METHODS AND COMPOSITIONS FOR TREATING HEMATOLOGICAL DISORDERS USING 252, 304, 19870, 14717, 9941, 19310 and 17832

The present invention relates to methods for the diagnosis and treatment of hematological disorders. Specifically, the present invention identifies the differential expression of 252, 304, 1980, 14717, 9941, 19310, OR 17832 genes in tissues relating to hematological disorders disease states, and/or in response to manipulations relevant to evaluation and prognosis of various hematological disorders, and for the identification of subjects exhibiting a predisposition to such conditions. The invention also provides methods for identifying a compound capable of modulating hematological disorders. The present invention also provides methods for the identification and therapeutic use of compounds as treatments of hematological disorders.
METHODS AND COMPOSITIONS FOR TREATING HEMATOLOGICAL
DISORDERS USING 252, 304, 1980, 14717, 9941, 19310 AND 17832

This application claims priority to U.S. provisional application number 60/34, 1606, filed December 17, 2001, the entire contents of which are herein incorporated by reference.

Targets involved in the regulation of bone marrow development provide novel therapeutic approaches to the treatment of primary bone marrow failure and bone marrow dysfunction secondary to toxic insults, most notably chemotherapy-induced cytopenias. There is a severe unmet medical need in this arena as the few therapies currently available are recombinant proteins and all act at a relatively late stage of lineage differentiation.

Marrow populations of human and murine origin enriched for hematopoietic stem cells as well as bone marrow stromal cell populations provide useful sources of material for gene discovery and annotation of targets involved in proliferation and maturation of precursor populations. Hematopoietic cells cultured under various circumstances, isolated from humans in vivo, or from animal models in vivo provide a rich source of raw material for gene expression analysis leading to the identification of novel therapeutics useful for hematological disorders.

Detailed Description of the Invention

The present invention provides methods and compositions for the diagnosis and treatment of patients with hematological disorders.

"Treatment", as used herein, is defined as the application or administration of a therapeutic agent to a patient, or application or administration of a therapeutic agent to an isolated tissue or cell line from a patient, who has a disease or disorder, a symptom of disease or disorder or a predisposition toward a disease or disorder, with the purpose of curing, healing, alleviating, relieving, altering, remedying, ameliorating, improving or affecting the disease or disorder, at least one symptom of disease or disorder or the predisposition toward a disease or disorder. A therapeutic agent includes, but is not limited to, small molecules, peptides, antibodies, ribozymes and antisense oligonucleotides. Representative molecules are described herein.

A hematological disorder as used herein includes, but is not limited to erythroid-associated disorders. As used herein, the term "erythroid associated disorders" include disorders involving aberrant (increased or deficient) erythroblast proliferation, e.g., an erythroleukemia, and aberrant (increased or deficient) erythroblast differentiation, e.g., an...
anemia. Erythrocyte-associated disorders include anemias such as, for example, drug-
(chemotherapy-) induced anemias, hemolytic anemias due to hereditary cell membrane
abnormalities, such as hereditary spherocytosis, hereditary elliptocytosis, and hereditary
pyropoikilocytosis; hemolytic anemias due to acquired cell membrane defects, such as
paroxysmal nocturnal hemoglobinuria and spur cell anemia; hemolytic anemias caused by
antibody reactions, for example to the RBC antigens, or antigens of the ABO system,
Lewis system, Li system, Rh system, Kidd system, Duffy system, and Kell system;
methemoglobinemia; a failure of erythropoiesis, for example, as a result of aplastic anemia,
pure red cell aplasia, myelodysplastic syndromes, sideroblastic anemias, and congenital
dyserythropoietic anemia; secondary anemia in non-hematologic disorders, for example, as a
result of chemotherapy, alcoholism, or liver disease; anemia of chronic disease, such as
chronic renal failure; and endocrine deficiency diseases.

Agents that modulate the polypeptides of the present invention or nucleic acid
activity or expression can be used to treat anemias, in particular, drug-induced anemias or
anemias associated with cancer chemotherapy, chronic renal failure, malignancies, adult
and juvenile rheumatoid arthritis, disorders of hemoglobin synthesis, prematurity, and
zidovudine treatment of HIV infection. A subject receiving the treatment can be
additionally treated with a second agent, e.g., erythropoietin, to further at least one
symptom of the condition. The order of the treatments can be reversed. The two
treatments can be administered simultaneously. The timing between treatments can be
varied.

As used herein, the term "erythropoietin" or "EPO" refers to a glycoprotein
produced in the kidney, which is the principal hormone responsible for stimulating red
blood cell production (erythropoiesis). EPO stimulates the division and differentiation of
committed erythroid progenitors in the bone marrow. Normal plasma erythropoietin levels
range from 0.01 to 0.03 Units/mL, and can increase up to 100 to 1,000-fold during hypoxia
*Kidney Intl.* 28:1 (1985). Recombinant human erythropoietin (rHuEpo or epoetin alpha) is
commercially available as EPOGEN.RTM. (epoetin alpha, recombinant human
erythropoietin) (Amgen Inc., Thousand Oaks, Calif.) and as PROCRIT.RTM. (epoetin

Another example of an erythroid-associated disorder is erythrocytosis.
Erythrocytosis, a disorder of red blood cell overproduction caused by excessive and/or
ectopic erythropoietin production, can be caused by cancers, e.g., a renal cell cancer, a
hepatocarcinoma, and a central nervous system cancer. Diseases associated with erythrocytosis include polycythemias, e.g., polycythemia vera, secondary polycythemia, and relative polycythemia.

A hematological disorder as used herein includes disorders involving B-cells which include, but are not limited to precursor B-cell neoplasms, such as lymphoblastic leukemia/lymphoma. Peripheral B-cell neoplasms include, but are not limited to, chronic lymphocytic leukemia/small lymphocytic lymphoma, follicular lymphoma, diffuse large B-cell lymphoma, Burkitt lymphoma, plasma cell neoplasms, multiple myeloma, and related entities, lymphoplasmacytic lymphoma (Waldenström macroglobulinemia), mantle cell lymphoma, marginal zone lymphoma (MALToma), and hairy cell leukemia.

A hematological disorder as used herein includes disorders of the bone marrow which include but are not limited to: diseases involving hematopoietic stem cells; committed lymphoid progenitor cells; lymphoid cells including B and T-cells; committed myeloid progenitors, including monocytes, granulocytes, and megakaryocytes; and committed erythroid progenitors. These include but are not limited to the leukemias, including B-lymphoid leukemias, T-lymphoid leukemias, undifferentiated leukemias; erythroleukemia, megakaryoblastic leukemia, monocytic; [leukemias are encompassed with and without differentiation; chronic and acute lymphoblastic leukemia, chronic and acute lymphocytic leukemia, chronic and acute myelogenous leukemia, lymphoma, myelo dysplastic syndrome, chronic and acute myeloid leukemia, myelomonocytic leukemia; chronic and acute myeloblastic leukemia, chronic and acute myelogenous leukemia, chronic and acute promyelocytic leukemia, chronic and acute myelocytic leukemia, hematologic malignancies of monocyte-macrophage lineage, such as juvenile chronic myelogenous leukemia; secondary AML, antecedent hematological disorder; refractory anemia; aplastic anemia; reactive cutaneous angioendotheliomatosis; fibrosing disorders involving altered expression in dendritic cells, disorders including systemic sclerosis, E-M syndrome, epidemic toxic oil syndrome, eosinophilic fasciitis localized forms of scleroderma, keloid, and fibrosing colonopathy; angiomatoid malignant fibrous histiocytoma; carcinoma, including primary head and neck squamous cell carcinoma; sarcoma, including kaposi’s sarcoma; fibroadenoma and phyllodes tumors, including mammary fibroadenoma; stromal tumors; phyllodes tumors, including histiocytoma; erythroblastosis; neurofibromatosis; diseases of the vascular endothelium; demyelinating, particularly in old lesions; gliosis, vasogenic edema, vascular disease, Alzheimer’s and Parkinson’s disease; T-cell lymphomas; B-cell lymphomas.
A hematological disorder as used herein can include platelet disorders including but not limited to disorders related to reduced platelet number, thrombocytopenia, include idiopathic thrombocytopenic purpura, including acute idiopathic thrombocytopenic purpura, drug-induced thrombocytopenia, HIV-associated thrombocytopenia, and thrombotic microangiopathies: thrombotic thrombocytopenic purpura and hemolytic-uremic syndrome.

A hematological disorder can also include thrombosis. Thrombosis can result from platelet dysfunction, e.g. seen in myocardial infarction, angina, hypertension, lipid disorders, diabetes mellitus; myelodysplastic syndromes; myeloproliferative syndromes (including polycythemia vera and thrombocytocemia); thrombotic thrombocytopenic purpuras; HIV-induced platelet disorders (AIDS-Thrombocytopenia); heparin induced thrombocytopenia; mural cell alterations/interactions leading to platelet aggregation/degranulation, vascular endothelial cell activation/injury, monocyte/macrophage extravasation and smooth muscle cell proliferation; autoimmune disorders such as, but not limited to vasculitis, antiphospholipid syndromes, systemic lupus erythromatosi; inflammatory diseases, such as, but not limited to immune activation; graft vs host disease; radiation induced hypercoagulation; clotting factor dysregulation either hereditary (autosomal dominant or recessive) such as, but not limited to clotting factor pathways including protein C/S, Anti-thrombin III deficiency, and the Factor V Leiden mutation or acquired such as but not limited to autoimmune, cancer-associated and drug-induced dysregulation of clotting factors.

A hematological disorder as used herein can include red cell disorders including but not limited to, anemias, such as hemolytic anemias, including hereditary spherocytosis, hemolytic disease due to erythrocyte enzyme defects: glucose-6-phosphate dehydrogenase deficiency, sickle cell disease, thalassemia syndromes, paroxysmal nocturnal hemoglobinuria, immunohemolytic anemia, and hemolytic anemia resulting from trauma to red cells; and anemias of diminished erythropoiesis, including megaloblastic anemias, such as anemias of vitamin B12 deficiency: pernicious anemia, and anemia of folate deficiency, iron deficiency anemia, anemia of chronic disease, aplastic anemia, pure red cell aplasia, and other forms of marrow failure.

A hematological disorder as used herein can include disease of T cells including but not limited to, cell-mediated hypersensitivity, such as delayed type hypersensitivity and T-cell-mediated cytotoxicity, and transplant rejection; autoimmune diseases, such as systemic lupus erythematosus, Sjögren syndrome, systemic sclerosis, inflammatory myopathies, mixed connective tissue disease, and polyarteritis nodosa and other vasculitides;
immunologic deficiency syndromes, including but not limited to, primary
immunodeficiencies, such as thymic hypoplasia, severe combined immunodeficiency
diseases, and AIDS; leukopenia; reactive (inflammatory) proliferations of white cells,
including but not limited to, leukocytosis, acute nonspecific lymphadenitis, and chronic
nonspecific lymphadenitis; neoplastic proliferations of white cells, including but not
limited to lymphoid neoplasms, such as precursor T-cell neoplasms, such as acute
lymphoblastic leukemia/lymphoma, peripheral T-cell and natural killer cell neoplasms that
include peripheral T-cell lymphoma, unspecified, adult T-cell leukemia/lymphoma,
mycosis fungoides and Sézary syndrome, and Hodgkin disease.

One aspect of the invention features 252, 304, 1980, 14717, 9941, 19310 OR
17832 polypeptides and biologically active or antigenic fragments thereof that are useful,
e.g., as reagents or targets in assays applicable to treatment and diagnosis of 252, 304,
1980, 14717, 9941, 19310 OR 17832 mediated or related disorders, e.g., hematopoietic
disorders (e.g., erythroid associated disorders). In another embodiment, the invention
provides 252, 304, 1980, 14717, 9941, 19310 OR 17832 polypeptides having a 252, 304,
1980, 14717, 9941, 19310 OR 17832 activity. Preferred polypeptides are 252, 304, 1980,
14717, 9941, 19310 OR 17832 proteins including at least one dual specificity phosphatase
catalytic domain, and, preferably, having a 252, 304, 1980, 14717, 9941, 19310 OR 17832
activity, e.g., a 252, 304, 1980, 14717, 9941, 19310 OR 17832 activity as described herein.

In a related aspect, the invention provides 252, 304, 1980, 14717, 9941, 19310 OR
17832 polypeptides or fragments operatively linked to non- 252, 304, 1980, 14717, 9941,
19310 OR 17832 polypeptides to form fusion proteins.

In another aspect, the invention features antibodies and antigen-binding fragments
thereof, that react with, or more preferably, specifically bind 252, 304, 1980, 14717, 9941,
19310 OR 17832 polypeptides.

In another aspect, the invention provides methods of screening for compounds that
modulate the expression or activity of the 252, 304, 1980, 14717, 9941, 19310 OR 17832
polypeptides or nucleic acids.

In still another aspect, the invention provides a process for modulating 252, 304,
1980, 14717, 9941, 19310 OR 17832 polypeptide or nucleic acid expression or activity,
e.g. using the screened compounds. In certain embodiments, the methods involve
treatment of conditions related to decreased activity or expression of the 252, 304, 1980,
14717, 9941, 19310 OR 17832 polypeptides or nucleic acids, such as conditions involving
aberrant cellular proliferation of a 252, 304, 1980, 14717, 9941, 19310 OR 17832 -
expressing cell, e.g., a hematopoietic cell (e.g., a myeloid (neutrophil) cell, a monocyte, an erythroid cell, a bone marrow cell, a CD34-expressing cell, a megakaryocyte). The condition may involve increased hematopoietic cell activity or proliferation as in the case of leukemia, e.g., an erythroleukemia; or decreased hematopoietic cell differentiation as in the case of, e.g., an anemia.

In still another aspect, the invention features a method of modulating (e.g., enhancing or inhibiting) the proliferation, survival, and/or differentiation of a cell, e.g., a 252, 304, 1980, 14717, 9941, 19310 OR 17832-expressing cell, e.g., a hematopoietic cell (e.g., a myeloid (neutrophil) cell, a monocyte, an erythroid cell, a bone marrow cell, a CD34-expressing cell, a megakaryocyte). The method includes contacting the cell with an agent that modulates the activity or expression of a 252, 304, 1980, 14717, 9941, 19310 OR 17832 polypeptide or nucleic acid, in an amount effective to modulate the proliferation and/or differentiation of the cell.

In a preferred embodiment, the 252, 304, 1980, 14717, 9941, 19310 OR 17832 polypeptide has an amino acid sequence identical to, or substantially identical to, SEQ ID NO:3, 6, 9, 12, 15, 18 or 21. In other embodiments, the 252, 304, 1980, 14717, 9941, 19310 OR 17832 polypeptide is a fragment of at least 15, 20, 50, 100, 150, 180, 200, or more contiguous amino acids of SEQ ID NO: 3, 6, 9, 12, 15, 18 or 21.

In a preferred embodiment, the 252, 304, 1980, 14717, 9941, 19310 OR 17832 nucleic acid has a nucleotide sequence identical to, or substantially identical to, SEQ ID NOs:1, 4, 7, 10, 13, 16 or 19. In other embodiments, the 252, 304, 1980, 14717, 9941, 19310 OR 17832 nucleic acid is a fragment of at least 50, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, or more contiguous nucleotides of SEQ ID NOs: 1, 4, 7, 10, 13, 16 or 19.

In a preferred embodiment, an agent modulates (e.g., increases or decreases) expression of the 252, 304, 1980, 14717, 9941, 19310 OR 17832 nucleic acid by, e.g., modulating transcription, mRNA stability, etc.

In preferred embodiments, the agent is a peptide, a phosphopeptide, a small molecule, e.g., a member of a combinatorial library, or an antibody, or any combination thereof. The antibody can be conjugated to a therapeutic moiety selected from the group consisting of a cytotoxin, a cytotoxic agent and a radioactive metal ion.

In additional preferred embodiments, the agent is an antisense, a ribozyme, or a triple helix molecule, or an 252, 304, 1980, 14717, 9941, 19310 OR 17832 nucleic acid, or any combination thereof.
In a preferred embodiment, the agent is administered in combination with a cytotoxic agent.

In a preferred embodiment, the cell, e.g., the 252, 304, 1980, 14717, 9941, 19310 OR 17832-expressing cell, is a hematopoietic cell, e.g., a myeloid, lymphoid or erythroid cell, or a precursor cell thereof. Examples of such cells include myelocytic cells (polymorphonuclear cells), erythrocytic cells, lymphocytes, monocytes, reticular cells, plasma cells and megakaryocytes, as well as stem cells for the different lineages, and precursors for the committed progenitor cells, for example, precursors of red blood cells (erythroblasts), macrophages (monoblasts), platelets (megakaryocytes), polymorphonuclear leucocytes (myeloblasts), and lymphocytes (lymphoblasts).

In a preferred embodiment, the cell, e.g., the 252, 304, 1980, 14717, 9941, 19310 OR 17832-expressing cell, is a bone marrow cell, e.g., a bone marrow CD34-expressing cell. Examples of CD34-expressing cells include immature haematopoietic precursor cells, haematopoietic colony-forming cells in bone marrow, including unipotent (CFU-GM, BFU-E) and pluripotent progenitors (CFU-GEMM, CFU-Mix and CFU-blast); as well as stromal cell precursors, terminal deoxy nucleotidyl transferase (TdT) expressing B- and T-lymphoid precursors, early myeloid cells and early erythroid cells.

In a preferred embodiment, the cell, e.g., the 252, 304, 1980, 14717, 9941, 19310 OR 17832-expressing cell, is a bone marrow erythroid cell, e.g., an erythroid progenitor (e.g., a GPA(low)CD71+ cell) or a differentiated cell, e.g., an erythrocyte or a megakaryocyte.

In a preferred embodiment, the cell, e.g., the 252, 304, 1980, 14717, 9941, 19310 OR 17832-expressing cell, is further contacted with a protein, e.g., a cytokine. Preferably, the protein is selected from the group consisting of G-CSF, GM-CSF, stem cell factor, and preferably erythropoietin. The protein contacting step can occur before, at the same time, or after the agent is contacted. The protein contacting step can be effected in vitro or ex vivo. For example, the cell, e.g., the 252, 304, 1980, 14717, 9941, 19310 OR 17832-expressing cell is obtained from a subject, e.g., a patient, and contacted with the protein ex vivo. The treated cell can be re-introduced into the subject. Alternatively, the protein contacting step can occur in vivo.

In a preferred embodiment, the agent and the 252, 304, 1980, 14717, 9941, 19310 OR 17832-polypeptide or nucleic acid are contacted in vitro or ex vivo.

In a preferred embodiment, the contacting step is effected in vivo in a subject, e.g., as part of a therapeutic or prophylactic protocol. Preferably, the subject is a human, e.g., a
patient with a hematopoietic disorder, e.g., a leukemia or an erythroid-associated disorder. For example, the subject can be a patient with an anemia, e.g., hemolytic anemia, aberrant erythropoiesis, secondary anemia in non-hematologic disorders, anemia of chronic disease such as chronic renal failure; endocrine deficiency disease; and/or erythrocytosis (e.g., polycythemia). Alternatively, the subject can be a cancer patient, e.g., a patient with leukemic cancer, e.g., an erythroid leukemia, or a carcinoma, e.g., a renal carcinoma. In other embodiments, the subject is a non-human animal, e.g., an experimental animal.

The contacting step(s) can be repeated.

In a preferred embodiment, the agent decreases the proliferation and/or enhances the differentiation of the cell, e.g., the 252, 304, 1980, 14717, 9941, 19310 OR 17832-expressing cell, e.g., the hematopoietic cell (e.g., the myeloid (neutrophil) cell, the monocyte, the erythroid cell, the bone marrow cell, the CD34-expressing cell, the megakaryocyte). Such agents can be used to treat or prevent cancers, e.g., leukemic cancers.

In a preferred embodiment, the agent increases the number of hematopoietic cells (e.g., myeloid (neutrophil) cells, monocytes, erythroid cells, bone marrow cells, CD34-expressing cells, megakaryocytes), by e.g., increasing the proliferation, survival, and/or stimulating the differentiation, of progenitor cells. Such agents can be used to treat or prevent hematopoietic disorders, e.g., anemias (e.g., hemolytic anemias, aberrant erythropoiesis, secondary anemias in nonhematologic disorders, anemias of chronic diseases such as chronic renal failure; endocrine deficiency diseases; and/or erythrocytosis, e.g., polycythemias).

In another aspect, the invention features a method of modulating hematopoiesis, e.g., erythropoiesis, comprising contacting a 252, 304, 1980, 14717, 9941, 19310 OR 17832-expressing cell, e.g., a hematopoietic cell, (e.g., a myeloid (neutrophil) cell, a monocyte, an erythroid cell, a bone marrow cell, a CD34-expressing cell, a megakaryocyte), with a agent that increases or decreases the activity or expression of a 252, 304, 1980, 14717, 9941, 19310 OR 17832 polypeptide or nucleic acid, thereby modulating the differentiation of the hematopoietic cell.

In yet another aspect, the invention features a method of treating or preventing a hematopoietic disorder, e.g., an erythroid-associated disorder, in a subject. The method includes administering to the subject an effective amount of a agent that modulates the activity or expression of a 252, 304, 1980, 14717, 9941, 19310 OR 17832 polypeptide or
nucleic acid such that the hematopoietic disorder is ameliorated or at least one symptom of the hematological disorder is decreased.

**Molecules of the Present Invention**

5  
**Gene ID 252**

The human 252 sequence (SEQ ID NO:1), (GI:1486234, known also as prostaglandin E receptor (EP4)) which is approximately 1554 nucleotides long including untranslated regions, contains a predicted methionine-initiated coding sequence of about 1467 nucleotides, including the termination codon (nucleotides indicated as coding of SEQ ID NO:1, SEQ ID NO:2). The coding sequence encodes a 488 amino acid protein (SEQ ID NO:3) (GI:1486235).

As assessed by TaqMan analysis, 252 mRNA was expressed ubiquitously at low levels in various human tissues, but at higher levels in CD34+ progenitor cells of the bone marrow. High levels of expression were also seen in the CD11b-/CD15+ neutrophil population in vivo, and levels of 252 mRNA were constant during myelopoiesis. 252 is a receptor for Prostaglandin E, a molecule known to have stimulatory effects on various progenitor cells and colony forming unit(CFU)-mix. Agonizing this pathway will stimulate CD34+ cell proliferation and expand the number of cells in all lineages. Due to the expression pattern, agents which modulate 252 activity would be useful therapeutics for hematological disorders as disclosed herein.

**Gene ID 304**

The human 304 sequence (SEQ ID NO:4), (GI:563981), known also vasopressin V1b receptor) which is approximately 1824 nucleotides long including untranslated regions, contains a predicted methionine-initiated coding sequence of about 1275 nucleotides, including the termination codon (nucleotides indicated as coding of SEQ ID NO:4, SEQ ID NO:5). The coding sequence encodes a 424 amino acid protein (SEQ ID NO:6) (GI:563982).

As assessed by TaqMan analysis expression of 304 mRNA was expressed in very few tissues, CD34+ progenitor cells of the bone marrow being among the most prominent. Expression was also seen in kidney and a single lung tumor. Expression of 304 mRNA was not observed in differentiated lineages of the bone marrow.
304 is the Vasopressin V3 receptor. Signalling through this receptor is known to have effects on endothelial cells. By agonizing this receptor, we will be able to stimulate progenitor cell proliferation and increase the numbers of cells in all lineages. Therefore, agents which modulate 304 activity would be useful therapeutics for hematological disorders as described herein.

**Gene ID 1980**

The human 1980 sequence (SEQ ID NO:7), (GI:251839, known also as glutamate receptor subunit) which is approximately 2943 nucleotides long including untranslated regions, contains a predicted methionine-initiated coding sequence of about 2943 nucleotides, including the termination codon (nucleotides indicated as coding of SEQ ID NO:7, SEQ ID NO:8). The coding sequence encodes a 980 amino acid protein (SEQ ID NO:9) (GI:251840).

Taqman expression analysis shows high levels of 1980 mRNA expression in CD34+ progenitor cells. Expression was absent in every other cell lineage in vitro and in vivo. Expression was also seen in normal human brain and kidney, as well as human umbilical vein endothelial cells (HUVEC). 1980 is one of the subunits of the glutamate receptor. Apart from its well-described effects on synaptic signalling, it is known to play a variety of roles in different cell types. It is known to enhance proliferation and survival in a number of cell types, and its inhibition can lead to cellular differentiation. By agonizing this receptor, we will be able to stimulate progenitor cell proliferation and increase the numbers of cells in all lineages. Therefore, agents which modulate 1980 activity would be useful therapeutics for hematological disorders as described herein.

**Gene ID 14717**

The human 14717 sequence (SEQ ID NO:10), (GI:2204346, known also as human manic fringe precursor) which is approximately 2042 nucleotides long including untranslated regions, contains a predicted methionine-initiated coding sequence of about 966 nucleotides, including the termination codon (nucleotides indicated as coding of SEQ ID NO:10, SEQ ID NO:11). The coding sequence encodes a 321 amino acid protein (SEQ ID NO:12) (GI:2204347).

As assessed by TaqMan analysis 14717 mRNA was highly expressed in CD34+ progenitor cells and in erythroid cell cultures in vitro, with levels of 14717 mRNA increasing during differentiation. Expression of 14717 mRNA at high levels was also seen
in GPA-lo cells, consistent with its having a role in erythropoiesis. Additional TaqMan analyses indicated that 14717 mRNA was low in all normal human tissues tested although high levels of expression in HUVEC cells was noted.

14717 or manic fringe is a glycosyltransferase involved in the processing of Notch. 14717 is required for the translocation of Notch to the cell surface. 14717 is known to be required for signalling through Notch. Modulating the activity of 14717 will stimulate erythropoiesis by inhibiting the function of Notch. Notch is known to inhibit differentiation by holding cells in an uncommitted state. Decreased signaling through Notch will result in a burst of erythroid commitment and differentiation. Therefore, modulators of 14717 activity would be useful as therapeutics in hematological disorders as disclosed herein.

Gene ID 9941

The human 9941 sequence (SEQ ID NO:13), (GI:407005, known also as calcium/calmodulin-dependent protein kinase) which is approximately 1740 nucleotides long including untranslated regions, contains a predicted methionine-initiated coding sequence of about 1422 nucleotides, including the termination codon (nucleotides indicated as coding of SEQ ID NO:13, SEQ ID NO:14). The coding sequence encodes a 473 amino acid protein (SEQ ID NO:15) (GI:407006).

As assessed by TaqMan analysis 9941 mRNA was found in normal human brain, T cells (CD3+, CD4+ and CD8+), CD61+ bone marrow megakaryocytes and in platelet RNA and at increased levels in platelets from patients with ischemic heart disease and unstable angina.

9941 calcium/calmodulin-dependent protein kinase IV, which is essential for cerebellar development and function, as determined by gene targeting [J Neurosci 2000 Nov 15;20(22):RC107]. Calcium mobilization is a critical component of platelet activation and degranulation. Calcium-dependent signal transduction has been implicated in the cytoskeletal translocation of signaling enzymes during platelet activation. Antagonizing 9941 activity would inhibit platelet reactivity. Therefore, modulators of 9941 activity would be useful therapeutics for hematological disorders as disclosed herein, including but not limited to thrombosis.

Gene ID 19310

The human 19310 sequence (SEQ ID NO:16), (GI: 7021036, known also as an oxidase) which is approximately 2291 nucleotides long including untranslated regions,
contains a predicted methionine-initiated coding sequence of about 1668 nucleotides, including the termination codon (nucleotides indicated as coding of SEQ ID NO:16, SEQ ID NO:17). The coding sequence encodes a 555 amino acid protein (SEQ ID NO:18) (GI: 7021037).

As assessed by TaqMan analysis, 19310 mRNA was found to be expressed in normal human brain, megakaryocyte progenitors, bone marrow megakaryocytes (CD61+) and platelets. Slightly elevated levels of 19310 mRNA was detected in the platelets from patients following myocardial infarct as compared with normal volunteer platelets.

Polyamines inhibit platelet aggregation. Polyamine oxidase is the second of two enzymes mediating the catabolism of polyamines. In addition, polyamines can stimulate synthesis of polyamine oxidase. Inhibiting 19310 enzymatic activity would prevent the enzymatic degradation of polyamines thereby promote anti-platelet activity and maintain hemostasis. Therefore, modulators of 19310 activity would be useful as therapeutics for hematological disorders as disclosed herein, including but not limited to thrombosis.

Gene ID 17832

The human 17832 sequence (SEQ ID NO:19), (GI:13325107, known also potassium channel, subfamily K (TWIK-2)) which is approximately 2649 nucleotides long including untranslated regions, contains a predicted methionine-initiated coding sequence of about 942 nucleotides, including the termination codon (nucleotides indicated as coding of SEQ ID NO:19, SEQ ID NO:20). The coding sequence encodes a 313 amino acid protein (SEQ ID NO:21) GI:13325108).

As assessed by TaqMan analysis, expression of 17832 mRNA was found at high levels in cultured megakaryocytes, more moderately in bone marrow megakaryocytes (CD61+ cells) and at increased levels in platelets from myocardial infarct and polycythemia vera patients as compared to normal volunteer platelets.

17832 or TWIK-2, contains a potential PKC phosphorylation site (Ser 158) and is able to be stimulated with the protein kinase C activator, PMA [JBC 1999; 274(12):7887-92]. Pathologic shear stress and thrombin activate PKC and stimulate platelet aggregation [JBC 1993; 268(5):3520-4]. 17832 function in megakaryocytes and platelets may be a compensatory mechanism or feed-back inhibition of platelet aggregation at sites of high shear stress and/or thrombin activation. Therefore, modulators of 17832 would be useful
as therapeutics for hematological disorders as disclosed herein, including but not limited to thrombosis.

Various aspects of the invention are described in further detail in the following subsections:

I. **Screening Assays:**

The invention provides a method (also referred to herein as a "screening assay") for identifying modulators, *i.e.*, candidate or test compounds or agents (*e.g.*, peptides, peptidomimetics, small molecules (organic or inorganic) or other drugs) which bind to 252, 304, 1980, 14717, 9941, 19310 OR 17832 proteins, have a stimulatory or inhibitory effect on, for example, 252, 304, 1980, 14717, 9941, 19310 OR 17832 expression or 252, 304, 1980, 14717, 9941, 19310 OR 17832 activity, or have a stimulatory or inhibitory effect on, for example, the expression or activity of a 252, 304, 1980, 14717, 9941, 19310 OR 17832 substrate. Compounds identified using the assays described herein may be useful for treating hematological disorders.

These assays are designed to identify compounds that bind to a 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein, bind to other intracellular or extracellular proteins that interact with a 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein, and interfere with the interaction of the 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein with other intercellular or extracellular proteins. For example, in the case of the 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein, which is a transmembrane receptor-type protein, such techniques can identify ligands for such a receptor. A 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein ligand or substrate can, for example, be used to at least one symptom of a hematological disorder. Such compounds may include, but are not limited to peptides, antibodies, or small organic or inorganic compounds. Such compounds may also include other cellular proteins.

Compounds identified via assays such as those described herein may be useful, for example, for treating hematological disorders. In instances whereby a hematological disorder condition results from an overall lower level of 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene expression and/or 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein in a cell or tissue, compounds that interact with the 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein may include compounds which accentuate or amplify the activity of the bound 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein. Such compounds
would bring about an effective increase in the level of 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein activity, thus ameliorating symptoms.

In other instances, mutations within the 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene may cause aberrant types or excessive amounts of 252, 304, 1980, 14717, 9941, 19310 OR 17832 proteins to be made which have a deleterious effect that leads to a hematological disorder. Similarly, physiological conditions may cause an excessive increase in 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene expression leading hematological disorders. In such cases, compounds that bind to a 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein may be identified that inhibit the activity of the 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein. Assays for testing the effectiveness of compounds identified by techniques such as those described in this section are discussed herein.

In one embodiment, the invention provides assays for screening candidate or test compounds which are substrates of a 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein or polypeptide or biologically active portion thereof. In another embodiment, the invention provides assays for screening candidate or test compounds which bind to or modulate the activity of a 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein or polypeptide or biologically active portion thereof. The test compounds of the present invention can be obtained using any of the numerous approaches in combinatorial library methods known in the art, including: biological libraries; spatially addressable parallel solid phase or solution phase libraries; synthetic library methods requiring deconvolution; the 'one-bead one-compound' library method; and synthetic library methods using affinity chromatography selection. The biological library approach is limited to peptide libraries, while the other four approaches are applicable to peptide, non-peptide oligomer or small molecule libraries of compounds (Lam, K.S. (1997) Anticancer Drug Des. 12:145).


Libraries of compounds may be presented in solution (e.g., Houghten (1992), Biotechniques 13:412-421), or on beads (Lam (1991) Nature 354:82-84), chips (Fodor (1993) Nature 364:555-556), bacteria (Ladner USP 5,223,409), spores (Ladner USP '409),

In one embodiment, an assay is a cell-based assay in which a cell which expresses a 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein or biologically active portion thereof is contacted with a test compound and the ability of the test compound to modulate 252, 304, 1980, 14717, 9941, 19310 OR 17832 activity is determined. Determining the ability of the test compound to modulate 252, 304, 1980, 14717, 9941, 19310 OR 17832 activity can be accomplished by monitoring, for example, intracellular calcium, IP$_3$, cAMP, or diacylglycerol concentration, the phosphorylation profile of intracellular proteins, cell proliferation and/or migration, gene expression of, for example, cell surface adhesion molecules or genes associated with hematopoiesis, or the activity of a 252, 304, 1980, 14717, 9941, 19310 OR 17832-regulated transcription factor. The cell can be of mammalian origin, e.g., a neural cell. In one embodiment, compounds that interact with a receptor domain can be screened for their ability to function as ligands, i.e., to bind to the receptor and modulate a signal transduction pathway. Identification of ligands, and measuring the activity of the ligand-receptor complex, leads to the identification of modulators (e.g., antagonists) of this interaction. Such modulators may be useful in the treatment of hematological disorders.

The ability of the test compound to modulate 252, 304, 1980, 14717, 9941, 19310 OR 17832 binding to a substrate or to bind to 252, 304, 1980, 14717, 9941, 19310 OR 17832 can also be determined. Determining the ability of the test compound to modulate 252, 304, 1980, 14717, 9941, 19310 OR 17832 binding to a substrate can be accomplished, for example, by coupling the 252, 304, 1980, 14717, 9941, 19310 OR 17832 substrate with a radioisotope or enzymatic label such that binding of the 252, 304, 1980, 14717, 9941, 19310 OR 17832 substrate to 252, 304, 1980, 14717, 9941, 19310 OR 17832 can be determined by detecting the labeled 252, 304, 1980, 14717, 9941, 19310 OR 17832 substrate in a complex. 252, 304, 1980, 14717, 9941, 19310 OR 17832 could also be coupled with a radioisotope or enzymatic label to monitor the ability of a test compound to modulate 252, 304, 1980, 14717, 9941, 19310 OR 17832 binding to a 252, 304, 1980, 14717, 9941, 19310 OR 17832 substrate in a complex. Determining the ability of the test compound to bind 252, 304, 1980, 14717, 9941, 19310 OR 17832 can be accomplished, for example, by coupling the compound with a radioisotope or enzymatic label such that
binding of the compound to 252, 304, 1980, 14717, 9941, 19310 OR 17832 can be
determined by detecting the labeled 252, 304, 1980, 14717, 9941, 19310 OR 17832
compound in a complex. For example, compounds (e.g., 252, 304, 1980, 14717, 9941,
19310 OR 17832 ligands or substrates) can be labeled with $^{125}$I, $^{35}$S, $^{14}$C, or $^{3}$H, either
directly or indirectly, and the radioisotope detected by direct counting of radioemmission
or by scintillation counting. Compounds can further be enzymatically labeled with, for
example, horseradish peroxidase, alkaline phosphatase, or luciferase, and the enzymatic
label detected by determination of conversion of an appropriate substrate to product.

It is also within the scope of this invention to determine the ability of a compound
(e.g., a 252, 304, 1980, 14717, 9941, 19310 OR 17832 ligand or substrate) to interact with
252, 304, 1980, 14717, 9941, 19310 OR 17832 without the labeling of any of the
interactants. For example, a microphysiometer can be used to detect the interaction of a
compound with 252, 304, 1980, 14717, 9941, 19310 OR 17832 without the labeling of
either the compound or the 252, 304, 1980, 14717, 9941, 19310 OR 17832 (McConnell, H.
M. et al. (1992) Science 257:1906-1912. As used herein, a "microphysiometer" (e.g.,
Cytosensor) is an analytical instrument that measures the rate at which a cell acidifies its
environment using a light-addressable potentiometric sensor (LAPS). Changes in this
acidification rate can be used as an indicator of the interaction between a compound and
252, 304, 1980, 14717, 9941, 19310 OR 17832.

In another embodiment, an assay is a cell-based assay comprising contacting a cell
expressing a 252, 304, 1980, 14717, 9941, 19310 OR 17832 target molecule (e.g., a 252,
304, 1980, 14717, 9941, 19310 OR 17832 substrate) with a test compound and determining
the ability of the test compound to modulate (e.g., stimulate or inhibit) the activity of the
252, 304, 1980, 14717, 9941, 19310 OR 17832 target molecule. Determining the ability of the
test compound to modulate the activity of a 252, 304, 1980, 14717, 9941, 19310 OR
17832 target molecule can be accomplished, for example, by determining the ability of the
252, 304, 1980, 14717, 9941, 19310 OR 17832 protein to bind to or interact with the 252,
304, 1980, 14717, 9941, 19310 OR 17832 target molecule.

Determining the ability of the 252, 304, 1980, 14717, 9941, 19310 OR 17832
protein or a biologically active fragment thereof, to bind to or interact with a 252, 304,
1980, 14717, 9941, 19310 OR 17832 target molecule can be accomplished by one of the
methods described above for determining direct binding. In a preferred embodiment,
determining the ability of the 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein to
bind to or interact with a 252, 304, 1980, 14717, 9941, 19310 OR 17832 target molecule
can be accomplished by determining the activity of the target molecule. For example, the activity of the target molecule can be determined by detecting induction of a cellular second messenger of the target (i.e., intracellular Ca\(^{2+}\), diacylglycerol, IP\(_3\), cAMP), detecting catalytic/enzymatic activity of the target on an appropriate substrate, detecting the induction of a reporter gene (comprising a target-responsive regulatory element operatively linked to a nucleic acid encoding a detectable marker, e.g., luciferase), or detecting a target-regulated cellular response (e.g., gene expression).

In yet another embodiment, an assay of the present invention is a cell-free assay in which a 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein or biologically active portion thereof, is contacted with a test compound and the ability of the test compound to bind to the 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein or biologically active portion thereof is determined. Preferred biologically active portions of the 252, 304, 1980, 14717, 9941, 19310 OR 17832 proteins to be used in assays of the present invention include fragments which participate in interactions with non-252, 304, 1980, 14717, 9941, 19310 OR 17832 molecules, e.g., fragments with high surface probability scores. Binding of the test compound to the 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein can be determined either directly or indirectly as described above. In a preferred embodiment, the assay includes contacting the 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein or biologically active portion thereof with a known compound which binds 252, 304, 1980, 14717, 9941, 19310 OR 17832 to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with a 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein, wherein determining the ability of the test compound to interact with a 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein comprises determining the ability of the test compound to preferentially bind to 252, 304, 1980, 14717, 9941, 19310 OR 17832 or biologically active portion thereof as compared to the known compound. Compounds that modulate the interaction of 252, 304, 1980, 14717, 9941, 19310 OR 17832 with a known target protein may be useful in regulating the activity of a 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein, especially a mutant 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein.

In another embodiment, the assay is a cell-free assay in which a 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein or biologically active portion thereof is contacted with a test compound and the ability of the test compound to modulate (e.g., stimulate or inhibit) the activity of the 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein or biologically active portion thereof is determined. Determining the ability of the test
compound to modulate the activity of a 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein can be accomplished, for example, by determining the ability of the 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein to bind to a 252, 304, 1980, 14717, 9941, 19310 OR 17832 target molecule by one of the methods described above for determining direct binding. Determining the ability of the 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein to bind to a 252, 304, 1980, 14717, 9941, 19310 OR 17832 target molecule can also be accomplished using a technology such as real-time Biomolecular Interaction Analysis (BIA) (Sjolander, S. and Urbaniczky, C. (1991) Anal. Chem. 63:2338-2345 and Szabo et al. (1995) Curr. Opin. Struct. Biol. 5:699-705). As used herein, "BIA" is a technology for studying biospecific interactions in real time, without labeling any of the interactants (e.g., BIAcore). Changes in the optical phenomenon of surface plasmon resonance (SPR) can be used as an indication of real-time reactions between biological molecules.

In another embodiment, determining the ability of the test compound to modulate the activity of a 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein can be accomplished by determining the ability of the 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein to further modulate the activity of a downstream effector of a 252, 304, 1980, 14717, 9941, 19310 OR 17832 target molecule. For example, the activity of the effector molecule on an appropriate target can be determined or the binding of the effector to an appropriate target can be determined as previously described.

In yet another embodiment, the cell-free assay involves contacting a 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein or biologically active portion thereof with a known compound which binds the 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with the 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein, wherein determining the ability of the test compound to interact with the 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein comprises determining the ability of the 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein to preferentially bind to or modulate the activity of a 252, 304, 1980, 14717, 9941, 19310 OR 17832 target molecule.

In more than one embodiment of the above assay methods of the present invention, it may be desirable to immobilize either 252, 304, 1980, 14717, 9941, 19310 OR 17832 or its target molecule to facilitate separation of complexed from uncomplexed forms of one or both of the proteins, as well as to accommodate automation of the assay. Binding of a test compound to a 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein, or interaction of a
252, 304, 1980, 14717, 9941, 19310 OR 17832 protein with a target molecule in the presence and absence of a candidate compound, can be accomplished in any vessel suitable for containing the reactants. Examples of such vessels include microtitre plates, test tubes, and micro-centrifuge tubes. In one embodiment, a fusion protein can be provided which adds a domain that allows one or both of the proteins to be bound to a matrix. For example, glutathione-S-transferase/252, 304, 1980, 14717, 9941, 19310 OR 17832 fusion proteins or glutathione-S-transferase/target fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, MO) or glutathione derivatized microtitre plates, which are then combined with the test compound or the test compound and either the non-adsorbed target protein or 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein, and the mixture incubated under conditions conducive to complex formation (e.g., at physiological conditions for salt and pH). Following incubation, the beads or microtitre plate wells are washed to remove any unbound components, the matrix immobilized in the case of beads, complex determined either directly or indirectly, for example, as described above. Alternatively, the complexes can be dissociated from the matrix, and the level of 252, 304, 1980, 14717, 9941, 19310 OR 17832 binding or activity determined using standard techniques.

Other techniques for immobilizing proteins on matrices can also be used in the screening assays of the invention. For example, either a 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein or a 252, 304, 1980, 14717, 9941, 19310 OR 17832 target molecule can be immobilized utilizing conjugation of biotin and streptavidin. Biotinylated 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein or target molecules can be prepared from biotin-NHS (N-hydroxy-succinimide) using techniques known in the art (e.g., biotinylation kit, Pierce Chemicals, Rockford, IL), and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical). Alternatively, antibodies reactive with 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein or target molecules but which do not interfere with binding of the 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein to its target molecule can be derivatized to the wells of the plate, and unbound target or 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein trapped in the wells by antibody conjugation. Methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies reactive with the 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein or target molecule, as well as enzyme-linked assays which rely on detecting an enzymatic
activity associated with the 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein or target molecule.

In another embodiment, modulators of 252, 304, 1980, 14717, 9941, 19310 OR 17832 expression are identified in a method wherein a cell is contacted with a candidate compound and the expression of 252, 304, 1980, 14717, 9941, 19310 OR 17832 mRNA or protein in the cell is determined. The level of expression of 252, 304, 1980, 14717, 9941, 19310 OR 17832 mRNA or protein in the presence of the candidate compound is compared to the level of expression of 252, 304, 1980, 14717, 9941, 19310 OR 17832 mRNA or protein in the absence of the candidate compound. The candidate compound can then be identified as a modulator of 252, 304, 1980, 14717, 9941, 19310 OR 17832 expression based on this comparison. For example, when expression of 252, 304, 1980, 14717, 9941, 19310 OR 17832 mRNA or protein is greater (statistically significantly greater) in the presence of the candidate compound than in its absence, the candidate compound is identified as a stimulator of 252, 304, 1980, 14717, 9941, 19310 OR 17832 mRNA or protein expression. Alternatively, when expression of 252, 304, 1980, 14717, 9941, 19310 OR 17832 mRNA or protein is less (statistically significantly less) in the presence of the candidate compound than in its absence, the candidate compound is identified as an inhibitor of 252, 304, 1980, 14717, 9941, 19310 OR 17832 mRNA or protein expression. The level of 252, 304, 1980, 14717, 9941, 19310 OR 17832 mRNA or protein expression in the cells can be determined by methods described herein for detecting 252, 304, 1980, 14717, 9941, 19310 OR 17832 mRNA or protein.

In yet another aspect of the invention, the 252, 304, 1980, 14717, 9941, 19310 OR 17832 proteins can be used as "bait proteins" in a two-hybrid assay or three-hybrid assay (see, e.g., U.S. Patent No. 5,283,317; Zervos et al. (1993) Cell 72:223-232; Madura et al. (1993) J. Biol. Chem. 268:12046-12054; Bartel et al. (1993) Biotechniques 14:920-924; Iwabuchi et al. (1993) Oncogene 8:1693-1696; and Brent WO94/10300), to identify other proteins, which bind to or interact with 252, 304, 1980, 14717, 9941, 19310 OR 17832 ("252, 304, 1980, 14717, 9941, 19310 OR 17832 -binding proteins" or "252, 304, 1980, 14717, 9941, 19310 OR 17832 -bp") and are involved in 252, 304, 1980, 14717, 9941, 19310 OR 17832 activity. Such 252, 304, 1980, 14717, 9941, 19310 OR 17832 -binding proteins are also likely to be involved in the propagation of signals by the 252, 304, 1980, 14717, 9941, 19310 OR 17832 proteins or 252, 304, 1980, 14717, 9941, 19310 OR 17832 targets as, for example, downstream elements of a 252, 304, 1980, 14717, 9941, 19310 OR 17832 -mediated signaling pathway. Alternatively, such 252, 304, 1980, 14717, 9941,
19310 OR 17832-binding proteins are likely to be 252, 304, 1980, 14717, 9941, 19310 OR 17832 inhibitors.

The two-hybrid system is based on the modular nature of most transcription factors, which consist of separable DNA-binding and activation domains. Briefly, the assay utilizes two different DNA constructs. In one construct, the gene that codes for a 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein is fused to a gene encoding the DNA binding domain of a known transcription factor (e.g., GAL-4). In the other construct, a DNA sequence, from a library of DNA sequences, that encodes an unidentified protein ("prey" or "sample") is fused to a gene that codes for the activation domain of the known transcription factor. If the "bait" and the "prey" proteins are able to interact, in vivo, forming a 252, 304, 1980, 14717, 9941, 19310 OR 17832-dependent complex, the DNA-binding and activation domains of the transcription factor are brought into close proximity. This proximity allows transcription of a reporter gene (e.g., LacZ) which is operably linked to a transcriptional regulatory site responsive to the transcription factor. Expression of the reporter gene can be detected and cell colonies containing the functional transcription factor can be isolated and used to obtain the cloned gene which encodes the protein which interacts with the 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein.

In another aspect, the invention pertains to a combination of two or more of the assays described herein. For example, a modulating agent can be identified using a cell-based or a cell free assay, and the ability of the agent to modulate the activity of a 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein can be confirmed in vivo, e.g., in an animal such as an animal model for hematological disorders, as described herein.

This invention further pertains to novel agents identified by the above-described screening assays. Accordingly, it is within the scope of this invention to further use an agent identified as described herein in an appropriate animal model. For example, an agent identified as described herein (e.g., a 252, 304, 1980, 14717, 9941, 19310 OR 17832 modulating agent, an antisense 252, 304, 1980, 14717, 9941, 19310 OR 17832 nucleic acid molecule, a 252, 304, 1980, 14717, 9941, 19310 OR 17832-specific antibody, or a 252, 304, 1980, 14717, 9941, 19310 OR 17832-binding partner) can be used in an animal model to determine the efficacy, toxicity, or side effects of treatment with such an agent. Alternatively, an agent identified as described herein can be used in an animal model to determine the mechanism of action of such an agent. Furthermore, this invention pertains to uses of novel agents identified by the above-described screening assays for treatments as described herein.
Any of the compounds, including but not limited to compounds such as those identified in the foregoing assay systems, may be tested for the ability to treat hematological disorders. Cell-based and animal model-based assays for the identification of compounds exhibiting such an ability to at least one symptom of hematological disorders are described herein.

In addition, animal-based models of hematological disorders, such as those described herein, may be used to identify compounds capable of treating hematological disorders. Such animal models may be used as test substrates for the identification of drugs, pharmaceuticals, therapies, and interventions which may be effective in treating hematological disorders. For example, animal models may be exposed to a compound, suspected of exhibiting an ability to treat hematological disorders, at a sufficient concentration and for a time sufficient to elicit such an amelioration of hematological disorders in the exposed animals. The response of the animals to the exposure may be monitored by assessing the reversal of the symptoms of hematological disorders before and after treatment.

With regard to intervention, any treatments which reverse any aspect of hematological disorders should be considered as candidates for human hematological disorders therapeutic intervention. Dosages of test agents may be determined by deriving dose-response curves.

Additionally, gene expression patterns may be utilized to assess the ability of a compound to at least one symptom of hematological disorders. For example, the expression pattern of one or more genes may form part of a "gene expression profile" or "transcriptional profile" which may be then be used in such an assessment. "Gene expression profile" or "transcriptional profile", as used herein, includes the pattern of mRNA expression obtained for a given tissue or cell type under a given set of conditions. Gene expression profiles may be generated, for example, by utilizing a differential display procedure, Northern analysis and/or RT-PCR. In one embodiment, 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene sequences may be used as probes and/or PCR primers for the generation and corroboration of such gene expression profiles.

Gene expression profiles may be characterized for known states, either cardiovascular disease or normal, within the cell- and/or animal-based model systems. Subsequently, these known gene expression profiles may be compared to ascertain the effect a test compound has to modify such gene expression profiles, and to cause the profile to more closely resemble that of a more desirable profile.
For example, administration of a compound may cause the gene expression profile of a hematological disorder disease model system to more closely resemble the control system. Administration of a compound may, alternatively, cause the gene expression profile of a control system to begin to mimic hematological disorders or a hematological disorder disease state. Such a compound may, for example, be used in further characterizing the compound of interest, or may be used in the generation of additional animal models.

II. Cell- and Animal-Based Model Systems

Described herein are cell- and animal-based systems which act as models for hematological disorders. These systems may be used in a variety of applications. For example, the cell- and animal-based model systems may be used to further characterize differentially expressed genes associated with cardiovascular disease, e.g., 252, 304, 1980, 14717, 9941, 19310 OR 17832. In addition, animal- and cell-based assays may be used as part of screening strategies designed to identify compounds which are capable of ameliorating hematological disorders, as described, below. Thus, the animal- and cell-based models may be used to identify drugs, pharmaceuticals, therapies and interventions which may be effective in treating a hematological disorder. Furthermore, such animal models may be used to determine the LD50 and the ED50 in animal subjects, and such data can be used to determine the in vivo efficacy of potential hematological disorders treatments.

A. Animal-Based Systems

Animal-based model systems of hematological disorders may include, but are not limited to, non-recombinant and engineered transgenic animals.

Non-recombinant animal models for hematological disorders may include, for example, genetic models.

Additionally, animal models exhibiting hematological disorders may be engineered by using, for example, 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene sequences described above, in conjunction with techniques for producing transgenic animals that are well known to those of skill in the art. For example, 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene sequences may be introduced into, and overexpressed in, the genome of the animal of interest, or, if endogenous 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene sequences are present, they may either be overexpressed or, alternatively, be disrupted in
order to underexpress or inactivate 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene expression.

The host cells of the invention can also be used to produce non-human transgenic animals. For example, in one embodiment, a host cell of the invention is a fertilized oocyte or an embryonic stem cell into which 252, 304, 1980, 14717, 9941, 19310 OR 17832 -coding sequences have been introduced. Such host cells can then be used to create non-human transgenic animals in which exogenous 252, 304, 1980, 14717, 9941, 19310 OR 17832 sequences have been introduced into their genome or homologous recombinant animals in which endogenous 252, 304, 1980, 14717, 9941, 19310 OR 17832 sequences have been altered. Such animals are useful for studying the function and/or activity of a 252, 304, 1980, 14717, 9941, 19310 OR 17832 and for identifying and/or evaluating modulators of 252, 304, 1980, 14717, 9941, 19310 OR 17832 activity. As used herein, a "transgenic animal" is a non-human animal, preferably a mammal, more preferably a rodent such as a rat or mouse, in which one or more of the cells of the animal includes a transgene. Other examples of transgenic animals include non-human primates, sheep, dogs, cows, goats, chickens, amphibians, and the like. A transgene is exogenous DNA which is integrated into the genome of a cell from which a transgenic animal develops and which remains in the genome of the mature animal, thereby directing the expression of an encoded gene product in one or more cell types or tissues of the transgenic animal. As used herein, a "homologous recombinant animal" is a non-human animal, preferably a mammal, more preferably a mouse, in which an endogenous 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene has been altered by homologous recombination between the endogenous gene and an exogenous DNA molecule introduced into a cell of the animal, e.g., an embryonic cell of the animal, prior to development of the animal.

A transgenic animal used in the methods of the invention can be created by introducing a 252, 304, 1980, 14717, 9941, 19310 OR 17832 -encoding nucleic acid into the male pronuclei of a fertilized oocyte, e.g., by microinjection, retroviral infection, and allowing the oocyte to develop in a pseudopregnant female foster animal. The 252, 304, 1980, 14717, 9941, 19310 OR 17832 cDNA sequence can be introduced as a transgene into the genome of a non-human animal. Alternatively, a nonhuman homologue of a human 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene, such as a mouse or rat 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene, can be used as a transgene. Alternatively, a 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene homologue, such as another 252, 304, 1980, 14717, 9941, 19310 OR 17832 family member, can be isolated based on
hybridization to the 252, 304, 1980, 14717, 9941, 19310 OR 17832 cDNA sequences and used as a transgene. Intronic sequences and polyadenylation signals can also be included in the transgene to increase the efficiency of expression of the transgene. A tissue-specific regulatory sequence(s) can be operably linked to a 252, 304, 1980, 14717, 9941, 19310 OR 17832 transgene to direct expression of a 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein to particular cells. Methods for generating transgenic animals via embryo manipulation and microinjection, particularly animals such as mice, have become conventional in the art and are described, for example, in U.S. Patent Nos. 4,736,866 and 4,870,009, both by Leder et al., U.S. Patent No. 4,873,191 by Wagner et al. and in Hogan, B., *Manipulating the Mouse Embryo*, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1986). Similar methods are used for production of other transgenic animals. A transgenic founder animal can be identified based upon the presence of a 252, 304, 1980, 14717, 9941, 19310 OR 17832 transgene in its genome and/or expression of 252, 304, 1980, 14717, 9941, 19310 OR 17832 mRNA in tissues or cells of the animals. A transgenic founder animal can then be used to breed additional animals carrying the transgene. Moreover, transgenic animals carrying a transgene encoding a 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein can further be bred to other transgenic animals carrying other transgenes.

To create a homologous recombinant animal, a vector is prepared which contains at least a portion of a 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene into which a deletion, addition or substitution has been introduced to thereby alter, e.g., functionally disrupt, the 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene. The 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene can be a human gene but more preferably, is a non-human homologue of a human 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene. For example, a rat 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene can be used to construct a homologous recombination nucleic acid molecule, e.g., a vector, suitable for altering an endogenous 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene in the mouse genome. In a preferred embodiment, the homologous recombination nucleic acid molecule is designed such that, upon homologous recombination, the endogenous 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene is functionally disrupted (i.e., no longer encodes a functional protein; also referred to as a "knock out" vector). Alternatively, the homologous recombination nucleic acid molecule can be designed such that, upon homologous recombination, the endogenous 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene is mutated or otherwise altered but still encodes functional protein (e.g., the upstream
regulatory region can be altered to thereby alter the expression of the endogenous 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein). In the homologous recombination nucleic acid molecule, the altered portion of the 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene is flanked at its 5’ and 3’ ends by additional nucleic acid sequence of the 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene to allow for homologous recombination to occur between the exogenous 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene carried by the homologous recombination nucleic acid molecule and an endogenous 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene in a cell, e.g., an embryonic stem cell. The additional flanking 252, 304, 1980, 14717, 9941, 19310 OR 17832 nucleic acid sequence is of sufficient length for successful homologous recombination with the endogenous gene.

Typically, several kilobases of flanking DNA (both at the 5’ and 3’ ends) are included in the homologous recombination nucleic acid molecule (see, e.g., Thomas, K.R. and Capecchi, M. R. (1987) Cell 51:503 for a description of homologous recombination vectors). The homologous recombination nucleic acid molecule is introduced into a cell, e.g., an embryonic stem cell line (e.g., by electroporation) and cells in which the introduced 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene has homologously recombined with the endogenous 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene are selected (see e.g., Li, E. et al. (1992) Cell 69:915). The selected cells can then injected into a blastocyst of an animal (e.g., a mouse) to form aggregation chimeras (see e.g., Bradley, A. in *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach*, E.J. Robertson, ed. (IRL, Oxford, 1987) pp. 113-152). A chimeric embryo can then be implanted into a suitable pseudopregnant female foster animal and the embryo brought to term. Progeny harboring the homologously recombined DNA in their germ cells can be used to breed animals in which all cells of the animal contain the homologously recombined DNA by germline transmission of the transgene. Methods for constructing homologous recombination nucleic acid molecules, e.g., vectors, or homologous recombinant animals are described further in Bradley, A. (1991) *Current Opinion in Biotechnology* 2:823-829 and in PCT International Publication Nos.: WO 90/11354 by Le Mouellec et al.; WO 91/01140 by Smithies et al.; WO 92/0968 by Zijlstra et al.; and WO 93/04169 by Berns et al.

In another embodiment, transgenic non-human animals for use in the methods of the invention can be produced which contain selected systems which allow for regulated expression of the transgene. One example of such a system is the cre/loxP recombinase system of bacteriophage P1. For a description of the cre/loxP recombinase system, see,
e.g., Lakso et al. (1992) Proc. Natl. Acad. Sci. USA 89:6232-6236. Another example of a recombinase system is the FLP recombinase system of *Saccharomyces cerevisiae* (O'Gorman et al. 1991) Science 251:1351-1355. If a cre/loxP recombinase system is used to regulate expression of the transgene, animals containing transgenes encoding both the Cre recombinase and a selected protein are required. Such animals can be provided through the construction of "double" transgenic animals, e.g., by mating two transgenic animals, one containing a transgene encoding a selected protein and the other containing a transgene encoding a recombinase.

Clones of the non-human transgenic animals described herein can also be produced according to the methods described in Wilmut, I. et al. (1997) *Nature* 385:810-813 and PCT International Publication Nos. WO 97/07668 and WO 97/07669. In brief, a cell, e.g., a somatic cell, from the transgenic animal can be isolated and induced to exit the growth cycle and enter G_0_ phase. The quiescent cell can then be fused, e.g., through the use of electrical pulses, to an enucleated oocyte from an animal of the same species from which the quiescent cell is isolated. The reconstructed oocyte is then cultured such that it develops to morula or blastocyst and then transferred to pseudopregnant female foster animal. The offspring borne of this female foster animal will be a clone of the animal from which the cell, e.g., the somatic cell, is isolated.

The 252, 304, 1980, 14717, 9941, 19310 OR 17832 transgenic animals that express 252, 304, 1980, 14717, 9941, 19310 OR 17832 mRNA or a 252, 304, 1980, 14717, 9941, 19310 OR 17832 peptide (detected immunocytochemically, using antibodies directed against 252, 304, 1980, 14717, 9941, 19310 OR 17832 epitopes) at easily detectable levels should then be further evaluated to identify those animals which display characteristic hematological disorders.

B. Cell-Based Systems

Cells that contain and express 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene sequences which encode a 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein, and, further, exhibit cellular phenotypes associated with e.g. hematopoiesis, may be used to identify compounds that exhibit an effect. Such cells may include non-recombinant monocyte cell lines, such as U937 (ATCC# CRL-1593), THP-1 (ATCC#TIB-202), and P388D1 (ATCC# TIB-63); endothelial cells such as human umbilical vein endothelial cells (HUVECs), human microvascular endothelial cells (HMVEC), and bovine aortic endothelial cells (BAECs); as well as generic mammalian cell lines such as HeLa cells and COS cells, e.g., COS-7 (ATCC# CRL-1651), cells described supra which constitute those...
cells relevant to hematology. Further, such cells may include recombinant, transgenic cell
lines. For example, the hematological disorders animal models of the invention, discussed
above, may be used to generate cell lines, containing one or more cell types involved in e.g.
hematopoeisis, that can be used as cell culture models for this disorder. While primary
cultures derived from the hematological disorders model transgenic animals of the
invention may be utilized, the generation of continuous cell lines is preferred. For
examples of techniques which may be used to derive a continuous cell line from the

Alternatively, cells of a cell type known to be involved in e.g. hematopoeisis may
be transfected with sequences capable of increasing or decreasing the amount of 252, 304,
1980, 14717, 9941, 19310 OR 17832 gene expression within the cell. For example, 252,
304, 1980, 14717, 9941, 19310 OR 17832 gene sequences may be introduced into, and
overexpressed in, the genome of the cell of interest, or, if endogenous 252, 304, 1980,
14717, 9941, 19310 OR 17832 gene sequences are present, they may be either
overexpressed or, alternatively disrupted in order to underexpress or inactivate 252, 304,
1980, 14717, 9941, 19310 OR 17832 gene expression.

In order to overexpress a 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene, the
coding portion of the 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene may be ligated
to a regulatory sequence which is capable of driving gene expression in the cell type of
interest, e.g., an endothelial cell. Such regulatory regions will be well known to those of
skill in the art, and may be utilized in the absence of undue experimentation. Recombinant
methods for expressing target genes are described above.

For underexpression of an endogenous 252, 304, 1980, 14717, 9941, 19310 OR
17832 gene sequence, such a sequence may be isolated and engineered such that when
reintroduced into the genome of the cell type of interest, the endogenous 252, 304, 1980,
14717, 9941, 19310 OR 17832 alleles will be inactivated. Preferably, the engineered 252,
304, 1980, 14717, 9941, 19310 OR 17832 sequence is introduced via gene targeting such
that the endogenous 252, 304, 1980, 14717, 9941, 19310 OR 17832 sequence is disrupted
upon integration of the engineered 252, 304, 1980, 14717, 9941, 19310 OR 17832
sequence into the cell's genome. Transfection of host cells with 252, 304, 1980, 14717,
9941, 19310 OR 17832 genes is discussed, above.

Cells treated with compounds or transfected with 252, 304, 1980, 14717, 9941,
19310 OR 17832 genes can be examined for phenotypes associated with e.g.
hematopoeisis.
Transfection of 252, 304, 1980, 14717, 9941, 19310 OR 17832 nucleic acid may be accomplished by using standard techniques (described in, for example, Ausubel (1989) supra). Transfected cells should be evaluated for the presence of the recombinant 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene sequences, for expression and accumulation of 252, 304, 1980, 14717, 9941, 19310 OR 17832 mRNA, and for the presence of recombinant 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein production. In instances wherein a decrease in 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene expression is desired, standard techniques may be used to demonstrate whether a decrease in endogenous 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene expression and/or in 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein production is achieved.

III. Predictive Medicine:

The present invention also pertains to the field of predictive medicine in which diagnostic assays, prognostic assays, and monitoring clinical trials are used for prognostic (predictive) purposes to thereby treat an individual prophylactically. Accordingly, one aspect of the present invention relates to diagnostic assays for determining 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein and/or nucleic acid expression as well as 252, 304, 1980, 14717, 9941, 19310 OR 17832 activity, in the context of a biological sample (e.g., blood, serum, cells, e.g., endothelial cells, or tissue, e.g., vascular tissue) to thereby determine whether an individual is afflicted with a predisposition or is experiencing hematological disorders. The invention also provides for prognostic (or predictive) assays for determining whether an individual is at risk of developing a hematological disorder. For example, mutations in a 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene can be assayed for in a biological sample. Such assays can be used for prognostic or predictive purpose to thereby phophylactically treat an individual prior to the onset of a hematological disorder.

Another aspect of the invention pertains to monitoring the influence of 252, 304, 1980, 14717, 9941, 19310 OR 17832 modulators (e.g., anti- 252, 304, 1980, 14717, 9941, 19310 OR 17832 antibodies or 252, 304, 1980, 14717, 9941, 19310 OR 17832 ribozymes) on the expression or activity of 252, 304, 1980, 14717, 9941, 19310 OR 17832 in clinical trials.

These and other agents are described in further detail in the following sections.

A. Diagnostic Assays
To determine whether a subject is afflicted with a disease, a biological sample may be obtained from a subject and the biological sample may be contacted with a compound or an agent capable of detecting a 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein or nucleic acid (e.g., mRNA or genomic DNA) that encodes a 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein, in the biological sample. A preferred agent for detecting 252, 304, 1980, 14717, 9941, 19310 OR 17832 mRNA or genomic DNA is a labeled nucleic acid probe capable of hybridizing to 252, 304, 1980, 14717, 9941, 19310 OR 17832 mRNA or genomic DNA. The nucleic acid probe can be, for example, the 252, 304, 1980, 14717, 9941, 19310 OR 17832 nucleic acid set forth in SEQ ID NO: 1, 4, 7, 10, 13, 16 or 19, or a portion thereof, such as an oligonucleotide of at least 15, 20, 25, 30, 25, 40, 45, 50, 100, 250 or 500 nucleotides in length and sufficient to specifically hybridize under stringent conditions to 252, 304, 1980, 14717, 9941, 19310 OR 17832 mRNA or genomic DNA. Other suitable probes for use in the diagnostic assays of the invention are described herein.

A preferred agent for detecting 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein in a sample is an antibody capable of binding to 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein, preferably an antibody with a detectable label. Antibodies can be polyclonal, or more preferably, monoclonal. An intact antibody, or a fragment thereof (e.g., Fab or F(ab')2) can be used. The term "labeled", with regard to the probe or antibody, is intended to encompass direct labeling of the probe or antibody by coupling (i.e., physically linking) a detectable substance to the probe or antibody, as well as indirect labeling of the probe or antibody by reactivity with another reagent that is directly labeled. Examples of indirect labeling include detection of a primary antibody using a fluorescently labeled secondary antibody and end-labeling of a DNA probe with biotin such that it can be detected with fluorescently labeled streptavidin.

The term "biological sample" is intended to include tissues, cells, and biological fluids isolated from a subject, as well as tissues, cells, and fluids present within a subject. That is, the detection method of the invention can be used to detect 252, 304, 1980, 14717, 9941, 19310 OR 17832 mRNA, protein, or genomic DNA in a biological sample in vitro as well as in vivo. For example, in vitro techniques for detection of 252, 304, 1980, 14717, 9941, 19310 OR 17832 mRNA include Northern hybridizations and in situ hybridizations. In vitro techniques for detection of 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein include enzyme linked immunosorbent assays (ELISAs), Western blots, immunoprecipitations and immunofluorescence. In vitro techniques for detection of 252,
304, 1980, 14717, 9941, 19310 OR 17832 genomic DNA include Southern hybridizations. Furthermore, in vivo techniques for detection of 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein include introducing into a subject a labeled anti- 252, 304, 1980, 14717, 9941, 19310 OR 17832 antibody. For example, the antibody can be labeled with a radioactive marker whose presence and location in a subject can be detected by standard imaging techniques.

In another embodiment, the methods further involve obtaining a control biological sample from a control subject, contacting the control sample with a compound or agent capable of detecting 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein, mRNA, or genomic DNA, such that the presence of 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein, mRNA or genomic DNA is detected in the biological sample, and comparing the presence of 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein, mRNA or genomic DNA in the control sample with the presence of 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein, mRNA or genomic DNA in the test sample.

B. Prognostic Assays

The present invention further pertains to methods for identifying subjects having or at risk of developing a disease associated with aberrant 252, 304, 1980, 14717, 9941, 19310 OR 17832 expression or activity.

As used herein, the term "aberrant" includes a 252, 304, 1980, 14717, 9941, 19310 OR 17832 expression or activity which deviates from the wild type 252, 304, 1980, 14717, 9941, 19310 OR 17832 expression or activity. Aberrant expression or activity includes increased or decreased expression or activity, as well as expression or activity which does not follow the wild type developmental pattern of expression or the subcellular pattern of expression. For example, aberrant 252, 304, 1980, 14717, 9941, 19310 OR 17832 expression or activity is intended to include the cases in which a mutation in the 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene causes the 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene to be under-expressed or over-expressed and situations in which such mutations result in a non-functional 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein or a protein which does not function in a wild-type fashion, e.g., a protein which does not interact with a 252, 304, 1980, 14717, 9941, 19310 OR 17832 substrate, or one which interacts with a non- 252, 304, 1980, 14717, 9941, 19310 OR 17832 substrate.

The assays described herein, such as the preceding diagnostic assays or the following assays, can be used to identify a subject having or at risk of developing a disease.
A biological sample may be obtained from a subject and tested for the presence or absence of a genetic alteration. For example, such genetic alterations can be detected by ascertaining the existence of at least one of 1) a deletion of one or more nucleotides from a 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene, 2) an addition of one or more nucleotides to a 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene, 3) a substitution of one or more nucleotides of a 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene, 4) a chromosomal rearrangement of a 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene, 5) an alteration in the level of a messenger RNA transcript of a 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene, 6) aberrant modification of a 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene, such as of the methylation pattern of the genomic DNA, 7) the presence of a non-wild type splicing pattern of a messenger RNA transcript of a 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene, 8) a non-wild type level of a 252, 304, 1980, 14717, 9941, 19310 OR 17832 -protein, 9) allelic loss of a 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene, and 10) inappropriate post-translational modification of a 252, 304, 1980, 14717, 9941, 19310 OR 17832 -protein.

As described herein, there are a large number of assays known in the art which can be used for detecting genetic alterations in a 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene. For example, a genetic alteration in a 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene may be detected using a probe/primer in a polymerase chain reaction (PCR) (see, e.g., U.S. Patent Nos. 4,683,195 and 4,683,202), such as anchor PCR or RACE PCR, or, alternatively, in a ligation chain reaction (LCR) (see, e.g., Landegran et al. (1988) Science 241:1077-1080; and Nakazawa et al. (1994) Proc. Natl. Acad. Sci. USA 91:360-364), the latter of which can be particularly useful for detecting point mutations in a 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene (see Abravaya et al. (1995) Nucleic Acids Res. 23:675-682). This method includes collecting a biological sample from a subject, isolating nucleic acid (e.g., genomic DNA, mRNA or both) from the sample, contacting the nucleic acid sample with one or more primers which specifically hybridize to a 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene under conditions such that hybridization and amplification of the 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene (if present) occurs, and detecting the presence or absence of an amplification product, or detecting the size of the amplification product and comparing the length to a control sample. It is anticipated that PCR and/or LCR may be desirable to use as a preliminary amplification step in conjunction with any of the techniques used for detecting mutations described herein.
Alternative amplification methods include: self sustained sequence replication
Q-Beta Replicase (Lizardi, P.M. et al. (1988) *Bio-Technology* 6:1197), or any other nucleic
acid amplification method, followed by the detection of the amplified molecules using
techniques well known to those of skill in the art. These detection schemes are especially
useful for the detection of nucleic acid molecules if such molecules are present in very low
numbers.

In an alternative embodiment, mutations in a 252, 304, 1980, 14717, 9941, 19310
OR 17832 gene from a biological sample can be identified by alterations in restriction
enzyme cleavage patterns. For example, sample and control DNA is isolated, amplified
(optionally), digested with one or more restriction endonucleases, and fragment length
sizes are determined by gel electrophoresis and compared. Differences in fragment length
sizes between sample and control DNA indicates mutations in the sample DNA.

Moreover, the use of sequence specific ribozymes (see, for example, U.S. Patent No.
5,498,531) can be used to score for the presence of specific mutations by development or
loss of a ribozyme cleavage site.

In other embodiments, genetic mutations in 252, 304, 1980, 14717, 9941, 19310
OR 17832 can be identified by hybridizing biological sample derived and control nucleic
acids, e.g., DNA or RNA, to high density arrays containing hundreds or thousands of
oligonucleotide probes (Cronin, M.T. et al. (1996) *Human Mutation* 7:244-255; Kozal,
M.J. et al. (1996) *Nature Medicine* 2:753-759). For example, genetic mutations in 252,
304, 1980, 14717, 9941, 19310 OR 17832 can be identified in two dimensional arrays
containing light-generated DNA probes as described in Cronin, M.T. et al. (1996) *supra*.

Briefly, a first hybridization array of probes can be used to scan through long stretches of
DNA in a sample and control to identify base changes between the sequences by making
linear arrays of sequential, overlapping probes. This step allows for the identification of
point mutations. This step is followed by a second hybridization array that allows for the
characterization of specific mutations by using smaller, specialized probe arrays
complementary to all variants or mutations detected. Each mutation array is composed of
parallel probe sets, one complementary to the wild-type gene and the other complementary
to the mutant gene.

In yet another embodiment, any of a variety of sequencing reactions known in the
art can be used to directly sequence the 252, 304, 1980, 14717, 9941, 19310 OR 17832

Other methods for detecting mutations in the 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene include methods in which protection from cleavage agents is used to detect mismatched bases in RNA/RNA or RNA/DNA heteroduplexes (Myers et al. (1985) *Science* 230:1242). In general, the art technique of "mismatch cleavage" starts by providing heteroduplexes formed by hybridizing (labeled) RNA or DNA containing the wild-type 252, 304, 1980, 14717, 9941, 19310 OR 17832 sequence with potentially mutant RNA or DNA obtained from a tissue sample. The double-stranded duplexes are treated with an agent which cleaves single-stranded regions of the duplex such as which will exist due to basepair mismatches between the control and sample strands. For instance, RNA/DNA duplexes can be treated with RNase and DNA/DNA hybrids treated with S1 nuclease to enzymatically digest the mismatched regions. In other embodiments, either DNA/DNA or RNA/DNA duplexes can be treated with hydroxylamine or osmium tetroxide and with piperidine in order to digest mismatched regions. After digestion of the mismatched regions, the resulting material is then separated by size on denaturing polyacrylamide gels to determine the site of mutation. See, for example, Cotton et al. (1988) *Proc. Natl Acad Sci USA* 85:4397 and Saleeba et al. (1992) *Methods Enzymol.* 217:286-295. In a preferred embodiment, the control DNA or RNA can be labeled for detection.

In still another embodiment, the mismatch cleavage reaction employs one or more proteins that recognize mismatched base pairs in double-stranded DNA (so called "DNA mismatch repair" enzymes) in defined systems for detecting and mapping point mutations in 252, 304, 1980, 14717, 9941, 19310 OR 17832 cDNAs obtained from samples of cells. For example, the mutY enzyme of *E. coli* cleaves A at G/A mismatches and the thymidine DNA glycosylase from HeLa cells cleaves T at G/T mismatches (Hsu et al. (1994)
Carcinogenesis 15:1657-1662). According to an exemplary embodiment, a probe based on a 252, 304, 1980, 14717, 9941, 19310 OR 17832 sequence, e.g., a wild-type 252, 304, 1980, 14717, 9941, 19310 OR 17832 sequence, is hybridized to a cDNA or other DNA product from a test cell(s). The duplex is treated with a DNA mismatch repair enzyme, and the cleavage products, if any, can be detected from electrophoresis protocols or the like. See, for example, U.S. Patent No. 5,459,039.

In other embodiments, alterations in electrophoretic mobility will be used to identify mutations in 252, 304, 1980, 14717, 9941, 19310 OR 17832 genes. For example, single strand conformation polymorphism (SSCP) may be used to detect differences in electrophoretic mobility between mutant and wild type nucleic acids (Orita et al. (1989) Proc Natl. Acad. Sci USA: 86:2766; see also Cotton (1993) Mutat. Res. 285:125-144 and Hayashi (1992) Genet. Anal. Tech. Appl. 9:73-79). Single-stranded DNA fragments of sample and control 252, 304, 1980, 14717, 9941, 19310 OR 17832 nucleic acids will be denatured and allowed to renature. The secondary structure of single-stranded nucleic acids varies according to sequence, the resulting alteration in electrophoretic mobility enables the detection of even a single base change. The DNA fragments may be labeled or detected with labeled probes. The sensitivity of the assay may be enhanced by using RNA (rather than DNA), in which the secondary structure is more sensitive to a change in sequence. In a preferred embodiment, the subject method utilizes heteroduplex analysis to separate double stranded heteroduplex molecules on the basis of changes in electrophoretic mobility (Keen et al. (1991) Trends Genet 7:5).

In yet another embodiment the movement of mutant or wild-type fragments in polyacrylamide gels containing a gradient of denaturant is assayed using denaturing gradient gel electrophoresis (DGGE) (Myers et al. (1985) Nature 313:495). When DGGE is used as the method of analysis, DNA will be modified to ensure that it does not completely denature, for example by adding a GC clamp of approximately 40 bp of high-melting GC-rich DNA by PCR. In a further embodiment, a temperature gradient is used in place of a denaturing gradient to identify differences in the mobility of control and sample DNA (Rosenbaum and Reissner (1987) Biophys Chem 265:12753).

Examples of other techniques for detecting point mutations include, but are not limited to, selective oligonucleotide hybridization, selective amplification, or selective primer extension. For example, oligonucleotide primers may be prepared in which the known mutation is placed centrally and then hybridized to target DNA under conditions which permit hybridization only if a perfect match is found (Saiki et al. (1986) Nature
324:163); Saiki et al. (1989) Proc. Natl Acad. Sci USA 86:6230). Such allele specific oligonucleotides are hybridized to PCR amplified target DNA or a number of different mutations when the oligonucleotides are attached to the hybridizing membrane and hybridized with labeled target DNA.

Alternatively, allele specific amplification technology which depends on selective PCR amplification may be used in conjunction with the instant invention. Oligonucleotides used as primers for specific amplification may carry the mutation of interest in the center of the molecule (so that amplification depends on differential hybridization) (Gibbs et al. (1989) Nucleic Acids Res. 17:2437-2448) or at the extreme 3' end of one primer where, under appropriate conditions, mismatch can prevent, or reduce polymerase extension (Prossner (1993) Tibtech 11:238). In addition it may be desirable to introduce a novel restriction site in the region of the mutation to create cleavage-based detection (Gasparini et al. (1992) Mol. Cell Probes 6:1). It is anticipated that in certain embodiments amplification may also be performed using Taq ligase for amplification (Barany (1991) Proc. Natl. Acad. Sci USA 88:189). In such cases, ligation will occur only if there is a perfect match at the 3' end of the 5' sequence making it possible to detect the presence of a known mutation at a specific site by looking for the presence or absence of amplification.

Furthermore, the prognostic assays described herein can be used to determine whether a subject can be administered a 252, 304, 1980, 14717, 9941, 19310 OR 17832 modulator (e.g., an agonist, antagonist, peptidomimetic, protein, peptide, nucleic acid, or small molecule) to effectively treat a disease.

C. Monitoring of Effects During Clinical Trials

The present invention further provides methods for determining the effectiveness of a 252, 304, 1980, 14717, 9941, 19310 OR 17832 modulator (e.g., a 252, 304, 1980, 14717, 9941, 19310 OR 17832 modulator identified herein) in treating a disease. For example, the effectiveness of a 252, 304, 1980, 14717, 9941, 19310 OR 17832 modulator in increasing 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene expression, protein levels, or in upregulating 252, 304, 1980, 14717, 9941, 19310 OR 17832 activity, can be monitored in clinical trials of subjects exhibiting decreased 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene expression, protein levels, or downregulated 252, 304, 1980, 14717, 9941, 19310 OR 17832 activity. Alternatively, the effectiveness of a 252, 304, 1980, 14717, 9941, 19310 OR 17832 modulator in decreasing 252, 304, 1980, 14717, 9941, 19310 OR
17832 gene expression, protein levels, or in downregulating 252, 304, 1980, 14717, 9941, 19310 OR 17832 activity, can be monitored in clinical trials of subjects exhibiting increased 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene expression, protein levels, or 252, 304, 1980, 14717, 9941, 19310 OR 17832 activity. In such clinical trials, the expression or activity of a 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene, and preferably, other genes that have been implicated in e.g. hematopoiesis can be used as a "read out" or marker of the phenotype of a particular cell.

For example, and not by way of limitation, genes, including 252, 304, 1980, 14717, 9941, 19310 OR 17832, that are modulated in cells by treatment with an agent which modulates 252, 304, 1980, 14717, 9941, 19310 OR 17832 activity (e.g., identified in a screening assay as described herein) can be identified. Thus, to study the effect of agents which modulate 252, 304, 1980, 14717, 9941, 19310 OR 17832 activity on subjects suffering from a hematological disorder in, for example, a clinical trial, cells can be isolated and RNA prepared and analyzed for the levels of expression of 252, 304, 1980, 14717, 9941, 19310 OR 17832 and other genes implicated in the hematological disorders disorder. The levels of gene expression (e.g., a gene expression pattern) can be quantified by Northern blot analysis or RT-PCR, as described herein, or alternatively by measuring the amount of protein produced, by one of the methods described herein, or by measuring the levels of activity of 252, 304, 1980, 14717, 9941, 19310 OR 17832 or other genes. In this way, the gene expression pattern can serve as a marker, indicative of the physiological response of the cells to the agent which modulates 252, 304, 1980, 14717, 9941, 19310 OR 17832 activity. This response state may be determined before, and at various points during treatment of the individual with the agent which modulates 252, 304, 1980, 14717, 9941, 19310 OR 17832 activity.

In a preferred embodiment, the present invention provides a method for monitoring the effectiveness of treatment of a subject with an agent which modulates 252, 304, 1980, 14717, 9941, 19310 OR 17832 activity (e.g., an agonist, antagonist, peptidomimetic, protein, peptide, nucleic acid, or small molecule identified by the screening assays described herein) including the steps of (i) obtaining a pre-administration sample from a subject prior to administration of the agent; (ii) detecting the level of expression of a 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein, mRNA, or genomic DNA in the pre-administration sample; (iii) obtaining one or more post-administration samples from the subject; (iv) detecting the level of expression or activity of the 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein, mRNA, or genomic DNA in the post-administration
samples; (v) comparing the level of expression or activity of the 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein, mRNA, or genomic DNA in the pre-administration sample with the 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein, mRNA, or genomic DNA in the post administration sample or samples; and (vi) altering the administration of the agent to the subject accordingly. For example, increased administration of the agent may be desirable to increase the expression or activity of 252, 304, 1980, 14717, 9941, 19310 OR 17832 to higher levels than detected, i.e., to increase the effectiveness of the agent. Alternatively, decreased administration of the agent may be desirable to decrease expression or activity of 252, 304, 1980, 14717, 9941, 19310 OR 17832 to lower levels than detected, i.e., to decrease the effectiveness of the agent. According to such an embodiment, 252, 304, 1980, 14717, 9941, 19310 OR 17832 expression or activity may be used as an indicator of the effectiveness of an agent, even in the absence of an observable phenotypic response.

IV. Methods of Treatment:

The present invention provides for both prophylactic and therapeutic methods of treating a subject, e.g., a human, at risk of (or susceptible to) a disease. With regard to both prophylactic and therapeutic methods of treatment, such treatments may be specifically tailored or modified, based on knowledge obtained from the field of pharmacogenomics. "Pharmacogenomics," as used herein, refers to the application of genomics technologies such as gene sequencing, statistical genetics, and gene expression analysis to drugs in clinical development and on the market. More specifically, the term refers to the study of how a patient's genes determine his or her response to a drug (e.g., a patient's "drug response phenotype", or "drug response genotype").

Thus, another aspect of the invention provides methods for tailoring an subject's prophylactic or therapeutic treatment with either the 252, 304, 1980, 14717, 9941, 19310 OR 17832 molecules of the present invention or 252, 304, 1980, 14717, 9941, 19310 OR 17832 modulators according to that individual's drug response genotype. Pharmacogenomics allows a clinician or physician to target prophylactic or therapeutic treatments to patients who will most benefit from the treatment and to avoid treatment of patients who will experience toxic drug-related side effects.

A. Prophylactic Methods

- 38 -
In one aspect, the invention provides a method for preventing in a subject, a disease by administering to the subject an agent which modulates 252, 304, 1980, 14717, 9941, 19310 OR 17832 expression or 252, 304, 1980, 14717, 9941, 19310 OR 17832 activity. Subjects at risk for a hematological disorder, can be identified by, for example, any or a combination of the diagnostic or prognostic assays described herein. Administration of a prophylactic agent can occur prior to the manifestation of symptoms characteristic of aberrant 252, 304, 1980, 14717, 9941, 19310 OR 17832 expression or activity, such that a disease is prevented or, alternatively, delayed in its progression. Depending on the type of 252, 304, 1980, 14717, 9941, 19310 OR 17832 aberrancy, for example, a 252, 304, 1980, 14717, 9941, 19310 OR 17832 agonist or 252, 304, 1980, 14717, 9941, 19310 OR 17832 antagonist agent can be used for treating the subject. The appropriate agent can be determined based on screening assays described herein.

B. Therapeutic Methods

Described herein are methods and compositions whereby hematological disorders may be ameliorated. Certain hematological disorders disorders are brought about, at least in part, by an excessive level of a gene product, or by the presence of a gene product exhibiting an abnormal or excessive activity. As such, the reduction in the level and/or activity of such gene products would bring about the amelioration of hematological disorders. Techniques for the reduction of gene expression levels or the activity of a protein are discussed below.

Alternatively, certain other hematological disorders disorders are brought about, at least in part, by the absence or reduction of the level of gene expression, or a reduction in the level of a protein’s activity. As such, an increase in the level of gene expression and/or the activity of such proteins would bring about the amelioration of hematological disorders.

In some cases, the up-regulation of a gene in a disease state reflects a protective role for that gene product in responding to the disease condition. Enhancement of such a gene’s expression, or the activity of the gene product, will reinforce the protective effect it exerts. Some hematological disorders states may result from an abnormally low level of activity of such a protective gene. In these cases also, an increase in the level of gene expression and/or the activity of such gene products would bring about the amelioration of hematological disorders. Techniques for increasing target gene expression levels or target gene product activity levels are discussed herein.
Accordingly, another aspect of the invention pertains to methods of modulating 252, 304, 1980, 14717, 9941, 19310 OR 17832 expression or activity for therapeutic purposes. Accordingly, in an exemplary embodiment, the modulatory method of the invention involves contacting a cell with a 252, 304, 1980, 14717, 9941, 19310 OR 17832 or agent that modulates one or more of the activities of 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein activity associated with the cell (e.g., an endothelial cell or an ovarian cell). An agent that modulates 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein activity can be an agent as described herein, such as a nucleic acid or a protein, a naturally-occurring target molecule of a 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein (e.g., a 252, 304, 1980, 14717, 9941, 19310 OR 17832 ligand or substrate), a 252, 304, 1980, 14717, 9941, 19310 OR 17832 antibody, a 252, 304, 1980, 14717, 9941, 19310 OR 17832 agonist or antagonist, a peptidomimetic of a 252, 304, 1980, 14717, 9941, 19310 OR 17832 agonist or antagonist, or other small molecule. In one embodiment, the agent stimulates one or more 252, 304, 1980, 14717, 9941, 19310 OR 17832 activities. Examples of such stimulatory agents include active 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein and a nucleic acid molecule encoding 252, 304, 1980, 14717, 9941, 19310 OR 17832 that has been introduced into the cell. In another embodiment, the agent inhibits one or more 252, 304, 1980, 14717, 9941, 19310 OR 17832 activities. Examples of such inhibitory agents include antisense 252, 304, 1980, 14717, 9941, 19310 OR 17832 nucleic acid molecules, anti- 252, 304, 1980, 14717, 9941, 19310 OR 17832 antibodies, and 252, 304, 1980, 14717, 9941, 19310 OR 17832 inhibitors. These modulatory methods can be performed in vitro (e.g., by culturing the cell with the agent) or, alternatively, in vivo (e.g., by administering the agent to a subject). As such, the present invention provides methods of treating an individual afflicted with a disease or disorder characterized by aberrant or unwanted expression or activity of a 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein or nucleic acid molecule. In one embodiment, the method involves administering an agent (e.g., an agent identified by a screening assay described herein), or combination of agents that modulates (e.g., upregulates or downregulates) 252, 304, 1980, 14717, 9941, 19310 OR 17832 expression or activity. In another embodiment, the method involves administering a 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein or nucleic acid molecule as therapy to compensate for reduced, aberrant, or unwanted 252, 304, 1980, 14717, 9941, 19310 OR 17832 expression or activity.

Stimulation of 252, 304, 1980, 14717, 9941, 19310 OR 17832 activity is desirable in situations in which 252, 304, 1980, 14717, 9941, 19310 OR 17832 is abnormally
downregulated and/or in which increased 252, 304, 1980, 14717, 9941, 19310 OR 17832 activity is likely to have a beneficial effect. Likewise, inhibition of 252, 304, 1980, 14717, 9941, 19310 OR 17832 activity is desirable in situations in which 252, 304, 1980, 14717, 9941, 19310 OR 17832 is abnormally upregulated and/or in which decreased 252, 304, 1980, 14717, 9941, 19310 OR 17832 activity is likely to have a beneficial effect.

(i) Methods for Inhibiting Target Gene Expression, Synthesis, or Activity

As discussed above, genes involved in cardiovascular disorders may cause such disorders via an increased level of gene activity. In some cases, such up-regulation may have a causative or exacerbating effect on the disease state. A variety of techniques may be used to inhibit the expression, synthesis, or activity of such genes and/or proteins.

For example, compounds such as those identified through assays described above, which exhibit inhibitory activity, may be used in accordance with the invention to at least one symptom of hematological disorders. Such molecules may include, but are not limited to, small organic molecules, peptides, antibodies, and the like.

For example, compounds can be administered that compete with endogenous ligand for the 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein. The resulting reduction in the amount of ligand-bound 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein will modulate endothelial cell physiology. Compounds that can be particularly useful for this purpose include, for example, soluble proteins or peptides, such as peptides comprising one or more of the extracellular domains, or portions and/or analogs thereof, of the 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein, including, for example, soluble fusion proteins such as Ig-tailed fusion proteins. (For a discussion of the production of Ig-tailed fusion proteins, see, for example, U.S. Patent No. 5,116,964). Alternatively, compounds, such as ligand analogs or antibodies, that bind to the 252, 304, 1980, 14717, 9941, 19310 OR 17832 receptor site, but do not activate the protein, (e.g., receptor-ligand antagonists) can be effective in inhibiting 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein activity.

Further, antisense and ribozyme molecules which inhibit expression of the 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene may also be used in accordance with the invention to inhibit aberrant 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene activity. Still further, triple helix molecules may be utilized in inhibiting aberrant 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene activity.
The antisense nucleic acid molecules used in the methods of the invention are typically administered to a subject or generated in situ such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding a 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein to thereby inhibit expression of the protein, e.g., by inhibiting transcription and/or translation. The hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule which binds to DNA duplexes, through specific interactions in the major groove of the double helix. An example of a route of administration of antisense nucleic acid molecules of the invention include direct injection at a tissue site.

Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then administered systemically. For example, for systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface, e.g., by linking the antisense nucleic acid molecules to peptides or antibodies which bind to cell surface receptors or antigens. The antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. To achieve sufficient intracellular concentrations of the antisense molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a strong pol II or pol III promoter are preferred.

In yet another embodiment, an antisense nucleic acid molecule used in the methods of the invention is an α-anomeric nucleic acid molecule. An α-anomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual β-units, the strands run parallel to each other (Gaultier et al. (1987) Nucleic Acids. Res. 15:6625-6641). The antisense nucleic acid molecule can also comprise a 2'-o-methylribonucleotide (Inoue et al. (1987) Nucleic Acids Res. 15:6131-6148) or a chimeric RNA-DNA analogue (Inoue et al. (1987) FEBS Lett. 215:327-330).

In still another embodiment, an antisense nucleic acid used in the methods of the invention is a ribozyme. Ribozymes are catalytic RNA molecules with ribonuclease activity which are capable of cleaving a single-stranded nucleic acid, such as an mRNA, to which they have a complementary region. Thus, ribozymes (e.g., hammerhead ribozymes (described in Haselhoff and Gerlach (1988) Nature 334:585-591)) can be used to catalytically cleave 252, 304, 1980, 14717, 9941, 19310 OR 17832 mRNA transcripts to thereby inhibit translation of 252, 304, 1980, 14717, 9941, 19310 OR 17832 mRNA. A ribozyme having specificity for a 577, 20739 or 57145-encoding nucleic acid can be designed based upon the nucleotide sequence of a 252, 304, 1980, 14717, 9941, 19310 OR
17832 cDNA disclosed herein (i.e., SEQ ID NO:1 or 3). For example, a derivative of a
Tetrahymena L-19 IVS RNA can be constructed in which the nucleotide sequence of the
active site is complementary to the nucleotide sequence to be cleaved in a 577, 20739 or
57145-encoding mRNA (see, for example, Cech et al. U.S. Patent No. 4,987,071; and Cech
17832 mRNA can be used to select a catalytic RNA having a specific ribonuclease activity
from a pool of RNA molecules (see, for example, Bartel, D. and Szostak, J.W. (1993)
Science 261:1411-1418).

252, 304, 1980, 14717, 9941, 19310 OR 17832 gene expression can also be
inhibited by targeting nucleotide sequences complementary to the regulatory region of the
252, 304, 1980, 14717, 9941, 19310 OR 17832 (e.g., the 252, 304, 1980, 14717, 9941,
19310 OR 17832 promoter and/or enhancers) to form triple helical structures that prevent
transcription of the 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene in target cells (see,

Antibodies that are both specific for the 252, 304, 1980, 14717, 9941, 19310 OR
17832 protein and interfere with its activity may also be used to modulate or inhibit 252,
304, 1980, 14717, 9941, 19310 OR 17832 protein function. Such antibodies may be
generated using standard techniques described herein, against the 252, 304, 1980, 14717,
9941, 19310 OR 17832 protein itself or against peptides corresponding to portions of the
protein. Such antibodies include but are not limited to polyclonal, monoclonal, Fab
fragments, single chain antibodies, or chimeric antibodies.

In instances where the target gene protein is intracellular and whole antibodies are
used, internalizing antibodies may be preferred. Lipofectin liposomes may be used to
deliver the antibody or a fragment of the Fab region which binds to the target epitope into
cells. Where fragments of the antibody are used, the smallest inhibitory
fragment which binds to the target protein’s binding domain is preferred. For example,
peptides having an amino acid sequence corresponding to the domain of the variable region
of the antibody that binds to the target gene protein may be used. Such peptides may be
synthesized chemically or produced via recombinant DNA technology using
methods well known in the art (described in, for example, Creighton (1983), supra; and
Sambrook et al. (1989) supra). Single chain neutralizing antibodies which bind to
intracellular target gene epitopes may also be administered. Such single chain antibodies
may be administered, for example, by expressing nucleotide sequences encoding single-
chain antibodies within the target cell population by utilizing, for example, techniques such as those described in Marasco et al. (1993) Proc. Natl. Acad. Sci. USA 90:7889-7893).

In some instances, the target gene protein is extracellular, or is a transmembrane protein, such as the 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein. Antibodies that are specific for one or more extracellular domains of the 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein, for example, and that interfere with its activity, are particularly useful in treating hematological disorders or a hematological disorder. Such antibodies are especially efficient because they can access the target domains directly from the bloodstream. Any of the administration techniques described below which are appropriate for peptide administration may be utilized to effectively administer inhibitory target gene antibodies to their site of action.

(ii) Methods for Restoring or Enhancing Target Gene Activity

Genes that cause hematological disorders may be underexpressed within disease situations. Alternatively, the activity of the protein products of such genes may be decreased, leading to the development of hematological disorders. Such down-regulation of gene expression or decrease of protein activity might have a causative or exacerbating effect on the disease state.

In some cases, genes that are up-regulated in the disease state might be exerting a protective effect. A variety of techniques may be used to increase the expression, synthesis, or activity of genes and/or proteins that exert a protective effect in response to hematological disorders conditions.

Described in this section are methods whereby the level 252, 304, 1980, 14717, 9941, 19310 OR 17832 activity may be increased to levels wherein hematological disorders are ameliorated. The level of 252, 304, 1980, 14717, 9941, 19310 OR 17832 activity may be increased, for example, by either increasing the level of 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene expression or by increasing the level of active 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein which is present.

For example, a 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein, at a level sufficient to at least one symptom of hematological disorders may be administered to a patient exhibiting such symptoms. Any of the techniques discussed below may be used for such administration. One of skill in the art will readily know how to determine the concentration of effective, non-toxic doses of the 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein, utilizing techniques such as those described below.
Additionally, RNA sequences encoding a 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein may be directly administered to a patient exhibiting hematological disorders, at a concentration sufficient to produce a level of 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein such that hematological disorders are ameliorated. Any of the techniques discussed below, which achieve intracellular administration of compounds, such as, for example, liposome administration, may be used for the administration of such RNA molecules. The RNA molecules may be produced, for example, by recombinant techniques such as those described herein.

Further, subjects may be treated by gene replacement therapy. One or more copies of a 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene, or a portion thereof, that directs the production of a normal 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein with 252, 304, 1980, 14717, 9941, 19310 OR 17832 function, may be inserted into cells using vectors which include, but are not limited to adenovirus, adeno-associated virus, and retrovirus vectors, in addition to other particles that introduce DNA into cells, such as liposomes. Additionally, techniques such as those described above may be used for the introduction of 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene sequences into human cells.

Cells, preferably, autologous cells, containing 252, 304, 1980, 14717, 9941, 19310 OR 17832 expressing gene sequences may then be introduced or reintroduced into the subject at positions which allow for the amelioration of hematological disorders. Such cell replacement techniques may be preferred, for example, when the gene product is a secreted, extracellular gene product.

C. Pharmaceutical Compositions

Another aspect of the invention pertains to methods for treating a subject suffering from a disease. These methods involve administering to a subject an agent which modulates 252, 304, 1980, 14717, 9941, 19310 OR 17832 expression or activity (e.g., an agent identified by a screening assay described herein), or a combination of such agents. In another embodiment, the method involves administering to a subject a 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein or nucleic acid molecule as therapy to compensate for reduced, aberrant, or unwanted 252, 304, 1980, 14717, 9941, 19310 OR 17832 expression or activity.

Stimulation of 252, 304, 1980, 14717, 9941, 19310 OR 17832 activity is desirable in situations in which 252, 304, 1980, 14717, 9941, 19310 OR 17832 is abnormally
downregulated and/or in which increased 252, 304, 1980, 14717, 9941, 19310 OR 17832
activity is likely to have a beneficial effect. Likewise, inhibition of 252, 304, 1980, 14717,
9941, 19310 OR 17832 activity is desirable in situations in which 252, 304, 1980, 14717,
9941, 19310 OR 17832 is abnormally upregulated and/or in which decreased 252, 304,
1980, 14717, 9941, 19310 OR 17832 activity is likely to have a beneficial effect.

The agents which modulate 252, 304, 1980, 14717, 9941, 19310 OR 17832 activity
can be administered to a subject using pharmaceutical compositions suitable for such
administration. Such compositions typically comprise the agent (e.g., nucleic acid
molecule, protein, or antibody) and a pharmaceutically acceptable carrier. As used herein
the language "pharmaceutically acceptable carrier" is intended to include any and all
solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and
absorption delaying agents, and the like, compatible with pharmaceutical administration.
The use of such media and agents for pharmaceutically active substances is well known in
the art. Except insofar as any conventional media or agent is incompatible with the active
compound, use thereof in the compositions is contemplated. Supplementary active
compounds can also be incorporated into the compositions.

A pharmaceutical composition used in the therapeutic methods of the invention is
formulated to be compatible with its intended route of administration. Examples of routes
of administration include parenteral, e.g., intravenous, intradermal, subcutaneous, oral
(e.g., inhalation), transdermal (topical), transmucosal, and rectal administration. Solutions
or suspensions used for parenteral, intradermal, or subcutaneous application can include the
following components: a sterile diluent such as water for injection, saline solution, fixed
oils, polyethylene glycols, glycerine, propylene glycol or other synthetic solvents;
antibacterial agents such as benzyl alcohol or methyl parabens; antioxidants such as
ascorbic acid or sodium bisulfite; chelating agents such as ethylenediaminetetraacetic acid;
buffers such as acetates, citrates or phosphates and agents for the adjustment of tonicity
such as sodium chloride or dextrose. pH can be adjusted with acids or bases, such as
hydrochloric acid or sodium hydroxide. The parenteral preparation can be enclosed in
ampoules, disposable syringes or multiple dose vials made of glass or plastic.

Pharmaceutical compositions suitable for injectable use include sterile aqueous
solutions (where water soluble) or dispersions and sterile powders for the extemporaneous
preparation of sterile injectable solutions or dispersion. For intravenous administration,
suitable carriers include physiological saline, bacteriostatic water, Cremophor EL™
(BASF, Parsippany, NJ) or phosphate buffered saline (PBS). In all cases, the composition
must be sterile and should be fluid to the extent that easy syringability exists. It must be
stable under the conditions of manufacture and storage and must be preserved against the
contaminating action of microorganisms such as bacteria and fungi. The carrier can be a
solvent or dispersion medium containing, for example, water, ethanol, polyol (for example,
glycerol, propylene glycol, and liquid polyethylene glycol, and the like), and suitable
mixtures thereof. The proper fluidity can be maintained, for example, by the use of a
coating such as lecithin, by the maintenance of the required particle size in the case of
dispersion and by the use of surfactants. Prevention of the action of microorganisms can be
achieved by various antibacterial and antifungal agents, for example, parabens,
chlorobutanol, phenol, ascorbic acid, thimerosal, and the like. In many cases, it will be
preferable to include isotonic agents, for example, sugars, polyalcohols such as manitol,
sorbitol, and sodium chloride in the composition. Prolonged absorption of the injectable
compositions can be brought about by including in the composition an agent which delays
absorption, for example, aluminum monostearate and gelatin.

Sterile injectable solutions can be prepared by incorporating the agent that
modulates 252, 304, 1980, 14717, 9941, 19310 OR 17832 activity (e.g., a fragment of a
252, 304, 1980, 14717, 9941, 19310 OR 17832 protein or an anti- 252, 304, 1980, 14717,
9941, 19310 OR 17832 antibody) in the required amount in an appropriate solvent with one
or a combination of ingredients enumerated above, as required, followed by filtered
sterilization. Generally, dispersions are prepared by incorporating the active compound
into a sterile vehicle which contains a basic dispersion medium and the required other
ingredients from those enumerated above. In the case of sterile powders for the preparation
of sterile injectable solutions, the preferred methods of preparation are vacuum drying and
freeze-drying which yields a powder of the active ingredient plus any additional desired
ingredient from a previously sterile-filtered solution thereof.

Oral compositions generally include an inert diluent or an edible carrier. They can
be enclosed in gelatin capsules or compressed into tablets. For the purpose of oral
therapeutic administration, the active compound can be incorporated with excipients and
used in the form of tablets, troches, or capsules. Oral compositions can also be prepared
using a fluid carrier for use as a mouthwash, wherein the compound in the fluid carrier is
applied orally and swished and expectorated or swallowed. Pharmaceutically compatible
binding agents, and/or adjuvant materials can be included as part of the composition. The
tablets, pills, capsules, troches and the like can contain any of the following ingredients, or
compounds of a similar nature: a binder such as microcrystalline cellulose, gum tragacanth
or gelatin; an excipient such as starch or lactose, a disintegrating agent such as alginic acid, Primogel, or corn starch; a lubricant such as magnesium stearate or Sterotes; a glidant such as colloidal silicon dioxide; a sweetening agent such as sucrose or saccharin; or a flavoring agent such as peppermint, methyl salicylate, or orange flavoring.

For administration by inhalation, the compounds are delivered in the form of an aerosol spray from pressurized container or dispenser which contains a suitable propellant, e.g., a gas such as carbon dioxide, or a nebulizer.

Systemic administration can also be by transmucosal or transdermal means. For transmucosal or transdermal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration, detergents, bile salts, and fusidic acid derivatives. Transmucosal administration can be accomplished through the use of nasal sprays or suppositories. For transdermal administration, the active compounds are formulated into ointments, salves, gels, or creams as generally known in the art.

The agents that modulate 252, 304, 1980, 14717, 9941, 19310 OR 17832 activity can also be prepared in the form of suppositories (e.g., with conventional suppository bases such as cocoa butter and other glycerides) or retention enemas for rectal delivery.

In one embodiment, the agents that modulate 252, 304, 1980, 14717, 9941, 19310 OR 17832 activity are prepared with carriers that will protect the compound against rapid elimination from the body, such as a controlled release formulation, including implants and microencapsulated delivery systems. Biodegradable, biocompatible polymers can be used, such as ethylene vinyl acetate, polyvinylalcohol, polyglycolic acid, collagen, polyorthoesters, and polyglyclic acid. Methods for preparation of such formulations will be apparent to those skilled in the art. The materials can also be obtained commercially from Alza Corporation and Nova Pharmaceuticals, Inc. Liposomal suspensions (including liposomes targeted to infected cells with monoclonal antibodies to viral antigens) can also be used as pharmaceutically acceptable carriers. These can be prepared according to methods known to those skilled in the art, for example, as described in U.S. Patent No. 4,522,811.

It is especially advantageous to formulate oral or parenteral compositions in dosage unit form for ease of administration and uniformity of dosage. Dosage unit form as used herein refers to physically discrete units suited as unitary dosages for the subject to be treated; each unit containing a predetermined quantity of active compound calculated to produce the desired therapeutic effect in association with the required pharmaceutical
carrier. The specification for the dosage unit forms of the invention are dictated by and
directly dependent on the unique characteristics of the agent that modulates 252, 304, 1980,
14717, 9941, 19310 OR 17832 activity and the particular therapeutic effect to be achieved,
and the limitations inherent in the art of compounding such an agent for the treatment of
subjects.

Toxicity and therapeutic efficacy of such agents can be determined by standard
pharmaceutical procedures in cell cultures or experimental animals, e.g., for determining
the LD50 (the dose lethal to 50% of the population) and the ED50 (the dose therapeutically
effective in 50% of the population). The dose ratio between toxic and therapeutic effects is
the therapeutic index and can be expressed as the ratio LD50/ED50. Agents which exhibit
large therapeutic indices are preferred. While agents that exhibit toxic side effects may be
used, care should be taken to design a delivery system that targets such agents to the site of
affected tissue in order to minimize potential damage to uninfected cells and, thereby,
reduce side effects.

The data obtained from the cell culture assays and animal studies can be used in
formulating a range of dosage for use in humans. The dosage of such 252, 304, 1980,
14717, 9941, 19310 OR 17832 modulating agents lies preferably within a range of
circulating concentrations that include the ED50 with little or no toxicity. The dosage may
vary within this range depending upon the dosage form employed and the route of
administration utilized. For any agent used in the therapeutic methods of the invention, the
therapeutically effective dose can be estimated initially from cell culture assays. A dose
may be formulated in animal models to achieve a circulating plasma concentration range
that includes the IC50 (i.e., the concentration of the test compound which achieves a half-
maximal inhibition of symptoms) as determined in cell culture. Such information can be
used to more accurately determine useful doses in humans. Levels in plasma may be
measured, for example, by high performance liquid chromatography.

As defined herein, a therapeutically effective amount of protein or polypeptide (i.e.,
an effective dosage) ranges from about 0.001 to 30 mg/kg body weight, preferably about
0.01 to 25 mg/kg body weight, more preferably about 0.1 to 20 mg/kg body weight, and
even more preferably about 1 to 10 mg/kg, 2 to 9 mg/kg, 3 to 8 mg/kg, 4 to 7 mg/kg, or 5 to
6 mg/kg body weight. The skilled artisan will appreciate that certain factors may influence
the dosage required to effectively treat a subject, including but not limited to the severity of
the disease or disorder, previous treatments, the general health and/or age of the subject,
and other diseases present. Moreover, treatment of a subject with a therapeutically
effective amount of a protein, polypeptide, or antibody can include a single treatment or, preferably, can include a series of treatments.

In a preferred example, a subject is treated with antibody, protein, or polypeptide in the range of between about 0.1 to 20 mg/kg body weight, one time per week for between about 1 to 10 weeks, preferably between 2 to 8 weeks, more preferably between about 3 to 7 weeks, and even more preferably for about 4, 5, or 6 weeks. It will also be appreciated that the effective dosage of antibody, protein, or polypeptide used for treatment may increase or decrease over the course of a particular treatment. Changes in dosage may result and become apparent from the results of diagnostic assays as described herein.

The present invention encompasses agents which modulate expression or activity. An agent may, for example, be a small molecule. For example, such small molecules include, but are not limited to, peptides, peptidomimetics, amino acids, amino acid analogs, polynucleotides, polynucleotide analogs, nucleotides, nucleotide analogs, organic or inorganic compounds (i.e., including heteroorganic and organometallic compounds) having a molecular weight less than about 10,000 grams per mole, organic or inorganic compounds having a molecular weight less than about 5,000 grams per mole, organic or inorganic compounds having a molecular weight less than about 1,000 grams per mole, organic or inorganic compounds having a molecular weight less than about 500 grams per mole, and salts, esters, and other pharmaceutically acceptable forms of such compounds. It is understood that appropriate doses of small molecule agents depends upon a number of factors within the ken of the ordinarily skilled physician, veterinarian, or researcher. The dose(s) of the small molecule will vary, for example, depending upon the identity, size, and condition of the subject or sample being treated, further depending upon the route by which the composition is to be administered, if applicable, and the effect which the practitioner desires the small molecule to have upon the nucleic acid or polypeptide of the invention.

Exemplary doses include milligram or microgram amounts of the small molecule per kilogram of subject or sample weight (e.g., about 1 microgram per kilogram to about 500 milligrams per kilogram, about 100 micrograms per kilogram to about 5 milligrams per kilogram, or about 1 microgram per kilogram to about 50 micrograms per kilogram). It is furthermore understood that appropriate doses of a small molecule depend upon the potency of the small molecule with respect to the expression or activity to be modulated. Such appropriate doses may be determined using the assays described herein. When one or more of these small molecules is to be administered to an animal (e.g., a human) in order to modulate expression or activity of a polypeptide or nucleic acid of the invention, a
physician, veterinarian, or researcher may, for example, prescribe a relatively low dose at first, subsequently increasing the dose until an appropriate response is obtained. In addition, it is understood that the specific dose level for any particular animal subject will depend upon a variety of factors including the activity of the specific compound employed, the age, body weight, general health, gender, and diet of the subject, the time of administration, the route of administration, the rate of excretion, any drug combination, and the degree of expression or activity to be modulated.

Further, an antibody (or fragment thereof) may be conjugated to a therapeutic moiety such as a cytotoxin, a therapeutic agent or a radioactive metal ion. A cytotoxin or cytotoxic agent includes any agent that is detrimental to cells. Examples include taxol, cytochalasin B, gramicidin D, ethidium bromide, emetine, mitomycin, etoposide, tenoposide, vincristine, vinblastine, colchicine, doxorubicin, daunorubicin, dihydroxy anthracycin dione, mitoxantrone, mithramycin, actinomycin D, 1-dehydrotestosterone, glucocorticoids, procaine, tetracaine, lidocaine, propranolol, and puromycin and analogs or homologs thereof. Therapeutic agents include, but are not limited to, antimitabolites (e.g., methotrexate, 6-mercaptopurine, 6-thioguanine, cytarabine, 5-fluorouracil decarbazine), alkylating agents (e.g., mechlorethamine, thioepa chlorambucil, melphalan, carmustine (BSNU) and lomustine (CCNU), cyclophosphamide, busulfan, dibromomannitol, streptozotocin, mitomycin C, and cis-dichlorodiamine platinum (II) (DDP) cisplatin), anthracyclines (e.g., daunorubicin (formerly daunomycin) and doxorubicin), antibiotics (e.g., dactinomycin (formerly actinomycin), bleomycin, mithramycin, and anthramycin (AMC)), and anti-mitotic agents (e.g., vincristine and vinblastine).

The conjugates of the invention can be used for modifying a given biological response, the drug moiety is not to be construed as limited to classical chemical therapeutic agents. For example, the drug moiety may be a protein or polypeptide possessing a desired biological activity. Such proteins may include, for example, a toxin such as abrin, ricin A, pseudomonas exotoxin, or diphtheria toxin; a protein such as tumor necrosis factor, alpha-interferon, beta-interferon, nerve growth factor, platelet derived growth factor, tissue plasminogen activator; or biological response modifiers such as, for example, lymphokines, interleukin-1 ("IL-1"), interleukin-2 ("IL-2"), interleukin-6 ("IL-6"), granulocyte macrophage colony stimulating factor ("GM-CSF"), granulocyte colony stimulating factor ("G-CSF"), or other growth factors.

Techniques for conjugating such therapeutic moiety to antibodies are well known, see, e.g., Arnon et al., "Monoclonal Antibodies For Immunotargeting Of Drugs In Cancer

The nucleic acid molecules used in the methods of the invention can be inserted into vectors and used as gene therapy vectors. Gene therapy vectors can be delivered to a subject by, for example, intravenous injection, local administration (see U.S. Patent 5,328,470) or by stereotactic injection (see, e.g., Chen et al. (1994) Proc. Natl. Acad. Sci. USA 91:3054-3057). The pharmaceutical preparation of the gene therapy vector can include the gene therapy vector in an acceptable diluent, or can comprise a slow release matrix in which the gene delivery vehicle is imbedded. Alternatively, where the complete gene delivery vector can be produced intact from recombinant cells, e.g., retroviral vectors, the pharmaceutical preparation can include one or more cells which produce the gene delivery system.

D. Pharmacogenomics

In conjunction with the therapeutic methods of the invention, pharmacogenomics (i.e., the study of the relationship between a subject's genotype and that subject's response to a foreign compound or drug) may be considered. Differences in metabolism of therapeutics can lead to severe toxicity or therapeutic failure by altering the relation between dose and blood concentration of the pharmacologically active drug. Thus, a physician or clinician may consider applying knowledge obtained in relevant pharmacogenomics studies in determining whether to administer an agent which modulates 252, 304, 1980, 14717, 9941, 19310 OR 17832 activity, as well as tailoring the dosage and/or therapeutic regimen of treatment with an agent which modulates 252, 304, 1980, 14717, 9941, 19310 OR 17832 activity.
Pharmacogenomics deals with clinically significant hereditary variations in the response to drugs due to altered drug disposition and abnormal action in affected persons. See, for example, Eichelbaum, M. et al. (1996) *Clin. Exp. Pharmacol. Physiol.* 23(10-11): 983-985 and Linder, M.W. et al. (1997) *Clin. Chem.* 43(2):254-266. In general, two types of pharmacogenetic conditions can be differentiated. Genetic conditions transmitted as a single factor altering the way drugs act on the body (altered drug action) or genetic conditions transmitted as single factors altering the way the body acts on drugs (altered drug metabolism). These pharmacogenetic conditions can occur either as rare genetic defects or as naturally-occurring polymorphisms. For example, glucose-6-phosphate aminopeptidase deficiency (G6PD) is a common inherited enzymopathy in which the main clinical complication is haemolysis after ingestion of oxidant drugs (anti-malarials, sulfonamides, analgesics, nitrofurans) and consumption of fava beans.

One pharmacogenomics approach to identifying genes that predict drug response, known as "a genome-wide association", relies primarily on a high-resolution map of the human genome consisting of already known gene-related markers (*e.g.*, a "bi-allelic" gene marker map which consists of 60,000-100,000 polymorphic or variable sites on the human genome, each of which has two variants). Such a high-resolution genetic map can be compared to a map of the genome of each of a statistically significant number of patients taking part in a Phase II/III drug trial to identify markers associated with a particular observed drug response or side effect. Alternatively, such a high resolution map can be generated from a combination of some ten million known single nucleotide polymorphisms (SNPs) in the human genome. As used herein, a "SNP" is a common alteration that occurs in a single nucleotide base in a stretch of DNA. For example, a SNP may occur once per every 1000 bases of DNA. A SNP may be involved in a disease process, however, the vast majority may not be disease-associated. Given a genetic map based on the occurrence of such SNPs, individuals can be grouped into genetic categories depending on a particular pattern of SNPs in their individual genome. In such a manner, treatment regimens can be tailored to groups of genetically similar individuals, taking into account traits that may be common among such genetically similar individuals.

Alternatively, a method termed the "candidate gene approach" can be utilized to identify genes that predict drug response. According to this method, if a gene that encodes a drug target is known (*e.g.*, a 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein used in the methods of the present invention), all common variants of that gene can be fairly
easily identified in the population and it can be determined if having one version of the
gene versus another is associated with a particular drug response.

As an illustrative embodiment, the activity of drug metabolizing enzymes is a major
determinant of both the intensity and duration of drug action. The discovery of genetic
polymorphisms of drug metabolizing enzymes (e.g., N-acetyltransferase 2 (NAT 2) and the
cytochrome P450 enzymes CYP2D6 and CYP2C19) has provided an explanation as to why
some patients do not obtain the expected drug effects or show exaggerated drug response
and serious toxicity after taking the standard and safe dose of a drug. These
polymorphisms are expressed in two phenotypes in the population, the extensive
metabolizer (EM) and poor metabolizer (PM). The prevalence of PM is different among
different populations. For example, the gene coding for CYP2D6 is highly polymorphic
and several mutations have been identified in PM, which all lead to the absence of
functional CYP2D6. Poor metabolizers of CYP2D6 and CYP2C19 quite frequently
experience exaggerated drug response and side effects when they receive standard doses.

If a metabolite is the active therapeutic moiety, PM show no therapeutic response, as
demonstrated for the analgesic effect of codeine mediated by its CYP2D6-formed
metabolite morphine. The other extreme are the so called ultra-rapid metabolizers who do
not respond to standard doses. Recently, the molecular basis of ultra-rapid metabolism has
been identified to be due to CYP2D6 gene amplification.

Alternatively, a method termed the "gene expression profiling" can be utilized to
identify genes that predict drug response. For example, the gene expression of an animal
dosed with a drug (e.g., a 252, 304, 1980, 14717, 9941, 19310 OR 17832 molecule or 252,
304, 1980, 14717, 9941, 19310 OR 17832 modulator used in the methods of the present
invention) can give an indication whether gene pathways related to toxicity have been
turned on.

Information generated from more than one of the above pharmacogenomics
approaches can be used to determine appropriate dosage and treatment regimens for
prophylactic or therapeutic treatment of a subject. This knowledge, when applied to dosing
or drug selection, can avoid adverse reactions or therapeutic failure and, thus, enhance
therapeutic or prophylactic efficiency when treating a subject suffering from a
cardiovascular disease, e.g., atherosclerosis, with an agent which modulates 252, 304,
1980, 14717, 9941, 19310 OR 17832 activity.
V. Recombinant Expression Vectors and Host Cells Used in the Methods of the Invention

The methods of the invention (e.g., the screening assays described herein) include the use of vectors, preferably expression vectors, containing a nucleic acid encoding a 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein (or a portion thereof). As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is a "plasmid", which refers to a circular double stranded DNA loop into which additional DNA segments can be ligated. Another type of vector is a viral vector, wherein additional DNA segments can be ligated into the viral genome. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (e.g., bacterial vectors having a bacterial origin of replication and episomal mammalian vectors). Other vectors (e.g., non-episomal mammalian vectors) are integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated along with the host genome. Moreover, certain vectors are capable of directing the expression of genes to which they are operatively linked. Such vectors are referred to herein as "expression vectors". In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids. In the present specification, "plasmid" and "vector" can be used interchangeably as the plasmid is the most commonly used form of vector. However, the invention is intended to include such other forms of expression vectors, such as viral vectors (e.g., replication defective retroviruses, adenoviruses and adeno-associated viruses), which serve equivalent functions.

The recombinant expression vectors to be used in the methods of the invention comprise a nucleic acid of the invention in a form suitable for expression of the nucleic acid in a host cell, which means that the recombinant expression vectors include one or more regulatory sequences, selected on the basis of the host cells to be used for expression, which is operatively linked to the nucleic acid sequence to be expressed. Within a recombinant expression vector, "operably linked" is intended to mean that the nucleotide sequence of interest is linked to the regulatory sequence(s) in a manner which allows for expression of the nucleotide sequence (e.g., in an in vitro transcription/translation system or in a host cell when the vector is introduced into the host cell). The term "regulatory sequence" is intended to include promoters, enhancers and other expression control elements (e.g., polyadenylation signals). Such regulatory sequences are described, for example, in Goeddel (1990) Methods Enzymol. 185:3-7. Regulatory sequences include those which direct constitutive expression of a nucleotide sequence in many types of host
cells and those which direct expression of the nucleotide sequence only in certain host cells
(e.g., tissue-specific regulatory sequences). It will be appreciated by those skilled in the art
that the design of the expression vector can depend on such factors as the choice of the host
cell to be transformed, the level of expression of protein desired, and the like. The
expression vectors of the invention can be introduced into host cells to thereby produce
proteins or peptides, including fusion proteins or peptides, encoded by nucleic acids as
described herein (e.g., 252, 304, 1980, 14717, 9941, 19310 OR 17832 proteins, mutant
forms of 252, 304, 1980, 14717, 9941, 19310 OR 17832 proteins, fusion proteins, and the
like).

The recombinant expression vectors to be used in the methods of the invention can
be designed for expression of 252, 304, 1980, 14717, 9941, 19310 OR 17832 proteins in
prokaryotic or eukaryotic cells. For example, 252, 304, 1980, 14717, 9941, 19310 OR
17832 proteins can be expressed in bacterial cells such as *E. coli*, insect cells (using
baculovirus expression vectors), yeast cells, or mammalian cells. Suitable host cells are
discussed further in Goeddel (1990) *supra*. Alternatively, the recombinant expression
vector can be transcribed and translated *in vitro*, for example using T7 promoter regulatory
sequences and T7 polymerase.

Expression of proteins in prokaryotes is most often carried out in *E. coli* with
vectors containing constitutive or inducible promoters directing the expression of either
fusion or non-fusion proteins. Fusion vectors add a number of amino acids to a protein
encoded therein, usually to the amino terminus of the recombinant protein. Such fusion
vectors typically serve three purposes: 1) to increase expression of recombinant protein; 2)
to increase the solubility of the recombinant protein; and 3) to aid in the purification of the
recombinant protein by acting as a ligand in affinity purification. Often, in fusion
expression vectors, a proteolytic cleavage site is introduced at the junction of the fusion
moiety and the recombinant protein to enable separation of the recombinant protein from
the fusion moiety subsequent to purification of the fusion protein. Such enzymes, and their
cognate recognition sequences, include Factor Xa, thrombin and enterokinase. Typical
fusion expression vectors include pGEX (Pharmacia Biotech Inc; Smith, D.B. and Johnson,
(Pharmacia, Piscataway, NJ) which fuse glutathione S-transferase (GST), maltose E
binding protein, or protein A, respectively, to the target recombinant protein.

Purified fusion proteins can be utilized in 252, 304, 1980, 14717, 9941, 19310 OR
17832 activity assays, (e.g., direct assays or competitive assays described in detail below),

- 56 -
or to generate antibodies specific for 252, 304, 1980, 14717, 9941, 19310 OR 17832 proteins. In a preferred embodiment, a 252, 304, 1980, 14717, 9941, 19310 OR 17832 fusion protein expressed in a retroviral expression vector of the present invention can be utilized to infect bone marrow cells which are subsequently transplanted into irradiated recipients. The pathology of the subject recipient is then examined after sufficient time has passed (e.g., six weeks).

In another embodiment, a nucleic acid of the invention is expressed in mammalian cells using a mammalian expression vector. Examples of mammalian expression vectors include pCDM8 (Seed, B. (1987) Nature 329:840) and pMT2PC (Kaufman et al. (1987) EMBO J. 6:187-195). When used in mammalian cells, the expression vector's control functions are often provided by viral regulatory elements. For example, commonly used promoters are derived from polyoma, Adenovirus 2, cytomegalovirus and Simian Virus 40. For other suitable expression systems for both prokaryotic and eukaryotic cells see chapters 16 and 17 of Sambrook, J. et al., Molecular Cloning: A Laboratory Manual. 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.

In another embodiment, the recombinant mammalian expression vector is capable of directing expression of the nucleic acid preferentially in a particular cell type (e.g., tissue-specific regulatory elements are used to express the nucleic acid).

The methods of the invention may further use a recombinant expression vector comprising a DNA molecule of the invention cloned into the expression vector in an antisense orientation. That is, the DNA molecule is operatively linked to a regulatory sequence in a manner which allows for expression (by transcription of the DNA molecule) of an RNA molecule which is antisense to 252, 304, 1980, 14717, 9941, 19310 OR 17832 mRNA. Regulatory sequences operatively linked to a nucleic acid cloned in the antisense orientation can be chosen which direct the continuous expression of the antisense RNA molecule in a variety of cell types, for instance viral promoters and/or enhancers, or regulatory sequences can be chosen which direct constitutive, tissue specific, or cell type specific expression of antisense RNA. The antisense expression vector can be in the form of a recombinant plasmid, phagemid, or attenuated virus in which antisense nucleic acids are produced under the control of a high efficiency regulatory region, the activity of which can be determined by the cell type into which the vector is introduced. For a discussion of the regulation of gene expression using antisense genes, see Weintraub, H. et al., Antisense

Another aspect of the invention pertains to the use of host cells into which a 252, 304, 1980, 14717, 9941, 19310 OR 17832 nucleic acid molecule of the invention is introduced, *e.g.*, a 252, 304, 1980, 14717, 9941, 19310 OR 17832 nucleic acid molecule within a recombinant expression vector or a 252, 304, 1980, 14717, 9941, 19310 OR 17832 nucleic acid molecule containing sequences which allow it to homologously recombine into a specific site of the host cell's genome. The terms "host cell" and "recombinant host cell" are used interchangeably herein. It is understood that such terms refer not only to the particular subject cell but to the progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein.

A host cell can be any prokaryotic or eukaryotic cell. For example, a 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein can be expressed in bacterial cells such as *E. coli*, insect cells, yeast or mammalian cells (such as Chinese hamster ovary cells (CHO) or COS cells). Other suitable host cells are known to those skilled in the art.

Vector DNA can be introduced into prokaryotic or eukaryotic cells via conventional transformation or transfection techniques. As used herein, the terms "transformation" and "transfection" are intended to refer to a variety of art-recognized techniques for introducing foreign nucleic acid (*e.g.*, DNA) into a host cell, including calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection, lipofection, or electroporation. Suitable methods for transforming or transfecting host cells can be found in Sambrook *et al.* (Molecular Cloning: A Laboratory Manual. 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989), and other laboratory manuals.

A host cell used in the methods of the invention, such as a prokaryotic or eukaryotic host cell in culture, can be used to produce (*i.e.*, express) a 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein. Accordingly, the invention further provides methods for producing a 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein using the host cells of the invention. In one embodiment, the method comprises culturing the host cell of the invention (into which a recombinant expression vector encoding a 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein has been introduced) in a suitable medium such that a 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein is produced. In another embodiment,
the method further comprises isolating a 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein from the medium or the host cell.

VI. Isolated Nucleic Acid Molecules Used in the Methods of the Invention

The methods of the invention include the use of isolated nucleic acid molecules that encode 252, 304, 1980, 14717, 9941, 19310 OR 17832 proteins or biologically active portions thereof, as well as nucleic acid fragments sufficient for use as hybridization probes to identify 252, 304, 1980, 14717, 9941, 19310 OR 17832 -encoding nucleic acid molecules (e.g., 252, 304, 1980, 14717, 9941, 19310 OR 17832 mRNA) and fragments for use as PCR primers for the amplification or mutation of 252, 304, 1980, 14717, 9941, 19310 OR 17832 nucleic acid molecules. As used herein, the term "nucleic acid molecule" is intended to include DNA molecules (e.g., cDNA or genomic DNA) and RNA molecules (e.g., mRNA) and analogs of the DNA or RNA generated using nucleotide analogs. The nucleic acid molecule can be single-stranded or double-stranded, but preferably is double-stranded DNA.

A nucleic acid molecule used in the methods of the present invention, e.g., a nucleic acid molecule having the nucleotide sequence of SEQ ID NO1, 4, 7, 10, 13, 16 or 19, or a portion thereof, can be isolated using standard molecular biology techniques and the sequence information provided herein. Using all or portion of the nucleic acid sequence of SEQ ID NO:1, 4, 7, 10, 13, 16 or 19, as a hybridization probe, 252, 304, 1980, 14717, 9941, 19310 OR 17832 nucleic acid molecules can be isolated using standard hybridization and cloning techniques (e.g., as described in Sambrook, J., Fritsh, E. F., and Maniatis, T. Molecular Cloning: A Laboratory Manual. 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989).

Moreover, a nucleic acid molecule encompassing all or a portion of SEQ ID NO:1, 4, 7, 10, 13, 16 or 19 can be isolated by the polymerase chain reaction (PCR) using synthetic oligonucleotide primers designed based upon the sequence of SEQ ID NO:1, 4, 7, 10, 13, 16 or 19.

A nucleic acid used in the methods of the invention can be amplified using cDNA, mRNA or, alternatively, genomic DNA as a template and appropriate oligonucleotide primers according to standard PCR amplification techniques. Furthermore, oligonucleotides corresponding to 252, 304, 1980, 14717, 9941, 19310 OR 17832 nucleotide sequences can be prepared by standard synthetic techniques, e.g., using an automated DNA synthesizer.
In a preferred embodiment, the isolated nucleic acid molecules used in the methods of the invention comprise the nucleotide sequence shown in SEQ ID NO:1, 4, 7, 10, 13, 16 or 19, a complement of the nucleotide sequence shown in SEQ ID NO:1, 4, 7, 10, 13, 16 or 19, or a portion of any of these nucleotide sequences. A nucleic acid molecule which is complementary to the nucleotide sequence shown in SEQ ID NO:1, 4, 7, 10, 13, 16 or 19, is one which is sufficiently complementary to the nucleotide sequence shown in SEQ ID NO:1, 4, 7, 10, 13, 16 or 19 such that it can hybridize to the nucleotide sequence shown in SEQ ID NO:1, 4, 7, 10, 13, 16 or 19 thereby forming a stable duplex.

In still another preferred embodiment, an isolated nucleic acid molecule used in the methods of the present invention comprises a nucleotide sequence which is at least about 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more identical to the entire length of the nucleotide sequence shown in SEQ ID NO:1, 4, 7, 10, 13, 16 or 19, or a portion of any of this nucleotide sequence.

Moreover, the nucleic acid molecules used in the methods of the invention can comprise only a portion of the nucleic acid sequence of SEQ ID NO:1, 4, 7, 10, 13, 16 or 19, for example, a fragment which can be used as a probe or primer or a fragment encoding a portion of a 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein, e.g., a biologically active portion of a 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein. The probe/primer typically comprises substantially purified oligonucleotide. The oligonucleotide typically comprises a region of nucleotide sequence that hybridizes under stringent conditions to at least about 12 or 15, preferably about 20 or 25, more preferably about 30, 35, 40, 45, 50, 55, 60, 65, or 75 consecutive nucleotides of a sense sequence of SEQ ID NO:1, 4, 7, 10, 13, 16 or 19, of an anti-sense sequence of SEQ ID NO:1, 4, 7, 10, 13, 16 or 19, or of a naturally occurring allelic variant or mutant of SEQ ID NO:1, 4, 7, 10, 13, 16 or 19. In one embodiment, a nucleic acid molecule used in the methods of the present invention comprises a nucleotide sequence which is greater than 100, 100-200, 200-300, 300-400, 400-500, 500-600, 600-700, 700-800, 800-900, 900-1000, 1000-1100, 1100-1200, 1200-1300, or more nucleotides in length and hybridizes under stringent hybridization conditions to a nucleic acid molecule of SEQ ID NO:1, 4, 7, 10, 13, 16 or 19.

As used herein, the term "hybridizes under stringent conditions" is intended to describe conditions for hybridization and washing under which nucleotide sequences that are significantly identical or homologous to each other remain hybridized to each other. Preferably, the conditions are such that sequences at least about 70%, more preferably at least about 80%, even more preferably at least about 85% or 90% identical to each other.
remain hybridized to each other. Such stringent conditions are known to those skilled in the art and can be found in *Current Protocols in Molecular Biology*, Ausubel *et al.*, eds., John Wiley & Sons, Inc. (1995), sections 2, 4 and 6. Additional stringent conditions can be found in *Molecular Cloning: A Laboratory Manual*, Sambrook *et al.*, Cold Spring Harbor Press, Cold Spring Harbor, NY (1989), chapters 7, 9 and 11. A preferred, non-limiting example of stringent hybridization conditions includes hybridization in 4X sodium chloride/sodium citrate (SSC), at about 65-70°C (or hybridization in 4X SSC plus 50% formamide at about 42-50°C) followed by one or more washes in 1X SSC, at about 65-70°C. A preferred, non-limiting example of highly stringent hybridization conditions includes hybridization in 1X SSC, at about 65-70°C (or hybridization in 1X SSC plus 50% formamide at about 42-50°C) followed by one or more washes in 0.3X SSC, at about 65-70°C. A preferred, non-limiting example of reduced stringency hybridization conditions includes hybridization in 4X SSC, at about 50-60°C (or alternatively hybridization in 6X SSC plus 50% formamide at about 40-45°C) followed by one or more washes in 2X SSC, at about 50-60°C. Ranges intermediate to the above-recited values, e.g., at 65-70°C or at 42-50°C are also intended to be encompassed by the present invention. SSPE (1xSSPE is 0.15M NaCl, 10mM NaH₂PO₄, and 1.25mM EDTA, pH 7.4) can be substituted for SSC (1xSSC is 0.15M NaCl and 15mM sodium citrate) in the hybridization and wash buffers; washes are performed for 15 minutes each after hybridization is complete. The hybridization temperature for hybrids anticipated to be less than 50 base pairs in length should be 5-10°C less than the melting temperature (Tₘ) of the hybrid, where Tₘ is determined according to the following equations. For hybrids less than 18 base pairs in length, Tₘ(°C) = 2(# of A + T bases) + 4(# of G + C bases). For hybrids between 18 and 49 base pairs in length, Tₘ(°C) = 81.5 + 16.6(log₁₀[Na⁺]) + 0.41(%G+C) - (600/N), where N is the number of bases in the hybrid, and [Na⁺] is the concentration of sodium ions in the hybridization buffer ([Na⁺] for 1xSSC = 0.165 M). It will also be recognized by the skilled practitioner that additional reagents may be added to hybridization and/or wash buffers to decrease non-specific hybridization of nucleic acid molecules to membranes, for example, nitrocellulose or nylon membranes, including but not limited to blocking agents (e.g., BSA or salmon or herring sperm carrier DNA), detergents (e.g., SDS), chelating agents (e.g., EDTA), Ficoll, PVP and the like. When using nylon membranes, in particular, an additional preferred, non-limiting example of stringent hybridization conditions is hybridization in 0.25-0.5M NaH₂PO₄, 7% SDS at about 65°C, followed by one or more
washes at 0.02M NaH₂PO₄, 1% SDS at 85°C, see *e.g.*, Church and Gilbert (1984) *Proc. Natl. Acad. Sci. USA* 81:1991-1995, (or alternatively 0.2X SSC, 1% SDS).

In preferred embodiments, the probe further comprises a label group attached thereto, *e.g.*, the label group can be a radioisotope, a fluorescent compound, an enzyme, or an enzyme co-factor. Such probes can be used as a part of a diagnostic test kit for identifying cells or tissue which misexpress a 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein, such as by measuring a level of a 252, 304, 1980, 14717, 9941, 19310 OR 17832 -encoding nucleic acid in a sample of cells from a subject *e.g.*, detecting 252, 304, 1980, 14717, 9941, 19310 OR 17832 mRNA levels or determining whether a genomic 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene has been mutated or deleted.

The methods of the invention further encompass the use of nucleic acid molecules that differ from the nucleotide sequence shown in SEQ ID NO:1, 4, 7, 10, 13, 16 or 19, due to degeneracy of the genetic code and thus encode the same 252, 304, 1980, 14717, 9941, 19310 OR 17832 proteins as those encoded by the nucleotide sequence shown in SEQ ID NO:1, 4, 7, 10, 13, 16 or 19, . In another embodiment, an isolated nucleic acid molecule included in the methods of the invention has a nucleotide sequence encoding a protein having an amino acid sequence shown in SEQ ID NO:3, 6, 9, 12, 15, 18 or 21.

The methods of the invention further include the use of allelic variants of human 252, 304, 1980, 14717, 9941, 19310 OR 17832 , *e.g.*, functional and non-functional allelic variants. Functional allelic variants are naturally occurring amino acid sequence variants of the human 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein that maintain a 252, 304, 1980, 14717, 9941, 19310 OR 17832 activity. Functional allelic variants will typically contain only conservative substitution of one or more amino acids of SEQ ID NO: NO:3, 6, 9, 12, 15, 18 or 21, or substitution, deletion or insertion of non-critical residues in non-critical regions of the protein.

Non-functional allelic variants are naturally occurring amino acid sequence variants of the human 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein that do not have a 252, 304, 1980, 14717, 9941, 19310 OR 17832 activity. Non-functional allelic variants will typically contain a non-conservative substitution, deletion, or insertion or premature truncation of the amino acid sequence of SEQ ID NO: NO:3, 6, 9, 12, 15, 18 or 21, or a substitution, insertion or deletion in critical residues or critical regions of the protein.

The methods of the present invention may further use non-human orthologues of the human 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein. Orthologues of the human 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein are proteins that are isolated
from non-human organisms and possess the same 252, 304, 1980, 14717, 9941, 19310 OR 17832 activity.

The methods of the present invention further include the use of nucleic acid molecules comprising the nucleotide sequence of SEQ ID NO:1, 4, 7, 10, 13, 16 or 19, or a portion thereof, in which a mutation has been introduced. The mutation may lead to amino acid substitutions at "non-essential" amino acid residues or at "essential" amino acid residues. A "non-essential" amino acid residue is a residue that can be altered from the wild-type sequence of 252, 304, 1980, 14717, 9941, 19310 OR 17832 (e.g., the sequence of SEQ ID NO: 3, 6, 9, 12, 15, 18 or 21) without altering the biological activity, whereas an "essential" amino acid residue is required for biological activity. For example, amino acid residues that are conserved among the 252, 304, 1980, 14717, 9941, 19310 OR 17832 proteins of the present invention are not likely to be amenable to alteration.

Mutations can be introduced into SEQ ID NO:1, 4, 7, 10, 13, 16 or 19, by standard techniques, such as site-directed mutagenesis and PCR-mediated mutagenesis. Preferably, conservative amino acid substitutions are made at one or more predicted non-essential amino acid residues. A "conservative amino acid substitution" is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art. These families include amino acids with basic side chains (e.g., lysine, arginine, histidine), acidic side chains (e.g., aspartic acid, glutamic acid), uncharged polar side chains (e.g., asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (e.g., glycine, alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (e.g., threonine, valine, isoleucine) and aromatic side chains (e.g., tyrosine, phenylalanine, tryptophan, histidine). Thus, a predicted nonessential amino acid residue in a 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein is preferably replaced with another amino acid residue from the same side chain family. Alternatively, in another embodiment, mutations can be introduced randomly along all or part of a 252, 304, 1980, 14717, 9941, 19310 OR 17832 coding sequence, such as by saturation mutagenesis, and the resultant mutants can be screened for 252, 304, 1980, 14717, 9941, 19310 OR 17832 biological activity to identify mutants that retain activity. Following mutagenesis of SEQ ID NO:1, 4, 7, 10, 13, 16 or 19, the encoded protein can be expressed recombinantly and the activity of the protein can be determined using the assay described herein.

Another aspect of the invention pertains to the use of isolated nucleic acid molecules which are antisense to the nucleotide sequence of SEQ ID NO:1, 4, 7, 10, 13, 16
or 19. An "antisense" nucleic acid comprises a nucleotide sequence which is complementary to a "sense" nucleic acid encoding a protein, e.g., complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence. Accordingly, an antisense nucleic acid can hydrogen bond to a sense nucleic acid. The antisense nucleic acid can be complementary to an entire 252, 304, 1980, 14717, 9941, 19310 OR 17832 coding strand, or to only a portion thereof. In one embodiment, an antisense nucleic acid molecule is antisense to a "coding region" of the coding strand of a nucleotide sequence encoding a 252, 304, 1980, 14717, 9941, 19310 OR 17832. The term "coding region" refers to the region of the nucleotide sequence comprising codons which are translated into amino acid residues. In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence encoding 252, 304, 1980, 14717, 9941, 19310 OR 17832. The term "noncoding region" refers to 5' and 3' sequences which flank the coding region that are not translated into amino acids (also referred to as 5' and 3' untranslated regions).

Given the coding strand sequences encoding 252, 304, 1980, 14717, 9941, 19310 OR 17832 disclosed herein, antisense nucleic acids of the invention can be designed according to the rules of Watson and Crick base pairing. The antisense nucleic acid molecule can be complementary to the entire coding region of 252, 304, 1980, 14717, 9941, 19310 OR 17832 mRNA, but more preferably is an oligonucleotide which is antisense to only a portion of the coding or noncoding region of 252, 304, 1980, 14717, 9941, 19310 OR 17832 mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of 252, 304, 1980, 14717, 9941, 19310 OR 17832 mRNA. An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An antisense nucleic acid of the invention can be constructed using chemical synthesis and enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (e.g., an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, e.g., phosphorothioate derivatives and acridine substituted nucleotides can be used. Examples of modified nucleotides which can be used to generate the antisense nucleic acid include 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xantine, 4-acetylcytosine, 5-(carboxyhydroxymethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylnaminomethyluracil,
dihydouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-
methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine,
3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-
methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine,
5'-methoxyacarbocymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine,
uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-
methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5- oxyacetic acid
methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-
-carboxypropyl)uracil, (acp3)w, and 2,6-diaminopurine. Alternatively, the antisense
nucleic acid can be produced biologically using an expression vector into which a nucleic
acid has been subcloned in an antisense orientation (i.e., RNA transcribed from the inserted
nucleic acid will be of an antisense orientation to a target nucleic acid of interest).
Antisense nucleic acid molecules used in the methods of the invention are further described
above, in section IV.

In yet another embodiment, the 252, 304, 1980, 14717, 9941, 19310 OR 17832
nucleic acid molecules used in the methods of the present invention can be modified at the
base moiety, sugar moiety or phosphate backbone to improve, e.g., the stability,
hybridization, or solubility of the molecule. For example, the deoxyribose phosphate
backbone of the nucleic acid molecules can be modified to generate peptide nucleic acids
(see Hyrup B. et al. (1996) Bioorganic & Medicinal Chemistry 4 (1): 5-23). As used
herein, the terms "peptide nucleic acids" or "PNAs" refer to nucleic acid mimics, e.g., DNA
mimics, in which the deoxyribose phosphate backbone is replaced by a pseudopeptide
backbone and only the four natural nucleobases are retained. The neutral backbone of
PNAs has been shown to allow for specific hybridization to DNA and RNA under
conditions of low ionic strength. The synthesis of PNA oligomers can be performed using
standard solid phase peptide synthesis protocols as described in Hyrup B. et al. (1996)

PNAs of 252, 304, 1980, 14717, 9941, 19310 OR 17832 nucleic acid molecules can be
used in the therapeutic and diagnostic applications described herein. For example,

PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene
expression by, for example, inducing transcription or translation arrest or inhibiting
replication. PNAs of 252, 304, 1980, 14717, 9941, 19310 OR 17832 nucleic acid
molecules can also be used in the analysis of single base pair mutations in a gene, (e.g., by
PNA-directed PCR clamping); as 'artificial restriction enzymes' when used in combination
with other enzymes, (e.g., S1 nucleases (Hyrup B. et al. (1996) supra)); or as probes or
primers for DNA sequencing or hybridization (Hyrup B. et al. (1996) supra; Perry-O’Keefe
et al. (1996) supra).
In another embodiment, PNAs of 252, 304, 1980, 14717, 9941, 19310 OR 17832
5 can be modified, (e.g., to enhance their stability or cellular uptake), by attaching lipophilic
or other helper groups to PNA, by the formation of PNA-DNA chimeras, or by the use of
liposomes or other techniques of drug delivery known in the art. For example, PNA-DNA
chimeras of 252, 304, 1980, 14717, 9941, 19310 OR 17832 nucleic acid molecules can be
generated which may combine the advantageous properties of PNA and DNA. Such
10 chimeras allow DNA recognition enzymes, (e.g., RNAse H and DNA polymerases), to
interact with the DNA portion while the PNA portion would provide high binding affinity
and specificity. PNA-DNA chimeras can be linked using linkers of appropriate lengths
selected in terms of base stacking, number of bonds between the nucleobases, and
orientation (Hyrup B. et al. (1996) supra). The synthesis of PNA-DNA chimeras can be
15 performed as described in Hyrup B. et al. (1996) supra and Finn P.J. et al. (1996) Nucleic
Acids Res. 24 (17): 3357-63. For example, a DNA chain can be synthesized on a solid
support using standard phosphoramidite coupling chemistry and modified nucleoside
analogs, e.g., 5’-(4-methoxytrityl)amino-5’-deoxy-thymidine phosphoramidite, can be used
as a between the PNA and the 5’ end of DNA (Mag, M. et al. (1989) Nucleic Acid Res. 17:
20 5973-88). PNA monomers are then coupled in a stepwise manner to produce a chimeric
molecule with a 5’ PNA segment and a 3’ DNA segment (Finn P.J. et al. (1996) supra).
Alternatively, chimeric molecules can be synthesized with a 5’DNA segment and a 3’ PNA
In other embodiments, the oligonucleotide used in the methods of the invention
25 may include other appended groups such as peptides (e.g., for targeting host cell receptors
in vivo), or agents facilitating transport across the cell membrane (see, e.g., Letsinger et al.
Sci. USA 84:648-652; PCT Publication No. W088/09810) or the blood-brain barrier (see,
e.g., PCT Publication No. W089/10134). In addition, oligonucleotides can be modified
30 with hybridization-triggered cleavage agents (See, e.g., Krol et al. (1988) Bio-Techniques 6:958-976) or intercalating agents. (See, e.g., Zon (1988) Pharm. Res. 5:539-549). To this
end, the oligonucleotide may be conjugated to another molecule, (e.g., a peptide,
hybridization triggered cross-linking agent, transport agent, or hybridization-triggered
cleavage agent).
VII. Isolated 252, 304, 1980, 14717, 9941, 19310 OR 17832 Proteins and Anti-
252, 304, 1980, 14717, 9941, 19310 OR 17832 Antibodies Used in the Methods of the
Invention

The methods of the invention include the use of isolated 252, 304, 1980, 14717,
9941, 19310 OR 17832 proteins, and biologically active portions thereof, as well as
polypeptide fragments suitable for use as immunogens to raise anti- 252, 304, 1980, 14717,
9941, 19310 OR 17832 antibodies. In one embodiment, native 252, 304, 1980, 14717,
9941, 19310 OR 17832 proteins can be isolated from cells or tissue sources by an
appropriate purification scheme using standard protein purification techniques. In another
embodiment, 252, 304, 1980, 14717, 9941, 19310 OR 17832 proteins are produced by
recombinant DNA techniques. Alternative to recombinant expression, a 252, 304, 1980,
14717, 9941, 19310 OR 17832 protein or polypeptide can be synthesized chemically using
standard peptide synthesis techniques.

As used herein, a "biologically active portion" of a 252, 304, 1980, 14717, 9941,
19310 OR 17832 protein includes a fragment of a 252, 304, 1980, 14717, 9941, 19310 OR
17832 protein having a 252, 304, 1980, 14717, 9941, 19310 OR 17832 activity.
Biologically active portions of a 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein
include peptides comprising amino acid sequences sufficiently identical to or derived from
the amino acid sequence of the 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein,
*e.g.*, the amino acid sequence shown in SEQ ID NO:3, 6, 9, which include fewer amino
acids than the full length 252, 304, 1980, 14717, 9941, 19310 OR 17832 proteins, and
exhibit at least one activity of a 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein.
Typically, biologically active portions comprise a domain or motif with at least one activity
of the 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein (*e.g.*, the N-terminal region
of the 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein that is believed to be
involved in the regulation of apoptotic activity). A biologically active portion of a 252,
304, 1980, 14717, 9941, 19310 OR 17832 protein can be a polypeptide which is, for
example, 25, 50, 75, 100, 125, 150, 175, 200, 250, 300 or more amino acids in length.

Biologically active portions of a 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein
can be used as targets for developing agents which modulate a 252, 304, 1980, 14717,
9941, 19310 OR 17832 activity.

In a preferred embodiment, the 252, 304, 1980, 14717, 9941, 19310 OR 17832
protein used in the methods of the invention has an amino acid sequence shown in SEQ ID
NO: 3, 6, 9, 12, 15, 18 or 21. In other embodiments, the 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein is substantially identical to SEQ ID NO: 3, 6, 9, 12, 15, 18 or 21, and retains the functional activity of the protein of SEQ ID NO: 3, 6, 9, 12, 15, 18 or 21, yet differs in amino acid sequence due to natural allelic variation or mutagenesis, as described in detail in subsection V above. Accordingly, in another embodiment, the 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein used in the methods of the invention is a protein which comprises an amino acid sequence at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more identical to SEQ ID NO: 3, 6, 9, 12, 15, 18 or 21.

To determine the percent identity of two amino acid sequences or of two nucleic acid sequences, the sequences are aligned for optimal comparison purposes (e.g., gaps can be introduced in one or both of a first and a second amino acid or nucleic acid sequence for optimal alignment and non-identical sequences can be disregarded for comparison purposes). In a preferred embodiment, the length of a reference sequence aligned for comparison purposes is at least 30%, preferably at least 40%, more preferably at least 50%, even more preferably at least 60%, and even more preferably at least 70%, 80%, or 90% of the length of the reference sequence (e.g., when aligning a second sequence to the 252, 304, 1980, 14717, 9941, 19310 OR 17832 amino acid sequence of SEQ ID NO: 3, 6, 9, 12, 15, 18 or 21 having 500 amino acid residues, at least 75, preferably at least 150, more preferably at least 225, even more preferably at least 300, and even more preferably at least 400 or more amino acid residues are aligned). The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are identical at that position (as used herein amino acid or nucleic acid "identity" is equivalent to amino acid or nucleic acid "homology"). The percent identity between the two sequences is a function of the number of identical positions shared by the sequences, taking into account the number of gaps, and the length of each gap, which need to be introduced for optimal alignment of the two sequences.

The comparison of sequences and determination of percent identity between two sequences can be accomplished using a mathematical algorithm. In a preferred embodiment, the percent identity between two amino acid sequences is determined using the Needleman and Wunsch (J. Mol. Biol. 48:444-453 (1970)) algorithm which has been incorporated into the GAP program in the GCG software package (available at
http://www.gcg.com), using either a Blosum 62 matrix or a PAM250 matrix, and a gap weight of 16, 14, 12, 10, 8, 6, or 4 and a length weight of 1, 2, 3, 4, 5, or 6. In yet another preferred embodiment, the percent identity between two nucleotide sequences is determined using the GAP program in the GCG software package (available at http://www.gcg.com), using a NWSgapdnaCMP matrix and a gap weight of 40, 50, 60, 70, or 80 and a length weight of 1, 2, 3, 4, 5, or 6. In another embodiment, the percent identity between two amino acid or nucleotide sequences is determined using the algorithm of E. Meyers and W. Miller (Comput. Appl. Biosci. 4:11-17 (1988)) which has been incorporated into the ALIGN program (version 2.0 or 2.0U), using a PAM120 weight residue table, a gap length penalty of 12 and a gap penalty of 4.

The methods of the invention may also use 252, 304, 1980, 14717, 9941, 19310 or 17832 chimeric or fusion proteins. As used herein, a 252, 304, 1980, 14717, 9941, 19310 or 17832 "chimeric protein" or "fusion protein" comprises a 252, 304, 1980, 14717, 9941, 19310 OR 17832 polypeptide operatively linked to a non- 252, 304, 1980, 14717, 9941, 19310 OR 17832 polypeptide. An " 252, 304, 1980, 14717, 9941, 19310 OR 17832 polypeptide" refers to a polypeptide having an amino acid sequence corresponding to a 252, 304, 1980, 14717, 9941, 19310 OR 17832 molecule, whereas a "non- 252, 304, 1980, 14717, 9941, 19310 OR 17832 polypeptide" refers to a polypeptide having an amino acid sequence corresponding to a protein which is not substantially homologous to the 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein, e.g., a protein which is different from the 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein and which is derived from the same or a different organism. Within a 252, 304, 1980, 14717, 9941, 19310 OR 17832 fusion protein the 252, 304, 1980, 14717, 9941, 19310 OR 17832 polypeptide can correspond to all or a portion of a 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein. In a preferred embodiment, a 252, 304, 1980, 14717, 9941, 19310 OR 17832 fusion protein comprises at least one biologically active portion of a 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein. In another preferred embodiment, a 252, 304, 1980, 14717, 9941, 19310 OR 17832 fusion protein comprises at least two biologically active portions of a 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein. Within the fusion protein, the term "operatively linked" is intended to indicate that the 252, 304, 1980, 14717, 9941, 19310 OR 17832 polypeptide and the non- 252, 304, 1980, 14717, 9941, 19310 OR 17832 polypeptide are fused in-frame to each other. The non- 252, 304, 1980, 14717, 9941, 19310 OR 17832 polypeptide can be fused to the N-terminus or C-terminus of the 252, 304, 1980, 14717, 9941, 19310 OR 17832 polypeptide.
For example, in one embodiment, the fusion protein is a GST- 252, 304, 1980, 14717, 9941, 19310 OR 17832 fusion protein in which the 252, 304, 1980, 14717, 9941, 19310 OR 17832 sequences are fused to the C-terminus of the GST sequences. Such fusion proteins can facilitate the purification of recombinant 252, 304, 1980, 14717, 9941, 19310 OR 17832.

In another embodiment, this fusion protein is a 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein containing a heterologous signal sequence at its N-terminus. In certain host cells (e.g., mammalian host cells), expression and/or secretion of 252, 304, 1980, 14717, 9941, 19310 OR 17832 can be increased through use of a heterologous signal sequence.

The 252, 304, 1980, 14717, 9941, 19310 OR 17832 fusion proteins used in the methods of the invention can be incorporated into pharmaceutical compositions and administered to a subject in vivo. The 252, 304, 1980, 14717, 9941, 19310 OR 17832 fusion proteins can be used to affect the bioavailability of a 252, 304, 1980, 14717, 9941, 19310 OR 17832 substrate. Use of 252, 304, 1980, 14717, 9941, 19310 OR 17832 fusion proteins may be useful therapeutically for the treatment of disorders caused by, for example, (i) aberrant modification or mutation of a gene encoding a 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein; (ii) mis-regulation of the 252, 304, 1980, 14717, 9941, 19310 OR 17832 gene; and (iii) aberrant post-translational modification of a 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein.

Moreover, the 252, 304, 1980, 14717, 9941, 19310 OR 17832-fusion proteins used in the methods of the invention can be used as immunogens to produce anti- 252, 304, 1980, 14717, 9941, 19310 OR 17832 antibodies in a subject, to purify 252, 304, 1980, 14717, 9941, 19310 OR 17832 ligands and in screening assays to identify molecules which inhibit the interaction of 252, 304, 1980, 14717, 9941, 19310 OR 17832 with a 252, 304, 1980, 14717, 9941, 19310 OR 17832 substrate.

Preferably, a 252, 304, 1980, 14717, 9941, 19310 OR 17832 chimeric or fusion protein used in the methods of the invention is produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional techniques, for example by employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including
automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers which give rise to complementary overhangs between two consecutive gene fragments which can subsequently be annealed and reamplified to generate a chimeric gene sequence (see, for example, *Current Protocols in Molecular Biology*, eds. Ausubel *et al.* John Wiley & Sons: 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (*e.g.*, a GST polypeptide). A 252, 304, 1980, 14717, 9941, 19310 OR 17832 -encoding nucleic acid can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein.

The present invention also pertains to the use of variants of the 252, 304, 1980, 14717, 9941, 19310 OR 17832 proteins which function as either 252, 304, 1980, 14717, 9941, 19310 OR 17832 agonists (mimetics) or as 252, 304, 1980, 14717, 9941, 19310 OR 17832 antagonists. Variants of the 252, 304, 1980, 14717, 9941, 19310 OR 17832 proteins can be generated by mutagenesis, *e.g.*, discrete point mutation or truncation of a 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein. An agonist of the 252, 304, 1980, 14717, 9941, 19310 OR 17832 proteins can retain substantially the same, or a subset, of the biological activities of the naturally occurring form of a 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein. An antagonist of a 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein can inhibit one or more of the activities of the naturally occurring form of the 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein by, for example, competitively modulating a 252, 304, 1980, 14717, 9941, 19310 OR 17832 -mediated activity of a 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein. Thus, specific biological effects can be elicited by treatment with a variant of limited function. In one embodiment, treatment of a subject with a variant having a subset of the biological activities of the naturally occurring form of the protein has fewer side effects in a subject relative to treatment with the naturally occurring form of the 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein.

In one embodiment, variants of a 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein which function as either 252, 304, 1980, 14717, 9941, 19310 OR 17832 agonists (mimetics) or as 252, 304, 1980, 14717, 9941, 19310 OR 17832 antagonists can be identified by screening combinatorial libraries of mutants, *e.g.*, truncation mutants, of a 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein for 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein agonist or antagonist activity. In one embodiment, a variegated library of 252, 304, 1980, 14717, 9941, 19310 OR 17832 variants is generated by combinatorial mutagenesis at the nucleic acid level and is encoded by a variegated gene
library. A variegated library of 252, 304, 1980, 14717, 9941, 19310 OR 17832 variants can be produced by, for example, enzymatically ligating a mixture of synthetic oligonucleotides into gene sequences such that a degenerate set of potential 252, 304, 1980, 14717, 9941, 19310 OR 17832 sequences is expressible as individual polypeptides, or alternatively, as a set of larger fusion proteins (e.g., for phage display) containing the set of 252, 304, 1980, 14717, 9941, 19310 OR 17832 sequences therein. There are a variety of methods which can be used to produce libraries of potential 252, 304, 1980, 14717, 9941, 19310 OR 17832 variants from a degenerate oligonucleotide sequence. Chemical synthesis of a degenerate gene sequence can be performed in an automatic DNA synthesizer, and the synthetic gene then ligated into an appropriate expression vector. Use of a degenerate set of genes allows for the provision, in one mixture, of all of the sequences encoding the desired set of potential 252, 304, 1980, 14717, 9941, 19310 OR 17832 sequences. Methods for synthesizing degenerate oligonucleotides are known in the art (see, e.g., Narang, S.A. (1983) Tetrahedron 39:3; Itakura et al. (1984) Annu. Rev. Biochem. 53:323; Itakura et al. (1984) Science 198:1056; Ike et al. (1983) Nucleic Acid Res. 11:477).

In addition, libraries of fragments of a 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein coding sequence can be used to generate a variegated population of 252, 304, 1980, 14717, 9941, 19310 OR 17832 fragments for screening and subsequent selection of variants of a 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein. In one embodiment, a library of coding sequence fragments can be generated by treating a double stranded PCR fragment of a 252, 304, 1980, 14717, 9941, 19310 OR 17832 coding sequence with a nuclease under conditions wherein nicking occurs only about once per molecule, denaturing the double stranded DNA, renaturing the DNA to form double stranded DNA which can include sense/antisense pairs from different nicked products, removing single stranded portions from reformed duplexes by treatment with S1 nuclease, and ligating the resulting fragment library into an expression vector. By this method, an expression library can be derived which encodes N-terminal, C-terminal and internal fragments of various sizes of the 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein.

Several techniques are known in the art for screening gene products of combinatorial libraries made by point mutations or truncation, and for screening cDNA libraries for gene products having a selected property. Such techniques are adaptable for rapid screening of the gene libraries generated by the combinatorial mutagenesis of 252, 304, 1980, 14717, 9941, 19310 OR 17832 proteins. The most widely used techniques, which are amenable to high throughput analysis, for screening large gene libraries
typically include cloning the gene library into replicable expression vectors, transforming appropriate cells with the resulting library of vectors, and expressing the combinatorial genes under conditions in which detection of a desired activity facilitates isolation of the vector encoding the gene whose product was detected. Recursive ensemble mutagenesis (REM), a new technique which enhances the frequency of functional mutants in the libraries, can be used in combination with the screening assays to identify 252, 304, 1980, 14717, 9941, 19310 OR 17832 variants (Arkin and Yourvan (1992) Proc. Natl. Acad. Sci. USA 89:7811-7815; Delgrave et al. (1993) Protein Engineering 6(3):327-331).

The methods of the present invention further include the use of anti-252, 304, 1980, 14717, 9941, 19310 OR 17832 antibodies. An isolated 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein, or a portion or fragment thereof, can be used as an immunogen to generate antibodies that bind 252, 304, 1980, 14717, 9941, 19310 OR 17832 using standard techniques for polyclonal and monoclonal antibody preparation. A full-length 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein can be used or, alternatively, antigenic peptide fragments of 252, 304, 1980, 14717, 9941, 19310 OR 17832 can be used as immunogens. The antigenic peptide of 252, 304, 1980, 14717, 9941, 19310 OR 17832 comprises at least 8 amino acid residues of the amino acid sequence shown in SEQ ID NO:3, 6, 9 and encompasses an epitope of 252, 304, 1980, 14717, 9941, 19310 OR 17832 such that an antibody raised against the peptide forms a specific immune complex with the 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein. Preferably, the antigenic peptide comprises at least 10 amino acid residues, more preferably at least 15 amino acid residues, even more preferably at least 20 amino acid residues, and most preferably at least 30 amino acid residues.

Preferred epitopes encompassed by the antigenic peptide are regions of 252, 304, 1980, 14717, 9941, 19310 OR 17832 that are located on the surface of the protein, e.g., hydrophilic regions, as well as regions with high antigenicity.

A 252, 304, 1980, 14717, 9941, 19310 OR 17832 immunogen is typically used to prepare antibodies by immunizing a suitable subject, (e.g., rabbit, goat, mouse, or other mammal) with the immunogen. An appropriate immunogenic preparation can contain, for example, recombinantly expressed 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein or a chemically synthesized 252, 304, 1980, 14717, 9941, 19310 OR 17832 polypeptide. The preparation can further include an adjuvant, such as Freund's complete or incomplete adjuvant, or similar immunostimulatory agent. Immunization of a suitable subject with an

The term "antibody" as used herein refers to immunoglobulin molecules and immunologically active portions of immunoglobulin molecules, i.e., molecules that contain an antigen binding site which specifically binds (immunoreacts with) an antigen, such as a 252, 304, 1980, 14717, 9941, 19310 OR 17832. Examples of immunologically active portions of immunoglobulin molecules include F(ab) and F(ab')2 fragments which can be generated by treating the antibody with an enzyme such as pepsin. The invention provides polyclonal and monoclonal antibodies that bind 252, 304, 1980, 14717, 9941, 19310 OR 17832 molecules. The term "monoclonal antibody" or "monoclonal antibody composition", as used herein, refers to a population of antibody molecules that contain only one species of an antigen binding site capable of immunoreacting with a particular epitope of 252, 304, 1980, 14717, 9941, 19310 OR 17832. A monoclonal antibody composition thus typically displays a single binding affinity for a particular 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein with which it immunoreacts.

Polyclonal anti-252, 304, 1980, 14717, 9941, 19310 OR 17832 antibodies can be prepared as described above by immunizing a suitable subject with a 252, 304, 1980, 14717, 9941, 19310 OR 17832 immunogen. The anti-252, 304, 1980, 14717, 9941, 19310 OR 17832 antibody titer in the immunized subject can be monitored over time by standard techniques, such as with an enzyme linked immunosorbent assay (ELISA) using immobilized 252, 304, 1980, 14717, 9941, 19310 OR 17832. If desired, the antibody molecules directed against 252, 304, 1980, 14717, 9941, 19310 OR 17832 can be isolated from the mammal (e.g., from the blood) and further purified by well known techniques, such as protein A chromatography to obtain the IgG fraction. At an appropriate time after immunization, e.g., when the anti-252, 304, 1980, 14717, 9941, 19310 OR 17832 antibody titers are highest, antibody-producing cells can be obtained from the subject and used to prepare monoclonal antibodies by standard techniques, such as the hybridoma technique originally described by Kohler and Milstein (1975) Nature 256:495-497 (see also, Brown et al. (1981) J. Immunol. 127:539-46; Brown et al. (1980) J. Biol. Chem. 255:4980-83; Yeh et al. (1976) Proc. Natl. Acad. Sci. USA 76:2927-31; and Yeh et al. (1982) Int. J. Cancer 29:269-75), the more recent human B cell hybridoma technique (Kozbor et al. (1983) Immunol Today 4:72), the EBV-hybridoma technique (Cole et al. (1985) Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, Inc., pp. 77-96) or trioma techniques. The technology for producing monoclonal antibody hybridomas is well known.

Any of the many well known protocols used for fusing lymphocytes and immortalized cell lines can be applied for the purpose of generating an anti- 252, 304, 1980, 14717, 9941, 19310 OR 17832 monoclonal antibody (see, e.g., G. Galfre *et al.* (1977) *Nature* 266:5505; Gefter *et al.* (1977) *supra*; Lerner (1981) *supra*; and Kenneth (1980) *supra*). Moreover, the ordinarily skilled worker will appreciate that there are many variations of such methods which also would be useful. Typically, the immortal cell line (e.g., a myeloma cell line) is derived from the same mammalian species as the lymphocytes. For example, murine hybridomas can be made by fusing lymphocytes from a mouse immunized with an immunogenic preparation of the present invention with an immortalized mouse cell line. Preferred immortal cell lines are mouse myeloma cell lines that are sensitive to culture medium containing hypoxanthine, aminopterin and thymidine ("HAT medium"). Any of a number of myeloma cell lines can be used as a fusion partner according to standard techniques, e.g., the P3-NS1/1-Ag4-1, P3-x63-Ag8.653 or Sp2/O-Ag14 myeloma lines. These myeloma lines are available from ATCC. Typically, HAT-sensitive mouse myeloma cells are fused to mouse splenocytes using polyethylene glycol ("PEG"). Hybridoma cells resulting from the fusion are then selected using HAT medium, which kills unfused and unproductively fused myeloma cells (unfused splenocytes die after several days because they are not transformed). Hybridoma cells producing a monoclonal antibody of the invention are detected by screening the hybridoma culture supernatants for antibodies that bind 252, 304, 1980, 14717, 9941, 19310 OR 17832, e.g., using a standard ELISA assay.

Alternative to preparing monoclonal antibody-secreting hybridomas, a monoclonal anti- 252, 304, 1980, 14717, 9941, 19310 OR 17832 antibody can be identified and isolated by screening a recombinant combinatorial immunoglobulin library (e.g., an antibody phage display library) with 252, 304, 1980, 14717, 9941, 19310 OR 17832 to thereby isolate immunoglobulin library members that bind 252, 304, 1980, 14717, 9941, 19310 OR 17832


An anti-252, 304, 1980, 14717, 9941, 19310 OR 17832 antibody can be used to detect 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein (e.g., in a cellular lysate or cell supernatant) in order to evaluate the abundance and pattern of expression of the 252, 304, 1980, 14717, 9941, 19310 OR 17832 protein. Anti-252, 304, 1980, 14717, 9941, 19310 OR 17832 antibodies can be used diagnostically to monitor protein levels in tissue as part of a clinical testing procedure, e.g., to, for example, determine the efficacy of a given treatment regimen. Detection can be facilitated by coupling (i.e., physically linking) the antibody to a detectable substance. Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, and radioactive materials. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, β-galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a luminescent material includes luminol; examples of bioluminescent materials include luciferase, luciferin, and aequorin, and examples of suitable radioactive material include $^{125}$I, $^{131}$I, $^{35}$S or $^3$H.

This invention is further illustrated by the following examples which should not be construed as limiting. The contents of all references, patents and published patent applications cited throughout this application, as well as the Figure and the Sequence Listing is incorporated herein by reference.

EXAMPLES

EXAMPLE 1: TISSUE DISTRIBUTION OF USING TAQMAN™ ANALYSIS

This example describes the TaqMan™ procedure. The Taqman™ procedure is a quantitative, reverse transcription PCR-based approach for detecting mRNA. The RT-PCR reaction exploits the 5' nuclease activity of AmpliTaq Gold™ DNA Polymerase to cleave a TaqMan™ probe during PCR. Briefly, cDNA was generated from the samples of interest, e.g., heart, kidney, liver, skeletal muscle, and various vessels, and used as the starting material for PCR amplification. In addition to the 5' and 3' gene-specific primers, a gene-
specific oligonucleotide probe (complementary to the region being amplified) was included in the reaction (i.e., the Taqman™ probe). The TaqMan™ probe includes the oligonucleotide with a fluorescent reporter dye covalently linked to the 5’ end of the probe (such as FAM (6-carboxyfluorescein), TET (6-carboxy-4,7,2’7’-tetrachlorofluorescein), JOE (6-carboxy-4,5-dichloro-2,7-dimethoxyfluorescein), or VIC) and a quencher dye (TAMRA (6-carboxy-N,N,N’,N’-tetramethylrhodamine) at the 3’ end of the probe.

During the PCR reaction, cleavage of the probe separates the reporter dye and the quencher dye, resulting in increased fluorescence of the reporter. Accumulation of PCR products is detected directly by monitoring the increase in fluorescence of the reporter dye. When the probe is intact, the proximity of the reporter dye to the quencher dye results in suppression of the reporter fluorescence. During PCR, if the target of interest is present, the probe specifically anneals between the forward and reverse primer sites. The 5’-3’ nucleolytic activity of the AmpliTaq™ Gold DNA Polymerase cleaves the probe between the reporter and the quencher only if the probe hybridizes to the target. The probe fragments are then displaced from the target, and polymerization of the strand continues. The 3’ end of the probe is blocked to prevent extension of the probe during PCR. This process occurs in every cycle and does not interfere with the exponential accumulation of product. RNA was prepared using the trizol method and treated with DNase to remove contaminating genomic DNA. cDNA was synthesized using standard techniques. Mock cDNA synthesis in the absence of reverse transcriptase resulted in samples with no detectable PCR amplification of the control gene confirms efficient removal of genomic DNA contamination.

Equivalents

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following claims.
What is claimed:

1. A method for identifying a compound capable of treating a hematologic disorder, comprising assaying the ability of the compound to modulate 252, 304, 1980, 14717, 9941, 19310 or 17832 nucleic acid expression or 252, 304, 1980, 14717, 9941, 19310 or 17832 polypeptide activity, thereby identifying a compound capable of treating a hematologic disorder.

2. A method for identifying a compound capable of modulating hematopoiesis comprising:
   a) contacting a cell which expresses 252, 304, 1980, 14717, 9941, 19310 or 17832 with a test compound; and
   b) assaying the ability of the test compound to modulate the expression of a 252, 304, 1980, 14717, 9941, 19310 or 17832 nucleic acid or the activity of a 252, 304, 1980, 14717, 9941, 19310 or 17832 polypeptide, thereby identifying a compound capable of modulating hematopoiesis.

3. A method for modulating hematopoiesis in a cell comprising contacting a cell with a 252, 304, 1980, 14717, 9941, 19310 or 17832 modulator, thereby modulating hematopoiesis in the cell.

4. The method of claim 2, wherein the cell is a hematopoietic cell.

5. The method of claim 3, wherein the 252, 304, 1980, 14717, 9941, 19310 or 17832 modulator is a small organic molecule, peptide, antibody or antisense nucleic acid molecule.

6. The method of claim 3, wherein the 252, 304, 1980, 14717, 9941, 19310 or 17832 modulator is capable of modulating 252, 304, 1980, 14717, 9941, 19310 or 17832 polypeptide activity.

7. The method of claim 6, wherein the 252, 304, 1980, 14717, 9941, 19310 or 17832 modulator is a small organic molecule, peptide, antibody or antisense nucleic acid molecule.
8. The method of claim 6, wherein the 252, 304, 1980, 14717, 9941, 19310 or 17832 modulator is capable of modulating 252, 304, 1980, 14717, 9941, 19310 or 17832 nucleic acid expression.

9. A method for treating a subject having a hematologic disorder characterized by aberrant 252, 304, 1980, 14717, 9941, 19310 or 17832 polypeptide activity or aberrant 252, 304, 1980, 14717, 9941, 19310 or 17832 nucleic acid expression comprising administering to the subject a 252, 304, 1980, 14717, 9941, 19310 or 17832 modulator, thereby treating said subject having a hematologic disorder.

10. The method of claim 9, wherein said hematological disorder is selected from the group consisting of disorders resulting from bone marrow irradiation or chemotherapy treatments for cancer, Pernicious Anemia, Hemorrhagic Anemia, Hemolytic Anemia, Aplastic Anemia, Sickle Cell Anemia, Sideroblastic Anemia, Anemia associated with chronic infections such as Malaria, Trypanosomiasis, HIV, Hepatitis virus or other viruses, Myelophthisic Anemias caused by marrow deficiencies, renal failure resulting from Anemia, Anemia, Polycythemia, Infectious Mononucleosis (IM), Acute Non-Lymphocytic Leukemia (ANLL), Acute Myeloid Leukemia (AML), Acute Promyelocytic Leukemia (APL), Acute Myelomonocytic Leukemia (AMMoL), Polycythemia Vera, Lymphoma, Acute Lymphocytic Leukemia (ALL), Chronic Lymphocytic Leukemia, Wilm’s Tumor, Ewing’s Sarcoma, Retinoblastoma, Hemophilia, disorders associated with an increased risk of Thrombosis, Herpes, Thalessemia, antibody-mediated disorders such as transfusion reactions and Erythroblastosis, mechanical trauma to red blood cells such as microangiopathic hemolytic anemias, Thrombotic Thrombocytopenic Purpura and disseminated intravascular coagulation, infections by parasites such as Plasmodium, chemical injuries from, e.g., lead poisoning, and Hypersplenism.

11. The method of claim 9, wherein said 252, 304, 1980, 14717, 9941, 19310 or 17832 modulator is administered in a pharmaceutically acceptable formulation.

12. The method of claim 9, wherein the 252, 304, 1980, 14717, 9941, 19310 or 17832 modulator is a small organic molecule, peptide, antibody or antisense nucleic acid molecule.
13. The method of claim 9, wherein the 252, 304, 1980, 14717, 9941, 19310 or 17832 modulator is capable of modulating 252, 304, 1980, 14717, 9941, 19310 or 17832 polypeptide activity.
SEQUENCE LISTING

<110> Carroll, Joseph M.
        Healy, Aileen

<120> METHODS AND COMPOSITIONS FOR TREATING
        HEMATOLOGICAL DISORDERS USING 252, 304, 1980, 14717, 9941,
        19310 AND 17832

<130> MPI01-292P1RWOM

<150> 60/341,606
<151> 2001-12-17

<160> 21

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 1554
<212> DNA
<213> Homo sapiens

<400> 1

cttgggggct cgtgaggctg ccacccgctgc tgcgctacata gcccaacgctc tgcactccac 60
ggcctccgac ccggccagcgc tgcgctacata gcccaacgctg cggcctccac 120
ggcctccgac ccggccagcgc tgcgctacata gcccaacgctg cggcctccac 180
ggcctccgac ccggccagcgc tgcgctacata gcccaacgctg cggcctccac 240
ggcctccgac ccggccagcgc tgcgctacata gcccaacgctg cggcctccac 300
ggcctccgac ccggccagcgc tgcgctacata gcccaacgctg cggcctccac 360
ggcctccgac ccggccagcgc tgcgctacata gcccaacgctg cggcctccac 420
ggcctccgac ccggccagcgc tgcgctacata gcccaacgctg cggcctccac 480
ggcctccgac ccggccagcgc tgcgctacata gcccaacgctg cggcctccac 540
ggcctccgac ccggccagcgc tgcgctacata gcccaacgctg cggcctccac 600
ggcctccgac ccggccagcgc tgcgctacata gcccaacgctg cggcctccac 660
ggcctccgac ccggccagcgc tgcgctacata gcccaacgctg cggcctccac 720
ggcctccgac ccggccagcgc tgcgctacata gcccaacgctg cggcctccac 780
ggcctccgac ccggccagcgc tgcgctacata gcccaacgctg cggcctccac 840
ggcctccgac ccggccagcgc tgcgctacata gcccaacgctg cggcctccac 900
ggcctccgac ccggccagcgc tgcgctacata gcccaacgctg cggcctccac 960
ggcctccgac ccggccagcgc tgcgctacata gcccaacgctg cggcctccac 1020
ggcctccgac ccggccagcgc tgcgctacata gcccaacgctg cggcctccac 1080
ggcctccgac ccggccagcgc tgcgctacata gcccaacgctg cggcctccac 1140
ggcctccgac ccggccagcgc tgcgctacata gcccaacgctg cggcctccac 1200
ggcctccgac ccggccagcgc tgcgctacata gcccaacgctg cggcctccac 1260
ggcctccgac ccggccagcgc tgcgctacata gcccaacgctg cggcctccac 1320
ggcctccgac ccggccagcgc tgcgctacata gcccaacgctg cggcctccac 1380
ggcctccgac ccggccagcgc tgcgctacata gcccaacgctg cggcctccac 1440
ggcctccgac ccggccagcgc tgcgctacata gcccaacgctg cggcctccac 1500
ggcctccgac ccggccagcgc tgcgctacata gcccaacgctg cggcctccac 1554
<210> 2
<211> 1554
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (85)...(1551)
Methionine Serine Threonine Proline Glycine Valine Asparagine Serine

Valine Methionine Phenylalanine Isoleucine Proline Alanine

Cysteine Leucine Serine Arginine Leucine Valine Leucine Valine

Valine Cysteine Glycine Leucine Alanine Valine Asparagine Leucine Glycine Thrreonine Leucine Leucine Valine

Proline Valine Threonine Isoleucine Threonine Methionine Glycine Glutamine Tryptophan Proline Glycine Glutamine

Proline Leucine Valine Isoleucine Threonine Proline Glycine Leucine Leucine Valine Proline Cysteine Isoleucine

Glycine Cysteine Asparagine Cysteine Arginine Leucine Gag Cysteine Cysteine Leucine Cysteine Arginine Cysteine

Asparagine Leucine Aspartic Acid Threonine Serine His Thrreonine Valine Asparagine Leucine Alanine Glycine

Cysteine Cysteine Threonine Cysteamine Cysteine Cysteine Leucine Cysteine Cysteine Arginine Cysteine

Proline Aspartic Acid Methionine Glycine Serine Serine Leucine Arginine Leucine Glycine Typtophan Proline Aspartic Acid

Threonine Cysteine Proline Isoleucine Phenylalanine Valine Asparaginealanine Threonine Serine

Valine Methionine Typtophan Alanine Valine Leucine Arginine Hisalanine Threonine Serine

Valine Arginine Threonine Serine Leucine Glutamine Glutamine Histidine Alanine Leucine Ala

Arginine Threonine Serine Leucine Glutamine Glutamine Histidine Alanine Leucine Ala
gcc tcg gtt gcc tcc cgg gcc cac ccc gct gcc tcc cca gcc ttg ccg
Ala Ser Val Ala Ser Arg Gly His Pro Ala Ala Ser Pro Ala Leu Pro
235 240 245

cgc ctc agc gac ttt cgg cgc cgc cgg agc ttc cgc cgc atc gcg ggc
Arg Leu Ser Asp Phe Arg Arg Arg Ser Phe Arg Arg Ile Ala Gly
255 260 265

gcc gag atc cag atg gtc atc tta ctc att gcc acc tcc ctg gtg gtg
Ala Glu Ile Gln Met Val Ile Leu Leu Ile Ala Thr Ser Leu Val Val
270 275 280

ctc atc tgc tcc atc cag ctc gtc gtg cga gta ttc gtc aac cag tta
Leu Ile Cys Ser Ile Pro Leu Val Val Arg Val Ser Val Phe Val Asn Gin Leu
285 290 295

tat cag cca agt ttt gag cga gaa gtc agt aag aat cca gat ttg cag
Tyr Gln Pro Ser Leu Glu Arg Glu Arg Val Ser Lys Asn Pro Asp Leu Gin
300 305 310

gcc atc cga att gct tct gtt aac ccc att cta gac ccc tgg ata tat
Ala Ile Arg Ile Ala Ser Pro Ile Asp Pro Ile Thr Tyr
315 320 325

atc tct ctg aga aag aca ctg ctc agt aag gca ata gag aag atc aag
Ile Leu Leu Arg Lys Thr Val Leu Ser Lys Ala Ile Gly Lys Ile Lys
330 340 345

tgc tct ttc tgc ctc att ggc ggg tcc cgc cgg gag cgc cgg cag
ttg Cys Leu Phe Cys Arg Ile Gly Gly Ser Arg Arg Gly Arg Ser Gly Gin
350 355 360

cac tgc tca gac agt cca agg aca tct tct gcc atg tca ggc cac tct
His Cys Ser Asp Ser Gin Arg Thr Ser Ser Ala Met Ser Gly His Ser
365 370 375

cgc tcc ttc atc tcc cgg gag ctg aag gag atc cag atg aca tct cag
Arg Ser Phe Ile Ser Arg Glu Leu Lys Glu Ile Ser Thr Ser Gin
380 385 390

acc tct ctg cca gac ctc tca ctg cca gac ctc agt gaa aat ggc ctt
Thr Leu Leu Pro Asp Leu Ser Leu Pro Asp Leu Ser Glu Asn Gly Leu
395 400 405

gga gcg agg aat ttg ctt cca ggt gtg cct ggc atg ggc ctc ggc cag
Gly Gly Arg Asn Leu Pro Gly Val Pro Gly Met Gly Leu Ala Gin
410 415 420 425

gaa gac acc acc tca ctg agg act ttg cga ata tca gag acc tca gag
Glu Asp Thr Thr Ser Leu Arg Thr Leu Arg Ile Ser Glu Thr Ser Asp
430 435 440

tct tca cag gtt cag gac tca gag agt gtc tta ctg gtg gat gag cag
Ser Ser Gin Gly Gin Asp Ser Ser Glu Ser Val Leu Val Asp Glu Ala
445 450 455

ggt ggg agc ggc agg gct ggg cct gcc cct aag ggg agc tcc ctg caa
Gly Gly Ser Gly Arg Ala Gin Pro Ala Pro Lys Gly Ser Ser Leu Gin
460 465 470

3
gtc aca ttt ccc agt gaa aca ctg aac tta tca gaa aaa tgt ata taa
Val Thr Phe Pro Ser Glu Thr Leu Asn Leu Ser Glu Lys Cys Ile *
475 480 485
tag

<210> 3
<211> 488
<212> PRT
<213> Homo sapiens

<400> 3
Met Ser Thr Pro Gly Val Asn Ser Ser Ala Ser Leu Ser Pro Asp Arg
1  5 10  15
Leu Asn Ser Pro Val Thr Ile Pro Ala Val Met Phe Ile Phe Gly Val
20  25  30
Val Gly Asn Leu Val Ala Ile Val Val Leu Cys Lys Ser Arg Lys Glu
35  40  45
Gln Lys Glu Thr Thr Phe Tyr Thr Leu Val Cys Gly Leu Ala Val Thr
50  55  60
Asp Leu Leu Gly Thr Leu Leu Ser Pro Val Thr Ile Ala Thr Tyr
65  70  75  80
Met Lys Gly Gln Trp Pro Gly Gly Gln Pro Leu Cys Glu Tyr Ser Thr
85  90  95
Phe Ile Leu Leu Phe Phe Ser Leu Ser Gly Leu Ser Ile Ile Cys Ala
100 105 110
Met Ser Val Glu Arg Tyr Leu Ala Ile Asn His Ala Tyr Phe Tyr Ser
115 120 125
His Tyr Val Asp Lys Arg Leu Ala Gly Leu Thr Leu Phe Ala Val Tyr
130 135 140
Ala Ser Asn Val Leu Phe Cys Ala Leu Pro Asn Met Gly Leu Gly Ser
145 150 155 160
Ser Arg Leu Gln Tyr Pro Asp Thr Trp Cys Phe Ile Asp Trp Thr Thr
165 170 175
Asn Val Thr Ala His Ala Ala Tyr Ser Tyr Met Tyr Ala Gly Phe Ser
180 185 190
Ser Phe Leu Ile Leu Ala Thr Val Leu Cys Asn Val Leu Val Cys Gly
195 200 205
Ala Leu Leu Arg Met His Arg Gln Phe Met Arg Arg Thr Ser Leu Gly
210 215 220
Thr Glu Gln His His Ala Ala Ala Ala Ser Val Ala Ser Arg Gly
225 230 235 240
His Pro Ala Ala Ser Pro Ala Leu Pro Arg Leu Ser Asp Phe Arg Arg
245 250 255
Arg Arg Ser Phe Arg Arg Ile Ala Gly Ala Glu Ile Gin Met Val Ile
260 265 270
Leu Leu Ile Ala Thr Ser Leu Val Val Leu Ile Cys Ser Ile Pro Leu
275 280 285
Val Val Arg Val Phe Val Asn Glu Leu Tyr Gln Pro Ser Leu Glu Arg
290 295 300
Glu Val Ser Lys Asn Pro Asp Leu Gln Ala Ile Arg Ile Ala Ser Val
305 310 315 320
Asn Pro Ile Leu Asp Pro Trp Ile Tyr Ile Leu Leu Arg Lys Thr Val
325 330 335
Leu Ser Lys Ala Ile Glu Lys Ile Lys Cys Leu Phe Cys Arg Ile Gly
340 345 350
Gly Ser Arg Arg Glu Arg Ser Gly Gln His Cys Ser Asp Ser Gin Arg
355 360 365
Thr Ser Ser Ala Met Ser Gly His Ser Arg Ser Phe Ile Ser Arg Glu
370 375 380
Leu Lys Glu Ile Ser Ser Thr Ser Gin Thr Leu Leu Pro Asp Leu Ser
<210> 4
<211> 1824
<212> DNA
<213> Homo sapiens

<400> 4
agccgaaagc ctgcctgtat aagatgtgca attactttgg tgcttgaagtt cctctgaacg 60
atatgccctt atttgcagac cttctctcct tcattctcaaa cgcttctctt ttctctctcatc 120
tctcctgcgc acctcatttt atcattcataac cttctctcaaa cctctccttc acctctcttc 180
atctccccctt cccagcagac ccctctcttc ccattctcagtt agttcctggaag tgggaatctc 240
accctctggg gcaacctcttc tggcccctaat gcacacac ccctctcttcc gcctctcttc 300
agctcttggcca aggtgaggagt cggagcttctc tgccaccccttgctgctt tccgccggttccc 360
aacctctcttc ctggctgtatc ccctctctctc tggcgcctcc gctgcaatatt cattcatttc 420
ctggctgctgc acctcagccg acagacctca gcctctcttgctctccagtt agtctgctgaag 480
cctctctcttc gggactgtgata cagggtcttca ggggtgtctg gatctgtgctg aagatgtgca 540
tctctctcttc gggactgtgata cagggtcttca ggggtgtctg gatctgtgctg aagatgtgca 600
tctctctcttc gggactgtgata cagggtcttca ggggtgtctg gatctgtgctg aagatgtgca 660
tctctctcttc gggactgtgata cagggtcttca ggggtgtctg gatctgtgctg aagatgtgca 720
tctctctcttc gggactgtgata cagggtcttca ggggtgtctg gatctgtgctg aagatgtgca 780
tctctctcttc gggactgtgata cagggtcttca ggggtgtctg gatctgtgctg aagatgtgca 840
tctctctcttc gggactgtgata cagggtcttca ggggtgtctg gatctgtgctg aagatgtgca 900
<210> 5
<211> 1824
<212> DNA
<213> Homo sapiens

<220> CDS
<221> (208)...(1482)
<222> 4
<table>
<thead>
<tr>
<th>Code</th>
<th>Sequence</th>
<th>Amino Acid(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>aac ccg acc cct ggg ggc acc ctc tct gcc ccc aat gcc aca aca ccc</td>
<td>Asn Pro Thr Pro Arg Gly Thr Leu Ser Ala Pro Asn Ala Thr Thr Pro</td>
</tr>
<tr>
<td>B</td>
<td>tgttgaagt cctctgaagc atttccgctt atttgaacc ttcttaccttctt ttctctccac ttctctggc acctcattt attcatcasa cctcttacct tgcattccaca cctctccttc atcctcttctt cccagcaaacc cttgctc atg gat tct ggg cct ctc tgt ggg gat gcc</td>
<td>Met Asp Ser Gly Pro Leu Trp Asp Ala</td>
</tr>
<tr>
<td>C</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>D</td>
<td>tgg ctc ggc cgg gat gag gatg cac aag gtt gag atccg gga gtc tgt</td>
<td>Trp Leu Gly Arg Asp Glu Leu Ala Lys Val Glu Ile Gly Val Leu</td>
</tr>
<tr>
<td>E</td>
<td>10</td>
<td>25</td>
</tr>
<tr>
<td>F</td>
<td>gcc act gtc ctc tgt gtc ggc acc ggg ggc aac ctc gct tgt ctg ctg</td>
<td>Ala Thr Val Leu Val Leu Ala Thr Val Gly Asn Ala Val Leu Leu</td>
</tr>
<tr>
<td>G</td>
<td>15</td>
<td>20</td>
</tr>
<tr>
<td>H</td>
<td>gcc act gtc ctc ggc cag ctc ggcc aag cgc tcc cgc atg cac ctc ttc gtc</td>
<td>Thr Leu Gly Gln Gly Arg Asp Arg Ser Arg Met His Leu Phe Val Leu</td>
</tr>
<tr>
<td>I</td>
<td>20</td>
<td>25</td>
</tr>
<tr>
<td>J</td>
<td>ctg cac tta gcc ctc aca gac ctc ggc tgt ggc ctc ttc cag tgt ctg</td>
<td>Leu His Leu Ala Ala Thr Leu Ala Val Ala Val Leu Val Leu</td>
</tr>
<tr>
<td>K</td>
<td>25</td>
<td>30</td>
</tr>
<tr>
<td>L</td>
<td>cca cag ctc tgg gac atc acc tac cgc ttc cag ggc ccc gac ctc</td>
<td>Pro Gln Leu Leu Trp Asp Ile Thr Tyr Arg Phe Gln Gly Pro Asp Leu</td>
</tr>
<tr>
<td>M</td>
<td>30</td>
<td>35</td>
</tr>
<tr>
<td>N</td>
<td>ctc tgg agg gcc gtc aag tac ctc ggc tgt ggc ctc atg ttt ggc ctc</td>
<td>Leu Cys Arg Ala Val Lys Tyr Leu Gln Val Leu Ser Met Phe Ala Ser</td>
</tr>
<tr>
<td>O</td>
<td>35</td>
<td>40</td>
</tr>
<tr>
<td>P</td>
<td>acc tac atg ctc ggc atg aag cgg ctc gac cgc tac ctc gct gtc tgt</td>
<td>Thr Tyr Met Leu Leu Ala Met Thr Leu Asp Arg Tyr Leu Ala Val Cys</td>
</tr>
<tr>
<td>Q</td>
<td>40</td>
<td>45</td>
</tr>
<tr>
<td>R</td>
<td>cac ctc ctc gc cag ctc cag cca ggc cag tcc acc tac ctc gtc</td>
<td>His Pro Leu Arg Ser Leu Gln Gln Pro Gly Gln Ser Thr Tyr Leu Leu</td>
</tr>
<tr>
<td>S</td>
<td>45</td>
<td>50</td>
</tr>
<tr>
<td>T</td>
<td>atc gct tgtt gcctt cgctt cgctt ggc gcc atgt tcgc acc ctc cct caa gtc</td>
<td>Ile Ala Ala Pro Trp Leu Leu Ala Ala Ile Phe Ser Leu Pro Gln Val</td>
</tr>
<tr>
<td>U</td>
<td>50</td>
<td>55</td>
</tr>
<tr>
<td>V</td>
<td>ttc att tgtcgc cgg gat gac atc ggc tca ggg gtt gtc gac</td>
<td>Phe Ile Phe Ser Leu Arg Glu Val Ile Gln Gly Ser Gly Val Leu Asp</td>
</tr>
<tr>
<td>W</td>
<td>55</td>
<td>60</td>
</tr>
<tr>
<td>X</td>
<td>tgg acc acc ctc gct tgt ctc ctc ctc ctc ctc aac ccg ggtg acc atg ctc aac gcc</td>
<td>Trp Thr Thr Leu Ala Ile Phe Val Leu Pro Val Thr Met Leu Thr Ala</td>
</tr>
<tr>
<td>Y</td>
<td>60</td>
<td>65</td>
</tr>
<tr>
<td>Z</td>
<td>65</td>
<td>70</td>
</tr>
<tr>
<td></td>
<td>70</td>
<td>75</td>
</tr>
<tr>
<td></td>
<td>75</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td>80</td>
<td>85</td>
</tr>
<tr>
<td></td>
<td>85</td>
<td></td>
</tr>
<tr>
<td></td>
<td>90</td>
<td></td>
</tr>
<tr>
<td></td>
<td>90</td>
<td></td>
</tr>
<tr>
<td></td>
<td>95</td>
<td></td>
</tr>
<tr>
<td></td>
<td>95</td>
<td></td>
</tr>
<tr>
<td></td>
<td>100</td>
<td></td>
</tr>
<tr>
<td></td>
<td>100</td>
<td></td>
</tr>
<tr>
<td></td>
<td>105</td>
<td></td>
</tr>
<tr>
<td></td>
<td>105</td>
<td></td>
</tr>
<tr>
<td></td>
<td>110</td>
<td></td>
</tr>
<tr>
<td></td>
<td>110</td>
<td></td>
</tr>
<tr>
<td></td>
<td>115</td>
<td></td>
</tr>
<tr>
<td></td>
<td>115</td>
<td></td>
</tr>
<tr>
<td></td>
<td>120</td>
<td></td>
</tr>
<tr>
<td></td>
<td>120</td>
<td></td>
</tr>
<tr>
<td></td>
<td>125</td>
<td></td>
</tr>
<tr>
<td></td>
<td>125</td>
<td></td>
</tr>
<tr>
<td></td>
<td>130</td>
<td></td>
</tr>
<tr>
<td></td>
<td>130</td>
<td></td>
</tr>
<tr>
<td></td>
<td>135</td>
<td></td>
</tr>
<tr>
<td></td>
<td>135</td>
<td></td>
</tr>
<tr>
<td></td>
<td>140</td>
<td></td>
</tr>
<tr>
<td></td>
<td>140</td>
<td></td>
</tr>
<tr>
<td></td>
<td>145</td>
<td></td>
</tr>
<tr>
<td></td>
<td>145</td>
<td></td>
</tr>
<tr>
<td></td>
<td>150</td>
<td></td>
</tr>
<tr>
<td></td>
<td>150</td>
<td></td>
</tr>
<tr>
<td></td>
<td>155</td>
<td></td>
</tr>
<tr>
<td></td>
<td>155</td>
<td></td>
</tr>
<tr>
<td></td>
<td>160</td>
<td></td>
</tr>
<tr>
<td></td>
<td>160</td>
<td></td>
</tr>
<tr>
<td></td>
<td>165</td>
<td></td>
</tr>
<tr>
<td></td>
<td>165</td>
<td></td>
</tr>
<tr>
<td></td>
<td>170</td>
<td></td>
</tr>
<tr>
<td></td>
<td>170</td>
<td></td>
</tr>
<tr>
<td></td>
<td>180</td>
<td></td>
</tr>
<tr>
<td></td>
<td>180</td>
<td></td>
</tr>
<tr>
<td></td>
<td>185</td>
<td></td>
</tr>
<tr>
<td></td>
<td>185</td>
<td></td>
</tr>
<tr>
<td></td>
<td>190</td>
<td></td>
</tr>
<tr>
<td></td>
<td>190</td>
<td></td>
</tr>
<tr>
<td></td>
<td>195</td>
<td></td>
</tr>
<tr>
<td></td>
<td>195</td>
<td></td>
</tr>
<tr>
<td></td>
<td>200</td>
<td></td>
</tr>
<tr>
<td></td>
<td>200</td>
<td></td>
</tr>
<tr>
<td></td>
<td>205</td>
<td></td>
</tr>
<tr>
<td></td>
<td>205</td>
<td></td>
</tr>
<tr>
<td></td>
<td>210</td>
<td></td>
</tr>
<tr>
<td></td>
<td>210</td>
<td></td>
</tr>
<tr>
<td></td>
<td>215</td>
<td></td>
</tr>
<tr>
<td></td>
<td>215</td>
<td></td>
</tr>
</tbody>
</table>
tgc tac agc ctc atc tgc cat gag atc tgt aac aac cta aac gtc aag
Cys Tyr Ser Leu Ile Cys His Glu Ile Cys Lys Asn Leu Lys Val Lys
220 225 230
aca cag gcc tgg cgg gtg gga gga ggg ggc tgg agg act tgt gac agg
Thr Glu Ala Trp Arg Val Gly Gly Gly Gly Trp Arg Thr Trp Asp Arg
235 240 245
ccc tca cct tcc acc tta gct gcc acc act cgg ggg ctc cca tct cgg
Pro Ser Pro Ser Thr Leu Ala Ala Thr Thr Arg Gly Leu Pro Ser Arg
250 255 260 265
gtc agc agc atc aac acc atc tca cgg gcc aag atc cga aca gtg aag
Val Ser Ser Ile Thr Ile Ser Arg Ala Lys Ile Arg Thr Val Lys
270 275 280
atg acc ttt gtc atc gtg ctc gcc tac atc gct tgc tgg gct ccc ttc
Met Thr Phe Val Ile Val Leu Ala Tyr Ile Ala Cys Trp Ala Pro Phe
285 290 295
ttc agt gtc cag atg tgg tcc gtg tgg gac aag aat gcc cct gat gaa
Phe Ser Val Glu Met Trp Ser Val Trp Asp Lys Asn Ala Pro Asp Glu
300 305 310
gat tcc acc aat gtg gct ttc acc atc tct atg ctt tgg ggc aac ctc
Asp Ser Thr Asn Val Ala Phe Thr Ile Ser Met Leu Leu Gly Asn Leu
315 320 325
aac agc tgc aac ccc tgg atc tac atg ggc ttc aac agc cac ctc
Asn Ser Cys Cys Asn Pro Thr Ile Tyr Met Gly Phe Asn Ser His Leu
330 335 340 345
tta ccg cgg ccc ctc ggt cac ctt gcc tgc tgt ggg ggt ccc cag ccc
Leu Pro Arg Pro Pro Arg Pro Arg His Leu Ala Cys Cys Gly Gly Pro Glu Pro
350 355 360
agg atg cgc cgg ctc tcc gac ggc agc ctc tgg agc cgc cac acc
Arg Met Arg Arg Arg Leu Ser Asp Gly Ser Leu Ser Ser Arg His Thr
365 370 375
acg ctg ctg acc cgc tcc agc tgc cgg gcc acc ctc agc ctc agc ctc
Thr Leu Thr Leu Thr Ser Cys Pro Thr Arg Ser Pro Arg Arg Pro Arg Ser Arg
380 385 390
agc cta acc ctc agt ggg agg ccc agg ctt gaa gag tca cca agg gag
Ser Leu Thr Leu Ser Gly Arg Pro Arg Pro Arg Glu Ser Arg Asp
395 400 405
ttg gag ctg gca gat ggg gaa ggc acc gct gag acc atc atc ttt tag
Leu Glu Leu Ala Asp Gly Glu Gly Thr Ala Glu Thr Ile Ile Phe *
410 415 420
gaaagactcg ctggggtctg tgaactgccc cagagactgt gggagtttct tgcctcacctc 1542
gggcactgga aatgagagct gggaggttaa gggttgaggt tagagagacc cttgctaaag 1602
cggagcaaga aggcccgaat gggtccccca cccctgggtgc acacgctgcc ctaggttgag 1662
ggtgctctca taaagttcaca atctcaagaca ctgctagctaa gggagaatca aacgctcctg 1722
cctcctgctgctcgcatatat ccagggggtg ccactgccac acaggtgcctgcc agatctagcgc* 1782
agggctaggag gtgtctgtctgtactagggggtg cacaggtgc cc 1824

<210> 6
<211> 424
<212>  PRT
<213>  Homo sapiens

<400>  6
Met  Asp  Ser  Gly  Pro  Leu  Trp  Asp  Ala  Asn  Pro  Thr  Pro  Arg  Gly  Thr
  1    5    10    15
Leu  Ser  Ala  Asn  Ala  Thr  Thr  Pro  Trp  Leu  Gly  Arg  Asp  Glu  Glu
  20   25    30
Leu  Ala  Lys  Val  Glu  Ile  Gly  Val  Leu  Ala  Thr  Val  Leu  Val  Leu  Ala
  35   40    45
Thr  Gly  Gly  Asn  Leu  Ala  Val  Leu  Leu  Thr  Leu  Gly  Gln  Leu  Gly  Arg
  50   55    60
Lys  Arg  Ser  Arg  Met  His  Leu  Phe  Val  Leu  His  Leu  Ala  Leu  Thr  Asp
  65    70    75    80
Leu  Ala  Val  Ala  Leu  Phe  Gln  Val  Leu  Pro  Gln  Leu  Leu  Trp  Asp  Ile
  85   90    95
Thr  Tyr  Arg  Phe  Gln  Gly  Pro  Asp  Leu  Leu  Cys  Arg  Ala  Val  Lys  Tyr
 100  105   110
Leu  Gln  Val  Leu  Ser  Met  Phe  Ala  Ser  Thr  Tyr  Met  Leu  Leu  Ala  Met
 115   120   125
Thr  Leu  Asp  Arg  Tyr  Leu  Ala  Val  Cys  His  Pro  Leu  Arg  Ser  Leu  Gln
 130   135   140
Gln  Pro  Gly  Gln  Ser  Thr  Tyr  Leu  Leu  Ile  Ala  Ala  Pro  Trp  Leu  Leu
 145   150   155   160
 Ala  Ala  Ile  Phe  Ser  Leu  Pro  Gln  Val  Phe  Ile  Phe  Ser  Arg  Leu  Gln
 165   170   175
Val  Ile  Gln  Gly  Ser  Gly  Val  Leu  Asp  Cys  Trp  Ala  Asp  Phe  Gly  Phe
 180   185   190
Pro  Trp  Gly  Pro  Arg  Ala  Tyr  Leu  Thr  Trp  Thr  Thr  Leu  Ala  Ile  Phe
 195   200   205
Val  Leu  Pro  Val  Thr  Met  Leu  Thr  Ala  Cys  Tyr  Ser  Leu  Ile  Cys  His
 210   215   220
Glu  Ile  Cys  Lys  Asn  Leu  Lys  Val  Lys  Thr  Gln  Ala  Trp  Arg  Val  Gly
 225   230   235   240
Gly  Gly  Gly  Trp  Arg  Thr  Trp  Asp  Arg  Pro  Ser  Pro  Ser  Thr  Leu  Ala
 245   250   255
 Ala  Thr  Thr  Arg  Gly  Leu  Pro  Ser  Arg  Val  Ser  Ile  Asn  Thr  Ile
 260   265   270
Ser  Arg  Ala  Lys  Ile  Arg  Thr  Val  Lys  Met  Thr  Phe  Val  Ile  Val  Leu
 275   280   285
 Ala  Tyr  Ile  Ala  Cys  Trp  Ala  Pro  Phe  Phe  Ser  Val  Gln  Met  Trp  Ser
 290   295   300
Val  Trp  Asp  Lys  Ala  Pro  Asp  Glu  Asp  Ser  Thr  Asn  Val  Ala  Phe
 305   310   315   320
Thr  Ile  Ser  Met  Leu  Leu  Gly  Asn  Leu  Asn  Ser  Cys  Cys  Asn  Pro  Trp
 325   330   335
 Ile  Tyr  Met  Gly  Phe  Asn  Ser  His  Leu  Leu  Pro  Arg  Pro  Leu  Arg  His
 340   345   350
Leu  Ala  Cys  Cys  Gly  Gly  Pro  Gln  Pro  Arg  Met  Arg  Arg  Arg  Leu  Ser
 355   360   365
 Asp  Gly  Ser  Leu  Ser  Ser  Arg  His  Thr  Thr  Leu  Leu  Thr  Arg  Ser  Ser
 370   375   380
Cys  Pro  Ala  Thr  Leu  Ser  Leu  Ser  Leu  Thr  Leu  Ser  Gly  Arg
 385   390   395   400
 Pro  Arg  Pro  Glu  Glu  Ser  Pro  Arg  Asp  Leu  Glu  Leu  Ala  Asp  Gly  Glu
 405   410   415
Gly  Thr  Ala  Glu  Thr  Ile  Ile  Phe
 420
<220>
<221> CDS
<222> (1)...(2943)

<400> 8
atg ccg gct gag ctg ctg ctg att gtt gcc ttc gcc agc ccc
Met Pro Ala Glu Leu Leu Leu Leu Leu Leu Ile Val Ala Phe Ala Ser Pro
1  5 10 15

gcg tgg cag gtg tca tca ctg cgct gct gcc atc ctg gat gat
Ser Cys Glu Val Leu Ser Ser Leu Arg Met Ala Ala Ile Leu Asp Asp
20 25 30

cag acg tgg tgt ggc cgc ggt gag cgt ctg gcc ttg gcc ttg gcc cgg
Gln Thr Val Cys Gly Arg Gly Glu Arg Leu Ala Leu Ala Arg
35 40 45

gag cag atc aac ggg atc atc gag gtc cca gcc aag gcc cga gtt gaa
Glu Gln Ile Asn Gly Ile Ile Glu Val Pro Ala Lys Ala Arg Val Glu
50 55 60

gta gac atc tgt gct tgt cag cgg gag agc gag tac gac acc acg gac
Val Asp Ile Phe Glu Leu Gln Arg Asp Ser Glu Tyr Glu Thr Thr Asp
65 70 75 80

acc atg tgt cag atc tta ccc aag tgt tgt ctc tct tgt ggt gcc ccc
Thr Met Cys Glu Ile Pro Lys Gly Val Val Ser Val Leu Gly Pro
85 90 95

tcc tct agc cca gca tct gcc tcc acc tgt agc cat atc tgt gga gag
Ser Ser Ser Pro Ala Ser Ala Ser Thr Val Ser His Ile Cys Gly Glu
100 105 110

aag gag atc ccc cac atc aag tgt tgt cc ggc cag gag aga ccc cgc ctt
Lys Glu Ile Pro His Ile Lys Val Gly Pro Glu Glu Thr Pro Arg Leu
115 120 125

cag tac ctt cgc ttc gcc tct gtc agc ctg tac ccc aqt aac gag cag
Gln Tyr Leu Arg Phe Ala Ser Val Ser Leu Tyr Pro Ser Asn Glu Asp
130 135 140

gtc agc tgt gcc ctc cga atc ctc aag tcc ttc aac tcc tcc tgg ctc
Val Ser Leu Ala Val Ser Arg Ile Leu Lys Ser Phe Asn Tyr Pro Ser
145 150 155 160

gcc agc ctc atc tgc cag aac ggt tag tcc ctc cag cta tgg tag gaa
gaa Ala Ser Leu Ile Cys Ala Lys Ala Glu Cys Leu Leu Arg Glu Glu
165 170 175

cgg tgt ggt gcc ttc gcc ttc atc aag gat agc ctg tca ggt gaa
Leu Val Arg Gly Phe Leu Ile Ser Lys Glu Thr Leu Ser Val Arg Met
180 185 190

ttg gac gac ccg gcc ccc aca cca ctc atg gat
Leu Asp Asp Ser Arg Asp Pro Thr Pro Leu Leu Lys Glu Ile Arg Glu
195 200 205

gac aag tgt tcc atc atc atc gac gcc aac gcc tcc atc tcc cac
Asp Lys Val Ser Thr Ile Ile Ile Asp Ala Asn Ala Ser Ile Ser His
210 215 220
ctc atc ctc cg t aag gcc tcg gaa ctg gga atg acc tca gcg ttt tac
Leu Ile Leu Arg Lys Ala Ser Glu Leu Gly Met Thr Ser Ala Phe Tyr
225 230 235 240

aag tac atc ctc acc acc atg gac ttc ccc atc ctg cat ctg gac ggt
Lys Tyr Ile Leu Thr Thr Met Asp Phe Pro Ile Leu His Leu Asp Gly
245 250 255

att gtg gag gac tcc ttc aac atc ctg gcc tcc atg ttc aac acg
Ile Val Glu Asp Ser Ser Asn Ile Leu Gly Phe Ser Met Phe Asn Thr
260 265 270 816

ctc cac ccc ttc tac cct gag ttt gtc cgc agc ctc aac atg tcc tgg
Ser His Pro Phe Tyr Pro Glu Phe Val Arg Ser Leu Asn Met Ser Trp
275 280 285 864

agg gag aac tgt gaa gcc agc acc tac ctg gcc cct gcg ctg tca gcc
Arg Glu Asn Cys Glu Ala Ser Thr Tyr Leu Gly Pro Ala Leu Ser Ala
290 295 300 912

gcc ctg atg ttt gac gcc ctg cac gtc gtt gcg agc gct gtc cga gag
Ala Leu Met Phe Asp Ala Val His Val Val Ser Ala Val Arg Glu
305 310 315 320 960

ctg aac cgc agc cag gag atc gtt gtc aag cct ctg gcc tgt aca ctg
Leu Asn Arg Ser Gln Glu Ile Gly Val Lys Pro Leu Ala Cys Thr Ser
325 330 335 1008

gcc aac att tgg ccc cac ggg acc agc ctc atg aac tac ctg cgc atg
Ala Asn Ile Trp Pro His Gly Thr Ser Leu Met Asn Tyr Leu Arg Met
340 345 350 1056

gta gag tat gat ggg ctg acc ggg cgg gtc gag ttc aac agc aac aag
Val Glu Tyr Asp Gly Leu Thr Gly Arg Val Glu Phe Asn Ser Lys Gly
355 360 365 1104

cag aga acc aac tac acc ctg cgc atc tga aag tcc cgg cag gcc
Gln Arg Thr Asn Tyr Leu Arg Ile Leu Glu Lys Ser Arg Gln Gly
370 375 380 1152

cac cg t gat ggg gtg tac tct aac cgc acc ctg gcc atg aat
His Arg Glu Ile Gly Val Trp Tyr Ser Asn Arg Thr Leu Ala Met Asn
385 390 395 400 1200

gcc acc acc ctg gac atc aac ctg ctc cag aca ctg gcc aac aag acc
Ala Thr Thr Leu Asp Ile Leu Ser Glu Thr Leu Ala Asn Lys Thr
405 410 415 1248

ctg gtg gtc aca acc atc ctg gag aac cca tac gtc atg cgc cgg ccc
Leu Val Val Thr Ile Leu Glu Asn Pro Tyr Val Met Arg Arg Pro
420 425 430 1296

aac ttc cag ggc ctg cgg aac gaa cgc ttc gag ggc ttc tgc gtg
Asn Phe Gln Gly Leu Ser Gly Asn Glu Arg Phe Glu Gly Phe Cys Val
435 440 445 1344

gac atg ctg cgg gag ctg gcc gag ctg ctc ccc tac cgc ctg
Asp Met Leu Arg Glu Leu Ala Glu Leu Leu Pro Phe Pro Tyr Arg Leu
450 455 460 1392

cgg tgt gat gaa cgg ctg tac ggg gcg ccc gag ccc aac gcc tcc
1440
<table>
<thead>
<tr>
<th>465</th>
<th>Arg</th>
<th>Leu</th>
<th>Val</th>
<th>Glu</th>
<th>Asp</th>
<th>Gly</th>
<th>Leu</th>
<th>Tyr</th>
<th>Gly</th>
<th>Ala</th>
<th>Pro</th>
<th>Glu</th>
<th>Pro</th>
<th>Asn</th>
<th>Gly</th>
<th>Ser</th>
<th>480</th>
</tr>
</thead>
<tbody>
<tr>
<td>tgg</td>
<td>acg</td>
<td>ggc</td>
<td>atg</td>
<td>gtt</td>
<td>ggc</td>
<td>gag</td>
<td>ctc</td>
<td>atc</td>
<td>aac</td>
<td>cgg</td>
<td>aag</td>
<td>gca</td>
<td>gac</td>
<td>ctg</td>
<td>gct</td>
<td>1488</td>
<td></td>
</tr>
<tr>
<td>Trp</td>
<td>Thr</td>
<td>Gly</td>
<td>Met</td>
<td>Val</td>
<td>Gly</td>
<td>Glu</td>
<td>Leu</td>
<td>Ile</td>
<td>Asn</td>
<td>Arg</td>
<td>Lys</td>
<td>Ala</td>
<td>Asp</td>
<td>Leu</td>
<td>Ala</td>
<td>485</td>
<td></td>
</tr>
<tr>
<td>490</td>
<td>gtt</td>
<td>gcc</td>
<td>gcc</td>
<td>ttc</td>
<td>acc</td>
<td>atc</td>
<td>aca</td>
<td>gct</td>
<td>gag</td>
<td>cgg</td>
<td>gag</td>
<td>aag</td>
<td>gtc</td>
<td>atc</td>
<td>gac</td>
<td>ttt</td>
<td>1536</td>
</tr>
<tr>
<td>Val</td>
<td>Ala</td>
<td>Ala</td>
<td>Phe</td>
<td>Thr</td>
<td>Ile</td>
<td>Thr</td>
<td>Ala</td>
<td>Glu</td>
<td>Arg</td>
<td>Glu</td>
<td>Lys</td>
<td>Val</td>
<td>Ile</td>
<td>Asp</td>
<td>Phe</td>
<td>500</td>
<td></td>
</tr>
<tr>
<td>510</td>
<td>tcc</td>
<td>aag</td>
<td>ccc</td>
<td>ttt</td>
<td>atg</td>
<td>acc</td>
<td>ctg</td>
<td>ggg</td>
<td>atc</td>
<td>agc</td>
<td>atc</td>
<td>ttc</td>
<td>tac</td>
<td>cga</td>
<td>gtc</td>
<td>cac</td>
<td>1584</td>
</tr>
<tr>
<td>Ser</td>
<td>Lys</td>
<td>Pro</td>
<td>Phe</td>
<td>Met</td>
<td>Thr</td>
<td>Leu</td>
<td>Gly</td>
<td>Ile</td>
<td>Ser</td>
<td>Ile</td>
<td>Leu</td>
<td>Tyr</td>
<td>Arg</td>
<td>Val</td>
<td>His</td>
<td>515</td>
<td></td>
</tr>
<tr>
<td>525</td>
<td>atg</td>
<td>gcc</td>
<td>cgc</td>
<td>aag</td>
<td>cct</td>
<td>ggc</td>
<td>tac</td>
<td>ttc</td>
<td>ttc</td>
<td>ctg</td>
<td>gac</td>
<td>ccc</td>
<td>ttc</td>
<td>ttc</td>
<td>cct</td>
<td>1632</td>
<td></td>
</tr>
<tr>
<td>Met</td>
<td>Gly</td>
<td>Arg</td>
<td>Lys</td>
<td>Pro</td>
<td>Gly</td>
<td>Tyr</td>
<td>Phe</td>
<td>Ser</td>
<td>Phe</td>
<td>Ser</td>
<td>Phe</td>
<td>Pro</td>
<td>Phe</td>
<td>Ser</td>
<td>Pro</td>
<td>Pro</td>
<td>530</td>
</tr>
<tr>
<td>540</td>
<td>gct</td>
<td>gtg</td>
<td>tgg</td>
<td>ctc</td>
<td>ttc</td>
<td>atg</td>
<td>ctt</td>
<td>ctt</td>
<td>gcc</td>
<td>tac</td>
<td>tgc</td>
<td>gct</td>
<td>agc</td>
<td>tgc</td>
<td>gtc</td>
<td>1680</td>
<td></td>
</tr>
<tr>
<td>Ala</td>
<td>Val</td>
<td>Trp</td>
<td>Leu</td>
<td>Phe</td>
<td>Met</td>
<td>Leu</td>
<td>Ala</td>
<td>Leu</td>
<td>Tyr</td>
<td>Leu</td>
<td>Ala</td>
<td>Val</td>
<td>Ser</td>
<td>Cys</td>
<td>Val</td>
<td>545</td>
<td></td>
</tr>
<tr>
<td>550</td>
<td>ttt</td>
<td>ctg</td>
<td>gct</td>
<td>gcc</td>
<td>agg</td>
<td>ctg</td>
<td>agc</td>
<td>ccc</td>
<td>tat</td>
<td>gag</td>
<td>tgg</td>
<td>tat</td>
<td>aac</td>
<td>cca</td>
<td>cac</td>
<td>1728</td>
<td></td>
</tr>
<tr>
<td>Leu</td>
<td>Phe</td>
<td>Leu</td>
<td>Ala</td>
<td>Ala</td>
<td>Ala</td>
<td>Ala</td>
<td>Leu</td>
<td>Arg</td>
<td>Leu</td>
<td>Gly</td>
<td>Ser</td>
<td>Leu</td>
<td>Thr</td>
<td>Phe</td>
<td>Pro</td>
<td>Val</td>
<td>Gly</td>
</tr>
<tr>
<td>570</td>
<td>cca</td>
<td>tgc</td>
<td>ctg</td>
<td>cgg</td>
<td>gca</td>
<td>cgc</td>
<td>ccc</td>
<td>cac</td>
<td>atc</td>
<td>ctg</td>
<td>gag</td>
<td>aac</td>
<td>cag</td>
<td>tac</td>
<td>acg</td>
<td>ctg</td>
<td>1776</td>
</tr>
<tr>
<td>Pro</td>
<td>Cys</td>
<td>Leu</td>
<td>Arg</td>
<td>Ala</td>
<td>Arg</td>
<td>Pro</td>
<td>His</td>
<td>Ile</td>
<td>Leu</td>
<td>Glu</td>
<td>Asn</td>
<td>Gln</td>
<td>Tyr</td>
<td>Thr</td>
<td>Leu</td>
<td>580</td>
<td></td>
</tr>
<tr>
<td>590</td>
<td>ggc</td>
<td>aac</td>
<td>agc</td>
<td>ctg</td>
<td>tgg</td>
<td>ttt</td>
<td>ccc</td>
<td>gtg</td>
<td>ggg</td>
<td>gcc</td>
<td>ttc</td>
<td>atg</td>
<td>cag</td>
<td>cag</td>
<td>gcc</td>
<td>tgg</td>
<td>1824</td>
</tr>
<tr>
<td>Gly</td>
<td>Asn</td>
<td>Ser</td>
<td>Leu</td>
<td>Trp</td>
<td>Phe</td>
<td>Pro</td>
<td>Val</td>
<td>Gly</td>
<td>Gly</td>
<td>Phe</td>
<td>Met</td>
<td>Gln</td>
<td>Gln</td>
<td>Gln</td>
<td>Gly</td>
<td>Ser</td>
<td>595</td>
</tr>
<tr>
<td>605</td>
<td>gag</td>
<td>atc</td>
<td>atg</td>
<td>ccc</td>
<td>cgg</td>
<td>ggc</td>
<td>ctg</td>
<td>tcc</td>
<td>acg</td>
<td>cgc</td>
<td>tgt</td>
<td>gtc</td>
<td>agc</td>
<td>gga</td>
<td>gtc</td>
<td>tgt</td>
<td>1872</td>
</tr>
<tr>
<td>Glu</td>
<td>Ile</td>
<td>Met</td>
<td>Pro</td>
<td>Arg</td>
<td>Ala</td>
<td>Leu</td>
<td>Ser</td>
<td>Thr</td>
<td>Arg</td>
<td>Cys</td>
<td>Val</td>
<td>Ser</td>
<td>Gly</td>
<td>Val</td>
<td>Trp</td>
<td>610</td>
<td></td>
</tr>
<tr>
<td>620</td>
<td>tgg</td>
<td>gcc</td>
<td>ttc</td>
<td>acc</td>
<td>tgg</td>
<td>atc</td>
<td>atc</td>
<td>ttc</td>
<td>ttc</td>
<td>tac</td>
<td>acg</td>
<td>gcc</td>
<td>aac</td>
<td>ctg</td>
<td>gcc</td>
<td>1920</td>
<td></td>
</tr>
<tr>
<td>Trp</td>
<td>Ala</td>
<td>Phe</td>
<td>Thr</td>
<td>Leu</td>
<td>Ile</td>
<td>Ile</td>
<td>Ile</td>
<td>Ser</td>
<td>Ser</td>
<td>Tyr</td>
<td>Thr</td>
<td>Ala</td>
<td>Asn</td>
<td>Leu</td>
<td>Ala</td>
<td>625</td>
<td></td>
</tr>
<tr>
<td>640</td>
<td>gcc</td>
<td>ttc</td>
<td>ctc</td>
<td>acc</td>
<td>gtg</td>
<td>cag</td>
<td>cgc</td>
<td>atg</td>
<td>gag</td>
<td>tgt</td>
<td>cct</td>
<td>tgt</td>
<td>gag</td>
<td>tgg</td>
<td>gcc</td>
<td>gat</td>
<td>1968</td>
</tr>
<tr>
<td>Ala</td>
<td>Phe</td>
<td>Leu</td>
<td>Thr</td>
<td>Val</td>
<td>Gln</td>
<td>Arg</td>
<td>Met</td>
<td>Glu</td>
<td>Val</td>
<td>Pro</td>
<td>Val</td>
<td>Glu</td>
<td>Ser</td>
<td>Ala</td>
<td>Asp</td>
<td>645</td>
<td></td>
</tr>
<tr>
<td>655</td>
<td>gac</td>
<td>ctg</td>
<td>gca</td>
<td>gat</td>
<td>cag</td>
<td>acc</td>
<td>aac</td>
<td>atc</td>
<td>gag</td>
<td>tat</td>
<td>ggc</td>
<td>acc</td>
<td>atc</td>
<td>cac</td>
<td>gcc</td>
<td>ggc</td>
<td>2016</td>
</tr>
<tr>
<td>Asp</td>
<td>Leu</td>
<td>Ala</td>
<td>Asp</td>
<td>Gln</td>
<td>Thr</td>
<td>Asn</td>
<td>Ile</td>
<td>Glu</td>
<td>Tyr</td>
<td>Gly</td>
<td>Thr</td>
<td>Ile</td>
<td>His</td>
<td>Ala</td>
<td>Gly</td>
<td>660</td>
<td></td>
</tr>
<tr>
<td>670</td>
<td>tcc</td>
<td>acc</td>
<td>atg</td>
<td>acc</td>
<td>ttc</td>
<td>cag</td>
<td>aat</td>
<td>tca</td>
<td>cgg</td>
<td>tac</td>
<td>caa</td>
<td>acg</td>
<td>tac</td>
<td>cag</td>
<td>cag</td>
<td>2064</td>
<td></td>
</tr>
<tr>
<td>Ser</td>
<td>Thr</td>
<td>Met</td>
<td>Thr</td>
<td>Phe</td>
<td>Phe</td>
<td>Glu</td>
<td>Asn</td>
<td>Ser</td>
<td>Arg</td>
<td>Tyr</td>
<td>Gln</td>
<td>Thr</td>
<td>Tyr</td>
<td>Gln</td>
<td>Arg</td>
<td>675</td>
<td></td>
</tr>
<tr>
<td>685</td>
<td>atg</td>
<td>tgg</td>
<td>aac</td>
<td>tac</td>
<td>atg</td>
<td>cag</td>
<td>tcc</td>
<td>aag</td>
<td>cag</td>
<td>ccc</td>
<td>agc</td>
<td>tgt</td>
<td>ttc</td>
<td>gtc</td>
<td>aag</td>
<td>agc</td>
<td>2112</td>
</tr>
<tr>
<td>Met</td>
<td>Trp</td>
<td>Asn</td>
<td>Tyr</td>
<td>Met</td>
<td>Gln</td>
<td>Ser</td>
<td>Lys</td>
<td>Gln</td>
<td>Pro</td>
<td>Ser</td>
<td>Val</td>
<td>Phe</td>
<td>Val</td>
<td>Lys</td>
<td>Ser</td>
<td>690</td>
<td></td>
</tr>
<tr>
<td>700</td>
<td>aca</td>
<td>gaa</td>
<td>gag</td>
<td>gcc</td>
<td>gtc</td>
<td>ctc</td>
<td>aac</td>
<td>tcc</td>
<td>cgc</td>
<td>tac</td>
<td>gcc</td>
<td>ttc</td>
<td>ctg</td>
<td>Thr</td>
<td>Glu</td>
<td>Gly</td>
<td>Ile</td>
</tr>
<tr>
<td></td>
<td>acc</td>
<td>agc</td>
<td>ccg</td>
<td>ccc</td>
<td>cgG</td>
<td>CCT</td>
<td>GCC</td>
<td>GCC</td>
<td>GGC</td>
<td>GGC</td>
<td>CGG</td>
<td>GAG</td>
<td>CTG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thr Ser Pro Pro Arg Pro Arg Pro Gly Pro Ala Gly Pro Arg Glu Leu</td>
<td>965</td>
<td>970</td>
<td>975</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>gcg</th>
<th>gag</th>
<th>cac</th>
<th>gag</th>
<th>tga</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ala Glu His Glu *</td>
<td>980</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>2928</th>
<th>2943</th>
</tr>
</thead>
</table>

<210> 9
<211> 980
<212> PRT
<213> Homo sapiens

<table>
<thead>
<tr>
<th></th>
<th>9</th>
<th>5</th>
<th>10</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Met Pro Ala Glu Leu Leu Leu Leu Leu Leu Leu Ile Val Ala Phe Ala Ser Pro</td>
<td>1</td>
<td>5</td>
<td>10</td>
<td>15</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>20</th>
<th>25</th>
<th>30</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ser Cys Gln Val Leu Ser Ser Leu Arg Met Ala Ala Ile Leu Asp Asp</td>
<td>20</td>
<td>25</td>
<td>30</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>35</th>
<th>40</th>
<th>45</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gln Thr Val Cys Gly Arg Gly Glu Arg Leu Ala Leu Ala Leu Ala Arg</td>
<td>35</td>
<td>40</td>
<td>45</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>50</th>
<th>55</th>
<th>60</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glu Gln Ile Asn Gly Ile Ile Glu Val Pro Ala Lys Ala Arg Val Glu</td>
<td>50</td>
<td>55</td>
<td>60</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>65</th>
<th>70</th>
<th>75</th>
<th>80</th>
</tr>
</thead>
<tbody>
<tr>
<td>Val Asp Ile Phe Glu Leu Glu Glu Arg Asp Ser Glu Tyr Glu Thr Thr Asp</td>
<td>65</td>
<td>70</td>
<td>75</td>
<td>80</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>85</th>
<th>90</th>
<th>95</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thr Met Cys Gln Ile Leu Pro Lys Gly Val Val Ser Val Leu Gly Pro</td>
<td>85</td>
<td>90</td>
<td>95</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>100</th>
<th>105</th>
<th>110</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ser Ser Ser Pro Ala Ser Ala Ser Thr Val Ser His Ile Cys Gly Glu</td>
<td>100</td>
<td>105</td>
<td>110</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>115</th>
<th>120</th>
<th>125</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lys Glu Ile Pro His Ile Lys Val Gly Pro Glu Glu Thr Pro Arg Leu</td>
<td>115</td>
<td>120</td>
<td>125</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>130</th>
<th>135</th>
<th>140</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gln Tyr Leu Arg Phe Ala Ser Val Ser Leu Tyr Pro Ser Asn Glu Asp</td>
<td>130</td>
<td>135</td>
<td>140</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>145</th>
<th>150</th>
<th>155</th>
<th>160</th>
</tr>
</thead>
<tbody>
<tr>
<td>Val Ser Leu Ala Val Ser Arg Ile Leu Lys Ser Phe Asn Tyr Pro Ser</td>
<td>145</td>
<td>150</td>
<td>155</td>
<td>160</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>165</th>
<th>170</th>
<th>175</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ala Ser Leu Ile Cys Ala Lys Ala Glu Cys Leu Leu Arg Leu Glu Glu</td>
<td>165</td>
<td>170</td>
<td>175</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>180</th>
<th>185</th>
<th>190</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leu Val Arg Gly Phe Leu Ile Ser Lys Glu Thr Leu Ser Val Arg Met</td>
<td>180</td>
<td>185</td>
<td>190</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>195</th>
<th>200</th>
<th>205</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leu Asp Asp Ser Arg Asp Pro Thr Pro Leu Leu Lys Glu Ile Arg Asp</td>
<td>195</td>
<td>200</td>
<td>205</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>210</th>
<th>215</th>
<th>220</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asp Lys Val Ser Thr Ile Ile Ile Asp Ala Asn Ala Ser Ile Ser His</td>
<td>210</td>
<td>215</td>
<td>220</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>225</th>
<th>230</th>
<th>235</th>
<th>240</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leu Ile Leu Arg Lys Ala Ser Glu Leu Gly Met Thr Ser Ala Phe Tyr</td>
<td>225</td>
<td>230</td>
<td>235</td>
<td>240</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>245</th>
<th>250</th>
<th>255</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lys Tyr Ile’Leu Thr Thr Met Asp Phe Pro Ile Leu His Leu Asp Gly</td>
<td>245</td>
<td>250</td>
<td>255</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>260</th>
<th>265</th>
<th>270</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ile Val Glu Asp Ser Ser Asn Ile Leu Gly Phe Ser Met Phe Asn Thr</td>
<td>260</td>
<td>265</td>
<td>270</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>275</th>
<th>280</th>
<th>285</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ser His Pro Phe Tyr Pro Glu Phe Val Arg Ser Leu Asn Met Ser Trp</td>
<td>275</td>
<td>280</td>
<td>285</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>290</th>
<th>295</th>
<th>300</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arg Glu Asn Cys Glu Ala Ser Thr Tyr Leu Gly Pro Ala Leu Ser Ala</td>
<td>290</td>
<td>295</td>
<td>300</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>305</th>
<th>310</th>
<th>315</th>
<th>320</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ala Leu Met Phe Asp Ala Val His Val Val Val Ser Ala Val Arg Glu</td>
<td>305</td>
<td>310</td>
<td>315</td>
<td>320</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>325</th>
<th>330</th>
<th>335</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leu Asn Arg Ser Glu Ile Gly Val Lys Pro Leu Ala Cys Thr Ser</td>
<td>325</td>
<td>330</td>
<td>335</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>340</th>
<th>345</th>
<th>350</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ala Asn Ile Trp Pro His Gly Thr Ser Leu Met Asn Tyr Leu Arg Met</td>
<td>340</td>
<td>345</td>
<td>350</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>360</th>
<th>365</th>
<th>370</th>
<th>375</th>
</tr>
</thead>
<tbody>
<tr>
<td>Val Glu Tyr Asp Gly Leu Thr Gly Arg Val Glu Phe Asn Ser Lys Gly</td>
<td>360</td>
<td>365</td>
<td>370</td>
<td>375</td>
</tr>
</tbody>
</table>
Gln Arg Thr Asn Tyr Thr Leu Arg Ile Leu Glu Lys Ser Arg Gln Gly
His Arg Glu Ile Gly Val Trp Tyr Ser Asn Arg Thr Leu Ala Met Asn
385 390 395 400
Ala Thr Thr Leu Asp Ile Asn Leu Ser Gln Thr Leu Ala Asn Lys Thr
405 410 415
Leu Val Val Thr Thr Ile Leu Glu Asn Pro Tyr Val Met Arg Arg Pro
420 425 430
Asn Phe Glu Gly Leu Ser Gly Asn Glu Arg Phe Glu Gly Phe Cys Val
435 440 445
Asn Met Leu Arg Glu Leu Ala Glu Leu Pro Phe Pro Tyr Arg Leu
450 455 460
Arg Leu Val Glu Asp Gly Leu Tyr Gly Ala Pro Glu Pro Asn Gly Ser
465 470 475 480
Trp Thr Gly Met Val Gly Glu Leu Ile Asn Arg Lys Ala Asp Leu Ala
485 490 495
Val Ala Ala Phe Thr Ile Thr Ala Glu Arg Glu Lys Val Ile Asp Phe
500 505 510
Ser Lys Pro Phe Met Thr Leu Gly Ile Ser Ile Leu Tyr Arg Val His
515 520 525
Met Gly Arg Gly Pro Gly Tyr Phe Ser Phe Leu Asp Pro Phe Ser Pro
530 535 540
Ala Val Trp Leu Phe Met Leu Leu Ala Tyr Leu Ala Val Ser Cys Val
545 550 555 560
Leu Phe Leu Ala Ala Arg Leu Ser Pro Tyr Glu Trp Tyr Asn Pro His
565 570 575
Pro Cys Leu Arg Ala Arg Pro His Ile Leu Glu Asn Gln Tyr Thr Leu
580 585 590
Gly Asn Ser Leu Trp Phe Pro Val Gly Gly Phe Met Glu Gln Gly Ser
595 600 605
Glu Ile Met Pro Arg Ala Leu Ser Thr Arg Cys Val Ser Gly Val Trp
610 615 620
Trp Ala Phe Thr Leu Ile Ile Ser Ser Tyr Thr Ala Asn Leu Ala
625 630 635 640
Ala Phe Leu Thr Val Gln Arg Met Glu Val Pro Val Glu Ser Ala Asp
645 650 655
Asp Leu Ala Asp Glu Thr Asn Ile Glu Tyr Gly Thr Ile His Ala Gly
660 665 670
Ser Thr Met Thr Phe Phe Glu Asn Ser Arg Tyr Gln Thr Tyr Gln Arg
675 680 685
Met Trp Asn Tyr Met Gln Ser Lys Gln Pro Ser Val Phe Val Lys Ser
690 695 700
Thr Glu Gly Ile Ala Val Leu Asn Ser Arg Tyr Ala Phe Leu
705 710 715 720
Leu Glu Ser Thr Met Asn Glu Tyr His Arg Arg Leu Asn Cys Asn Leu
725 730 735
Thr Gln Ile Gly Leu Leu Asp Thr Lys Gly Tyr Gly Ile Gly Met
740 745 750
Pro Leu Gly Ser Pro Phe Arg Asp Glu Ile Thr Leu Ala Ile Leu Glu
755 760 765
Leu Glu Glu Asn Arg Leu Glu Ile Leu Lys Arg Lys Trp Trp Glu
770 775 780
Gly Gly Arg Cys Pro Lys Glu Glu Asp His Arg Ala Lys Gly Leu Gly
785 790 795 800
Met Glu Asn Ile Gly Ile Phe Ile Val Leu Ile Cys Gly Leu Ile
805 810 815
Ile Ala Val Phe Val Ala Val Met Glu Phe Ile Trp Ser Thr Arg Arg
820 825 830
Ser Ala Glu Ser Glu Glu Val Ser Val Cys Gln Glu Met Leu Gln Glu
835 840 845
ttaggggcta ccatgggcc agcttggtgcc aaggggaact aagcaaatcc aatatagatg 1860
ttttggtgga agggttttcga gcggcaacctc catttcctctg ctttaaatgt ggggttgtatg 1920
gctggtactg gggcagccac cttgctacca tgaagggagag gccaagcaca tcatccacag 1980
cattccctc cagcatcttgt ttcgtgtacca aaaaattaatg cttatattgt ttaagtcaaaa 2040
	2042

<210> 11
<211> 2042
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (171)...(1136)

<221> misc_feature
<222> (1)...(2042)
<223> n = A,T,C or G

<400> 11
ccatcgtgcca gctctcgggt ccagagcctc agctcctacc tcttcctctcc tgc cagcccc 60
gtctgtgctc ggcaccttct gcctctgctcg ggcaccttct gcctctcctt gtc cagcccc 120
gacacggttgg gatttggggt cttgagctgtc tgggtctcct cagggcaca 176

Met Gln

1

tgc cgg ctc ccg cgg cgc ctc gct gct gga gca ctc ctc acc ctc ctg tgc 224
Cys Arg Leu Pro Arg Gly Leu Ala Gly Ala Leu Leu Thr Leu Leu Cys
5 10 15

atg ggg ctc ctg tgt ctg cgg tac cac tgt gt ctc ccg cag cgg 272
Met Gly Leu Leu Cys Leu Arg Tyr His Leu Asn Leu Ser Pro Gln Arg
20 25 30

Gln Pro Asn Pro Gly Pro Pro Lys
35 40 45 50

gta caa ggg acc ccc gag ctg agc cag ccc acc ccc ggg ccc cct aag 320
Val Glu Gly Thr Pro Glu Leu Ser Gln Pro Asn Pro Gly Pro Pro Lys

35 40 45 50

ctg ccc tgg ctc cgc ctg gag ctg ctt ctc acc acg acc ccg ggt ctc 368
Leu Glu Leu His Val Phe Ile Ala Val Lys Thr Thr Arg Ala Phe
55 60 65

His Leu Arg Leu Leu Leu Leu Leu Leu Asp Thr Trp Val Ser Arg Thr
70 75 80

agg gaa ctg acu ttc tgc ttc acc gac agc cca gac aaa ggc ctt cag 416
Arg Glu Leu Thr Phe Val Phe Thr Asp Ser Asp Lys Gly Leu Gln
85 90 95

gag aga ctg ggg ctc cac ctt ggt gtc acc aac tgc tcc ggc gag cac 464
Glu Arg Leu Gly Ser His Leu Val Thr Asn Cys Ser Ala Glu His
100 105 110

agc cca gct ctg ttc aag atg gct gct gag ttc gac acc ttc 512
Ser His Pro Ala Leu Ser Cys Lys Met Ala Ala Glu Phe Asp Thr Phe
115 120 125 130

agg cag tgg ggg ctg agg tgg ttc tgc cag gat gac gat gac aac tgc 560
Leu Ala Ser Gly Leu Arg Trp Phe Cys His Val Asp Asp Asn Tyr
135 140 145

ttg gcc agt ggg ctt agg tgg ttc tgc cat gtt gac gat gac aac tat
gtg aac cca agg gcg ctg ctg cag ctt ctg aga gcc ttc cgg ctg gcc
Val Asn Pro Arg Ala Leu Leu Gln Leu Leu Arg Ala Phe Pro Leu Ala
150 155 160

cgc gac gtc tat gtg gga agg ccc agc ctg aac cgg ccc atc cat gcc
Arg Asp Val Tyr Val Gly Arg Pro Ser Leu Asn Arg Pro Ile His Ala
165 170 175

tca gag cca cag ccc cac aac cgc acg agg ctg gta cag ttc tgg ttt
Ser Glu Pro Glu Pro His Asn Arg Thr Arg Leu Val Glu Phe Trp Phe
180 185 190

gcc act ggg ggt gct ggc ttc tgc atc aat cgc aaa ctg gct ttg aag
Ala Thr Gly Gly Gly Cys Thr Met Asp Arg Cys Leu Leu Lys
195 200 205 210

atg gct ccg tgg gcc agt gcc tcc cgt ttc atg gac aca tct gct ctc
Met Ala Pro Trp Ala Ser Gly Arg Pro Asp Ser Arg Pro Ser Thr Ser Ala
215 220 225

atc cgg ctg cct gat gac tgc acc atg ggc tat atc att gag tgc aag
Ile Arg Leu Pro Asp Asp Cys Thr Met Gly Tyr Ile Ile Glu Cys Lys
230 235 240

crc ggc ggc cgc ctg cag ccc agc ccc ctc ttt cac ctc cac ctg gag
Leu Gly Gly Arg Leu Gln Pro Ser Cys His Ser His Leu Glu
245 250 255

acc ctg cag ctg ctg agg act gca cag ctc cca gaa cag gtc acc ctc
Thr Leu Gln Leu Leu Arg Ala Leu Gln Leu Pro Glu Val Thr Leu
260 265 270

agc tac ggt gtc ttt gag ggg aaa ctc aac gtc att aag cta cag gcc
Ser Tyr Gly Val Phe Glu Gly Lys Leu Asn Val Ile Lys Leu Gln Gly
275 280 285 290

ccc ttc ccc cgg gac gac ccc tcc aga ttt cgc tcc ctc cat tgt
Pro Phe Ser Pro Glu Asp Pro Ser Arg Phe Arg Ser Leu His Cys
295 300 305

crc ctc tat cca gat cca ccc tgt ctc cag ctg ggt gcc cga tga
Leu Leu Tyr Pro Asp Thr Pro Trp Cys Pro Glu Leu Gly Ala Arg
310 315 320

atcctgaact gctgggccaa ggtggccag aagtccttg gttgtgccttg gctccaaag

1196

1256

1316

1376

1436

1496

1556

1616

1736

1796

1916

1976

2036

2042
<210> 12
<211> 321
<212> PRT
<213> Homo sapiens

<400> 12
Met Gln Cys Arg Leu Pro Arg Gly Leu Ala Gly Ala Leu Leu Thr Leu
1      5        10       15
Leu Cys Met Gly Leu Leu Cys Leu Arg Tyr His Leu Asn Leu Ser Pro
20     25       30        35
Gln Arg Val Glu Gly Thr Pro Glu Leu Ser Gln Pro Asn Pro Gly Pro
35     40       45        50
Pro Lys Leu Gln Leu His Asp Val Phe Ile Ala Val Lys Thr Thr Arg
50     55       60        65
Ala Phe His Arg Leu Arg Leu Glu Leu Leu Leu Asp Thr Trp Val Ser
70     75       80        85
Arg Thr Arg Glu Leu Thr Phe Val Phe Thr Asp Ser Pro Asp Lys Gly
85     90       95        100
Leu Gln Glu Arg Leu Gly Ser His Leu Val Val Thr Asn Cys Ser Ala
100    105      110       115
Glu His Ser His Pro Ala Leu Ser Cys Lys Met Ala Ala Ala Glu Phe Asp
115    120      125       130
Thr Phe Leu Ala Ser Gly Leu Arg Trp Phe Cys His Val Asp Asp Asp
130    135      140       145
Asn Tyr Val Asn Pro Arg Ala Leu Leu Gln Leu Leu Arg Ala Phe Pro
145    150      155       160
Leu Ala Arg Asp Val Tyr Val Gly Arg Pro Ser Leu Asn Arg Pro Ile
165    170      175       180
His Ala Ser Glu Pro Gln Pro His Asn Arg Thr Arg Leu Val Gln Phe
180    185      190       195
Trp Phe Ala Thr Gly Gly Ala Gly Phe Cys Ile Asn Arg Lys Leu Ala
195    200      205       210
Leu Lys Met Ala Pro Trp Ala Ser Gly Ser Arg Phe Met Asp Thr Ser
210    215      220       225
Ala Leu Ile Arg Leu Pro Asp Asp Cys Thr Met Gly Tyr Ile Ile Glu
225    230      235       240
Cys Lys Leu Gly Gly Arg Leu Gln Pro Ser Pro Leu Phe His Ser His
245    250      255       260
Leu Glu Thr Leu Gln Leu Leu Arg Thr Ala Gln Leu Pro Glu Gln Val
260    265      270       275
Thr Leu Ser Tyr Gly Val Phe Gly Lys Leu Asn Val Ile Lys Leu
275    280      285       290
Gln Gly Pro Phe Ser Pro Glu Glu Asp Pro Ser Arg Phe Arg Ser Leu
290    295      300       305
His Cys Leu Leu Tyr Pro Asp Thr Pro Trp Cys Pro Gln Leu Gly Ala
305    310      315       320
Arg

<210> 13
<211> 1740
<212> DNA
<213> Homo sapiens

<400> 13
gccgccccgtg gccgccccgtg tactgcatacg gcagcgcggcg gcggcgccgg gcggcgcttcg gcggccggcc 60
gaggtccccgc gtcggatagtc cttcgcctac gcgtctgcgcct ctgtctgtctgcct 120
ccccctcgttc gccccgcccc gcggcgcccc gcggcgcttc gcggcgatagtc 180
ccccctcgttc 240
gtgctacatc catttgtctc agatgcaaac agaaggggac ccagaagccct tatgctctc caa 300
gaattgtaaaa gaaacagctg gacaaaaaa tcgtaaagac tgagatagga gttctctc ttctc 360
gctctcaaca tccaaacctt ataaaactcta aagagatatt tgaacccccct acagaaatac 420
gtcgggttctc agaactgtgc acaggaggg aactgtttgtg taggatttgtg gaaagggacag 480
attacagttga gcagagatgc gcagagccgcg ttaaaactat cctaggagcgt gtctttctc 540
tacatgaaaag tggaggttggc cattgctgtat ctaaaccagca gaatctctctt tatgcaacctc 600
cagcccccaag taacccattact aacctagctg atttgtgagt cttccaaactt agctataaggac 660
aagctgtcctt gaagacagta tgtggaaccc cagggtactgc gcacacctgca aaatcttagag 720
gtgtggctcatt tggagctgag atgactctgt ctgcgtttagg ataataacctc tatactctcatt 780
tttgtgatt tgaaccacttc tatgtagaaag gagcgcatca gtctctgcttc agggataactc 840
tgaaagaattga ttaaactctt acatccccct gttgtagatg aattattcctta aagtccaaagag 900
actctgtgcaag aaaaattataaatt gttttggtatc caaaaagaaac gcgtgactaca ttctcaacgctc 960
tctagctgccgt gttggctacaagt gtaaagacag ccaatcattgtt acacatggtgtt acctgtaaag 1020
agaagctccta gagatctaatg ggcgggctgta agctattaaag acgcgtgtaag gctgtggtgta 1080
ccctctctcct cgtgggaaagtc gcacagggag gcctcagacca catccagggag aacctgcaag 1140
ctagagcgag tcccttcctct atccaagactg gcacagggag atcagaaacgct atctcagagag 1200
agagaaatg tcaagcgccag gggccccaaag cccagttaaa ggggccccag gcgagccgtta 1260
tgaaagctgcag acgcctcaggg aaggttaaag gcctagatat aaatctctgtaa ggcgccccaaag 1320
aagaagctgag cagcgagcttg aagattgtgag atagagtctg gcagagagccg taagagtgagc 1380
gctctccagga ggaagacttgg aagactgttgag agagcgagcg acgtccccag gacgccccaaag 1440
gaagctctgtgc tgtgctgttaa gaagttccac acagagagtct gacgctcagg gactaactaag 1500
cagcgctctct cagatctctgg agcgtccccag cccatatatta gcgcctttgg ccttcgagcaag 1560
gaagattgtgg acagccatag atgtactata tgtatccctg tcctctgtgctc aaaaacactctg 1620
acatatatac cattggttaag ttctaaactc gttcgatgtcg actcgcatttat gaaaaatataa 1680
gtctgactta gtcagccgatg tgtattactta ataccgagctc agttaaacttcg caagtttaac 1740

<210> 14
<211> 1740
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (78)...(1499)

<400> 14

ggggccccgct gcgggcccctg tgtctcgtgcgc gcagcgccggcc gcgcgcctccct 60
mgagttcccgc tgcagaaag atgtctc aaa gttac gcgtgcc tcc tgc tcc tgc 110
met leu lys val thr val pro ser cys ser ala

<510>

tcg ttc tgc tct tct gcg acc gcc agt gcg gcc ccc ggg acc gcg acc

<550>

ser ser cys ser ser thr ala ala ala pro gly thr ala ser

<590>

ctc gtc ccc gat tac tgg atc gcg gcc tcc aac agg gat gcg ctc ggc

<630>

leu val pro asp tyr trp ile asp gly ser asn arg asp ala leu ser

<670>

gat ttc ttc gag gtg gag tgg cag gag gcg ggg acc caa tca ttc aaag

<710>

asp phe phe glu glu ser gly gly arg gly ala thr ser ile

<750>

<790>

<830>

<870>

<910>

<950>

<990>

20
Val Leu Leu Arg Leu Ser His Pro Asn Ile Ile Lys Leu Lys Glu Ile
95 100 105

Phe Glu Thr Pro Thr Glu Ile Ser Leu Val Leu Glu Leu Val Thr Gly
110 115 120

Gly Leu Phe Asp Arg Ile Val Lys Gly Tyr Tyr Ser Glu Arg
125 130 135

Asp Ala Ala Asp Ala Val Lys Glu Ile Leu Glu Ala Val Ala Tyr Leu
140 145 150 155

His Glu Asn Gly Ile Val His Arg Leu Lys Pro Asn Leu Arg
160 165 170

Tyr Ala Thr Pro Ala Pro Leu Ala Pro Leu Lys Ile Ala Tyr Phe
175 180 185

Leu Ser Lys Ile Val Glu His Gln Val Leu Met Lys Thr Cys Gly
190 195 200

Thr Pro Gly Tyr Cys Ala Pro Glu Ile Leu Arg Gly Cys Ala Tyr Gln
205 210 215

Pro Glu Val Asp Met Trp Ser Val Gly Ile Ile Thr Ile Leu Leu
220 225 230 235

Cys Gly Phe Glu Pro Phe Tyr Asp Glu Arg Gly Asp Glu Phe Met Phe
240 245 250

Arg Arg Ile Leu Asn Cys Glu Tyr Tyr Phe Ile Ser Pro Trp Trp Asp
255 260 265

Val Ser Leu Asn Ala Lys Asp Val Arg Lys Leu Ile Leu Leu
270 275 280

Asp Pro Lys Arg Leu Thr Thr Phe Glu Ala Leu Glu His Pro Trp
285 290 295

Val Thr Gly Ala Ala Asn Phe Val His Met Asp Thr Ala Glu Lys
300 305 310 315

Lys Leu Glu Glu Phe Asn Ala Arg Arg Lys Leu Lys Ala Ala Val Lys
320 325 330
gct gtg gtg gcc tct tcc cgc tgt gga agt gcc agc agc agc cat ggc
Ala Val Val Ala Ser Ser Arg Leu Gly Ser Ser Ser Ser Ser His Gly
335 340 345

agc atc cag gag agc cac aag gct agc cga gac cct tct cca atc caa
Ser Ile Gln Glu Ser His Lys Ala Ser Arg Asp Pro Ser Pro Ile Gln
350 355 360

gat ggc aac gag gac atg aaa gct att cca gaa gga gag aaa att cca
Asp Gly Asn Glu Asp Met Lys Ala Ile Pro Glu Gly Glu Lys Ile Gln
370 375

ggc gat ggg gcc caa gcc gca gtt aag ggg gca cag gct gag ctg atg
Gly Asp Gly Ala Gln Ala Ala Val Lys Gly Ala Gln Ala Glu Leu Met
380 385 390 395

aag gtg caa gcc tta gag aaa gtt aaa ggt gca gat ata aat gct gaa
Lys Val Gln Ala Leu Glu Lys Val Lys Gly Ala Asp Ile Asn Ala Glu
400 405 410

gag gcc ccc aaa atg gtg ccc aag gca gtg gag gat ggg ata aag gtg
Glu Ala Pro Lys Met Val Pro Lys Ala Val Glu Asp Gly Ile Lys Val
415 420 425

gct gac ctg gaa cta gag gag ggc cta gca gag gag aag ctg aag act
Ala Asp Leu Glu Leu Glu Gly Leu Ala Glu Gly Lys Leu Lys Thr
430 435 440

gtg gag gag gca gca gct ccc aga gaa ggg caa gga agc tct gct gtg
Val Glu Ala Ala Ala Ala Pro Arg Glu Gly Gin Gly Ser Ser Ser Ala Val
445 450 455

gtt cca gaa gtt cca cag caa gat gtg atc ctg cca gag tac taa
Gly Phe Glu Val Pro Glu Gin Asp Val Ile Leu Pro Glu Tyr *
460 465 470

acagcttcttc tcagatctgg aagcccaaca ccggcatattt atgtactttg tccttcagca 1559
agaagaaggg ggaagcagtg taagtctact aagtgtcatcg ttttttgagg tggaaaaaca 1619
tacatatata ccaagttcatc atctctacat cagtcatgtg gctcttcttt tagaaat c 1739
tagatcctttt atggcatgt cattctg ataaatgtt actgtcctaa tgcagtttaa 1740

<210> 15
<211> 473
<212> PRT
<213> Homo sapiens

<400> 15
Met Leu Lys Val Thr Val Pro Ser Cys Ser Ala Ser Ser Ser Cys Ser Ser 1
5 10 15
Val Thr Ala Ser Ala Ala Pro Gly Thr Ala Ser Leu Val Pro Asp Tyr
20 25 30
Trp Ile Asp Gly Ser Asn Arg Asp Ala Leu Ser Asp Phe Phe Glu Val
35 40 45
Glu Ser Glu Leu Gly Arg Gly Ala Thr Ser Ile Val Tyr Arg Cys Lys
50 55 60
Gln Lys Gly Thr Gln Lys Pro Tyr Ala Leu Lys Val Leu Lys Lys Thr
65 70 75 80
Val Asp Lys Lys Ile Val Arg Thr Glu Ile Gly Val Leu Leu Arg Leu
85 90 95
Ser His Pro Asn Ile Ile Lys Leu Lys Glu Ile Phe Glu Thr Pro Thr
100  105  110
Glu Ile Ser Leu Val Leu Glu Leu Val Thr Gly Gly Glu Leu Phe Asp
115  120  125
Arg Ile Val Glu Lys Gly Tyr Tyr Ser Glu Arg Asp Ala Ala Asp Ala
130  135  140
Val Lys Gln Ile Leu Glu Ala Val Ala Tyr Leu His Glu Asn Gly Ile
145  150  155  160
Val His Arg Asp Leu Lys Pro Glu Asn Leu Leu Tyr Ala Thr Pro Ala
165  170  175
Pro Asp Ala Pro Leu Lys Ile Ala Asp Phe Gly Leu Ser Lys Ile Val
180  185  190
Glu His Gln Val Leu Met Lys Thr Val Cys Gly Thr Pro Gly Tyr Cys
195  200  205
Ala Pro Glu Ile Leu Arg Gly Cys Ala Tyr Gly Pro Glu Val Asp Met
210  215  220
Trp Ser Val Gly Ile Thr Tyr Ile Leu Leu Cys Gly Phe Glu Pro
225  230  235  240
Phe Tyr Asp Glu Arg Gly Asp Gln Phe Met Phe Arg Arg Ile Leu Asn
245  250  255
Cys Glu Tyr Tyr Phe Ile Ser Pro Trp Trp Asp Glu Val Ser Leu Asn
260  265  270
Ala Lys Asp Leu Val Arg Lys Leu Ile Leu Val Leu Asp Pro Lys Lys Arg
275  280  285
Leu Thr Thr Phe Glu Ala Leu Gln His Pro Trp Val Thr Gly Lys Ala
290  295  300
Ala Asn Phe Val His Met Asp Thr Ala Gln Lys Leu Gln Glu Phe
305  310  315  320
Asn Ala Arg Arg Asp Leu Gly Ser Ala Ser Ser Ser His Gly Ser Ile Glu Glu Ser
325  330  335  340
Ser Arg Leu Gly Ser Ala Ser Ser Ser His Gly Ser Ile Glu Glu Ser
345  350
His Lys Ala Ser Arg Asp Pro Ser Pro Ile Gln Asp Gly Asn Glu Asp
355  360  365
Met Lys Ala Ile Pro Glu Gly Lys Ile Gln Gly Asp Gly Ala Gln
370  375  380
Ala Ala Val Lys Gly Ala Glu Leu Met Lys Val Gln Ala Leu
385  390  395  400
Glu Lys Val Lys Gly Ala Asp Ile Asn Ala Glu Glu Ala Pro Lys Met
405  410  415
Val Pro Lys Ala Val Glu Asp Gly Ile Lys Val Ala Asp Leu Glu Leu
420  425  430
Glu Gly Leu Ala Glu Gly Lys Leu Lys Thr Val Glu Glu Ala Ala
435  440  445
Ala Pro Arg Glu Gly Gly Ser Ser Ala Val Gly Phe Glu Val Pro
450  455  460
Gln Gln Asp Val Ile Leu Pro Glu Tyr
465  470

<210> 16
<211> 2291
<212> DNA
<213> Homo sapiens

<400> 16
cttttcgtgc tgatcttgga ggccagagtgt cgccgtctca tcatattccag gttcatctac 60
cagcgttggc caaagacagt ctgctctgta tgtagctctt gcactctgaga aactattttc 120
tagccacagg cccttggaaa tgtgagcagt ctttgacagc ccactgaggcc ttatattgga 180
cggtcacaag ggacgcgggct ctagaaaggtg agccggaccggt gatgcagaag ttgtagatcc 240
taggtttgaca gtagcggtgtga ccctctcagt cgcggcctcac ggagaaaggg acagctctgtg 300
gttggtgtcag tggggtgccggt cttgctgcagt cccaaagcactttgacagc 360
ggttccacgg atgtcactgt gcttgaggctt tccagccacaa tccagggcccg tgtgcagagt 420
gtgaaccttg gagacgccacc cttcaggctgt ggagccacct ggtatccatgg ctcccatgg 480
aacccctatct atcatcctagc aagaagccacag ggccctcttg ggagagacac cgtaggggaa 540
cggtcgcctgc gacagtgccttg ctttttttttt ctaagacatt tggaggaagc ctagcagagc 580
ccagagctgg ccaagctgctt gaaactgctgg atgcacctgg ctatgctgc gag ctggtgagcc 640
tgttggactg ccctccccct cgggctctcttggt gcgcggtgctt ggtggagtgc ggtgggtgag 700
gatgcctgcc ggcctctcttgcc gctatccctcc tcttggtcttg cttgaggtgc cttgagggctg 760
ggctggagcc ctcctctcttg ggcagggact ggaagctgcag ctgcagcagc ggtgggagcc 820
aatacagaga ggcggagagc ggtggagcttg ggtggaggttg tggagctgtg gttggctgag 880
catctgctcg cgcagttctgc gcctcttgcttc gcctcttgctt ccaggtgctgc ccctctcttg 940
atctgctctg ccctctcttg gtcctctcttg cctctctcttg gtcctctcttg ccctctctcttg 1000
aatactgcag tgtgtgtgtg tggaggttttt tggaggtgttt gtcctctcttg ccctctctcttg 1060
agaaaagatg tttttttttt ttgagatggtg cccctctcttg tgcctctcttg cttgaggtgc 1120
tagctttttt ttttttttttt gtccctctcttg cctctctcttg gtcctctcttg 1180
tcgtgtgtgc cggggagggc ggtggaggtttgtg ggtggaggtttgtg gtcctctcttg ccctctctcttg 1240
agaaaagatg tttttttttt ttgagatggtg cccctctcttg tgcctctcttg cttgaggtgc 1300
agaaaagatg tttttttttt ttgagatggtg cccctctcttg tgcctctcttg cttgaggtgc 1360
agaaaagatg tttttttttt ttgagatggtg cccctctcttg tgcctctcttg cttgaggtgc 1420
agaaaagatg tttttttttt ttgagatggtg cccctctcttg tgcctctcttg cttgaggtgc 1480
agaaaagatg tttttttttt ttgagatggtg cccctctcttg tgcctctcttg cttgaggtgc 1540
agaaaagatg tttttttttt ttgagatggtg cccctctcttg tgcctctcttg cttgaggtgc 1600
agaaaagatg tttttttttt ttgagatggtg cccctctcttg tgcctctcttg cttgaggtgc 1660
agaaaagatg tttttttttt ttgagatggtg cccctctcttg tgcctctcttg cttgaggtgc 1720
agaaaagatg tttttttttt ttgagatggtg cccctctcttg tgcctctcttg cttgaggtgc 1780
agaaaagatg tttttttttt ttgagatggtg cccctctcttg tgcctctcttg cttgaggtgc 1840
agaaaagatg tttttttttt ttgagatggtg cccctctcttg tgcctctcttg cttgaggtgc 1900
agaaaagatg tttttttttt ttgagatggtg cccctctcttg tgcctctcttg cttgaggtgc 1960
agaaaagatg tttttttttt ttgagatggtg cccctctcttg tgcctctcttg cttgaggtgc 2020
agaaaagatg tttttttttt ttgagatggtg cccctctcttg tgcctctcttg cttgaggtgc 2080
agaaaagatg tttttttttt ttgagatggtg cccctctcttg tgcctctcttg cttgaggtgc 2140
agaaaagatg tttttttttt ttgagatggtg cccctctcttg tgcctctcttg cttgaggtgc 2200
agaaaagatg tttttttttt ttgagatggtg cccctctcttg tgcctctcttg cttgaggtgc 2260
agaaaagatg tttttttttt ttgagatggtg cccctctcttg tgcctctcttg cttgaggtgc 2320
agaaaagatg tttttttttt ttgagatggtg cccctctcttg tgcctctcttg cttgaggtgc 2380

<210> 17
<211> 2291
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (223)...(1890)

<400> 17
catttctgtgc tgatcctggga ggccagacttg ccctctcttg tccatcccaag tccctcctcc 60
cagctgtgcct caaagacagt cgctctcttg cccacatgtc tgcctctcttg gcctctctcttg 120
atgcctcggag ccctctctcttt tggagctggt cccacatgtc tgcctctcttg gcctctctcttg 180
cgctccacag ggccagacttg ccctctctcttg tccatcccaag tccctcctcc 240

Glu Ser Ser Gly Asp Ser Ala Asp Asp Pro Leu Ser Arg Gly Leu Arg
5 10 15 20

Arg Arg Gly Gln Pro Arg Val Val Val Ile Gly Ala Gly Leu Ala Gly
25 30 35

Leu Ala Ala Ala Ala Ala Ala Ala Ala Ala Ala
378

Met Gln Ser Cys
WO 03/051180

PCT/US02/40194

```
40  45  50

426
gtg ctt gag gct tcc agc cac atc gga ggc cgt gtg cag aet gtg aaa
Val Leu Ala Ser Ser His Ile Gly Gly Arg Val Gln Ser Val Lys
55  60  65

cct gga cac gcc acc ttt gag ctg gga gcc acc tgg atc cat ggc tcc
Leu Gly His Ala Thr Phe Glu Leu Gly Ala Thr Trp Ile His Gly Ser
70  75  80

cat ggg aac cct atc tat cat cta gca gaa gcc aac ggc ctc ctg gaa
His Gly Asn Pro Ile Tyr His Leu Ala Glu Ala Asn Gly Leu Leu Glu
85  90  95  100

gag aca acc gat ggg gaa cgc aac gtc ggc ctc acc atg gtc tat tcc
Glu Thr Thr Asp Gly Glu Arg Ser Val Gly Arg Ile Ser Leu Tyr Ser
105 110 115

618
aag aat ggc tgt gcc tgc tac ctt acc aac cac gcc cgc agg atc ccc
Lys Asn Gly Val Ala Cys Tyr Leu Thr Asn His Gly Arg Arg Ile Pro
120 125 130

666
aag gac tgt gtt gag gaa tcc agc gat tta tac aac gag gtc tat aac
Lys Asp Val Val Glu Glu Phe Ser Asp Leu Tyr Asn Glu Val Tyr Asn
135 140 145

ttg acc cag gag tcc tcc cgg cac gat aaa cca gtc aat gct gaa aat
Leu Thr Glu Glu Phe Phe Arg His Asp Lys Pro Val Asn Ala Glu Ser
150 155 160

762
caa aat agc tgt ggg tgt tcc acc cga gag gac ggt cgt aac cgc atc
Gln Asn Ser Val Gly Val Phe Thr Arg Glu Glu Val Arg Asn Arg Ile
165 170 175 180

810
agg aat gac cct gac gag cac cca gag gct acc aag cgc tgt aag ctc gcc
Arg Asn Asp Pro Asp Pro Glu Ala Thr Lys Arg Leu Lys Leu Ala
185 190 195

858
atg atc cag cag tac ctg aag tgt gag agc tgt gag agc agc tca cac
Met Ile Gln Gln Tyr Leu Lys Val Glu Ser Cys Glu Ser Ser Ser His
200 205 210

906
agc atg gac gag gtg tcc ctg agc gcc ttc ggg gag tgg acc gag atc
Ser Met Asp Glu Val Ser Leu Ser Ala Phe Gly Glu Trp Thr Glu Ile
215 220 225

954
ccc ggc gct cac cac atc atc ccc tgg ctc att cgg gtt gtg gag
Pro Gly Ala His His Ile Pro Ser Gly Phe Met Arg Val Val Glu
230 235 240

cgg gtg cac gag gct gtc cac atc atc ccc tcc ggc ttc att cgg gtt gtg gag
Leu Leu Ala Glu Gly Ile Pro Ala His Val Ile Gln Leu Gly Lys Pro
245 250 255 260

1050
gtc cgc tgc att cac tgg gac cac gcc tca gcc cgc ccc aga ggc cct
Val Arg Cys Ile His Trp Asp Gln Ala Ser Ala Arg Pro Arg Gly Pro
265 270 275

gag att gag ccc cgg ggt gag ggc gac cac aat cac gac act cgg gag
Glu Ile Glu Pro Arg Gly Glu Gly Asp His Asn His Asp Thr Gly Glu
280 285 290

25
```
gtg ggc cag  cag ggt gga gag  gag ccc cgg ggg ggc agg tgg gat  gag gat
Gly  Gly  Gln  Gly  Gly  Glu  Glu  Pro  Arg  Gly  Gly  Arg  Trp  Asp  Glu  Asp
  295                      300                      305

gag cag tgg  tgc  tgc gac tgc gag  gac ttg gag  ctg  atc  cgc
Glu  Gln  Trp  Ser  Val  Val  Glu  Cys  Glu  Asp  Cys  Glu  Leu  Ile  Pro
  310                      315                      320

gcg gac cat  tgc  acc  gtt gtc  cta  ggt  tgt  aag  agg  cag
Ala  Asp  His  Val  Ile  Val  Thr  Val  Ser  Leu  Gly  Val  Leu  Lys  Arg  Gln
  325                      330                      335                      340

tac acc  agt  ttc  ttc  cgg  cca  ggc  ctg  ccc  aca  gag  aag  gtg  gct  gcc
Tyr  Thr  Ser  Phe  Phe  Arg  Pro  Gly  Leu  Pro  Thr  Glu  Val  Ala  Ala
  345                      350                      355

atc cac  cgc  ctg  att  ggc  acc  acc  gac  aag  atc  ttt  ctg  gaa  ttc
Ile  His  Arg  Leu  Gly  Ile  Thr  Thr  Asp  Lys  Ile  Phe  Leu  Glu  Phe
  360                      365                      370

gag  gag  ccc  ttc  tgt  ggc  cct  gag  tgc  aac  agc  cta  cag  ttt  gtg  tgg
Glu  Glu  Pro  Phe  Trp  Gly  Pro  Asn  Ser  Leu  Gln  Phe  Val  Trp
  375                      380                      385

gag  gag  gca  gag  agc  cac  acc  ctc  acc  tac  cca  cct  gag  ctc  tgg
Glu  Asp  Glu  Ala  Glu  Ser  His  Thr  Leu  Thr  Tyr  Pro  Glu  Leu  Trp
  390                      395                      400

tac  cgc  aag  atc  tgc  ggc  ttt  gat  gtc  tct  tac  ccg  cct  gag  cgcc  tac
Tyr  Arg  Lys  Ile  Cys  Gly  Phe  Asp  Val  Leu  Tyr  Pro  Glu  Arg  Tyr
  405                      410                      415                      420

ggc  cat  tgt  ctg  agc  ggc  tgg  atc  tgt  ggc  gac  ggc  ctc  gtc  atg
Gly  His  Val  Leu  Ser  Gly  Trp  Ile  Cys  Gly  Glu  Glu  Ala  Val  Met
  425                      430                      435

gag  aag  tgt  gat  gac  gag  gca  tgt  ggc  gag  atc  tgt  acg  gag  atg  ctg
Glu  Lys  Cys  Asp  Glu  Ala  Val  Ala  Glu  Ile  Cys  Thr  Glu  Met  Leu
  440                      445                      450

cgt  cag  ttc  aca  ggg  aac  ccc  aac  att  cca  aaa  cct  cgg  cga  atc  tgg
Arg  Gln  Phe  Thr  Gly  Asn  Pro  Asn  Ile  Pro  Lys  Pro  Arg  Arg  Ile  Leu
  455                      460                      465

cgc  tgc  gcc  tgg  ggc  agc  aac  cct  tac  ttc  cgc  ggc  tcc  tat  tca  tac
Arg  Ser  Ala  Trp  Gly  Ser  Asn  Pro  Tyr  Phe  Arg  Gly  Ser  Tyr
  470                      475                      480

cgc  tgg  tgc  agc  ggc  cgg  gaa  cgg  gat  ggc  tgt  aga  cgg  gag  atc  tgt
Thr  Gln  Val  Gly  Ser  Ser  Gly  Ala  Asp  Val  Glu  Leu  Ala  Lys  Pro
  485                      490                      495                      500

cgg  tac  cca  gag  agc  tca  aag  cga  ggc  gcc  atg  cgg  cct  tgt  ttt
Leu  Pro  Tyr  Thr  Ser  Lys  Thr  Ala  Pro  Met  Glu  Val  Leu  Phe
  505                      510                      515

tcc  ggt  gag  ccc  acc  cac  cgc  aag  tac  ttc  acc  acc  cac  ggt  gct
Ser  Gly  Glu  Ala  Thr  His  Arg  Lys  Tyr  Tyr  Ser  Thr  Thr  His  Gly  Ala
  520                      525                      530
ctg ctg tcc ggc cag cgt gag gct gcc cgc ctc att gag atg tac cga 1866
Leu Leu Ser Gly Gln Arg Glu Ala Ala Arg Leu Ile Glu Met Tyr Arg
535 540 545

gac ctc cag cag ggg acc tga gggctgtcct cgctgcagag aagagccact 1920
Asp Leu Phe Gln Gln Gly Thr *
550 555

aactcgtgac ctcagcctg ccctcttgctg ccgtgtgcctc ctgcccttctc gatcctctgt 1980
agaaagatt ttatcttct ctgagcttag cgcccttgac tgccttcaga ctggccttgct 2040
tagctttctc tttttctcag gctggggctg gacaggttg ggccgtgtag ttacctctgtg 2100
ctgtatcccg tgccccccact tgctttactcc ctgctctgcc ttgtttattga aagttgctctc 2160
aatctcttgcc attttgggt ataataaaaag gtcctctccc ctgccaataaa aaaaaaaaaaa 2220
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 2280
aaaaaaaaaa c
2291

<210> 18
<211> 555
<212> PRT
<213> Homo sapiens

<400> 18
Met Gln Ser Cys Glu Ser Ser Gly Asp Ser Ala Asp Asp Pro Leu Ser 1
      5
Arg Gly Leu Arg Arg Arg Gly Gln Pro Arg Val Val Val Ile Gly Ala 20
      25
Gly Leu Ala Gly Leu Ala Ala Ala Lys Ala Leu Leu Gln Gly Phe 35
      40
Thr Asp Val Thr Val Leu Ala Ser Ser Ser His Ile Gly Gly Arg Val 50
      55
Gln Ser Val Lys Leu Gly His Ala Thr Phe Glu Leu Gly Ala Thr Trp 65
      75
Ile His Gly Ser His Gly Asn Pro Ile Tyr His Leu Ala Glu Ala Asn 85
      90
Gly Leu Leu Glu Thr Thr Asp Gly Glu Arg Ser Val Gly Arg Ile 100
     105
Ser Leu Tyr Ser Lys Asn Gly Val Ala Cys Tyr Leu Thr Asn His Gly 115
     120
Arg Arg Ile Pro Lys Asp Val Val Glu Phe Ser Asp Leu Tyr Asn 130
     135
Glu Val Tyr Asn Leu Thr Glu Phe Phe Arg His Asp Lys Pro Val 145
     150
Asn Ala Glu Ser Gln Asn Ser Val Gly Val Phe Thr Arg Glu Glu Val 165
     170
Arg Asn Arg Ile Arg Asn Asp Pro Asp Asp Pro Glu Ala Thr Lys Arg 180
     185
Leu Lys Leu Ala Met Ile Gln Tyr Leu Lys Val Glu Ser Cys Glu 195
     200
Ser Ser Ser His Ser Met Asp Glu Val Ser Leu Ser Ala Phe Gly Glu 210
     215
Trp Thr Glu Ile Pro Gly Ala His His Ile Ile Pro Ser Gly Phe Met 225
     230
Arg Val Val Glu Leu Ala Glu Gly Ile Pro Ala His Val Ile Gln 245
     250
Leu Gly Lys Pro Val Arg Cys Ile His Trp Asp Gln Ala Ser Ala Arg 260
     265
Pro Arg Gly Pro Glu Ile Glu Pro Arg Gly Glu Gly Asp His Asn His 275
     280
Asp Thr Gly Glu Gly Gly Gly Gly Gly Gly Pro Arg Gly Arg Arg 290
     295
Trp Asp Glu Asp Glu Glu Trp Ser Val Val Val Glu Cys Glu Asp Cys

27
cttgtcctctc ggcattctcg ctgcctctgt cttccccctct tgtgtcctctc gtttctcatt 1380
ccttctctatg ttcgctctgt gtctctcaat taacctcag tcaactgtctg attctactgg 1440
ggtggtgggt cagaccccatc ttcagccacc agatgttgcg ctacacccctg gacaagtgcg 1500
tgccccggttgc tgcggtccgttg tgtcagacgtg cttggagacat agttggggtgt 1620
ggacacgaga tccgtagctg cttcagccgt gctgagggag acacagtcctg cccacagcag 1680
tgatgccacc cccatatcatc ggttaagcatc acactgtatg gacatcagcag 1740
ggacagatgt tgcgctctatg catttttttc aggccgctcc gcagcagcag ccaatgcttc 1800
ggacgcttg ccgtgccttc gggtatagct cttttccaca gcctgtgcag cggaaaacag 1860
acatagacc ccccccacag cttggacctg cttggtgggg ccagagtggg cccagcccccac 1920
atttttaagag cggcctccat ttttttctg tcggcttgctg ttcagacgctg cttgtctaaa 2040
tcccaaacat cggcccgagt gcgtgctggtg cgccctggaa tcggctgggt ttagcctaaa 2100
tggtcttatt gcgtgctggtg cttggtggcc tcaacacctc tctcttgatt 2160
tgttcttta ctgctgattg ggtgcatact ataatacaac tttcaaaagga tttcagcgtg 2220
agtggcacc cctccagcctc atgggccttg ccgaccagcg agttgggtag agttgggtag 2280
tgaccctccg aggtggcacc cagttgggag gaactgtatg cccatgacag cccagccccc 2340
cacaaacatc cccatgattg ggtgtgttctg ggtcagcagc gtctgctcag cttggtgggg 2400
agccagctgg ggacagatgt gcgtgctggtg cgggggctgg tcggctgggt tttagcctaaa 2460
cagctgctgc cccgctgcag cgggggctgg tcggctgggt tttagcctaaa 2520
aaaaagaaaa aagagagagag aagagagagag aagagagagag aagagagagag aagagagagag aagagagagag aagagagagag aaaaaaa 2640
a 2649

<210> 20
<211> 2649
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (120)...(1061)

<400> 20
tatagtgatt ggtttacagcc cccgccccggc tagtcggaaac tgaacttacct gcccagcgctg 60
cggccagcgcc gcggcagcctc gcggcgccg ccaggccttc gctggctgcc 119
atg cggagg aag ggc gcg ctt cttg gcc gcg gcc gcg tgc gac gcc gcg gcg tacgccgc 167
Met Arg Arg Gly Ala Leu Leu Ala Gly Ala Leu Ala Ala Tyr Ala Ala 1 5 10 15

tac ctt gtg ctt gcg gcg ctt cttg gtg gcg gcg ctt cttg cag gcgg ccgc cacc 215
Tyr Leu Val Leu Gly Ala Leu Leu Leu Val Ala Arg Leu Glu Gly Pro His 20 25 30

gaa gcc agg ctc cga gcc gag cttg caggc gcg cgg cag ctt ctt 263
Glu Ala Arg Leu Arg Leu Glu Glu Thr Leu Arg Ala Glu Val Leu Leu 35 40 45

cag cgc agc ccc tgt gtg gtct gcc gcc gcc ctt gcg cgc ctc gtg gag 311
Gln Arg Ser Pro Cys Val Ala Al asym Ala Arg Leu Asp Ala Phe Val Glu 50 55 60

cga gtg ctt gcg gcc gga cgg cgg gtc gtt ctt ctt gcc aac gct 359
Arg Val Leu Ala Gly Arg Leu Gly Arg Val Val Leu Ala Asn Ala 65 70 75 80

tcg cgg gtt gcg cgc cgg cgg gtc ctt ctt gcg ctcgt gcc cgg ctc gtg 407
Ser Gly Ser Ala Asn Ala Ser Asp Pro Ala Trp Asp Phe Ala Ser Ala 85 90 95

ttc ctc ctc gcg agc acg ctt atc acc acc gcg ttg gtc tat ggc tac aca 455

29
Leu Phe Phe Ala Ser Thr Leu Ile Thr Thr Val Gly Tyr Gly Tyr Thr
100 105 110

acg cca ctg act gat gcg ggc aag gcc ttc tcc atc gcc ttt gcg ctc
Thr Pro Leu Thr Asp Ala Gly Lys Ala Ser Thr Val Ala Phe Ala Leu
115 120 125

ttg gcc gtg ccg acc acc atg ctg ctg acc gcc tca gcc cag cgc
Leu Gly Val Pro Thr Thr Met Leu Leu Thr Ala Ser Ala Gln Arg
130 135 140

ttg tca ctg ctg act cac gtc gtc ctc tct tgt agc atg cgt
Leu Ser Leu Leu Leu Thr His Pro Leu Pro Leu Leu Thr Ala Ser Ala Leu
145 150 155 160

ttg ggc tgg gcc ccc gcc cgg gcg ggc tgg cac tgg gtc gcc ctc
Trp Gly Trp Asp Pro Arg Ala Ala Cys Thr His Leu Val Ala Leu
165 170 175

ttg ggg gtc gta gtc acc gtc tgg ctc gtt cgg gct gtc atc ttt
Leu Gly Val Val Val Val Leu Leu Pro Ala Ala Pro Ala Ile Phe
180 185 190

gcc cac ctc gag gag gcc tgg agc ttc tgt gag gcc ttc tac ttc tgc
Ala His Leu Glu Glu Ala Trp Ser Phe Leu Asp Ala Phe Tyr Phe Cys
195 200 205

ttt atc tct ctg tcc acc atc ggc ctg gcc gac tac gtc ccc ggg gag
Phe Ile Ser Leu Ser Thr Ile Gly Leu Gly Asp Tyr Val Pro Gly Glu
210 215 220

gcc cct ggc cag ccc tac cgg gcc ctc tac aag gtc ctg gtc gca gtc
Ala Pro Gly Glu Pro Tyr Arg Ala Leu Tyr Lys Val Leu Val Thr Val
225 230 235 240

tac ctc ttc ctg cgg gtc ctg atg gtc ctg gtc cag acc ttc
Tyr Leu Phe Leu Gly Leu Val Ala Met Val Leu Val Leu Gln Thr Phe
245 250 255

cgc cac gtg tcc gcc ctc cac gcc ctc acg gag ctc atc ctg ctg
Arg His Val Ser Asp Leu His Gly Leu Thr Glu Leu Ile Leu Pro
260 265 270

cct cgg tgc cct gcc agt ttc aat gcg gat gag gcg gag ggc gtc gcg
Pro Pro Cys Pro Ala Ser Phe Asn Ala Ala Glu Asp Asp Arg Val Asp
275 280 285

atc ctg gcc ccc cag cag tag tgc ccg gcc ctc tct gcc acc tcc
Ile Leu Gly Pro Glu Pro Glu Ser His Gln Gln Leu Ser Ala Ser Ser
290 295 300

cac acc gac tac gct tcc atc ccc agg tag cttggtgccgcc ctc tgg ccc
His Thr Asp Tyr Ala Ser Ile Pro Arg *
305 310

cttgggtgctg ccctgcctgg gactgagggg tccaggccag cagagctggc tgtacaggaa 1141
tgctccagag cacagccaggt tattgtagg ccctgtgcgc caccgtctct ctatttgcttc 1201
cacagcatgg gctgggtatg aaggggcaag atccctttgc cccatgctcc gggtcccact 1261
ggctggccac cac ataaccttg cttctgtctt tttcttcctc tcttttacac cttgttctcttc 1321
tggctctctgc gcatctctgc tccctctcgc tttcccctcct cgtctctctg ttctccctctc 1381
ttttcctcgt tcctgctcttg tcttccaatt aaccactcgt caactgctggt ttctactggg 1441
ctgtgggtcct agacccctatt tcagccaccca gattgtgctgc tacacccctgg acaagtgcct 1501
gccqgtctctt gagcccttgatt ttcctcagct gcacaaatgag aaagaataaag aatnttgccc 1561
tcaacactctc cttactgtcct gccccctgtgc tagacatgtgc tggacatac ccgactggctt 1621
gagaactgcc cttatgggac accttatagc cagcctgctg cagccttgatc cccagacagt 1681
ggtgtaaggcct ctaaaggtgct gactggtgac tgtggattcagg tctgctgctc cagcctgacg 1741
gggagaggctt cctagagggaa gggacatcagc actaagacct gaactatgagt aatattgccc 1801
gagaagagtgtg tacctgtacact atttttttctca cttgtgtccca cggagccagag ccctagggag 1861
gacccctgtgt gtacgctttcgc gcacatggcct atttttctac cttttcttctgg ggcctcaggg 1921
catagcaggct caccctccactg tgtatgtgtct gcctcgttcttc ttacctcagct cacatcactt 2011
cccaacttac gccacacccct ggggggtgcct cttggggcagct cacatcacttcc tctctgattt 2161
gttctcttatag ctgataacattc taataaacatc ttccaaaggtgt ttccaggtcttta 2221
gttgtgtgcct ctcagcctcttt ctttgagcagtt cgtggaggctt cggagctgttg gtcttctctt 2281
gagggcagaggt ggtgctgtagc atgctatagcag acacatggagt gcctttatcttc aaacaacacc 2341
caaaaattctt ctttgagcctt gcctctggtgtgcct gtacgctttc ctttgagaaggt ctttgagccttt 2391
gcttgaggtgt gaaggttcctgct tcctggcctgg aggaggaggtc gcttgaggtgt gcctctggt 2451
agctgtcgcac cttccagcttg gcacatggcct tccagctgctgc ggcctcaggg gcacatcactt 2521
aaaagaagagaga cagagagagagaa cagagagagagaa cagagagagagaa cagagagagagaa 2581
aaaagaagagaga gagaagagagagaa gagaagagagagaa 2641

<210> 21
<211> PRT
<212> Homo sapiens

<400> 21
Met Arg Arg Gly Ala Leu Leu Ala Gly Ala Leu Ala Ala Tyr Ala Ala
  1    5    10    15
Tyr Leu Val Leu Gly Ala Leu Leu Val Ala Arg Leu Glu Gly Pro His
  20   25   30
Glu Ala Arg Leu Arg Ala Glu Leu Thr Leu Arg Ala Gin Leu Leu
  35   40   45
Gln Arg Ser Pro Cys Val Ala Pro Ala Leu Asp Ala Phe Val Glu
  50   55   60
Arg Val Leu Ala Ala Gly Arg Leu Gly Arg Val Val Leu Ala Asn Ala
  65   70   75   80
Ser Gly Ser Ala Asn Ala Ser Asp Pro Ala Trp Asp Phe Ala Ser Ala
  85   90   95
Leu Phe Phe Ala Ser Thr Leu Ile Thr Thr Val Gly Tyr Gly Tyr Thr
 100  105  110
Thr Pro Leu Thr Asp Ala Gly Lys Ala Phe Ser Ile Ala Phe Ala Leu
 115  120  125
Leu Gly Val Pro Thr Thr Met Leu Leu Leu Thr Ala Ser Ala Gin Arg
 130  135  140
Leu Ser Leu Leu Leu Thr His Pro Leu Ser Thr Phe Ala Ser Met Arg
 145  150  155  160
Trp Gly Trp Asp Pro Arg Arg Ala Ala Cys Thr His Leu Val Ala Leu
 165  170  175
Leu Gly Val Val Thr Val Cys Phe Leu Val Pro Ala Val Ile Phe
 180  185  190
Ala His Leu Glu Glu Ala Trp Ser Phe Leu Asp Ala Phe Tyr Phe Cys
 195  200  205
Phe Ile Ser Leu Ser Thr Ile Gly Leu Gly Asp Tyr Val Pro Gly Glu
 210  215  220
Ala Pro Gly Glu Pro Tyr Arg Ala Leu Tyr Lys Val Leu Val Thr Val
 225  230  235  240
Tyr Leu Phe Leu Gly Leu Ala Met Val Leu Val Leu Gin Thr Phe
 245  250  255
Arg His Val Ser His Leu Ser Gly His Leu Thr Glu Leu Ile Leu Leu Pro
<table>
<thead>
<tr>
<th></th>
<th>260</th>
<th>265</th>
<th>270</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pro</td>
<td>Pro</td>
<td>Cys</td>
</tr>
<tr>
<td></td>
<td>275</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ile</td>
<td>Leu</td>
<td>Gly</td>
</tr>
<tr>
<td></td>
<td>290</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>His</td>
<td>Thr</td>
<td>Asp</td>
</tr>
<tr>
<td></td>
<td>305</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>