis, as well as for therapeutic and diagnostic applications. Immunoassays can be used to qualitatively or quantitatively analyze seryl-tRNA synthetase protein. A general overview of the applicable technology can be found in Harlow & Lane, *Antibodies: A Laboratory Manual* (1988).

A. Production of antibodies

Methods of producing polyclonal and monoclonal antibodies that react specifically with the seryl-tRNA synthetase proteins are known to those of skill in the art (see, e.g., Coligan, *Current Protocols in Immunology* (1991); Harlow & Lane, *supra*; Goding, *Monoclonal Antibodies: Principles and Practice* (2d ed. 1986); and Kohler & Milstein, *Nature* 256:495-497 (1975). Such techniques include antibody preparation by selection of antibodies from libraries of recombinant antibodies in phage or similar vectors, as well as preparation of polyclonal and monoclonal antibodies by immunizing rabbits or mice (see, e.g., Huse *et al.*, *Science* 246:1275-1281 (1989); Ward *et al.*, *Nature* 341:544-546 (1989)).

A number of immunogens comprising portions of seryl-tRNA synthetase protein may be used to produce antibodies specifically reactive with seryl-tRNA synthetase protein. For example, recombinant seryl-tRNA synthetase protein or an antigenic fragment thereof, can be isolated as described herein. Recombinant protein can be expressed in eukaryotic or prokaryotic cells as described above, and purified as generally described above. Recombinant protein is the preferred immunogen for the production of monoclonal or polyclonal antibodies. Alternatively, a synthetic peptide derived from the sequences disclosed herein and conjugated to a carrier protein can be used as an immunogen. Naturally occurring protein may also be used either in pure or impure form. The product is then injected into an animal capable of producing antibodies. Either monoclonal or polyclonal antibodies may be generated, for subsequent use in immunoassays to measure the protein.

Methods of production of polyclonal antibodies are known to those of skill in the art. An inbred strain of mice (e.g., BALB/C mice) or rabbits is immunized with the protein using a standard adjuvant, such as Freund's adjuvant, and a standard immunization protocol. The animal's immune response to the immunogen preparation is monitored by taking test bleeds and determining the titer of reactivity to the beta subunits. When appropriately high titers of antibody to the immunogen are obtained, blood is collected from the animal and antisera are prepared. Further fractionation of the antisera to enrich for antibodies reactive to the protein can be done if desired (see, Harlow & Lane, *supra*).

Monoclonal antibodies may be obtained by various techniques familiar to those skilled in the art. Briefly, spleen cells from an animal immunized with a desired antigen are immortalized, commonly by fusion with a myeloma cell (see, Kohler & Milstein, *Eur. J. Immunol.* 6:511-519 (1976)). Alternative methods of immortalization include transformation with Epstein Barr Virus, oncogenes, or retroviruses, or other methods well

known in the art. Colonies arising from single immortalized cells are screened for production of antibodies of the desired specificity and affinity for the antigen, and yield of the monoclonal antibodies produced by such cells may be enhanced by various techniques, including injection into the peritoneal cavity of a vertebrate host. Alternatively, one may
5 isolate DNA sequences which encode a monoclonal antibody or a binding fragment thereof by screening a DNA library from human B cells according to the general protocol outlined by Huse, *et al.*, *Science* 246:1275-1281 (1989).

Monoclonal antibodies and polyclonal sera are collected and titered against the immunogen protein in an immunoassay, for example, a solid phase immunoassay with the
10 immunogen immobilized on a solid support. Typically, polyclonal antisera with a titer of 10^4 or greater are selected and tested for their cross reactivity against non-seryl-tRNA synthetase proteins, using a competitive binding immunoassay. Specific polyclonal antisera and monoclonal antibodies will usually bind with a K_d of at least about 0.1 mM, more usually at least about 1 μ M, preferably at least about 0.1 μ M or better, and most preferably, 0.01 μ M or
15 better. Antibodies specific only for a particular seryl-tRNA synthetase ortholog, such as human seryl-tRNA synthetase, can also be made, by subtracting out other cross-reacting orthologs from a species such as a non-human mammal. In this manner, antibodies that bind only to seryl-tRNA synthetase protein may be obtained.

Once the specific antibodies against seryl-tRNA synthetase protein are
20 available, the protein can be detected by a variety of immunoassay methods. In addition, the antibody can be used therapeutically as a seryl-tRNA synthetase modulators. For a review of immunological and immunoassay procedures, see *Basic and Clinical Immunology* (Stites & Terr eds., 7th ed. 1991). Moreover, the immunoassays of the present invention can be performed in any of several configurations, which are reviewed extensively in *Enzyme*
25 *Immunoassay* (Maggio, ed., 1980); and Harlow & Lane, *supra*.

B. Immunological binding assays

Seryl-tRNA synthetase protein can be detected and/or quantified using any of a number of well recognized immunological binding assays (*see, e.g.*, U.S. Patents 4,366,241;
30 4,376,110; 4,517,288; and 4,837,168). For a review of the general immunoassays, see also *Methods in Cell Biology: Antibodies in Cell Biology*, volume 37 (Asai, ed. 1993); *Basic and Clinical Immunology* (Stites & Terr, eds., 7th ed. 1991). Immunological binding assays (or immunoassays) typically use an antibody that specifically binds to a protein or antigen of choice (in this case the seryl-tRNA synthetase protein or antigenic subsequence thereof). The

antibody (e.g., anti-seryl-tRNA synthetase) may be produced by any of a number of means well known to those of skill in the art and as described above.

Immunoassays also often use a labeling agent to specifically bind to and label the complex formed by the antibody and antigen. The labeling agent may itself be one of the moieties comprising the antibody/antigen complex. Thus, the labeling agent may be a labeled seryl-tRNA synthetase or a labeled anti-seryl-tRNA synthetase antibody. Alternatively, the labeling agent may be a third moiety, such a secondary antibody, that specifically binds to the antibody/seryl-tRNA synthetase complex (a secondary antibody is typically specific to antibodies of the species from which the first antibody is derived). Other proteins capable of specifically binding immunoglobulin constant regions, such as protein A or protein G may also be used as the label agent. These proteins exhibit a strong non-immunogenic reactivity with immunoglobulin constant regions from a variety of species (*see, e.g., Kronval et al., J. Immunol.* 111:1401-1406 (1973); *Akerstrom et al., J. Immunol.* 135:2589-2542 (1985)). The labeling agent can be modified with a detectable moiety, such as biotin, to which another molecule can specifically bind, such as streptavidin. A variety of detectable moieties are well known to those skilled in the art.

Throughout the assays, incubation and/or washing steps may be required after each combination of reagents. Incubation steps can vary from about 5 seconds to several hours, optionally from about 5 minutes to about 24 hours. However, the incubation time will depend upon the assay format, antigen, volume of solution, concentrations, and the like. Usually, the assays will be carried out at ambient temperature, although they can be conducted over a range of temperatures, such as 10°C to 40°C.

Non-competitive assay formats

Immunoassays for detecting seryl-tRNA synthetase in samples may be either competitive or noncompetitive. Noncompetitive immunoassays are assays in which the amount of antigen is directly measured. In one preferred "sandwich" assay, for example, the anti-seryl-tRNA synthetase antibodies can be bound directly to a solid substrate on which they are immobilized. These immobilized antibodies then capture seryl-tRNA synthetase present in the test sample. seryl-tRNA synthetase proteins thus immobilized are then bound by a labeling agent, such as a second seryl-tRNA synthetase antibody bearing a label. Alternatively, the second antibody may lack a label, but it may, in turn, be bound by a labeled third antibody specific to antibodies of the species from which the second antibody is

derived. The second or third antibody is typically modified with a detectable moiety, such as biotin, to which another molecule specifically binds, e.g., streptavidin, to provide a detectable moiety.

5 Competitive assay formats

In competitive assays, the amount of seryl-tRNA synthetase protein present in the sample is measured indirectly by measuring the amount of a known, added (exogenous) seryl-tRNA synthetase protein displaced (competed away) from an anti-seryl-tRNA synthetase antibody by the unknown seryl-tRNA synthetase protein present in a sample. In
10 one competitive assay, a known amount of seryl-tRNA synthetase protein is added to a sample and the sample is then contacted with an antibody that specifically binds to seryl-tRNA synthetase protein. The amount of exogenous seryl-tRNA synthetase protein bound to the antibody is inversely proportional to the concentration of seryl-tRNA synthetase protein present in the sample. In a particularly preferred embodiment, the antibody is immobilized
15 on a solid substrate. The amount of seryl-tRNA synthetase protein bound to the antibody may be determined either by measuring the amount of seryl-tRNA synthetase present in seryl-tRNA synthetase protein/antibody complex, or alternatively by measuring the amount of remaining uncomplexed protein. The amount of seryl-tRNA synthetase protein may be detected by providing a labeled seryl-tRNA synthetase molecule.

20 A hapten inhibition assay is another preferred competitive assay. In this assay the known seryl-tRNA synthetase protein is immobilized on a solid substrate. A known amount of anti-seryl-tRNA synthetase antibody is added to the sample, and the sample is then contacted with the immobilized seryl-tRNA synthetase. The amount of anti-seryl-tRNA synthetase antibody bound to the known immobilized seryl-tRNA synthetase is inversely
25 proportional to the amount of seryl-tRNA synthetase protein present in the sample. Again, the amount of immobilized antibody may be detected by detecting either the immobilized fraction of antibody or the fraction of the antibody that remains in solution. Detection may be direct where the antibody is labeled or indirect by the subsequent addition of a labeled moiety that specifically binds to the antibody as described above.

30 Cross-reactivity determinations

Immunoassays in the competitive binding format can also be used for crossreactivity determinations. For example, a seryl-tRNA synthetase protein can be immobilized to a solid support. Proteins (e.g., seryl-tRNA synthetase and homologs) are

added to the assay that compete for binding of the antisera to the immobilized antigen. The ability of the added proteins to compete for binding of the antisera to the immobilized protein is compared to the ability of the seryl-tRNA synthetase protein to compete with itself. The percent crossreactivity for the above proteins is calculated, using standard calculations.

5 Those antisera with less than 10% crossreactivity with each of the added proteins listed above are selected and pooled. The cross-reacting antibodies are optionally removed from the pooled antisera by immunoabsorption with the added considered proteins, e.g., distantly related homologs.

The immunoabsorbed and pooled antisera are then used in a competitive
10 binding immunoassay as described above to compare a second protein, thought to be perhaps an allele or polymorphic variant of a seryl-tRNA synthetase protein, to the immunogen protein. In order to make this comparison, the two proteins are each assayed at a wide range of concentrations and the amount of each protein required to inhibit 50% of the binding of the antisera to the immobilized protein is determined. If the amount of the second protein
15 required to inhibit 50% of binding is less than 10 times the amount of the seryl-tRNA synthetase protein that is required to inhibit 50% of binding, then the second protein is said to specifically bind to the polyclonal antibodies generated to seryl-tRNA synthetase immunogen.

20 Other assay formats

Western blot (immunoblot) analysis is used to detect and quantify the presence of seryl-tRNA synthetase in the sample. The technique generally comprises separating sample proteins by gel electrophoresis on the basis of molecular weight, transferring the separated proteins to a suitable solid support, (such as a nitrocellulose filter, a nylon filter, or
25 derivatized nylon filter), and incubating the sample with the antibodies that specifically bind seryl-tRNA synthetase. The anti-seryl-tRNA synthetase antibodies specifically bind to the seryl-tRNA synthetase on the solid support. These antibodies may be directly labeled or alternatively may be subsequently detected using labeled antibodies (e.g., labeled sheep anti-mouse antibodies) that specifically bind to the anti-seryl-tRNA synthetase antibodies.

30 Other assay formats include liposome immunoassays (LIA), which use liposomes designed to bind specific molecules (e.g., antibodies) and release encapsulated reagents or markers. The released chemicals are then detected according to standard techniques (*see Monroe et al., Amer. Clin. Prod. Rev. 5:34-41 (1986)*).

Reduction of non-specific binding

One of skill in the art will appreciate that it is often desirable to minimize non-specific binding in immunoassays. Particularly, where the assay involves an antigen or antibody immobilized on a solid substrate it is desirable to minimize the amount of non-specific binding to the substrate. Means of reducing such non-specific binding are well known to those of skill in the art. Typically, this technique involves coating the substrate with a proteinaceous composition. In particular, protein compositions such as bovine serum albumin (BSA), nonfat powdered milk, and gelatin are widely used with powdered milk being most preferred.

10

Labels

The particular label or detectable group used in the assay is not a critical aspect of the invention, as long as it does not significantly interfere with the specific binding of the antibody used in the assay. The detectable group can be any material having a detectable physical or chemical property. Such detectable labels have been well-developed in the field of immunoassays and, in general, most any label useful in such methods can be applied to the present invention. Thus, a label is any composition detectable by spectroscopic, photochemical, biochemical, immunochemical, electrical, optical or chemical means. Useful labels in the present invention include magnetic beads (e.g., DYNABEADSTM), fluorescent dyes (e.g., fluorescein isothiocyanate, Texas red, rhodamine, and the like), radiolabels (e.g., ³H, ¹²⁵I, ³⁵S, ¹⁴C, or ³²P), enzymes (e.g., horse radish peroxidase, alkaline phosphatase and others commonly used in an ELISA), and colorimetric labels such as colloidal gold or colored glass or plastic beads (e.g., polystyrene, polypropylene, latex, etc.).

15

20

The label may be coupled directly or indirectly to the desired component of the assay according to methods well known in the art. As indicated above, a wide variety of labels may be used, with the choice of label depending on sensitivity required, ease of conjugation with the compound, stability requirements, available instrumentation, and disposal provisions.

25

Non-radioactive labels are often attached by indirect means. Generally, a ligand molecule (e.g., biotin) is covalently bound to the molecule. The ligand then binds to another molecules (e.g., streptavidin) molecule, which is either inherently detectable or covalently bound to a signal system, such as a detectable enzyme, a fluorescent compound, or a chemiluminescent compound. The ligands and their targets can be used in any suitable

combination with antibodies that recognize seryl-tRNA synthetase protein, or secondary antibodies that recognize anti-seryl-tRNA synthetase.

The molecules can also be conjugated directly to signal generating compounds, e.g., by conjugation with an enzyme or fluorophore. Enzymes of interest as labels will primarily be hydrolases, particularly phosphatases, esterases and glycosidases, or oxidotases, particularly peroxidases. Fluorescent compounds include fluorescein and its derivatives, rhodamine and its derivatives, dansyl, umbelliferone, etc. Chemiluminescent compounds include luciferin, and 2,3-dihydrophthalazinediones, e.g., luminol. For a review of various labeling or signal producing systems that may be used, see U.S. Patent No. 4,391,904.

Means of detecting labels are well known to those of skill in the art. Thus, for example, where the label is a radioactive label, means for detection include a scintillation counter or photographic film as in autoradiography. Where the label is a fluorescent label, it may be detected by exciting the fluorochrome with the appropriate wavelength of light and detecting the resulting fluorescence. The fluorescence may be detected visually, or by the use of electronic detectors such as charge coupled devices (CCDs) or photomultipliers and the like. Similarly, enzymatic labels may be detected by providing the appropriate substrates for the enzyme and detecting the resulting reaction product. Finally simple colorimetric labels may be detected simply by observing the color associated with the label. Thus, in various dipstick assays, conjugated gold often appears pink, while various conjugated beads appear the color of the bead.

Some assay formats do not require the use of labeled components. For instance, agglutination assays can be used to detect the presence of the target antibodies. In this case, antigen-coated particles are agglutinated by samples comprising the target antibodies. In this format, none of the components need be labeled and the presence of the target antibody is detected by simple visual inspection.

Gene therapy

The present invention provides the nucleic acids of seryl-tRNA synthetase-angiogenesis associated protein for the transfection of cells *in vitro* and *in vivo*. These nucleic acids can be inserted into any of a number of well-known vectors for the transfection of target cells and organisms as described below. The nucleic acids are transfected into cells, *ex vivo* or *in vivo*, through the interaction of the vector and the target cell. The nucleic acid, under the control of a promoter, then expresses a seryl-tRNA synthetase protein of the

present invention, thereby mitigating the effects of absent, partial inactivation, or abnormal expression of the seryl-tRNA synthetase gene, particularly as it relates to angiogenesis. The compositions are administered to a patient in an amount sufficient to elicit a therapeutic response in the patient. An amount adequate to accomplish this is defined as “therapeutically effective dose or amount.”

Such gene therapy procedures have been used to correct acquired and inherited genetic defects, cancer, and other diseases in a number of contexts. The ability to express artificial genes in humans facilitates the prevention and/or cure of many important human diseases, including many diseases which are not amenable to treatment by other therapies (for a review of gene therapy procedures, *see* Anderson, *Science* 256:808-813 (1992); Nabel & Felgner, *TIBTECH* 11:211-217 (1993); Mitani & Caskey, *TIBTECH* 11:162-166 (1993); Mulligan, *Science* 926-932 (1993); Dillon, *TIBTECH* 11:167-175 (1993); Miller, *Nature* 357:455-460 (1992); Van Brunt, *Biotechnology* 6(10):1149-1154 (1998); Vigne, *Restorative Neurology and Neuroscience* 8:35-36 (1995); Kremer & Perricaudet, *British Medical Bulletin* 51(1):31-44 (1995); Haddada *et al.*, in *Current Topics in Microbiology and Immunology* (Doerfler & Böhm eds., 1995); and Yu *et al.*, *Gene Therapy* 1:13-26 (1994)).

The nucleic acids of the invention can also be used to make transgenic animals, such as transgenic mice, either by knock-out or overexpression. Such animals are useful, e.g., for testing modulators of angiogenesis.

Pharmaceutical compositions and administration

Pharmaceutically acceptable carriers are determined in part by the particular composition being administered (e.g., nucleic acid, protein, modulatory compounds or transduced cell), as well as by the particular method used to administer the composition. Accordingly, there are a wide variety of suitable formulations of pharmaceutical compositions of the present invention (*see, e.g., Remington's Pharmaceutical Sciences*, 17th ed., 1989). Administration can be in any convenient manner, e.g., by injection, oral administration, inhalation, transdermal application, or rectal administration.

Formulations suitable for oral administration can consist of (a) liquid solutions, such as an effective amount of the packaged nucleic acid suspended in diluents, such as water, saline or PEG 400; (b) capsules, sachets or tablets, each containing a predetermined amount of the active ingredient, as liquids, solids, granules or gelatin; (c) suspensions in an appropriate liquid; and (d) suitable emulsions. Tablet forms can include

one or more of lactose, sucrose, mannitol, sorbitol, calcium phosphates, corn starch, potato starch, microcrystalline cellulose, gelatin, colloidal silicon dioxide, talc, magnesium stearate, stearic acid, and other excipients, colorants, fillers, binders, diluents, buffering agents, moistening agents, preservatives, flavoring agents, dyes, disintegrating agents, and
5 pharmaceutically compatible carriers. Lozenge forms can comprise the active ingredient in a flavor, e.g., sucrose, as well as pastilles comprising the active ingredient in an inert base, such as gelatin and glycerin or sucrose and acacia emulsions, gels, and the like containing, in addition to the active ingredient, carriers known in the art.

The compound of choice, alone or in combination with other suitable
10 components, can be made into aerosol formulations (i.e., they can be “nebulized”) to be administered via inhalation. Aerosol formulations can be placed into pressurized acceptable propellants, such as dichlorodifluoromethane, propane, nitrogen, and the like.

Formulations suitable for parenteral administration, such as, for example, by
15 intraarticular (in the joints), intravenous, intramuscular, intradermal, intraperitoneal, and subcutaneous routes, include aqueous and non-aqueous, isotonic sterile injection solutions, which can contain antioxidants, buffers, bacteriostats, and solutes that render the formulation isotonic with the blood of the intended recipient, and aqueous and non-aqueous sterile suspensions that can include suspending agents, solubilizers, thickening agents, stabilizers, and preservatives. In the practice of this invention, compositions can be administered, for
20 example, by intravenous infusion, orally, topically, intraperitoneally, intravesically or intrathecally. Parenteral administration and intravenous administration are the preferred methods of administration. The formulations of commends can be presented in unit-dose or multi-dose sealed containers, such as ampules and vials.

Injection solutions and suspensions can be prepared from sterile powders,
25 granules, and tablets of the kind previously described. Cells transduced by nucleic acids for *ex vivo* therapy can also be administered intravenously or parenterally as described above.

The dose administered to a patient, in the context of the present invention should be sufficient to effect a beneficial therapeutic response in the patient over time. The dose will be determined by the efficacy of the particular vector employed and the condition
30 of the patient, as well as the body weight or surface area of the patient to be treated. The size of the dose also will be determined by the existence, nature, and extent of any adverse side-effects that accompany the administration of a particular vector, or transduced cell type in a particular patient.

In determining the effective amount of the vector to be administered in the treatment or prophylaxis of conditions owing to diminished or aberrant expression of the seryl-tRNA synthetase protein, the physician evaluates circulating plasma levels of the vector, vector toxicities, progression of the disease, and the production of anti-vector antibodies. In general, the dose equivalent of a naked nucleic acid from a vector is from about 1 μg to 100 μg for a typical 70 kilogram patient, and doses of vectors which include a retroviral particle are calculated to yield an equivalent amount of therapeutic nucleic acid.

For administration, compounds and transduced cells of the present invention can be administered at a rate determined by the LD-50 of the inhibitor, vector, or transduced cell type, and the side-effects of the inhibitor, vector or cell type at various concentrations, as applied to the mass and overall health of the patient. Administration can be accomplished via single or divided doses.

All publications and patent applications cited in this specification are herein incorporated by reference as if each individual publication or patent application were specifically and individually indicated to be incorporated by reference.

Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, it will be readily apparent to one of ordinary skill in the art in light of the teachings of this invention that certain changes and modifications may be made thereto without departing from the spirit or scope of the appended claims.

EXAMPLES

The following example is provided by way of illustration only and not by way of limitation. Those of skill in the art will readily recognize a variety of noncritical parameters that could be changed or modified to yield essentially similar results.

Example 1: Identification of a gene involved in modulation of angiogenesis

A genetic screening approach was designed to identify genes involved in regulating angiogenesis. cDNA libraries constructed in retroviral vectors were transduced into early passage endothelial cells. Cell clones were isolated, which displayed a phenotype that correlated with downregulation of angiogenesis *in vivo* (i.e., downregulation of the cell surface marker $\alpha\text{v}\beta\text{3}$ integrin). The loss-of-angiogenesis phenotype was demonstrated to be

dependent on a retrovirally-encoded gene by a phenotypic transfer assay. A candidate retrovirally-encoded gene sequence encoding seryl-tRNA synthetase was recovered by PCR. The sequence of the candidate gene sequence is provided in Figure 2.

WHAT IS CLAIMED IS:

- 1 1. A method for identifying a compound that modulates angiogenesis,
2 the method comprising the steps of:
3 (i) contacting the compound with a seryl-tRNA synthetase polypeptide, the
4 polypeptide encoded by a nucleic acid that hybridizes under stringent conditions to a
5 nucleic acid encoding a polypeptide comprising an amino acid sequence of SEQ ID
6 NO:2; and
7 (ii) determining the functional effect of the compound upon the seryl-
8 tRNA synthetase polypeptide.
- 1 2. The method of claim 1, wherein the functional effect is determined
2 *in vitro*.
- 1 3. The method of claim 2, wherein the functional effect is a physical
2 effect.
- 1 4. The method of claim 2, wherein the functional effect is determined
2 by measuring ligand or substrate binding to the polypeptide.
- 1 5. The method of claim 2, wherein the functional effect is a chemical
2 effect.
- 1 6. The method of claim 5, wherein the functional effect is determined
2 by measuring aminoacyl-tRNA synthetase activity of the polypeptide.
- 1 7. The method of claim 1, wherein the polypeptide is expressed in a
2 eukaryotic host cell.
- 1 8. The method of claim 7, wherein the functional effect is a physical
2 effect.
- 1 9. The method of claim 8, wherein the functional effect is determined
2 by measuring ligand or substrate binding to the polypeptide.
- 1 10. The method of claim 7, wherein the functional effect is a chemical
2 or phenotypic effect.

- 1 11. The method of claim 7, wherein the host cell is an endothelial cell.
- 1 12. The method of claim 11, wherein the functional effect is
2 determined by measuring $\alpha v\beta 3$ expression, haptotaxis, or chemotaxis.
- 1 13 . The method of claim 1, wherein modulation is inhibition of
2 angiogenesis.
- 1 14. The method of claim 1 wherein the polypeptide is recombinant.
- 1 15. The method of claim 1, wherein the polypeptide comprises a
2 sequence of SEQ ID NO:2.
- 1 16. The method of claim 1, wherein the compound is an antibody.
- 1 17. The method of claim 1, wherein the compound is an antisense
2 molecule.
- 1 18. The method of claim 1, wherein the compound is a small organic
2 molecule.
- 1 19 . The method of claim 1, wherein the compound is an RNAi
2 molecule.
- 1 20. A method of modulating angiogenesis in a subject, the method
2 comprising the step of administering to the subject a therapeutically effective amount of a
3 compound identified using the method of claim 1.
- 1 21. The method of claim 20, wherein the subject is a human.
- 1 22. The method of claim 20, wherein the compound is an antibody.
- 1 23. The method of claim 20, wherein the compound is an antisense
2 molecule.
- 1 24. The method of claim 20, wherein the compound is a small organic
2 molecule.

1 25. The method of claim 20, wherein the compound is an RNAi
2 molecule.

1 26. The method of claim 20, wherein the compound inhibits
2 angiogenesis.

1 27. A method of modulating angiogenesis in a subject, the method
2 comprising the step of administering to the subject a therapeutically effective amount of a
3 seryl-tRNA synthetase polypeptide, the polypeptide encoded by a nucleic acid that
4 hybridizes under stringent conditions to a nucleic acid encoding a polypeptide comprising
5 an amino acid sequence of SEQ ID NO:2.

1 28. A method of modulating angiogenesis in a subject, the method
2 comprising the step of administering to the subject a therapeutically effective amount of a
3 nucleic acid encoding a seryl-tRNA synthetase polypeptide, wherein the nucleic acid
4 hybridizes under stringent conditions to a nucleic acid encoding a polypeptide comprising an
5 amino acid sequence of SEQ ID NO:2.

```

FEATURES                                 Location/Qualifiers
  source                                 1..1942
                                          /organism="Homo sapiens"
                                          /mol_type="mRNA"
                                          /db_xref="taxon:9606"
                                          /chromosome="1"
                                          /map="1p13.3-p13.1"
  gene                                   1..1942
                                          /gene="SARS"
                                          /note="synonyms: SERS, SERRS"
                                          /db_xref="LocusID:6301"
                                          /db_xref="MIM:607529"
  CDS                                    76..1620
                                          /gene="SARS"
                                          /EC_number="6.1.1.11"
                                          /note="serine-tRNA ligase"
                                          /codon_start=1
                                          /product="seryl-tRNA synthetase"
                                          /protein_id="NP_006504.2"
                                          /db_xref="GI:16306548"
                                          /db_xref="LocusID:6301"
                                          /db_xref="MIM:607529"
                                          /translation="MVLDDLFRVDKGGDPALIRETQEKRFKDPGLVDQLVKADSEWR
RCRFRADNLNKLNLCSTIGEKMKKKEPVGDDSVPENVLVSFDDLADALANLKVSO
IKKVRLLIDEAILKCAERIKLEAERFENLREIGNLLHPSVPI SNDEDVDNKVERIWG
DCTVRKKYSHVDLVVMVDGFEGEKGAVVAGSRGYFLKGVLVFLEQALIQYALRTLGSR
GYIPIYTPFFMRKEVMQEVQLSQFDEELYKVIKKGSEKSDDNSYDEKYLIIATSEQPI
AALHRDEWLRPEDLP IKYAGLSTCFRQEVGSHGRDTRGIFRVHQFEKIEQFVYSSPHD
NKSWEEMFEEMITTAEEFYQSLGIPYHIVNIVSGSLNHAASKKLDLEAWFPGSGAFREL
VSCSNCTDYQARRLRIRYQTKKMMMDKVEFVHMLNATMCATTRTICAI LENYQTEKGI
TVPEKLEKFMPPGLQELIPFVKPAPIEQEPSKKQKQKQHEGSKKKAARDVTLENRLQN
MEVTDA"
  misc feature                           82..1482
                                          /gene="SARS"
                                          /note="Region: Seryl-tRNA synthetase [Translation,
ribosomal structure and biogenesis]"
                                          /db_xref="CDD:COG0172"
  polyA signal                           1831..1836
                                          /gene="SARS"
  polyA site                             1852
                                          /gene="SARS"
                                          /evidence=experimental
  polyA site                             1895
                                          /gene="SARS"
                                          /evidence=experimental

```

FIG. 1

BASE COUNT	530 a	460 c	538 g	414 t				
ORIGIN								
1	gcagtgcggc	ggtcacaggc	tgagtgctgc	ggcgcgatcc	ttgcttccct	gagcgttggc		
61	ccgggaggaa	agaagatggt	gctggatctg	gatttgtttc	gggtggataa	aggaggggac		
121	ccagccctca	tccgagagac	gcaggagaag	cgcttcaagg	accgaggact	agtggaccag		
181	ctggtgaagg	cagacagcga	gtggcgacga	tgtagatttc	gggcagacaa	cttgaacaag		
241	ctgaagaacc	tatgcagcaa	gacaatcgga	gagaaaatga	agaaaaaaga	gccagtggga		
301	gatgatgagt	ctgtcccaga	gaatgtgctg	agtttcgatg	accttactgc	agacgcttta		
361	gctaacctga	aagtctcaca	aatcaaaaaa	gtccgactcc	tcattgatga	agccatcctg		
421	aagtgtgacg	cggagcggat	aaagttggaa	gcagagcggg	ttgagaacct	ccgagagatt		
481	gggaaccttc	tgcacccttc	tgtaccatc	agtaacgatg	aggatgtgga	caacaaagta		
541	gagaggattt	ggggtgattg	tacagtcagg	aagaagtact	ctcatgtgga	cctggtgggtg		
601	atggtagatg	gctttgaagg	cgaaaagggg	gccgtgggtg	ctgggagtgc	agggtacttc		
661	ttgaaggggg	tcctggtggt	cctggaacag	gctctcatcc	agtatgccct	tcgcaccttg		
721	ggaagtcggg	gctacattcc	catttatacc	ccctttttca	tgaggaagga	ggtcatgcag		
781	gaggtggcac	agctcagcca	gtttgatgaa	gaactttata	aggatgattg	caaaggcagt		
841	gaaaagtctg	atgacaactc	ctatgatgag	aagtacctga	ttgccacctc	agagcagccc		
901	attgctgccc	tgcaccggga	tgagtggctc	cggccggagg	acctgcccac	caagtatgct		
961	ggcctgtcta	cctgcttccg	tcaggagggtg	ggctcccatg	gccgtgacac	ccgtggcatc		
1021	ttccgagtcc	atcagtttga	gaagattgaa	cagtttgtgt	actcatcacc	ccatgacaac		
1081	aagtcatggg	agatgtttga	agagatgatt	accaccgcag	aggagtctta	ccagtccctg		
1141	gggattcctt	accacattgt	gaatattgtc	tcaggttctt	tgaatcatgc	tgccagtaag		
1201	aagcttgacc	tggaggcctg	gtttccgggc	tcaggagcct	tcctgtagtt	ggtctcctgt		
1261	tctaattgca	cggattacca	ggctcgccgg	cttcgaatcc	gatatgggca	aaccaagaag		
1321	atgatggaca	agggtggagt	tgtccatgat	ctcaatgcta	ccatgtgcgc	cactaccctg		
1381	accatctgcg	ccatcctgga	gaactaccag	acagagaagg	gcatcactgt	gcctgagaaa		
1441	ttgaaggagt	tcattgccc	aggactgcaa	gaactgatcc	cccttgtgaa	gcctgccc		
1501	attgagcagg	agccatcaaa	gaagcagaag	aagcaacatg	agggcagcaa	aaagaaagca		
1561	gcagcaagag	acgtcaccct	agaaaacagg	ctgcagaaca	tggaggtcac	cgatgcttga		
1621	acattcctgc	ctcctatctt	gccaggcttt	catttctgtc	tgctgagatc	tcagagcctg		
1681	cccaacagca	gggaagccaa	gcacccatcc	atccccctgc	ccccatctga	ctgcgtagct		
1741	gagaggggaa	cagtgccatg	taccacacag	atgttctctg	ctcctcgcat	gggcataggg		
1801	accatcatt	gatgactgat	gaaaccatgt	aataaagcat	ctctggggag	ggcttaggac		
1861	tcttctcag	tcttcttccc	cgggcttgaa	ccccgaaaaa	aaaaaaaaaa	aaaaaaaaaa		
1921	aaaaaaaaaa	aaaaaaaaaa	aa					

FIG. 1 (cont'd)

cacaggctgagtgcggcgatccttggcttccctgagcgttggccccggaggaaagaag ATG GTG CTG GAT CTG GAT TTG TTT CGG 89
M V L D L D L F R 9
GTG GAT AAA GGA GGG GAC CCA GCC CTC ATC CGA GAG ACG CAG GAG AAG CGC TTC AAG GAC CCG GGA CTA GTG GAC 164
V D K G G D P A L I R E T Q E K R F K D P G L V D 34
CAG CTG GTG AAG GCA GAC AGC GAG TGG CGA CGA TGT AGA TTT CGG GCA GAC AAC TTG AAC AAG CTG AAG AAC CTA 239
Q L V K A D S E W R R C R F R A D N L N K L K N L 59
TGC AGC AAG ACA ATC GGA GAG AAA ATG AAG AAA GAG CCA CTG GGA GAT GAT GAT TCT GTC CCA GAG AAT GTG 314
C S K T I G E K M K K E P V G D E S V P E N V 84
CTG AGT TTC GAT GAC CTT ACT GCA GAC GCT TTA GCT AAC CTG AAA GTC TCA CAA ATC AAA AAA GTC CGA CTC CTC 389
L S F D D L T A D A L A N L K V S Q I K K V R L L 109
ATT GAT GAA GCC ATC CTG AAG TGT GAC GCG GAG CGG ATA AAG TTG GAA GCA GAG CGG TTT GAG AAC CTC CGA GAG 464
I D E A I L K C D A E R I K L E A E R F E N L R E 134
ATT GGG AAC CTT CTG CAC CCT TCT GTA CCC ATC AGT AAC GAT GAT GTG GAC AAC AAA GTA GAG AGG ATT TGG 539
I G N L L H P S V P I S N D E D V D N K V E R I W 159
GGT GAT TGT ACA GTC AGG AAG AAC TAC TCT CAT GTG GAC CTG GTG ATG GTA GAT GGC TTT GAA GGC GAA AAG 614
G D C T V R K K Y S H V D L V V M V D G F E G E K 184
GGG GCC GTG GCT GGG AGT CGA GGG TAC TTC TTG AAG GGG GTC CTG GTG TTC CTG GAA CAG GCT CTC ATC CAG 689
G A V V A G S R G Y F L K G V L V F L E Q A L I Q 209
TAT GCC CTT CGC ACC TTG GGA AGT CGG GGC TAC ATT CCC ATT TAT ACC CCC TTT TTC ATG AGG AAG GAG GTC ATG 764
Y A L R T L G S R G Y I P I Y T P F M R K E V M 234
CAG GAG GTG GCA CAG CTC AGC CAG TTT GAT GAA GAA CTT TAT AAG GTG ATT GGC AAA GGC AGT GAA AAG TCT GAT 839
Q E V A Q L S Q F D E E L Y K V I G K G S E K S D 259
GAC AAC TCC TAT GAT GAG AAG TAC CTG ATT GCC ACC TCA GAG CAG CCC ATT GCT GCC CTG CAC CGG GAT GAG TGG 914
D N S Y D E K Y L I A T S E Q P I A A L H R D E W 284
CTC CGG CCG GAG GAC CTC CCC ATC AAG TAT GCT GGC CTG TCT ACC TGC TTC CGT CAG GAG GTG GGC TCC CAT GGC 989
L R P E D L P I K Y A G L S T C F R Q E V G S H G 309

FIG. 2

CGT GAC ACC CGT GGC ATC TTC CGA GTC CAT CAG TTT GAG AAG ATT GAA CAG TTT GTG TAC TCA TCA CCC CAT GAC 1064
R D T R G I F R V H Q F E K I E Q F V Y S S S P H D 334
AAC AAG TCA TGG GAG ATG TTT GAA GAG ATG ATT ACC ACC GCA GAG GAG TTC TAC CAG TCC CTG GGG ATT CCT TAC 1139
N K S W E M F E E M I T T A E E F Y Q S L G I P Y 359
CAC ATT GTG AAT ATT GTC TCA GGT TCT TTG AAT CAT GCT GCC AGT AAG AAG CTT GAC CTG GAG GCC TGG TTT CCG 1214
H I V N I V S G S L N H A A S K K L D L E A W F P 384
GGC TCA GGA GCC TTC CGT GAG TTG GTC TCC TGT TCT AAT TGC ACG GAT TAC CAG GCT CGC CGG CTT CGA ATC CGA 1289
G S G A F R E L V S C S N C T D Y Q A R R L R I R 409
TAT GGG CAA ACC AAG AAG ATG ATG GAC AAG GTG GAG TTT GTC CAT ATG CTC AAT GCT ACC ATG TGC GCC ACT ACC 1364
Y G Q T K K M D K V E F V H M L N A T M C A T T 434
CGT ACC ATC TGC GCC ATC CTG GAG AAC TAC CAG ACA GAG AAG GGC ATC ACT GTG CCT GAG AAA TTG AAG GAG TTC 1439
R T I C A I L E N Y Q T E K G I T V P E K L K E F 459
ATG CCG CCA GGA CTG CAA GAA CTG ATC CCC TTT GTG AAG CCT GCG CCC ATT GAG CAG CCA TCA AAG AAG CAG 1514
M P P G L Q E L I P F V K P A P I E Q E P S K K Q 484
AAG AAG CAA CAT GAG GGC AGC AAA AAG AAA GCA GCA AGA GAC GTC ACC CTA GAA AAC AGG CTG CAG AAC ATG 1589
K K Q H E G S K K A A A R D V T L E N R L Q N M 509
GAG GTC ACC GAT GCT TGA acatttcctgcctccctatttccaggctttcatttctgtctgtcagagcctgccccacagcagggaag 1683
E V T D A *515

ccaagcaccattcatccccctgccccatctgactgctgagaggggaaacagtgccatgtaccacacagatgttcctgtctcctcgcgatgggcat 1783
agggacccatcattgatgactgatgaaaccatgtaataaagcatttttgggggggaaaccccccccccccccccccccccccccccccccc 1883
aa

FIG. 2 (cont'd)