

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

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
25 March 2004 (25.03.2004)

PCT

(10) International Publication Number  
WO 2004/024076 A2

(51) International Patent Classification<sup>7</sup>: A61K CA 94010 (US). WU, Thomas D. [US/US]; 113 Elsie Street, San Francisco, CA 94110 (US).

(21) International Application Number:  
PCT/US2003/028361

(74) Agent: CARPENTER, David A.; c/o GENENTECH, INC., MS 49, 1 DNA Way, South San Francisco, CA 94080-4990 (US).

(22) International Filing Date:  
10 September 2003 (10.09.2003)

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

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
60/410,174 11 September 2002 (11.09.2002) US

(71) Applicant (for all designated States except US): GENENTECH, INC. [US/US]; 1 DNA Way, SOUTH SAN FRANCISCO, CA 94080 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): BODARY, Sarah C. [US/US]; 1520 Sand Hill Road, Apt. 205, Palo Alto, CA 94066 (US). CLARK, Hilary [US/US]; 495 Harkness Avenue, San Francisco, CA 94134 (US). HUNTE, Brisdell [US/US]; 1133 Dolores Street, #16, San Francisco, CA 94110 (US). JACKMAN, Janet K. [US/US]; 94 Patrick Way, Half Moon Bay, CA 94019 (US). SCHOENFELD, Jill R. [US/US]; 680 Spring Creek Drive, Ashland, OR 97520 (US). WILLIAMS, P. Mickey [US/US]; 509 Alto Avenue, Half Moon Bay, CA 94019 (US). WOOD, William I. [US/US]; 35 Southdown Court, Hillsborough, CA 94010 (US).

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

Published:

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

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

WO 2004/024076 A2

(54) Title: NOVEL COMPOSITIONS AND METHODS FOR THE TREATMENT OF IMMUNE RELATED DISEASES

(57) Abstract: The present invention relates to compositions containing novel proteins and methods of using those compositions for the diagnosis and treatment of immune related diseases.

**COMPOSITIONS AND METHODS FOR THE TREATMENT OF IMMUNE RELATED DISEASES**

5

Field of the Invention

The present invention relates to compositions and methods useful for the diagnosis and treatment of immune related diseases.

Background of the Invention

10 Immune related and inflammatory diseases are the manifestation or consequence of fairly complex, often multiple interconnected biological pathways which in normal physiology are critical to respond to insult or injury, initiate repair from insult or injury, and mount innate and acquired defense against foreign organisms. Disease or pathology occurs when these normal physiological pathways cause additional insult or injury either as directly related to the intensity of the response, as a consequence of abnormal regulation 15 or excessive stimulation, as a reaction to self, or as a combination of these.

Though the genesis of these diseases often involves multistep pathways and often multiple different biological systems/pathways, intervention at critical points in one or more of these pathways can have an ameliorative or therapeutic effect. Therapeutic intervention can occur by either antagonism of a detrimental process/pathway or stimulation of a beneficial process/pathway.

20 Many immune related diseases are known and have been extensively studied. Such diseases include immune-mediated inflammatory diseases, non-immune-mediated inflammatory diseases, infectious diseases, immunodeficiency diseases, neoplasia, *etc.*

25 T lymphocytes (T cells) are an important component of a mammalian immune response. T cells recognize antigens which are associated with a self-molecule encoded by genes within the major histocompatibility complex (MHC). The antigen may be displayed together with MHC molecules on the 30 surface of antigen presenting cells, virus infected cells, cancer cells, grafts, *etc.* The T cell system eliminates these altered cells which pose a health threat to the host mammal. T cells include helper T cells and cytotoxic T cells. Helper T cells proliferate extensively following recognition of an antigen -MHC complex on an antigen presenting cell. Helper T cells also secrete a variety of cytokines, *i.e.*, lymphokines, which play a central role in the activation of B cells, cytotoxic T cells and a variety of other cells which participate in the immune response.

35 Immune related diseases could be treated by suppressing the immune response. Using neutralizing antibodies that inhibit molecules having immune stimulatory activity would be beneficial in the treatment of immune-mediated and inflammatory diseases. Molecules which inhibit the immune response can be utilized (proteins directly or via the use of antibody agonists) to inhibit the immune response and thus ameliorate 40 immune related disease.

CD4+ T cells are known to be important regulators of inflammation. Herein, CD4+ T cells were activated and the profile of genes differentially expressed upon activation was analyzed. As such, the activation specific genes may be potential therapeutic targets. *In vivo* co-stimulation is necessary for a productive immune proliferative response. The list of costimulatory molecules is quite extensive and it is

still unclear just which co-stimulatory molecules play critical roles in different types and stages of inflammation.

CD4+ T cells are known to be important regulators of inflammation. Herein, CD4+ T cells were activated and the profile of genes differentially expressed upon activation was analyzed. As such, the activation specific genes may be potential therapeutic targets. *In vivo* co-stimulation is necessary for a productive immune proliferative response. The list of costimulatory molecules is quite extensive and it is still unclear just which co-stimulatory molecules play critical roles in different types and stages of inflammation. In this application the focus is on genes which are specifically upregulated by stimulation with anti-CD3/ICAM, or anti-CD3/anti-CD28 and may be useful in targeting inflammatory processes which are associated with these different molecules.

Despite the above identified advances in T cell research, there is a great need for additional diagnostic and therapeutic agents capable of detecting the presence of a T cell mediated disorders in a mammal and for effectively reducing these disorders. Accordingly, it is an objective of the present invention to identify polypeptides that are overexpressed in activated T cells as compared to resting T cells, and to use those polypeptides, and their encoding nucleic acids, to produce compositions of matter useful in the therapeutic treatment and diagnostic detection of T cell mediated disorders in mammals.

#### Summary of the Invention

##### A. Embodiments

The present invention concerns compositions and methods useful for the diagnosis and treatment of immune related disease in mammals, including humans. The present invention is based on the identification of proteins (including agonist and antagonist antibodies) which are a result of stimulation of the immune response in mammals. Immune related diseases can be treated by suppressing or enhancing the immune response. Molecules that enhance the immune response stimulate or potentiate the immune response to an antigen. Molecules which stimulate the immune response can be used therapeutically where enhancement of the immune response would be beneficial. Alternatively, molecules that suppress the immune response attenuate or reduce the immune response to an antigen (e.g., neutralizing antibodies) can be used therapeutically where attenuation of the immune response would be beneficial (e.g., inflammation). Accordingly, the PRO polypeptides, agonists and antagonists thereof are also useful to prepare medicines and medicaments for the treatment of immune-related and inflammatory diseases. In a specific aspect, such medicines and medicaments comprise a therapeutically effective amount of a PRO polypeptide, agonist or antagonist thereof with a pharmaceutically acceptable carrier. Preferably, the admixture is sterile.

In a further embodiment, the invention concerns a method of identifying agonists or antagonists to a PRO polypeptide which comprises contacting the PRO polypeptide with a candidate molecule and monitoring a biological activity mediated by said PRO polypeptide. Preferably, the PRO polypeptide is a native sequence PRO polypeptide. In a specific aspect, the PRO agonist or antagonist is an anti-PRO antibody.

In another embodiment, the invention concerns a composition of matter comprising a PRO polypeptide or an agonist or antagonist antibody which binds the polypeptide in admixture with a carrier or excipient. In one aspect, the composition comprises a therapeutically effective amount of the polypeptide or

antibody. In another aspect, when the composition comprises an immune stimulating molecule, the composition is useful for: (a) increasing infiltration of inflammatory cells into a tissue of a mammal in need thereof, (b) stimulating or enhancing an immune response in a mammal in need thereof, (c) increasing the proliferation of T-lymphocytes in a mammal in need thereof in response to an antigen, (d) stimulating the 5 activity of T-lymphocytes or (e) increasing the vascular permeability. In a further aspect, when the composition comprises an immune inhibiting molecule, the composition is useful for: (a) decreasing infiltration of inflammatory cells into a tissue of a mammal in need thereof, (b) inhibiting or reducing an immune response in a mammal in need thereof, (c) decreasing the activity of T-lymphocytes or (d) decreasing the proliferation of T-lymphocytes in a mammal in need thereof in response to an antigen. In 10 another aspect, the composition comprises a further active ingredient, which may, for example, be a further antibody or a cytotoxic or chemotherapeutic agent. Preferably, the composition is sterile.

In another embodiment, the invention concerns a method of treating an immune related disorder in a mammal in need thereof, comprising administering to the mammal an effective amount of a PRO polypeptide, an agonist thereof, or an antagonist thereto. In a preferred aspect, the immune related disorder 15 is selected from the group consisting of: systemic lupus erythematosis, rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis, spondyloarthropathies, systemic sclerosis, idiopathic inflammatory myopathies, Sjögren's syndrome, systemic vasculitis, sarcoidosis, autoimmune hemolytic anemia, autoimmune thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal disease, demyelinating diseases of the central and peripheral nervous systems such as multiple sclerosis, idiopathic demyelinating 20 polyneuropathy or Guillain-Barré syndrome, and chronic inflammatory demyelinating polyneuropathy, hepatobiliary diseases such as infectious, autoimmune chronic active hepatitis, primary biliary cirrhosis, granulomatous hepatitis, and sclerosing cholangitis, inflammatory bowel disease, gluten-sensitive enteropathy, and Whipple's disease, autoimmune or immune-mediated skin diseases including bullous skin diseases, erythema multiforme and contact dermatitis, psoriasis, allergic diseases such as asthma, allergic 25 rhinitis, atopic dermatitis, food hypersensitivity and urticaria, immunologic diseases of the lung such as eosinophilic pneumonias, idiopathic pulmonary fibrosis and hypersensitivity pneumonitis, transplantation associated diseases including graft rejection and graft -versus-host-disease.

In another embodiment, the invention provides an antibody which specifically binds to any of the above or below described polypeptides. Optionally, the antibody is a monoclonal antibody, humanized 30 antibody, antibody fragment or single-chain antibody. In one aspect, the present invention concerns an isolated antibody which binds a PRO polypeptide. In another aspect, the antibody mimics the activity of a PRO polypeptide (an agonist antibody) or conversely the antibody inhibits or neutralizes the activity of a PRO polypeptide (an antagonist antibody). In another aspect, the antibody is a monoclonal antibody, which 35 preferably has nonhuman complementarity determining region (CDR) residues and human framework region (FR) residues. The antibody may be labeled and may be immobilized on a solid support. In a further aspect, the antibody is an antibody fragment, a monoclonal antibody, a single-chain antibody, or an anti-idiotypic antibody.

In yet another embodiment, the present invention provides a composition comprising an anti-PRO antibody in admixture with a pharmaceutically acceptable carrier. In one aspect, the composition comprises 40 a therapeutically effective amount of the antibody. Preferably, the composition is sterile. The composition

may be administered in the form of a liquid pharmaceutical formulation, which may be preserved to achieve extended storage stability. Alternatively, the antibody is a monoclonal antibody, an antibody fragment, a humanized antibody, or a single-chain antibody.

In a further embodiment, the invention concerns an article of manufacture, comprising:

5 (a) a composition of matter comprising a PRO polypeptide or agonist or antagonist thereof;  
(b) a container containing said composition; and  
(c) a label affixed to said container, or a package insert included in said container referring to the use of said PRO polypeptide or agonist or antagonist thereof in the treatment of an immune related disease. The composition may comprise a therapeutically effective amount of the PRO polypeptide or the  
10 agonist or antagonist thereof.

In yet another embodiment, the present invention concerns a method of diagnosing an immune related disease in a mammal, comprising detecting the level of expression of a gene encoding a PRO polypeptide (a) in a test sample of tissue cells obtained from the mammal, and (b) in a control sample of known normal tissue cells of the same cell type, wherein a higher or lower expression level in the test  
15 sample as compared to the control sample indicates the presence of immune related disease in the mammal from which the test tissue cells were obtained.

In another embodiment, the present invention concerns a method of diagnosing an immune disease in a mammal, comprising (a) contacting an anti-PRO antibody with a test sample of tissue cells obtained from the mammal, and (b) detecting the formation of a complex between the antibody and a PRO polypeptide, in the test sample; wherein the formation of said complex is indicative of the presence or absence of said disease. The detection may be qualitative or quantitative, and may be performed in comparison with monitoring the complex formation in a control sample of known normal tissue cells of the same cell type. A larger quantity of complexes formed in the test sample indicates the presence or absence of an immune disease in the mammal from which the test tissue cells were obtained. The antibody  
20 preferably carries a detectable label. Complex formation can be monitored, for example, by light microscopy, flow cytometry, fluorimetry, or other techniques known in the art. The test sample is usually obtained from an individual suspected of having a deficiency or abnormality of the immune system.  
25

In another embodiment, the invention provides a method for determining the presence of a PRO polypeptide in a sample comprising exposing a test sample of cells suspected of containing the PRO polypeptide to an anti-PRO antibody and determining the binding of said antibody to said cell sample. In a specific aspect, the sample comprises a cell suspected of containing the PRO polypeptide and the antibody binds to the cell. The antibody is preferably detectably labeled and/or bound to a solid support.

In another embodiment, the present invention concerns an immune-related disease diagnostic kit, comprising an anti-PRO antibody and a carrier in suitable packaging. The kit preferably contains  
35 instructions for using the antibody to detect the presence of the PRO polypeptide. Preferably the carrier is pharmaceutically acceptable.

In another embodiment, the present invention concerns a diagnostic kit, containing an anti-PRO antibody in suitable packaging. The kit preferably contains instructions for using the antibody to detect the PRO polypeptide.

40 In another embodiment, the invention provides a method of diagnosing an immune-related disease

in a mammal which comprises detecting the presence or absence or a PRO polypeptide in a test sample of tissue cells obtained from said mammal, wherein the presence or absence of the PRO polypeptide in said test sample is indicative of the presence of an immune-related disease in said mammal.

In another embodiment, the present invention concerns a method for identifying an agonist of a PRO polypeptide comprising:

(a) contacting cells and a test compound to be screened under conditions suitable for the induction of a cellular response normally induced by a PRO polypeptide; and

(b) determining the induction of said cellular response to determine if the test compound is an effective agonist, wherein the induction of said cellular response is indicative of said test compound being an effective agonist.

In another embodiment, the invention concerns a method for identifying a compound capable of inhibiting the activity of a PRO polypeptide comprising contacting a candidate compound with a PRO polypeptide under conditions and for a time sufficient to allow these two components to interact and determining whether the activity of the PRO polypeptide is inhibited. In a specific aspect, either the candidate compound or the PRO polypeptide is immobilized on a solid support. In another aspect, the non-immobilized component carries a detectable label. In a preferred aspect, this method comprises the steps of:

(a) contacting cells and a test compound to be screened in the presence of a PRO polypeptide under conditions suitable for the induction of a cellular response normally induced by a PRO polypeptide; and

(b) determining the induction of said cellular response to determine if the test compound is an effective antagonist.

In another embodiment, the invention provides a method for identifying a compound that inhibits the expression of a PRO polypeptide in cells that normally express the polypeptide, wherein the method comprises contacting the cells with a test compound and determining whether the expression of the PRO polypeptide is inhibited. In a preferred aspect, this method comprises the steps of:

(a) contacting cells and a test compound to be screened under conditions suitable for allowing expression of the PRO polypeptide; and

(b) determining the inhibition of expression of said polypeptide.

In yet another embodiment, the present invention concerns a method for treating an immune-related disorder in a mammal that suffers therefrom comprising administering to the mammal a nucleic acid molecule that codes for either (a) a PRO polypeptide, (b) an agonist of a PRO polypeptide or (c) an antagonist of a PRO polypeptide, wherein said agonist or antagonist may be an anti-PRO antibody. In a preferred embodiment, the mammal is human. In another preferred embodiment, the nucleic acid is administered via *ex vivo* gene therapy. In a further preferred embodiment, the nucleic acid is comprised within a vector, more preferably an adenoviral, adeno-associated viral, lentiviral or retroviral vector.

In yet another aspect, the invention provides a recombinant viral particle comprising a viral vector consisting essentially of a promoter, nucleic acid encoding (a) a PRO polypeptide, (b) an agonist polypeptide of a PRO polypeptide, or (c) an antagonist polypeptide of a PRO polypeptide, and a signal sequence for cellular secretion of the polypeptide, wherein the viral vector is in association with viral structural proteins. Preferably, the signal sequence is from a mammal, such as from a native PRO polypeptide.

In a still further embodiment, the invention concerns an *ex vivo* producer cell comprising a nucleic

acid construct that expresses retroviral structural proteins and also comprises a retroviral vector consisting essentially of a promoter, nucleic acid encoding (a) a PRO polypeptide, (b) an agonist polypeptide of a PRO polypeptide or (c) an antagonist polypeptide of a PRO polypeptide, and a signal sequence for cellular secretion of the polypeptide, wherein said producer cell packages the retroviral vector in association with the structural proteins to produce recombinant retroviral particles.

5 In a still further embodiment, the invention provides a method of increasing the activity of T-lymphocytes in a mammal comprising administering to said mammal (a) a PRO polypeptide, (b) an agonist of a PRO polypeptide, or (c) an antagonist of a PRO polypeptide, wherein the activity of T-lymphocytes in the mammal is increased.

10 In a still further embodiment, the invention provides a method of decreasing the activity of T-lymphocytes in a mammal comprising administering to said mammal (a) a PRO polypeptide, (b) an agonist of a PRO polypeptide, or (c) an antagonist of a PRO polypeptide, wherein the activity of T-lymphocytes in the mammal is decreased.

15 In a still further embodiment, the invention provides a method of increasing the proliferation of T-lymphocytes in a mammal comprising administering to said mammal (a) a PRO polypeptide, (b) an agonist of a PRO polypeptide, or (c) an antagonist of a PRO polypeptide, wherein the proliferation of T-lymphocytes in the mammal is increased.

20 In a still further embodiment, the invention provides a method of decreasing the proliferation of T-lymphocytes in a mammal comprising administering to said mammal (a) a PRO polypeptide, (b) an agonist of a PRO polypeptide, or (c) an antagonist of a PRO polypeptide, wherein the proliferation of T-lymphocytes in the mammal is decreased.

#### B. Additional Embodiments

In other embodiments of the present invention, the invention provides vectors comprising DNA encoding any of the herein described polypeptides. Host cell comprising any such vector are also provided.

25 By way of example, the host cells may be CHO cells, *E. coli*, or yeast. A process for producing any of the herein described polypeptides is further provided and comprises culturing host cells under conditions suitable for expression of the desired polypeptide and recovering the desired polypeptide from the cell culture.

30 In other embodiments, the invention provides chimeric molecules comprising any of the herein described polypeptides fused to a heterologous polypeptide or amino acid sequence. Example of such chimeric molecules comprise any of the herein described polypeptides fused to an epitope tag sequence or a Fc region of an immunoglobulin.

35 In another embodiment, the invention provides an antibody which specifically binds to any of the above or below described polypeptides. Optionally, the antibody is a monoclonal antibody, humanized antibody, antibody fragment or single-chain antibody.

In yet other embodiments, the invention provides oligonucleotide probes useful for isolating genomic and cDNA nucleotide sequences or as antisense probes, wherein those probes may be derived from any of the above or below described nucleotide sequences.

40 In other embodiments, the invention provides an isolated nucleic acid molecule comprising a nucleotide sequence that encodes a PRO polypeptide.

In one aspect, the isolated nucleic acid molecule comprises a nucleotide sequence having at least about 80% nucleic acid sequence identity, alternatively at least about 81% nucleic acid sequence identity, alternatively at least about 82% nucleic acid sequence identity, alternatively at least about 83% nucleic acid sequence identity, alternatively at least about 84% nucleic acid sequence identity, alternatively at least about 85% nucleic acid sequence identity, alternatively at least about 86% nucleic acid sequence identity, alternatively at least about 87% nucleic acid sequence identity, alternatively at least about 88% nucleic acid sequence identity, alternatively at least about 89% nucleic acid sequence identity, alternatively at least about 90% nucleic acid sequence identity, alternatively at least about 91% nucleic acid sequence identity, alternatively at least about 92% nucleic acid sequence identity, alternatively at least about 93% nucleic acid sequence identity, alternatively at least about 94% nucleic acid sequence identity, alternatively at least about 95% nucleic acid sequence identity, alternatively at least about 96% nucleic acid sequence identity, alternatively at least about 97% nucleic acid sequence identity, alternatively at least about 98% nucleic acid sequence identity and alternatively at least about 99% nucleic acid sequence identity to (a) a DNA molecule encoding a PRO polypeptide having a full-length amino acid sequence as disclosed herein, an amino acid sequence lacking the signal peptide as disclosed herein, an extracellular domain of a transmembrane protein, with or without the signal peptide, as disclosed herein or any other specifically defined fragment of the full-length amino acid sequence as disclosed herein, or (b) the complement of the DNA molecule of (a).

In other aspects, the isolated nucleic acid molecule comprises a nucleotide sequence having at least about 80% nucleic acid sequence identity, alternatively at least about 81% nucleic acid sequence identity, alternatively at least about 82% nucleic acid sequence identity, alternatively at least about 83% nucleic acid sequence identity, alternatively at least about 84% nucleic acid sequence identity, alternatively at least about 85% nucleic acid sequence identity, alternatively at least about 86% nucleic acid sequence identity, alternatively at least about 87% nucleic acid sequence identity, alternatively at least about 88% nucleic acid sequence identity, alternatively at least about 89% nucleic acid sequence identity, alternatively at least about 90% nucleic acid sequence identity, alternatively at least about 91% nucleic acid sequence identity, alternatively at least about 92% nucleic acid sequence identity, alternatively at least about 93% nucleic acid sequence identity, alternatively at least about 94% nucleic acid sequence identity, alternatively at least about 95% nucleic acid sequence identity, alternatively at least about 96% nucleic acid sequence identity, alternatively at least about 97% nucleic acid sequence identity, alternatively at least about 98% nucleic acid sequence identity and alternatively at least about 99% nucleic acid sequence identity to (a) a DNA molecule comprising the coding sequence of a full-length PRO polypeptide cDNA as disclosed herein, the coding sequence of a PRO polypeptide lacking the signal peptide as disclosed herein, the coding sequence of an extracellular domain of a transmembrane PRO polypeptide, with or without the signal peptide, as disclosed herein or the coding sequence of any other specifically defined fragment of the full-length amino acid sequence as disclosed herein, or (b) the complement of the DNA molecule of (a).

In a further aspect, the invention concerns an isolated nucleic acid molecule comprising a nucleotide sequence having at least about 80% nucleic acid sequence identity, alternatively at least about 81% nucleic acid sequence identity, alternatively at least about 82% nucleic acid sequence identity, alternatively at least about 83% nucleic acid sequence identity, alternatively at least about 84% nucleic acid sequence identity, alternatively at least about 85% nucleic acid sequence identity, alternatively at least about

86% nucleic acid sequence identity, alternatively at least about 87% nucleic acid sequence identity, alternatively at least about 88% nucleic acid sequence identity, alternatively at least about 89% nucleic acid sequence identity, alternatively at least about 90% nucleic acid sequence identity, alternatively at least about 91% nucleic acid sequence identity, alternatively at least about 92% nucleic acid sequence identity,  
5 alternatively at least about 93% nucleic acid sequence identity, alternatively at least about 94% nucleic acid sequence identity, alternatively at least about 95% nucleic acid sequence identity, alternatively at least about 96% nucleic acid sequence identity, alternatively at least about 97% nucleic acid sequence identity, alternatively at least about 98% nucleic acid sequence identity and alternatively at least about 99% nucleic acid sequence identity to (a) a DNA molecule that encodes the same mature polypeptide encoded by any of  
10 the human protein cDNAs as disclosed herein, or (b) the complement of the DNA molecule of (a).

Another aspect the invention provides an isolated nucleic acid molecule comprising a nucleotide sequence encoding a PRO polypeptide which is either transmembrane domain-deleted or transmembrane domain-inactivated, or is complementary to such encoding nucleotide sequence, wherein the transmembrane domain(s) of such polypeptide are disclosed herein. Therefore, soluble extracellular domains of the herein  
15 described PRO polypeptides are contemplated.

Another embodiment is directed to fragments of a PRO polypeptide coding sequence, or the complement thereof, that may find use as, for example, hybridization probes, for encoding fragments of a PRO polypeptide that may optionally encode a polypeptide comprising a binding site for an anti-PRO antibody or as antisense oligonucleotide probes. Such nucleic acid fragments are usually at least about 20 nucleotides in length, alternatively at least about 30 nucleotides in length, alternatively at least about 40 nucleotides in length, alternatively at least about 50 nucleotides in length, alternatively at least about 60 nucleotides in length, alternatively at least about 70 nucleotides in length, alternatively at least about 80 nucleotides in length, alternatively at least about 90 nucleotides in length, alternatively at least about 100 nucleotides in length, alternatively at least about 110 nucleotides in length, alternatively at least about 120 nucleotides in length, alternatively at least about 130 nucleotides in length, alternatively at least about 140 nucleotides in length, alternatively at least about 150 nucleotides in length, alternatively at least about 160 nucleotides in length, alternatively at least about 170 nucleotides in length, alternatively at least about 180 nucleotides in length, alternatively at least about 190 nucleotides in length, alternatively at least about 200 nucleotides in length, alternatively at least about 250 nucleotides in length, alternatively at least about 300 nucleotides in length, alternatively at least about 350 nucleotides in length, alternatively at least about 400 nucleotides in length, alternatively at least about 450 nucleotides in length, alternatively at least about 500 nucleotides in length, alternatively at least about 600 nucleotides in length, alternatively at least about 700 nucleotides in length, alternatively at least about 800 nucleotides in length, alternatively at least about 900 nucleotides in length and alternatively at least about 1000 nucleotides in length, wherein in this context the  
25 term "about" means the referenced nucleotide sequence length plus or minus 10% of that referenced length. It is noted that novel fragments of a PRO polypeptide-encoding nucleotide sequence may be determined in a routine manner by aligning the PRO polypeptide-encoding nucleotide sequence with other known nucleotide sequences using any of a number of well known sequence alignment programs and determining which PRO polypeptide-encoding nucleotide sequence fragment(s) are novel. All of such PRO polypeptide-encoding  
30 nucleotide sequences are contemplated herein. Also contemplated are the PRO polypeptide fragments  
35  
40

encoded by these nucleotide molecule fragments, preferably those PRO polypeptide fragments that comprise a binding site for an anti-PRO antibody.

In another embodiment, the invention provides isolated PRO polypeptide encoded by any of the isolated nucleic acid sequences herein above identified.

5 In a certain aspect, the invention concerns an isolated PRO polypeptide, comprising an amino acid sequence having at least about 80% amino acid sequence identity, alternatively at least about 81% amino acid sequence identity, alternatively at least about 82% amino acid sequence identity, alternatively at least about 83% amino acid sequence identity, alternatively at least about 84% amino acid sequence identity, alternatively at least about 85% amino acid sequence identity, alternatively at least about 86% amino acid  
10 sequence identity, alternatively at least about 87% amino acid sequence identity, alternatively at least about 88% amino acid sequence identity, alternatively at least about 89% amino acid sequence identity, alternatively at least about 90% amino acid sequence identity, alternatively at least about 91% amino acid sequence identity, alternatively at least about 92% amino acid sequence identity, alternatively at least about 93% amino acid sequence identity, alternatively at least about 94% amino acid sequence identity,  
15 alternatively at least about 95% amino acid sequence identity, alternatively at least about 96% amino acid sequence identity, alternatively at least about 97% amino acid sequence identity, alternatively at least about 98% amino acid sequence identity and alternatively at least about 99% amino acid sequence identity to a PRO polypeptide having a full-length amino acid sequence as disclosed herein, an amino acid sequence lacking the signal peptide as disclosed herein, an extracellular domain of a transmembrane protein, with or  
20 without the signal peptide, as disclosed herein or any other specifically defined fragment of the full-length amino acid sequence as disclosed herein.

In a further aspect, the invention concerns an isolated PRO polypeptide comprising an amino acid sequence having at least about 80% amino acid sequence identity, alternatively at least about 81% amino acid sequence identity, alternatively at least about 82% amino acid sequence identity, alternatively at least about 83% amino acid sequence identity, alternatively at least about 84% amino acid sequence identity, alternatively at least about 85% amino acid sequence identity, alternatively at least about 86% amino acid sequence identity, alternatively at least about 87% amino acid sequence identity, alternatively at least about 88% amino acid sequence identity, alternatively at least about 89% amino acid sequence identity, alternatively at least about 90% amino acid sequence identity, alternatively at least about 91% amino acid  
25 sequence identity, alternatively at least about 92% amino acid sequence identity, alternatively at least about 93% amino acid sequence identity, alternatively at least about 94% amino acid sequence identity, alternatively at least about 95% amino acid sequence identity, alternatively at least about 96% amino acid sequence identity, alternatively at least about 97% amino acid sequence identity, alternatively at least about 98% amino acid sequence identity and alternatively at least about 99% amino acid sequence identity to an  
30 amino acid sequence encoded by any of the human protein cDNAs as disclosed herein.

In a specific aspect, the invention provides an isolated PRO polypeptide without the N-terminal signal sequence and/or the initiating methionine and is encoded by a nucleotide sequence that encodes such an amino acid sequence as herein before described. Processes for producing the same are also herein described, wherein those processes comprise culturing a host cell comprising a vector which comprises the

appropriate encoding nucleic acid molecule under conditions suitable for expression of the PRO polypeptide and recovering the PRO polypeptide from the cell culture.

Another aspect the invention provides an isolated PRO polypeptide which is either transmembrane domain-deleted or transmembrane domain-inactivated. Processes for producing the same are also herein described, wherein those processes comprise culturing a host cell comprising a vector which comprises the appropriate encoding nucleic acid molecule under conditions suitable for expression of the PRO polypeptide and recovering the PRO polypeptide from the cell culture.

In yet another embodiment, the invention concerns agonists and antagonists of a native PRO polypeptide as defined herein. In a particular embodiment, the agonist or antagonist is an anti-PRO antibody or a small molecule.

In a further embodiment, the invention concerns a method of identifying agonists or antagonists to a PRO polypeptide which comprise contacting the PRO polypeptide with a candidate molecule and monitoring a biological activity mediated by said PRO polypeptide. Preferably, the PRO polypeptide is a native PRO polypeptide.

In a still further embodiment, the invention concerns a composition of matter comprising a PRO polypeptide, or an agonist or antagonist of a PRO polypeptide as herein described, or an anti-PRO antibody, in combination with a carrier. Optionally, the carrier is a pharmaceutically acceptable carrier.

Another embodiment of the present invention is directed to the use of a PRO polypeptide, or an agonist or antagonist thereof as herein before described, or an anti-PRO antibody, for the preparation of a medicament useful in the treatment of a condition which is responsive to the PRO polypeptide, an agonist or antagonist thereof or an anti-PRO antibody.

#### BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows a nucleotide sequence (SEQ ID NO:1) of a native sequence PRO83478 cDNA, wherein SEQ ID NO:1 is a clone designated herein as "DNA327205".

Figure 2 shows the amino acid sequence (SEQ ID NO:2) derived from the coding sequence of SEQ ID NO:1 shown in Figure 1.

Figure 3 shows a nucleotide sequence (SEQ ID NO:3) of a native sequence PRO69889 cDNA, wherein SEQ ID NO:3 is a clone designated herein as "DNA304780".

Figure 4 shows the amino acid sequence (SEQ ID NO:4) derived from the coding sequence of SEQ ID NO:3 shown in Figure 3.

Figure 5 shows a nucleotide sequence (SEQ ID NO:5) of a native sequence PRO37073 cDNA, wherein SEQ ID NO:5 is a clone designated herein as "DNA304459".

Figure 6 shows the amino acid sequence (SEQ ID NO:6) derived from the coding sequence of SEQ ID NO:5 shown in Figure 5.

Figure 7 shows a nucleotide sequence (SEQ ID NO:7) of a native sequence PRO4984 cDNA, wherein SEQ ID NO:7 is a clone designated herein as "DNA304460".

Figure 8 shows the amino acid sequence (SEQ ID NO:8) derived from the coding sequence of SEQ ID NO:7 shown in Figure 7.

Figure 9 shows a nucleotide sequence (SEQ ID NO:9) of a native sequence PRO71039 cDNA, wherein SEQ ID NO:9 is a clone designated herein as "DNA304661".

Figure 10 shows the amino acid sequence (SEQ ID NO:10) derived from the coding sequence of SEQ ID NO:9 shown in Figure 9.

5 Figure 11 shows a nucleotide sequence (SEQ ID NO:11) of a native sequence PRO71191 cDNA, wherein SEQ ID NO:11 is a clone designated herein as "DNA304781".

Figure 12 shows the amino acid sequence (SEQ ID NO:12) derived from the coding sequence of SEQ ID NO:11 shown in Figure 11.

10 Figure 13 shows a nucleotide sequence (SEQ ID NO:13) of a native sequence PRO271 cDNA, wherein SEQ ID NO:13 is a clone designated herein as "DNA327206".

Figure 14 shows the amino acid sequence (SEQ ID NO:14) derived from the coding sequence of SEQ ID NO:14 shown in Figure 14.

Figure 15A-B shows a nucleotide sequence (SEQ ID NO:15) of a native sequence PRO71042 cDNA, wherein SEQ ID NO:15 is a clone designated herein as "DNA304464".

15 Figure 16 shows the amino acid sequence (SEQ ID NO:16) derived from the coding sequence of SEQ ID NO:15 shown in Figure 15A-B.

Figure 17 shows a nucleotide sequence (SEQ ID NO:17) of a native sequence PRO71242 cDNA, wherein SEQ ID NO:17 is a clone designated herein as "DNA304835".

20 Figure 18 shows the amino acid sequence (SEQ ID NO:18) derived from the coding sequence of SEQ ID NO:17 shown in Figure 17.

Figure 19 shows a nucleotide sequence (SEQ ID NO:19) of a native sequence PRO6015 cDNA, wherein SEQ ID NO:19 is a clone designated herein as "DNA96866".

Figure 20 shows the amino acid sequence (SEQ ID NO:20) derived from the coding sequence of SEQ ID NO:19 shown in Figure 19.

25 Figure 21 shows a nucleotide sequence (SEQ ID NO:21) of a native sequence PRO34336 cDNA, wherein SEQ ID NO:21 is a clone designated herein as "DNA304466".

Figure 22 shows the amino acid sequence (SEQ ID NO:22) derived from the coding sequence of SEQ ID NO:21 shown in Figure 21.

30 Figure 23 shows a nucleotide sequence (SEQ ID NO:23) of a native sequence PRO71043 cDNA, wherein SEQ ID NO:23 is a clone designated herein as "DNA304467".

Figure 24 shows the amino acid sequence (SEQ ID NO:24) derived from the coding sequence of SEQ ID NO:23 shown in Figure 23.

Figure 25 shows a nucleotide sequence (SEQ ID NO:25) of a native sequence PRO71044 cDNA, wherein SEQ ID NO:25 is a clone designated herein as "DNA304468".

35 Figure 26 shows the amino acid sequence (SEQ ID NO:26) derived from the coding sequence of SEQ ID NO:25 shown in Figure 25.

Figure 27 shows a nucleotide sequence (SEQ ID NO:27) of a native sequence PRO71045 cDNA, wherein SEQ ID NO:27 is a clone designated herein as "DNA304469".

40 Figure 28 shows the amino acid sequence (SEQ ID NO:28) derived from the coding sequence of SEQ ID NO:27 shown in Figure 27.

Figure 29 shows a nucleotide sequence (SEQ ID NO:29) of a native sequence PRO71046 cDNA, wherein SEQ ID NO:29 is a clone designated herein as “DNA304470”.

Figure 30 shows the amino acid sequence (SEQ ID NO:30) derived from the coding sequence of SEQ ID NO:29 shown in Figure 29.

5 Figure 31 shows a nucleotide sequence (SEQ ID NO:31) of a native sequence PRO83479 cDNA, wherein SEQ ID NO:31 is a clone designated herein as “DNA327207”.

Figure 32 shows the amino acid sequence (SEQ ID NO:32) derived from the coding sequence of SEQ ID NO:31 shown in Figure 31.

10 Figure 33 shows a nucleotide sequence (SEQ ID NO:33) of a native sequence PRO535 cDNA, wherein SEQ ID NO:33 is a clone designated herein as “DNA304472”.

Figure 34 shows the amino acid sequence (SEQ ID NO:34) derived from the coding sequence of SEQ ID NO:33 shown in Figure 33.

Figure 35 shows a nucleotide sequence (SEQ ID NO:35) of a native sequence PRO4426 cDNA, wherein SEQ ID NO:35 is a clone designated herein as “DNA304783”.

15 Figure 36 shows the amino acid sequence (SEQ ID NO:36) derived from the coding sequence of SEQ ID NO:35 shown in Figure 35.

Figure 37 shows a nucleotide sequence (SEQ ID NO:37) of a native sequence PRO34447 cDNA, wherein SEQ ID NO:37 is a clone designated herein as “DNA218651”.

20 Figure 38 shows the amino acid sequence (SEQ ID NO:38) derived from the coding sequence of SEQ ID NO:37 shown in Figure 37.

Figure 39 shows a nucleotide sequence (SEQ ID NO:39) of a native sequence PRO2023 cDNA, wherein SEQ ID NO:39 is a clone designated herein as “DNA304473”.

Figure 40 shows the amino acid sequence (SEQ ID NO:40) derived from the coding sequence of SEQ ID NO:39 shown in Figure 39.

25 Figure 41 shows a nucleotide sequence (SEQ ID NO:41) of a native sequence PRO25349 cDNA, wherein SEQ ID NO:41 is a clone designated herein as “DNA189412”.

Figure 42 shows the amino acid sequence (SEQ ID NO:42) derived from the coding sequence of SEQ ID NO:41 shown in Figure 41

30 Figure 43 shows a nucleotide sequence (SEQ ID NO:43) of a native sequence PRO1265 cDNA, wherein SEQ ID NO:43 is a clone designated herein as “DNA304827”.

Figure 44 shows the amino acid sequence (SEQ ID NO:44) derived from the coding sequence of SEQ ID NO:43 shown in Figure 43.

Figure 45 shows a nucleotide sequence (SEQ ID NO:45) of a native sequence PRO34298 cDNA, wherein SEQ ID NO:45 is a clone designated herein as “DNA217256”.

35 Figure 46 shows the amino acid sequence (SEQ ID NO:46) derived from the coding sequence of SEQ ID NO:45 shown in Figure 45.

Figure 47 shows a nucleotide sequence (SEQ ID NO:47) of a native sequence PRO738 cDNA, wherein SEQ ID NO:47 is a clone designated herein as “DNA304784”.

40 Figure 48 shows the amino acid sequence (SEQ ID NO:48) derived from the coding sequence of SEQ ID NO:47 shown in Figure 47.

Figure 49 shows a nucleotide sequence (SEQ ID NO:49) of a native sequence PRO71049 cDNA, wherein SEQ ID NO:49 is a clone designated herein as "DNA304475".

Figure 50 shows the amino acid sequence (SEQ ID NO:50) derived from the coding sequence of SEQ ID NO:49 shown in Figure 49.

5 Figure 51 shows a nucleotide sequence (SEQ ID NO:51) of a native sequence PRO21341 cDNA, wherein SEQ ID NO:51 is a clone designated herein as "DNA287180".

Figure 52 shows the amino acid sequence (SEQ ID NO:52) derived from the coding sequence of SEQ ID NO:51 shown in Figure 51.

10 Figure 53 shows a nucleotide sequence (SEQ ID NO:53) of a native sequence PRO1125 cDNA, wherein SEQ ID NO:53 is a clone designated herein as "DNA304476".

Figure 54 shows the amino acid sequence (SEQ ID NO:54) derived from the coding sequence of SEQ ID NO:53 shown in Figure 53.

Figure 55 shows a nucleotide sequence (SEQ ID NO:55) of a native sequence PRO4369 cDNA, wherein SEQ ID NO:55 is a clone designated herein as "DNA327208".

15 Figure 56 shows the amino acid sequence (SEQ ID NO:56) derived from the coding sequence of SEQ ID NO:55 shown in Figure 55.

Figure 57 shows a nucleotide sequence (SEQ ID NO:57) of a native sequence PRO177 cDNA, wherein SEQ ID NO:57 is a clone designated herein as "DNA327209".

20 Figure 58 shows the amino acid sequence (SEQ ID NO:58) derived from the coding sequence of SEQ ID NO:57 shown in Figure 57.

Figure 59 shows a nucleotide sequence (SEQ ID NO:59) of a native sequence PRO83480 cDNA, wherein SEQ ID NO:59 is a clone designated herein as "DNA327210".

Figure 60 shows the amino acid sequence (SEQ ID NO:60) derived from the coding sequence of SEQ ID NO:59 shown in Figure 59.

25 Figure 61 shows a nucleotide sequence (SEQ ID NO:61) of a native sequence PRO71052 cDNA, wherein SEQ ID NO:61 is a clone designated herein as "DNA327211".

Figure 62 shows the amino acid sequence (SEQ ID NO:62) derived from the coding sequence of SEQ ID NO:61 shown in Figure 61.

30 Figure 63 shows a nucleotide sequence (SEQ ID NO:63) of a native sequence PRO37125 cDNA, wherein SEQ ID NO:63 is a clone designated herein as "DNA226662".

Figure 64 shows the amino acid sequence (SEQ ID NO:64) derived from the coding sequence of SEQ ID NO:63 shown in Figure 63.

Figure 65 shows a nucleotide sequence (SEQ ID NO:65) of a native sequence PRO83481 cDNA, wherein SEQ ID NO:65 is a clone designated herein as "DNA327212".

35 Figure 66 shows the amino acid sequence (SEQ ID NO:66) derived from the coding sequence of SEQ ID NO:65 shown in Figure 65.

Figure 67 shows a nucleotide sequence (SEQ ID NO:67) of a native sequence PRO71054 cDNA, wherein SEQ ID NO:67 is a clone designated herein as "DNA304482".

40 Figure 68 shows the amino acid sequence (SEQ ID NO:68) derived from the coding sequence of SEQ ID NO:67 shown in Figure 67.

Figure 69 shows a nucleotide sequence (SEQ ID NO:69) of a native sequence PRO69462 cDNA, wherein SEQ ID NO:69 is a clone designated herein as “DNA287171”.

Figure 70 shows the amino acid sequence (SEQ ID NO:70) derived from the coding sequence of SEQ ID NO:69 shown in Figure 69.

5 Figure 71 shows a nucleotide sequence (SEQ ID NO:71) of a native sequence PRO268 cDNA, wherein SEQ ID NO:71 is a clone designated herein as “DNA304484”.

Figure 72 shows the amino acid sequence (SEQ ID NO:72) derived from the coding sequence of SEQ ID NO:71 shown in Figure 71.

10 Figure 73 shows a nucleotide sequence (SEQ ID NO:73) of a native sequence PRO615 cDNA, wherein SEQ ID NO:73 is a clone designated herein as “DNA304485”.

Figure 74 shows the amino acid sequence (SEQ ID NO:74) derived from the coding sequence of SEQ ID NO:73 shown in Figure 73.

Figure 75 shows a nucleotide sequence (SEQ ID NO:75) of a native sequence PRO25402 cDNA, wherein SEQ ID NO:75 is a clone designated herein as “DNA189504”.

15 Figure 76 shows the amino acid sequence (SEQ ID NO:76) derived from the coding sequence of SEQ ID NO:75 shown in Figure 75.

Figure 77 shows a nucleotide sequence (SEQ ID NO:77) of a native sequence PRO71055 cDNA, wherein SEQ ID NO:77 is a clone designated herein as “DNA304486”.

20 Figure 78 shows the amino acid sequence (SEQ ID NO:78) derived from the coding sequence of SEQ ID NO:77 shown in Figure 77.

Figure 79 shows a nucleotide sequence (SEQ ID NO:79) of a native sequence PRO71056 cDNA, wherein SEQ ID NO:79 is a clone designated herein as “DNA304487”.

Figure 80 shows the amino acid sequence (SEQ ID NO:80) derived from the coding sequence of SEQ ID NO:79 shown in Figure 79.

25 Figure 81 shows a nucleotide sequence (SEQ ID NO:81) of a native sequence PRO34332 cDNA, wherein SEQ ID NO:81 is a clone designated herein as “DNA218280”.

Figure 82 shows the amino acid sequence (SEQ ID NO:82) derived from the coding sequence of SEQ ID NO:81 shown in Figure 81.

30 Figure 83 shows a nucleotide sequence (SEQ ID NO:83) of a native sequence PRO71057 cDNA, wherein SEQ ID NO:83 is a clone designated herein as “DNA304488”.

Figure 84 shows the amino acid sequence (SEQ ID NO:84) derived from the coding sequence of SEQ ID NO:83 shown in Figure 83.

Figure 85 shows a nucleotide sequence (SEQ ID NO:85) of a native sequence PRO71058 cDNA, wherein SEQ ID NO:85 is a clone designated herein as “DNA304489”.

35 Figure 86 shows the amino acid sequence (SEQ ID NO:86) derived from the coding sequence of SEQ ID NO:85 shown in Figure 85.

Figure 87 shows a nucleotide sequence (SEQ ID NO:87) of a native sequence PRO83482 cDNA, wherein SEQ ID NO:87 is a clone designated herein as “DNA327213”.

40 Figure 88 shows the amino acid sequence (SEQ ID NO:88) derived from the coding sequence of SEQ ID NO:87 shown in Figure 87.

Figure 89 shows a nucleotide sequence (SEQ ID NO:89) of a native sequence PRO34454 cDNA, wherein SEQ ID NO:89 is a clone designated herein as "DNA218676".

Figure 90 shows the amino acid sequence (SEQ ID NO:90) derived from the coding sequence of SEQ ID NO:89 shown in Figure 89.

5 Figure 91 shows a nucleotide sequence (SEQ ID NO:91) of a native sequence PRO6243 cDNA, wherein SEQ ID NO:91 is a clone designated herein as "DNA304491".

Figure 92 shows the amino acid sequence (SEQ ID NO:92) derived from the coding sequence of SEQ ID NO:91 shown in Figure 91.

10 Figure 93 shows a nucleotide sequence (SEQ ID NO:93) of a native sequence PRO1864 cDNA, wherein SEQ ID NO:93 is a clone designated herein as "DNA304492".

Figure 94 shows the amino acid sequence (SEQ ID NO:94) derived from the coding sequence of SEQ ID NO:93 shown in Figure 93.

Figure 95 shows a nucleotide sequence (SEQ ID NO:95) of a native sequence PRO71060 cDNA, wherein SEQ ID NO:95 is a clone designated herein as "DNA304493".

15 Figure 96 shows the amino acid sequence (SEQ ID NO:96) derived from the coding sequence of SEQ ID NO:95 shown in Figure 95.

Figure 97 shows a nucleotide sequence (SEQ ID NO:97) of a native sequence PRO71061 cDNA, wherein SEQ ID NO:97 is a clone designated herein as "DNA304494".

20 Figure 98 shows the amino acid sequence (SEQ ID NO:98) derived from the coding sequence of SEQ ID NO:97 shown in Figure 97.

Figure 99 shows a nucleotide sequence (SEQ ID NO:99) of a native sequence PRO793 cDNA, wherein SEQ ID NO:99 is a clone designated herein as "DNA304495".

Figure 100 shows the amino acid sequence (SEQ ID NO:100) derived from the coding sequence of SEQ ID NO:99 shown in Figure 99.

25 Figure 101 shows a nucleotide sequence (SEQ ID NO:101) of a native sequence PRO83483 cDNA, wherein SEQ ID NO:101 is a clone designated herein as "DNA327214".

Figure 102 shows the amino acid sequence (SEQ ID NO:102) derived from the coding sequence of SEQ ID NO:101 shown in Figure 101.

30 Figure 103 shows a nucleotide sequence (SEQ ID NO:103) of a native sequence PRO60929 cDNA, wherein SEQ ID NO:103 is a clone designated herein as "DNA272832".

Figure 104 shows the amino acid sequence (SEQ ID NO:104) derived from the coding sequence of SEQ ID NO:103 shown in Figure 103.

#### DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

35 I. Definitions

The terms "PRO polypeptide" and "PRO" as used herein and when immediately followed by a numerical designation refer to various polypeptides, wherein the complete designation (i.e., PRO/number) refers to specific polypeptide sequences as described herein. The terms "PRO/number polypeptide" and "PRO/number" wherein the term "number" is provided as an actual numerical designation as used herein encompass native sequence polypeptides and polypeptide variants (which are further defined herein). The

PRO polypeptides described herein may be isolated from a variety of sources, such as from human tissue types or from another source, or prepared by recombinant or synthetic methods. The term "PRO polypeptide" refers to each individual PRO/number polypeptide disclosed herein. All disclosures in this specification which refer to the "PRO polypeptide" refer to each of the polypeptides individually as well as jointly. For example, descriptions of the preparation of, purification of, derivation of, formation of antibodies to or against, administration of, compositions containing, treatment of a disease with, etc., pertain to each polypeptide of the invention individually. The term "PRO polypeptide" also includes variants of the PRO/number polypeptides disclosed herein.

A "native sequence PRO polypeptide" comprises a polypeptide having the same amino acid sequence as the corresponding PRO polypeptide derived from nature. Such native sequence PRO polypeptides can be isolated from nature or can be produced by recombinant or synthetic means. The term "native sequence PRO polypeptide" specifically encompasses naturally-occurring truncated or secreted forms of the specific PRO polypeptide (e.g., an extracellular domain sequence), naturally-occurring variant forms (e.g., alternatively spliced forms) and naturally-occurring allelic variants of the polypeptide. In various embodiments of the invention, the native sequence PRO polypeptides disclosed herein are mature or full-length native sequence polypeptides comprising the full-length amino acids sequences shown in the accompanying figures. Start and stop codons are shown in bold font and underlined in the figures. However, while the PRO polypeptide disclosed in the accompanying figures are shown to begin with methionine residues designated herein as amino acid position 1 in the figures, it is conceivable and possible that other methionine residues located either upstream or downstream from the amino acid position 1 in the figures may be employed as the starting amino acid residue for the PRO polypeptides.

The PRO polypeptide "extracellular domain" or "ECD" refers to a form of the PRO polypeptide which is essentially free of the transmembrane and cytoplasmic domains. Ordinarily, a PRO polypeptide ECD will have less than 1% of such transmembrane and/or cytoplasmic domains and preferably, will have less than 0.5% of such domains. It will be understood that any transmembrane domains identified for the PRO polypeptides of the present invention are identified pursuant to criteria routinely employed in the art for identifying that type of hydrophobic domain. The exact boundaries of a transmembrane domain may vary but most likely by no more than about 5 amino acids at either end of the domain as initially identified herein. Optionally, therefore, an extracellular domain of a PRO polypeptide may contain from about 5 or fewer amino acids on either side of the transmembrane domain/extracellular domain boundary as identified in the Examples or specification and such polypeptides, with or without the associated signal peptide, and nucleic acid encoding them, are contemplated by the present invention.

The approximate location of the "signal peptides" of the various PRO polypeptides disclosed herein are shown in the present specification and/or the accompanying figures. It is noted, however, that the C-terminal boundary of a signal peptide may vary, but most likely by no more than about 5 amino acids on either side of the signal peptide C-terminal boundary as initially identified herein, wherein the C-terminal boundary of the signal peptide may be identified pursuant to criteria routinely employed in the art for identifying that type of amino acid sequence element (e.g., Nielsen et al., *Prot. Eng.* 10:1-6 (1997) and von Heinje et al., *Nucl. Acids. Res.* 14:4683-4690 (1986)). Moreover, it is also recognized that, in some cases, cleavage of a signal sequence from a secreted polypeptide is not entirely uniform, resulting in more than one

secreted species. These mature polypeptides, where the signal peptide is cleaved within no more than about 5 amino acids on either side of the C-terminal boundary of the signal peptide as identified herein, and the polynucleotides encoding them, are contemplated by the present invention.

"PRO polypeptide variant" means an active PRO polypeptide as defined above or below having at

5 least about 80% amino acid sequence identity with a full-length native sequence PRO polypeptide sequence as disclosed herein, a PRO polypeptide sequence lacking the signal peptide as disclosed herein, an extracellular domain of a PRO polypeptide, with or without the signal peptide, as disclosed herein or any other fragment of a full-length PRO polypeptide sequence as disclosed herein. Such PRO polypeptide variants include, for instance, PRO polypeptides wherein one or more amino acid residues are added, or  
10 deleted, at the N- or C-terminus of the full-length native amino acid sequence. Ordinarily, a PRO polypeptide variant will have at least about 80% amino acid sequence identity, alternatively at least about 81% amino acid sequence identity, alternatively at least about 82% amino acid sequence identity, alternatively at least about 83% amino acid sequence identity, alternatively at least about 84% amino acid sequence identity, alternatively at least about 85% amino acid sequence identity, alternatively at least about 86% amino acid sequence identity, alternatively at least about 87% amino acid sequence identity, alternatively at least about 88% amino acid sequence identity, alternatively at least about 89% amino acid sequence identity, alternatively at least about 90% amino acid sequence identity, alternatively at least about 91% amino acid sequence identity, alternatively at least about 92% amino acid sequence identity, alternatively at least about 93% amino acid sequence identity, alternatively at least about 94% amino acid  
20 sequence identity, alternatively at least about 95% amino acid sequence identity, alternatively at least about 96% amino acid sequence identity, alternatively at least about 97% amino acid sequence identity, alternatively at least about 98% amino acid sequence identity and alternatively at least about 99% amino acid sequence identity to a full-length native sequence PRO polypeptide sequence as disclosed herein, a PRO polypeptide sequence lacking the signal peptide as disclosed herein, an extracellular domain of a PRO  
25 polypeptide, with or without the signal peptide, as disclosed herein or any other specifically defined fragment of a full-length PRO polypeptide sequence as disclosed herein. Ordinarily, PRO variant polypeptides are at least about 10 amino acids in length, alternatively at least about 20 amino acids in length, alternatively at least about 30 amino acids in length, alternatively at least about 40 amino acids in length, alternatively at least about 50 amino acids in length, alternatively at least about 60 amino acids in length,  
30 alternatively at least about 70 amino acids in length, alternatively at least about 80 amino acids in length, alternatively at least about 90 amino acids in length, alternatively at least about 100 amino acids in length, alternatively at least about 150 amino acids in length, alternatively at least about 200 amino acids in length, alternatively at least about 300 amino acids in length, or more.

"Percent (%) amino acid sequence identity" with respect to the PRO polypeptide sequences

35 identified herein is defined as the percentage of amino acid residues in a candidate sequence that are identical with the amino acid residues in the specific PRO polypeptide sequence, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity, and not considering any conservative substitutions as part of the sequence identity. Alignment for purposes of determining percent amino acid sequence identity can be achieved in various ways that are within the skill in  
40 the art, for instance, using publicly available computer software such as BLAST, BLAST-2, ALIGN or

Megalign (DNASTAR) software. Those skilled in the art can determine appropriate parameters for measuring alignment, including any algorithms needed to achieve maximal alignment over the full length of the sequences being compared. For purposes herein, however, % amino acid sequence identity values are generated using the sequence comparison computer program ALIGN-2, wherein the complete source code 5 for the ALIGN-2 program is provided in Table 1 below. The ALIGN-2 sequence comparison computer program was authored by Genentech, Inc. and the source code shown in Table 1 below has been filed with user documentation in the U.S. Copyright Office, Washington D.C., 20559, where it is registered under U.S. Copyright Registration No. TXU510087. The ALIGN-2 program is publicly available through Genentech, Inc., South San Francisco, California or may be compiled from the source code provided in Table 1 below. 10 The ALIGN-2 program should be compiled for use on a UNIX operating system, preferably digital UNIX V4.0D. All sequence comparison parameters are set by the ALIGN-2 program and do not vary.

In situations where ALIGN-2 is employed for amino acid sequence comparisons, the % amino acid sequence identity of a given amino acid sequence A to, with, or against a given amino acid sequence B (which can alternatively be phrased as a given amino acid sequence A that has or comprises a certain % 15 amino acid sequence identity to, with, or against a given amino acid sequence B) is calculated as follows:

$$100 \text{ times the fraction } X/Y$$

where X is the number of amino acid residues scored as identical matches by the sequence alignment 20 program ALIGN-2 in that program's alignment of A and B, and where Y is the total number of amino acid residues in B. It will be appreciated that where the length of amino acid sequence A is not equal to the length of amino acid sequence B, the % amino acid sequence identity of A to B will not equal the % amino acid sequence identity of B to A. As examples of % amino acid sequence identity calculations using this 25 method, Tables 2 and 3 demonstrate how to calculate the % amino acid sequence identity of the amino acid sequence designated "Comparison Protein" to the amino acid sequence designated "PRO", wherein "PRO" represents the amino acid sequence of a hypothetical PRO polypeptide of interest, "Comparison Protein" represents the amino acid sequence of a polypeptide against which the "PRO" polypeptide of interest is being compared, and "X", "Y" and "Z" each represent different hypothetical amino acid residues.

Unless specifically stated otherwise, all % amino acid sequence identity values used herein are 30 obtained as described in the immediately preceding paragraph using the ALIGN-2 computer program. However, % amino acid sequence identity values may also be obtained as described below by using the WU-BLAST-2 computer program (Altschul et al., *Methods in Enzymology* 266:460-480 (1996)). Most of the WU-BLAST-2 search parameters are set to the default values. Those not set to default values, i.e., the 35 adjustable parameters, are set with the following values: overlap span = 1, overlap fraction = 0.125, word threshold (T) = 11, and scoring matrix = BLOSUM62. When WU-BLAST-2 is employed, a % amino acid sequence identity value is determined by dividing (a) the number of matching identical amino acid residues between the amino acid sequence of the PRO polypeptide of interest having a sequence derived from the native PRO polypeptide and the comparison amino acid sequence of interest (i.e., the sequence against which the PRO polypeptide of interest is being compared which may be a PRO variant polypeptide) as 40 determined by WU-BLAST-2 by (b) the total number of amino acid residues of the PRO polypeptide of

interest. For example, in the statement "a polypeptide comprising an the amino acid sequence A which has or having at least 80% amino acid sequence identity to the amino acid sequence B", the amino acid sequence A is the comparison amino acid sequence of interest and the amino acid sequence B is the amino acid sequence of the PRO polypeptide of interest.

5 Percent amino acid sequence identity may also be determined using the sequence comparison program NCBI-BLAST2 (Altschul et al., *Nucleic Acids Res.* 25:3389-3402 (1997)). The NCBI-BLAST2 sequence comparison program may be downloaded from <http://www.ncbi.nlm.nih.gov> or otherwise obtained from the National Institute of Health, Bethesda, MD. NCBI-BLAST2 uses several search parameters, wherein all of those search parameters are set to default values including, for example, unmask = yes, strand 10 = all, expected occurrences = 10, minimum low complexity length = 15/5, multi-pass e-value = 0.01, constant for multi-pass = 25, dropoff for final gapped alignment = 25 and scoring matrix = BLOSUM62.

In situations where NCBI-BLAST2 is employed for amino acid sequence comparisons, the % amino acid sequence identity of a given amino acid sequence A to, with, or against a given amino acid sequence B (which can alternatively be phrased as a given amino acid sequence A that has or comprises a 15 certain % amino acid sequence identity to, with, or against a given amino acid sequence B) is calculated as follows:

$$100 \text{ times the fraction } X/Y$$

20 where X is the number of amino acid residues scored as identical matches by the sequence alignment program NCBI-BLAST2 in that program's alignment of A and B, and where Y is the total number of amino acid residues in B. It will be appreciated that where the length of amino acid sequence A is not equal to the length of amino acid sequence B, the % amino acid sequence identity of A to B will not equal the % amino acid sequence identity of B to A.

25 "PRO variant polynucleotide" or "PRO variant nucleic acid sequence" means a nucleic acid molecule which encodes an active PRO polypeptide as defined below and which has at least about 80% nucleic acid sequence identity with a nucleotide acid sequence encoding a full-length native sequence PRO polypeptide sequence as disclosed herein, a full-length native sequence PRO polypeptide sequence lacking the signal peptide as disclosed herein, an extracellular domain of a PRO polypeptide, with or without the 30 signal peptide, as disclosed herein or any other fragment of a full-length PRO polypeptide sequence as disclosed herein. Ordinarily, a PRO variant polynucleotide will have at least about 80% nucleic acid sequence identity, alternatively at least about 81% nucleic acid sequence identity, alternatively at least about 82% nucleic acid sequence identity, alternatively at least about 83% nucleic acid sequence identity, alternatively at least about 84% nucleic acid sequence identity, alternatively at least about 85% nucleic acid 35 sequence identity, alternatively at least about 86% nucleic acid sequence identity, alternatively at least about 87% nucleic acid sequence identity, alternatively at least about 88% nucleic acid sequence identity, alternatively at least about 89% nucleic acid sequence identity, alternatively at least about 90% nucleic acid sequence identity, alternatively at least about 91% nucleic acid sequence identity, alternatively at least about 92% nucleic acid sequence identity, alternatively at least about 93% nucleic acid sequence identity, 40 alternatively at least about 94% nucleic acid sequence identity, alternatively at least about 95% nucleic acid

sequence identity, alternatively at least about 96% nucleic acid sequence identity, alternatively at least about 97% nucleic acid sequence identity, alternatively at least about 98% nucleic acid sequence identity and alternatively at least about 99% nucleic acid sequence identity with a nucleic acid sequence encoding a full-length native sequence PRO polypeptide sequence as disclosed herein, a full-length native sequence PRO

5 polypeptide sequence lacking the signal peptide as disclosed herein, an extracellular domain of a PRO polypeptide, with or without the signal sequence, as disclosed herein or any other fragment of a full-length PRO polypeptide sequence as disclosed herein. Variants do not encompass the native nucleotide sequence.

Ordinarily, PRO variant polynucleotides are at least about 30 nucleotides in length, alternatively at least about 60 nucleotides in length, alternatively at least about 90 nucleotides in length, alternatively at least 10 about 120 nucleotides in length, alternatively at least about 150 nucleotides in length, alternatively at least about 180 nucleotides in length, alternatively at least about 210 nucleotides in length, alternatively at least about 240 nucleotides in length, alternatively at least about 270 nucleotides in length, alternatively at least about 300 nucleotides in length, alternatively at least about 450 nucleotides in length, alternatively at least about 600 nucleotides in length, alternatively at least about 900 nucleotides in length, or more.

15 "Percent (%) nucleic acid sequence identity" with respect to PRO-encoding nucleic acid sequences identified herein is defined as the percentage of nucleotides in a candidate sequence that are identical with the nucleotides in the PRO nucleic acid sequence of interest, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity. Alignment for purposes of determining percent nucleic acid sequence identity can be achieved in various ways that are within the skill 20 in the art, for instance, using publicly available computer software such as BLAST, BLAST-2, ALIGN or Megalign (DNASTAR) software. For purposes herein, however, % nucleic acid sequence identity values are generated using the sequence comparison computer program ALIGN-2, wherein the complete source code for the ALIGN-2 program is provided in Table 1 below. The ALIGN-2 sequence comparison computer program was authored by Genentech, Inc. and the source code shown in Table 1 below has been filed with 25 user documentation in the U.S. Copyright Office, Washington D.C., 20559, where it is registered under U.S. Copyright Registration No. TXU510087. The ALIGN-2 program is publicly available through Genentech, Inc., South San Francisco, California or may be compiled from the source code provided in Table 1 below. The ALIGN-2 program should be compiled for use on a UNIX operating system, preferably digital UNIX V4.0D. All sequence comparison parameters are set by the ALIGN-2 program and do not vary.

30 In situations where ALIGN-2 is employed for nucleic acid sequence comparisons, the % nucleic acid sequence identity of a given nucleic acid sequence C to, with, or against a given nucleic acid sequence D (which can alternatively be phrased as a given nucleic acid sequence C that has or comprises a certain % nucleic acid sequence identity to, with, or against a given nucleic acid sequence D) is calculated as follows:

35 100 times the fraction W/Z

where W is the number of nucleotides scored as identical matches by the sequence alignment program ALIGN-2 in that program's alignment of C and D, and where Z is the total number of nucleotides in D. It will be appreciated that where the length of nucleic acid sequence C is not equal to the length of nucleic acid 40 sequence D, the % nucleic acid sequence identity of C to D will not equal the % nucleic acid sequence

identity of D to C. As examples of % nucleic acid sequence identity calculations, Tables 4 and 5, demonstrate how to calculate the % nucleic acid sequence identity of the nucleic acid sequence designated "Comparison DNA" to the nucleic acid sequence designated "PRO-DNA", wherein "PRO-DNA" represents a hypothetical PRO-encoding nucleic acid sequence of interest, "Comparison DNA" represents the 5 nucleotide sequence of a nucleic acid molecule against which the "PRO-DNA" nucleic acid molecule of interest is being compared, and "N", "L" and "V" each represent different hypothetical nucleotides.

Unless specifically stated otherwise, all % nucleic acid sequence identity values used herein are obtained as described in the immediately preceding paragraph using the ALIGN-2 computer program.

However, % nucleic acid sequence identity values may also be obtained as described below by using the

10 WU-BLAST-2 computer program (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Most of the WU-BLAST-2 search parameters are set to the default values. Those not set to default values, i.e., the adjustable parameters, are set with the following values: overlap span = 1, overlap fraction = 0.125, word threshold (T) = 11, and scoring matrix = BLOSUM62. When WU-BLAST-2 is employed, a % nucleic acid sequence identity value is determined by dividing (a) the number of matching identical nucleotides between 15 the nucleic acid sequence of the PRO polypeptide-encoding nucleic acid molecule of interest having a sequence derived from the native sequence PRO polypeptide-encoding nucleic acid and the comparison nucleic acid molecule of interest (i.e., the sequence against which the PRO polypeptide-encoding nucleic acid molecule of interest is being compared which may be a variant PRO polynucleotide) as determined by WU-BLAST-2 by (b) the total number of nucleotides of the PRO polypeptide-encoding nucleic acid 20 molecule of interest. For example, in the statement "an isolated nucleic acid molecule comprising a nucleic acid sequence A which has or having at least 80% nucleic acid sequence identity to the nucleic acid sequence B", the nucleic acid sequence A is the comparison nucleic acid molecule of interest and the nucleic acid sequence B is the nucleic acid sequence of the PRO polypeptide-encoding nucleic acid molecule of interest.

25 Percent nucleic acid sequence identity may also be determined using the sequence comparison program NCBI-BLAST2 (Altschul et al., Nucleic Acids Res. 25:3389-3402 (1997)). The NCBI-BLAST2 sequence comparison program may be downloaded from <http://www.ncbi.nlm.nih.gov> or otherwise obtained from the National Institute of Health, Bethesda, MD. NCBI-BLAST2 uses several search parameters, wherein all of those search parameters are set to default values including, for example, unmask = yes, strand 30 = all, expected occurrences = 10, minimum low complexity length = 15/5, multi-pass e-value = 0.01, constant for multi-pass = 25, dropoff for final gapped alignment = 25 and scoring matrix = BLOSUM62.

In situations where NCBI-BLAST2 is employed for sequence comparisons, the % nucleic acid sequence identity of a given nucleic acid sequence C to, with, or against a given nucleic acid sequence D (which can alternatively be phrased as a given nucleic acid sequence C that has or comprises a certain % 35 nucleic acid sequence identity to, with, or against a given nucleic acid sequence D) is calculated as follows:

$$100 \text{ times the fraction W/Z}$$

where W is the number of nucleotides scored as identical matches by the sequence alignment program 40 NCBI-BLAST2 in that program's alignment of C and D, and where Z is the total number of nucleotides in

D. It will be appreciated that where the length of nucleic acid sequence C is not equal to the length of nucleic acid sequence D, the % nucleic acid sequence identity of C to D will not equal the % nucleic acid sequence identity of D to C.

In other embodiments, PRO variant polynucleotides are nucleic acid molecules that encode an active PRO polypeptide and which are capable of hybridizing, preferably under stringent hybridization and wash conditions, to nucleotide sequences encoding a full-length PRO polypeptide as disclosed herein. PRO variant polypeptides may be those that are encoded by a PRO variant polynucleotide.

"Isolated," when used to describe the various polypeptides disclosed herein, means polypeptide that has been identified and separated and/or recovered from a component of its natural environment.

10 Contaminant components of its natural environment are materials that would typically interfere with diagnostic or therapeutic uses for the polypeptide, and may include enzymes, hormones, and other proteinaceous or non-proteinaceous solutes. In preferred embodiments, the polypeptide will be purified (1) to a degree sufficient to obtain at least 15 residues of N-terminal or internal amino acid sequence by use of a spinning cup sequenator, or (2) to homogeneity by SDS-PAGE under non-reducing or reducing conditions 15 using Coomassie blue or, preferably, silver stain. Isolated polypeptide includes polypeptide *in situ* within recombinant cells, since at least one component of the PRO polypeptide natural environment will not be present. Ordinarily, however, isolated polypeptide will be prepared by at least one purification step.

An "isolated" PRO polypeptide-encoding nucleic acid or other polypeptide-encoding nucleic acid is a nucleic acid molecule that is identified and separated from at least one contaminant nucleic acid molecule 20 with which it is ordinarily associated in the natural source of the polypeptide-encoding nucleic acid. An isolated polypeptide-encoding nucleic acid molecule is other than in the form or setting in which it is found in nature. Isolated polypeptide-encoding nucleic acid molecules therefore are distinguished from the specific polypeptide-encoding nucleic acid molecule as it exists in natural cells. However, an isolated polypeptide-encoding nucleic acid molecule includes polypeptide-encoding nucleic acid molecules 25 contained in cells that ordinarily express the polypeptide where, for example, the nucleic acid molecule is in a chromosomal location different from that of natural cells.

The term "control sequences" refers to DNA sequences necessary for the expression of an operably linked coding sequence in a particular host organism. The control sequences that are suitable for prokaryotes, for example, include a promoter, optionally an operator sequence, and a ribosome binding site. 30 Eukaryotic cells are known to utilize promoters, polyadenylation signals, and enhancers.

Nucleic acid is "operably linked" when it is placed into a functional relationship with another nucleic acid sequence. For example, DNA for a presequence or secretory leader is operably linked to DNA for a polypeptide if it is expressed as a preprotein that participates in the secretion of the polypeptide; a promoter or enhancer is operably linked to a coding sequence if it affects the transcription of the sequence; 35 or a ribosome binding site is operably linked to a coding sequence if it is positioned so as to facilitate translation. Generally, "operably linked" means that the DNA sequences being linked are contiguous, and, in the case of a secretory leader, contiguous and in reading phase. However, enhancers do not have to be contiguous. Linking is accomplished by ligation at convenient restriction sites. If such sites do not exist, the synthetic oligonucleotide adaptors or linkers are used in accordance with conventional practice.

The term "antibody" is used in the broadest sense and specifically covers, for example, single anti-PRO monoclonal antibodies (including agonist, antagonist, and neutralizing antibodies), anti-PRO antibody compositions with polyepitopic specificity, single chain anti-PRO antibodies, and fragments of anti-PRO antibodies (see below). The term "monoclonal antibody" as used herein refers to an antibody obtained from 5 a population of substantially homogeneous antibodies, i.e., the individual antibodies comprising the population are identical except for possible naturally-occurring mutations that may be present in minor amounts.

"Stringency" of hybridization reactions is readily determinable by one of ordinary skill in the art, and generally is an empirical calculation dependent upon probe length, washing temperature, and salt 10 concentration. In general, longer probes require higher temperatures for proper annealing, while shorter probes need lower temperatures. Hybridization generally depends on the ability of denatured DNA to reanneal when complementary strands are present in an environment below their melting temperature. The higher the degree of desired homology between the probe and hybridizable sequence, the higher the relative temperature which can be used. As a result, it follows that higher relative temperatures would tend to make 15 the reaction conditions more stringent, while lower temperatures less so. For additional details and explanation of stringency of hybridization reactions, see Ausubel et al., Current Protocols in Molecular Biology, Wiley Interscience Publishers, (1995).

"Stringent conditions" or "high stringency conditions", as defined herein, may be identified by those that: (1) employ low ionic strength and high temperature for washing, for example 0.015 M sodium 20 chloride/0.0015 M sodium citrate/0.1% sodium dodecyl sulfate at 50°C; (2) employ during hybridization a denaturing agent, such as formamide, for example, 50% (v/v) formamide with 0.1% bovine serum albumin/0.1% Ficoll/0.1% polyvinylpyrrolidone/50mM sodium phosphate buffer at pH 6.5 with 750 mM sodium chloride, 75 mM sodium citrate at 42°C; or (3) employ 50% formamide, 5 x SSC (0.75 M NaCl, 0.075 M sodium citrate), 50 mM sodium phosphate (pH 6.8), 0.1% sodium pyrophosphate, 5 x Denhardt's 25 solution, sonicated salmon sperm DNA (50 µg/ml), 0.1% SDS, and 10% dextran sulfate at 42°C, with washes at 42°C in 0.2 x SSC (sodium chloride/sodium citrate) and 50% formamide at 55°C, followed by a high-stringency wash consisting of 0.1 x SSC containing EDTA at 55°C.

"Moderately stringent conditions" may be identified as described by Sambrook et al., Molecular Cloning: A Laboratory Manual, New York: Cold Spring Harbor Press, 1989, and include the use of washing 30 solution and hybridization conditions (e.g., temperature, ionic strength and %SDS) less stringent than those described above. An example of moderately stringent conditions is overnight incubation at 37°C in a solution comprising: 20% formamide, 5 x SSC (150 mM NaCl, 15 mM trisodium citrate), 50 mM sodium phosphate (pH 7.6), 5 x Denhardt's solution, 10% dextran sulfate, and 20 mg/ml denatured sheared salmon sperm DNA, followed by washing the filters in 1 x SSC at about 37-50°C. The skilled artisan will recognize 35 how to adjust the temperature, ionic strength, etc. as necessary to accommodate factors such as probe length and the like.

The term "epitope tagged" when used herein refers to a chimeric polypeptide comprising a PRO polypeptide fused to a "tag polypeptide". The tag polypeptide has enough residues to provide an epitope against which an antibody can be made, yet is short enough such that it does not interfere with activity of the 40 polypeptide to which it is fused. The tag polypeptide preferably also is fairly unique so that the antibody

does not substantially cross-react with other epitopes. Suitable tag polypeptides generally have at least six amino acid residues and usually between about 8 and 50 amino acid residues (preferably, between about 10 and 20 amino acid residues).

As used herein, the term "immunoadhesin" designates antibody-like molecules which combine the binding specificity of a heterologous protein (an "adhesin") with the effector functions of immunoglobulin constant domains. Structurally, the immunoadhesins comprise a fusion of an amino acid sequence with the desired binding specificity which is other than the antigen recognition and binding site of an antibody (i.e., is "heterologous"), and an immunoglobulin constant domain sequence. The adhesin part of an immunoadhesin molecule typically is a contiguous amino acid sequence comprising at least the binding site of a receptor or a ligand. The immunoglobulin constant domain sequence in the immunoadhesin may be obtained from any immunoglobulin, such as IgG-1, IgG-2, IgG-3, or IgG-4 subtypes, IgA (including IgA-1 and IgA-2), IgE, IgD or IgM.

"Active" or "activity" for the purposes herein refers to form(s) of a PRO polypeptide which retain a biological and/or an immunological activity of native or naturally-occurring PRO, wherein "biological" activity refers to a biological function (either inhibitory or stimulatory) caused by a native or naturally-occurring PRO other than the ability to induce the production of an antibody against an antigenic epitope possessed by a native or naturally-occurring PRO and an "immunological" activity refers to the ability to induce the production of an antibody against an antigenic epitope possessed by a native or naturally-occurring PRO.

The term "antagonist" is used in the broadest sense, and includes any molecule that partially or fully blocks, inhibits, or neutralizes a biological activity of a native PRO polypeptide disclosed herein. In a similar manner, the term "agonist" is used in the broadest sense and includes any molecule that mimics a biological activity of a native PRO polypeptide disclosed herein. Suitable agonist or antagonist molecules specifically include agonist or antagonist antibodies or antibody fragments, fragments or amino acid sequence variants of native PRO polypeptides, peptides, antisense oligonucleotides, small organic molecules, etc. Methods for identifying agonists or antagonists of a PRO polypeptide may comprise contacting a PRO polypeptide with a candidate agonist or antagonist molecule and measuring a detectable change in one or more biological activities normally associated with the PRO polypeptide.

"Treatment" refers to both therapeutic treatment and prophylactic or preventative measures, wherein the object is to prevent or slow down (lessen) the targeted pathologic condition or disorder. Those in need of treatment include those already with the disorder as well as those prone to have the disorder or those in whom the disorder is to be prevented.

"Chronic" administration refers to administration of the agent(s) in a continuous mode as opposed to an acute mode, so as to maintain the initial therapeutic effect (activity) for an extended period of time. "Intermittent" administration is treatment that is not consecutively done without interruption, but rather is cyclic in nature.

"Mammal" for purposes of treatment refers to any animal classified as a mammal, including humans, domestic and farm animals, and zoo, sports, or pet animals, such as dogs, cats, cattle, horses, sheep, pigs, goats, rabbits, etc. Preferably, the mammal is human.

Administration "in combination with" one or more further therapeutic agents includes simultaneous (concurrent) and consecutive administration in any order.

"Carriers" as used herein include pharmaceutically acceptable carriers, excipients, or stabilizers which are nontoxic to the cell or mammal being exposed thereto at the dosages and concentrations

5 employed. Often the physiologically acceptable carrier is an aqueous pH buffered solution. Examples of physiologically acceptable carriers include buffers such as phosphate, citrate, and other organic acids; antioxidants including ascorbic acid; low molecular weight (less than about 10 residues) polypeptide; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino acids such as glycine, glutamine, asparagine, arginine or lysine; 10 monosaccharides, disaccharides, and other carbohydrates including glucose, mannose, or dextrins; chelating agents such as EDTA; sugar alcohols such as mannitol or sorbitol; salt-forming counterions such as sodium; and/or nonionic surfactants such as TWEEN™, polyethylene glycol (PEG), and PLURONICS™.

15 "Antibody fragments" comprise a portion of an intact antibody, preferably the antigen binding or variable region of the intact antibody. Examples of antibody fragments include Fab, Fab', F(ab')<sub>2</sub>, and Fv fragments; diabodies; linear antibodies (Zapata et al., *Protein Eng.* 8(10): 1057-1062 [1995]); single-chain antibody molecules; and multispecific antibodies formed from antibody fragments.

20 Papain digestion of antibodies produces two identical antigen-binding fragments, called "Fab" fragments, each with a single antigen-binding site, and a residual "Fc" fragment, a designation reflecting the ability to crystallize readily. Pepsin treatment yields an F(ab')<sub>2</sub> fragment that has two antigen-combining sites and is still capable of cross-linking antigen.

25 "Fv" is the minimum antibody fragment which contains a complete antigen-recognition and - binding site. This region consists of a dimer of one heavy- and one light-chain variable domain in tight, non-covalent association. It is in this configuration that the three CDRs of each variable domain interact to define an antigen-binding site on the surface of the V<sub>H</sub>-V<sub>L</sub> dimer. Collectively, the six CDRs confer antigen- binding specificity to the antibody. However, even a single variable domain (or half of an Fv comprising only three CDRs specific for an antigen) has the ability to recognize and bind antigen, although at a lower affinity than the entire binding site.

30 The Fab fragment also contains the constant domain of the light chain and the first constant domain (CH1) of the heavy chain. Fab fragments differ from Fab' fragments by the addition of a few residues at the carboxy terminus of the heavy chain CH1 domain including one or more cysteines from the antibody hinge region. Fab'-SH is the designation herein for Fab' in which the cysteine residue(s) of the constant domains bear a free thiol group. F(ab')<sub>2</sub> antibody fragments originally were produced as pairs of Fab' fragments which have hinge cysteines between them. Other chemical couplings of antibody fragments are also known.

35 The "light chains" of antibodies (immunoglobulins) from any vertebrate species can be assigned to one of two clearly distinct types, called kappa and lambda, based on the amino acid sequences of their constant domains.

Depending on the amino acid sequence of the constant domain of their heavy chains, immunoglobulins can be assigned to different classes. There are five major classes of immunoglobulins:

IgA, IgD, IgE, IgG, and IgM, and several of these may be further divided into subclasses (isotypes), e.g., IgG1, IgG2, IgG3, IgG4, IgA, and IgA2.

"Single-chain Fv" or "sFv" antibody fragments comprise the  $V_H$  and  $V_L$  domains of antibody, wherein these domains are present in a single polypeptide chain. Preferably, the Fv polypeptide further comprises a polypeptide linker between the  $V_H$  and  $V_L$  domains which enables the sFv to form the desired structure for antigen binding. For a review of sFv, see Pluckthun in The Pharmacology of Monoclonal Antibodies, vol. 113, Rosenburg and Moore eds., Springer-Verlag, New York, pp. 269-315 (1994).

The term "diabodies" refers to small antibody fragments with two antigen-binding sites, which fragments comprise a heavy-chain variable domain ( $V_H$ ) connected to a light-chain variable domain ( $V_L$ ) in the same polypeptide chain ( $V_H-V_L$ ). By using a linker that is too short to allow pairing between the two domains on the same chain, the domains are forced to pair with the complementary domains of another chain and create two antigen-binding sites. Diabodies are described more fully in, for example, EP 404,097; WO 93/11161; and Hollinger et al., Proc. Natl. Acad. Sci. USA, 90:6444-6448 (1993).

An "isolated" antibody is one which has been identified and separated and/or recovered from a component of its natural environment. Contaminant components of its natural environment are materials which would interfere with diagnostic or therapeutic uses for the antibody, and may include enzymes, hormones, and other proteinaceous or nonproteinaceous solutes. In preferred embodiments, the antibody will be purified (1) to greater than 95% by weight of antibody as determined by the Lowry method, and most preferably more than 99% by weight, (2) to a degree sufficient to obtain at least 15 residues of N-terminal or internal amino acid sequence by use of a spinning cup sequenator, or (3) to homogeneity by SDS-PAGE under reducing or nonreducing conditions using Coomassie blue or, preferably, silver stain. Isolated antibody includes the antibody *in situ* within recombinant cells since at least one component of the antibody's natural environment will not be present. Ordinarily, however, isolated antibody will be prepared by at least one purification step.

An antibody that "specifically binds to" or is "specific for" a particular polypeptide or an epitope on a particular polypeptide is one that binds to that particular polypeptide or epitope on a particular polypeptide without substantially binding to any other polypeptide or polypeptide epitope.

The word "label" when used herein refers to a detectable compound or composition which is conjugated directly or indirectly to the antibody so as to generate a "labeled" antibody. The label may be detectable by itself (e.g. radioisotope labels or fluorescent labels) or, in the case of an enzymatic label, may catalyze chemical alteration of a substrate compound or composition which is detectable.

By "solid phase" is meant a non-aqueous matrix to which the antibody of the present invention can adhere. Examples of solid phases encompassed herein include those formed partially or entirely of glass (e.g., controlled pore glass), polysaccharides (e.g., agarose), polyacrylamides, polystyrene, polyvinyl alcohol and silicones. In certain embodiments, depending on the context, the solid phase can comprise the well of an assay plate; in others it is a purification column (e.g., an affinity chromatography column). This term also includes a discontinuous solid phase of discrete particles, such as those described in U.S. Patent No. 4,275,149.

A "liposome" is a small vesicle composed of various types of lipids, phospholipids and/or surfactant which is useful for delivery of a drug (such as a PRO polypeptide or antibody thereto) to a mammal. The components of the liposome are commonly arranged in a bilayer formation, similar to the lipid arrangement of biological membranes.

5 A "small molecule" is defined herein to have a molecular weight below about 500 Daltons.

The term "immune related disease" means a disease in which a component of the immune system of a mammal causes, mediates or otherwise contributes to a morbidity in the mammal. Also included are diseases in which stimulation or intervention of the immune response has an ameliorative effect on progression of the disease. Included within this term are immune-mediated inflammatory diseases, non-  
10 immune-mediated inflammatory diseases, infectious diseases, immunodeficiency diseases, neoplasia, *etc.*

The term "T cell mediated disease" means a disease in which T cells directly or indirectly mediate or otherwise contribute to a morbidity in a mammal. The T cell mediated disease may be associated with cell mediated effects, lymphokine mediated effects, *etc.*, and even effects associated with B cells if the B cells are stimulated, for example, by the lymphokines secreted by T cells.

15 Examples of immune-related and inflammatory diseases, some of which are immune or T cell mediated, which can be treated according to the invention include systemic lupus erythematosis, rheumatoid arthritis, juvenile chronic arthritis, spondyloarthropathies, systemic sclerosis (scleroderma), idiopathic inflammatory myopathies (dermatomyositis, polymyositis), Sjögren's syndrome, systemic vasculitis, sarcoidosis, autoimmune hemolytic anemia (immune pancytopenia, paroxysmal nocturnal hemoglobinuria),  
20 autoimmune thrombocytopenia (idiopathic thrombocytopenic purpura, immune-mediated thrombocytopenia), thyroiditis (Grave's disease, Hashimoto's thyroiditis, juvenile lymphocytic thyroiditis, atrophic thyroiditis), diabetes mellitus, immune-mediated renal disease (glomerulonephritis, tubulointerstitial nephritis), demyelinating diseases of the central and peripheral nervous systems such as multiple sclerosis, idiopathic demyelinating polyneuropathy or Guillain-Barré syndrome, and chronic inflammatory  
25 demyelinating polyneuropathy, hepatobiliary diseases such as infectious hepatitis (hepatitis A, B, C, D, E and other non-hepatotropic viruses), autoimmune chronic active hepatitis, primary biliary cirrhosis, granulomatous hepatitis, and sclerosing cholangitis, inflammatory bowel disease (ulcerative colitis: Crohn's disease), gluten-sensitive enteropathy, and Whipple's disease, autoimmune or immune-mediated skin diseases including bullous skin diseases, erythema multiforme and contact dermatitis, psoriasis, allergic  
30 diseases such as asthma, allergic rhinitis, atopic dermatitis, food hypersensitivity and urticaria, immunologic diseases of the lung such as eosinophilic pneumonias, idiopathic pulmonary fibrosis and hypersensitivity pneumonitis, transplantation associated diseases including graft rejection and graft -versus-host-disease. Infectious diseases including viral diseases such as AIDS (HIV infection), hepatitis A, B, C, D, and E, herpes, *etc.*, bacterial infections, fungal infections, protozoal infections and parasitic infections.

35 The term "effective amount" is a concentration or amount of a PRO polypeptide and/or agonist/antagonist which results in achieving a particular stated purpose. An "effective amount" of a PRO polypeptide or agonist or antagonist thereof may be determined empirically. Furthermore, a "therapeutically effective amount" is a concentration or amount of a PRO polypeptide and/or agonist/antagonist which is effective for achieving a stated therapeutic effect. This amount may also be determined empirically.

40 The term "cytotoxic agent" as used herein refers to a substance that inhibits or prevents the function

of cells and/or causes destruction of cells. The term is intended to include radioactive isotopes (e.g.,  $I^{131}$ ,  $I^{125}$ ,  $Y^{90}$  and  $Re^{186}$ ), chemotherapeutic agents, and toxins such as enzymatically active toxins of bacterial, fungal, plant or animal origin, or fragments thereof.

A "chemotherapeutic agent" is a chemical compound useful in the treatment of cancer. Examples of chemotherapeutic agents include adriamycin, doxorubicin, epirubicin, 5-fluorouracil, cytosine arabinoside ("Ara-C"), cyclophosphamide, thiotepa, busulfan, cytoxin, taxoids, e.g., paclitaxel (Taxol, Bristol-Myers Squibb Oncology, Princeton, NJ), and doxetaxel (Taxotere, Rhône-Poulenc Rorer, Antony, France), toxotere, methotrexate, cisplatin, melphalan, vinblastine, bleomycin, etoposide, ifosfamide, mitomycin C, mitoxantrone, vincristine, vinorelbine, carboplatin, teniposide, daunomycin, carminomycin, aminopterin, dactinomycin, mitomycins, esperamicins (see U.S. Pat. No. 4,675,187), melphalan and other related nitrogen mustards. Also included in this definition are hormonal agents that act to regulate or inhibit hormone action on tumors such as tamoxifen and onapristone.

A "growth inhibitory agent" when used herein refers to a compound or composition which inhibits growth of a cell, especially cancer cell overexpressing any of the genes identified herein, either *in vitro* or *in vivo*. Thus, the growth inhibitory agent is one which significantly reduces the percentage of cells overexpressing such genes in S phase. Examples of growth inhibitory agents include agents that block cell cycle progression (at a place other than S phase), such as agents that induce G1 arrest and M-phase arrest. Classical M-phase blockers include the vincas (vincristine and vinblastine), taxol, and topo II inhibitors such as doxorubicin, epirubicin, daunorubicin, etoposide, and bleomycin. Those agents that arrest G1 also spill over into S-phase arrest, for example, DNA alkylating agents such as tamoxifen, prednisone, dacarbazine, mechlorethamine, cisplatin, methotrexate, 5-fluorouracil, and ara-C. Further information can be found in *The Molecular Basis of Cancer*, Mendelsohn and Israel, eds., Chapter 1, entitled "Cell cycle regulation, oncogens, and antineoplastic drugs" by Murakami *et al.* (WB Saunders: Philadelphia, 1995), especially p. 13.

The term "cytokine" is a generic term for proteins released by one cell population which act on another cell as intercellular mediators. Examples of such cytokines are lymphokines, monokines, and traditional polypeptide hormones. Included among the cytokines are growth hormone such as human growth hormone, N-methionyl human growth hormone, and bovine growth hormone; parathyroid hormone; thyroxine; insulin; proinsulin; relaxin; prorelaxin; glycoprotein hormones such as follicle stimulating hormone (FSH), thyroid stimulating hormone (TSH), and luteinizing hormone (LH); hepatic growth factor; fibroblast growth factor; prolactin; placental lactogen; tumor necrosis factor- $\alpha$  and - $\beta$ ; mullerian-inhibiting substance; mouse gonadotropin-associated peptide; inhibin; activin; vascular endothelial growth factor; integrin; thrombopoietin (TPO); nerve growth factors such as NGF- $\beta$ ; platelet-growth factor; transforming growth factors (TGFs) such as TGF- $\alpha$  and TGF- $\beta$ ; insulin-like growth factor-I and -II; erythropoietin (EPO); osteoinductive factors; interferons such as interferon- $\alpha$ , - $\beta$ , and - $\gamma$ ; colony stimulating factors (CSFs) such as macrophage-CSF (M-CSF); granulocyte-macrophage-CSF (GM-CSF); and granulocyte-CSF (G-CSF); interleukins (ILs) such as IL-1, IL-1 $\alpha$ , IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-11, IL-12; a tumor necrosis factor such as TNF- $\alpha$  or TNF- $\beta$ ; and other polypeptide factors including LIF and kit ligand (KL). As used herein, the term cytokine includes proteins from natural sources or from recombinant cell culture

and biologically active equivalents of the native sequence cytokines.

As used herein, the term "immunoadhesin" designates antibody-like molecules which combine the binding specificity of a heterologous protein (an "adhesin") with the effector functions of immunoglobulin constant domains. Structurally, the immunoadhesins comprise a fusion of an amino acid sequence with the desired binding specificity which is other than the antigen recognition and binding site of an antibody (*i.e.*, is "heterologous"), and an immunoglobulin constant domain sequence. The adhesin part of an immunoadhesin molecule typically is a contiguous amino acid sequence comprising at least the binding site of a receptor or a ligand. The immunoglobulin constant domain sequence in the immunoadhesin may be obtained from any immunoglobulin, such as IgG-1, IgG-2, IgG-3, or IgG-4 subtypes, IgA (including IgA-1 and IgA-2), IgE, IgD or IgM.

As used herein, the term "inflammatory cells" designates cells that enhance the inflammatory response such as mononuclear cells, eosinophils, macrophages, and polymorphonuclear neutrophils (PMN).

**Table 1**

**Table 1 (cont')**

```

/*
*/
5  #include <stdio.h>
# include <ctype.h>

#define MAXJMP 16      /* max jumps in a diag */
#define MAXGAP 24      /* don't continue to penalize gaps larger than this */
10 #define J MPS 1024   /* max jmps in an path */
#define MX 4           /* save if there's at least MX-1 bases since last jmp */

#define DMAT 3          /* value of matching bases */
#define DMIS 0          /* penalty for mismatched bases */
15 #define DINS0 8       /* penalty for a gap */
#define DINS1 1          /* penalty per base */
#define PINS0 8          /* penalty for a gap */
#define PINS1 4          /* penalty per residue */

20 struct jmp {
    short n[MAXJMP];    /* size of jmp (neg for delay) */
    unsigned short x[MAXJMP]; /* base no. of jmp in seq x */
};                      /* limits seq to 2^16 -1 */

25 struct diag {
    int score;           /* score at last jmp */
    long offset;          /* offset of prev block */
    short ijmp;           /* current jmp index */
    struct jmp jp;        /* list of jmps */
30};

35 struct path {
    int spc;             /* number of leading spaces */
    short n[J MPS]; /* size of jmp (gap) */
    int x[J MPS]; /* loc of jmp (last elem before gap) */
};

40 char *ofile;           /* output file name */
char *namex[2];          /* seq names: getseqs() */
char *prog;               /* prog name for err msgs */
char *seqx[2];            /* seqs: getseqs() */
int dmax;                 /* best diag: nw() */
int dmax0;                /* final diag */
45 int dna;                /* set if dna: main() */
int endgaps;              /* set if penalizing end gaps */
int gapx, gapy;           /* total gaps in seqs */
int len0, len1;            /* seq lens */
int ngapx, ngapy;          /* total size of gaps */
int smax;                 /* max score: nw() */
50 int *xbm;                /* bitmap for matching */
long offset;               /* current offset in jmp file */
struct diag *dx;           /* holds diagonals */
struct path pp[2];          /* holds path for seqs */

55 char *calloc(), *malloc(), *index(), *strcpy();
char *getseq(), *g_malloc();

```

**Table 1 (cont')**

```

/* Needleman-Wunsch alignment program
*
* usage: progs file1 file2
* where file1 and file2 are two dna or two protein sequences.
* The sequences can be in upper- or lower-case and may contain ambiguity
* Any lines beginning with ';' or '<' are ignored
* Max file length is 65535 (limited by unsigned short x in the jmp struct)
* A sequence with 1/3 or more of its elements ACGTU is assumed to be DNA
* Output is in the file "align.out"
*
* The program may create a tmp file in /tmp to hold info about traceback.
* Original version developed under BSD 4.3 on a vax 8650
*/
15 #include "nw.h"
#include "day.h"

static _dbval[26] = {
20   1,14,2,13,0,0,4,11,0,0,12,0,3,15,0,0,0,5,6,8,8,7,9,0,10,0
};

static _pbval[26] = {
25   1, 2|(1<<('D'-'A'))|(1<<('N'-'A')), 4, 8, 16, 32, 64,
   128, 256, 0xFFFFFFFF, 1<<10, 1<<11, 1<<12, 1<<13, 1<<14,
   1<<15, 1<<16, 1<<17, 1<<18, 1<<19, 1<<20, 1<<21, 1<<22,
   1<<23, 1<<24, 1<<25|(1<<('E'-'A'))|(1<<('Q'-'A'))
};

main(ac, av)
30   main
      int      ac;
      char    *av[ ];
{
  prog = av[0];
35  if (ac != 3) {
      sprintf(stderr,"usage: %s file1 file2\n", prog);
      sprintf(stderr,"where file1 and file2 are two dna or two protein sequences.\n");
      sprintf(stderr,"The sequences can be in upper- or lower-case\n");
      sprintf(stderr,"Any lines beginning with ';' or '<' are ignored\n");
      sprintf(stderr,"Output is in the file \"align.out\"\n");
      exit(1);
    }
  namex[0] = av[1];
  namex[1] = av[2];
45  seqx[0] = getseq(namex[0], &len0);
  seqx[1] = getseq(namex[1], &len1);
  xbm = (dna)? _dbval : _pbval;

  endgaps = 0;           /* 1 to penalize endgaps */
50  ofile = "align.out";    /* output file */

  nw();                  /* fill in the matrix, get the possible jmps */
  readjmps();            /* get the actual jmps */
  print();                /* print stats, alignment */
55  cleanup(0);            /* unlink any tmp files */
}

```

**Table 1 (cont?)**

```

/* do the alignment, return best score: main()
 * dna: values in Fitch and Smith, PNAS, 80, 1382-1386, 1983
 * pro: PAM 250 values
5  * When scores are equal, we prefer mismatches to any gap, prefer
 * a new gap to extending an ongoing gap, and prefer a gap in seqx
 * to a gap in seq y.
 */
10 nw()
    nw
    {
        char      *px, *py;      /* seqs and ptrs */
        int       *ndely, *dely;  /* keep track of dely */
        int       ndelx, delx;  /* keep track of delx */
        int       *tmp;        /* for swapping row0, row1 */
        int       mis;         /* score for each type */
        int       ins0, ins1;  /* insertion penalties */
        register id;         /* diagonal index */
        register ij;         /* jmp index */
        register *col0, *col1; /* score for curr, last row */
        register xx, yy;     /* index into seqs */
15
20
25
30
35
40
45
50
55
60
    dx = (struct diag *)g_malloc("to get diags", len0+len1+1, sizeof(struct diag));
    ndely = (int *)g_malloc("to get ndely", len1+1, sizeof(int));
    dely = (int *)g_malloc("to get dely", len1+1, sizeof(int));
    col0 = (int *)g_malloc("to get col0", len1+1, sizeof(int));
    col1 = (int *)g_malloc("to get col1", len1+1, sizeof(int));
    ins0 = (dna)? DINS0 : PINS0;
    ins1 = (dna)? DINS1 : PINS1;

    smax = -10000;
    if (endgaps) {
        for (col0[0] = dely[0] = -ins0, yy = 1; yy <= len1; yy++) {
            col0[yy] = dely[yy] = col0[yy-1] - ins1;
            ndely[yy] = yy;
        }
        col0[0] = 0;      /* Waterman Bull Math Biol 84 */
    }
    else
        for (yy = 1; yy <= len1; yy++)
            dely[yy] = -ins0;

    /* fill in match matrix
     */
    for (px = seqx[0], xx = 1; xx <= len0; px++, xx++) {
        /* initialize first entry in col
         */
        if (endgaps) {
            if (xx == 1)
                col1[0] = delx = -(ins0+ins1);
            else
                col1[0] = delx = col0[0] - ins1;
            ndelx = xx;
        }
        else {
            col1[0] = 0;
            delx = -ins0;
            ndelx = 0;
        }
    }

```

**Table 1 (cont')**

...nw

```

5      for (py = seqx[1], yy = 1; yy <= len1; py++, yy++) {
          mis = col0[yy-1];
          if (dna)
              mis += (xbm[*px-'A']&xbm[*py-'A'])? DMAT : DMIS;
          else
              mis += _day[*px-'A'][*py-'A'];

10     /* update penalty for del in x seq;
          * favor new del over ongong del
          * ignore MAXGAP if weighting endgaps
          */
15     if (endgaps || ndely[yy] < MAXGAP) {
          if (col0[yy] - ins0 >= dely[yy]) {
              dely[yy] = col0[yy] - (ins0+ins1);
              ndely[yy] = 1;
          } else {
              dely[yy] -= ins1;
              ndely[yy]++;
          }
      } else {
          if (col0[yy] - (ins0+ins1) >= dely[yy]) {
              dely[yy] = col0[yy] - (ins0+ins1);
              ndely[yy] = 1;
          } else
              ndely[yy]++;
      }

20     /* update penalty for del in y seq;
          * favor new del over ongong del
          */
25     if (endgaps || ndelx < MAXGAP) {
          if (col1[yy-1] - ins0 >= delx) {
              delx = col1[yy-1] - (ins0+ins1);
              ndelx = 1;
          } else {
              delx -= ins1;
              ndelx++;
          }
      } else {
          if (col1[yy-1] - (ins0+ins1) >= delx) {
              delx = col1[yy-1] - (ins0+ins1);
              ndelx = 1;
          } else
              ndelx++;
      }

30     /* pick the maximum score; we're favoring
          * mis over any del and delx over dely
          */
35
40
45
50

```

55

60

**Table 1 (cont'')**

...nw

```

5      id = xx - yy + len1 - 1;
      if (mis >= delx && mis >= dely[yy])
      col1[yy] = mis;
      else if (delx >= dely[yy]) {
      col1[yy] = delx;
      ij = dx[id].ijmp;
      if (dx[id].jp.n[0] && (!dna || (ndelx >= MAXJMP
10      && xx > dx[id].jp.x[ij]+MX) || mis > dx[id].score+DINS0)) {
          dx[id].ijmp++;
          if (++ij >= MAXJMP) {
              writejmps(id);
              ij = dx[id].ijmp = 0;
              dx[id].offset = offset;
              offset += sizeof(struct jmp) + sizeof(offset);
          }
      }
      dx[id].jp.n[ij] = ndelx;
      dx[id].jp.x[ij] = xx;
      dx[id].score = delx;
  }
  else {
      col1[yy] = dely[yy];
      ij = dx[id].ijmp;
25      if (dx[id].jp.n[0] && (!dna || (ndely[yy] >= MAXJMP
          && xx > dx[id].jp.x[ij]+MX) || mis > dx[id].score+DINS0)) {
          dx[id].ijmp++;
          if (++ij >= MAXJMP) {
              writejmps(id);
              ij = dx[id].ijmp = 0;
              dx[id].offset = offset;
              offset += sizeof(struct jmp) + sizeof(offset);
          }
      }
      dx[id].jp.n[ij] = -ndely[yy];
      dx[id].jp.x[ij] = xx;
      dx[id].score = dely[yy];
  }
40      if (xx == len0 && yy < len1) {
          /* last col
          */
          if (endgaps)
              col1[yy] -= ins0+ins1*(len1-yy);
          if (col1[yy] > smax) {
              smax = col1[yy];
              dmax = id;
          }
      }
50      if (endgaps && xx < len0)
          col1[yy-1] -= ins0+ins1*(len0-xx);
      if (col1[yy-1] > smax) {
          smax = col1[yy-1];
          dmax = id;
      }
      tmp = col0; col0 = col1; col1 = tmp;
  }
60      (void) free((char *)ndely);
      (void) free((char *)dely);
      (void) free((char *)col0);
      (void) free((char *)col1);
  }
}

```

**Table 1 (cont')**

```

/*
 *
 * print() -- only routine visible outside this module
 *
 * static:
 * getmat() -- trace back best path, count matches: print()
 * pr_align() -- print alignment of described in array p[ ]; print()
 * dumpblock() -- dump a block of lines with numbers, stars: pr_align()
 * nums() -- put out a number line: dumpblock()
 * putline() -- put out a line (name, [num], seq, [num]): dumpblock()
 * stars() - -put a line of stars: dumpblock()
 * stripname() -- strip any path and prefix from a seqname
 */

15 #include "nw.h"

#define SPC      3
#define P_LINE   256      /* maximum output line */
20 #define P_SPC   3      /* space between name or num and seq */

extern _day[26][26];
int  olen;           /* set output line length */
FILE *fx;            /* output file */

25 print()
{
    print
{
    int    lx, ly, firstgap, lastgap;      /* overlap */
30    if ((fx = fopen(ofile, "w")) == 0) {
        fprintf(stderr,"%s: can't write %s\n", prog, ofile);
        cleanup(1);
    }
35    fprintf(fx, "<first sequence: %s (length = %d)\n", namex[0], len0);
    fprintf(fx, "<second sequence: %s (length = %d)\n", namex[1], len1);
    olen = 60;
    lx = len0;
    ly = len1;
40    firstgap = lastgap = 0;
    if (dmax < len1 - 1) {           /* leading gap in x */
        pp[0].spc = firstgap = len1 - dmax - 1;
        ly -= pp[0].spc;
    }
45    else if (dmax > len1 - 1) {    /* leading gap in y */
        pp[1].spc = firstgap = dmax - (len1 - 1);
        lx -= pp[1].spc;
    }
50    if (dmax0 < len0 - 1) {       /* trailing gap in x */
        lastgap = len0 - dmax0 - 1;
        lx -= lastgap;
    }
55    else if (dmax0 > len0 - 1) { /* trailing gap in y */
        lastgap = dmax0 - (len0 - 1);
        ly -= lastgap;
    }
    getmat(lx, ly, firstgap, lastgap);
    pr_align();
}
60

```

**Table 1 (cont')**

```

/*
 * trace back the best path, count matches
 */
5 static
    getmat(lx, ly, firstgap, lastgap)
        int      lx, ly;          /* "core" (minus endgaps) */
        int      firstgap, lastgap; /* leading/trailing overlap */
    {
10    int      nm, i0, i1, siz0, siz1;
    char     outx[32];
    double   pct;
    register int n0, n1;
    register char *p0, *p1;
15
    /* get total matches, score
     */
    i0 = i1 = siz0 = siz1 = 0;
    p0 = seqx[0] + pp[1].spc;
20    p1 = seqx[1] + pp[0].spc;
    n0 = pp[1].spc + 1;
    n1 = pp[0].spc + 1;
25
    nm = 0;
    while (*p0 && *p1) {
        if (siz0) {
30            p1++;
            n1++;
            siz0--;
        }
        else if (siz1) {
            p0++;
            n0++;
            siz1--;
        }
35        else {
            if (xbm[*p0-'A']&xbm[*p1-'A'])
                nm++;
            if (n0++ == pp[0].x[i0])
                siz0 = pp[0].n[i0++];
            if (n1++ == pp[1].x[i1])
                siz1 = pp[1].n[i1++];
            p0++;
            p1++;
45        }
    }
47
    /* pct homology:
     * if penalizing endgaps, base is the shorter seq
     * else, knock off overhangs and take shorter core
     */
50    if (endgaps)
        lx = (len0 < len1)? len0 : len1;
    else
55        lx = (lx < ly)? lx : ly;
    pct = 100.*(double)nm/(double)lx;
    fprintf(fx, "\n");
    fprintf(fx, "<%d match%s in an overlap of %d: %.2f percent similarity\n",
60        nm, (nm == 1)? ":" : "es", lx, pct);

```

**Table 1 (cont')**

```

    fprintf(fx, "<gaps in first sequence: %d", gapx);
    if (gapx) {
5      (void) sprintf(outx, " (%d %s%$s)",
                   ngapx, (dna)? "base":"residue", (ngapx == 1)? "" :"s");
      fprintf(fx,"%$s", outx);

    fprintf(fx, ", gaps in second sequence: %d", gapy);
10   if (gapy) {
      (void) sprintf(outx, " (%d %s%$s)",
                   ngapy, (dna)? "base":"residue", (ngapy == 1)? "" :"s");
      fprintf(fx,"%$s", outx);
    }
15   if (dna)
      fprintf(fx,
              "\n<score: %d (match = %d, mismatch = %d, gap penalty = %d + %d per base)\n",
              smax, DMAT, DMIS, DINS0, DINS1);
    else
20      fprintf(fx,
              "\n<score: %d (Dayhoff PAM 250 matrix, gap penalty = %d + %d per residue)\n",
              smax, PINS0, PINS1);
    if (endgaps)
      fprintf(fx,
25      "<endgaps penalized. left endgap: %d %s%$s, right endgap: %d %s%$s\n",
      firstgap, (dna)? "base" : "residue", (firstgap == 1)? "" : "s",
      lastgap, (dna)? "base" : "residue", (lastgap == 1)? "" : "s");
    else
30      fprintf(fx, "<endgaps not penalized\n");
}

static nm;          /* matches in core -- for checking */
static lmax;        /* lengths of stripped file names */
static ij[2];        /* jmp index for a path */
35 static nc[2];        /* number at start of current line */
static ni[2];        /* current elem number -- for gapping */
static siz[2];
static char *ps[2];    /* ptr to current element */
static char *po[2];    /* ptr to next output char slot */
40 static char out[2][P_LINE]; /* output line */
static char star[P_LINE]; /* set by stars() */

/*
 * print alignment of described in struct path pp[ ]
*/
45 static
pr_align
{
  int nn;          /* char count */
  int more;
50  register i;

  for (i = 0, lmax = 0; i < 2; i++) {
    nn = stripname(namex[i]);
    if (nn > lmax)
55      lmax = nn;

    nc[i] = 1;
    ni[i] = 1;
    siz[i] = ij[i] = 0;
    ps[i] = seqx[i];
    po[i] = out[i];
60  }
}

```

**Table 1 (cont')**

```

for (nn = nm = 0, more = 1; more; ) {
    for (i = more = 0; i < 2; i++) {
        /*
         * do we have more of this sequence?
         */
        if (!*ps[i])
            continue;
    10
        more++;

        if (pp[i].spc) { /* leading space */
            *po[i]++ = ' ';
            pp[i].spc--;
        }
        else if (siz[i]) { /* in a gap */
            *po[i]++ = ' ';
            siz[i]--;
        }
        else { /* we're putting a seq element
            /*
             *po[i] = *ps[i];
             if (islower(*ps[i]))
                 *ps[i] = toupper(*ps[i]);
            po[i]++;
            ps[i]++;
            /*
             * are we at next gap for this seq?
             */
            if (ni[i] == pp[i].x[ij[i]]) {
                /*
                 * we need to merge all gaps
                 * at this location
                 */
                siz[i] = pp[i].n[ij[i]++];
                while (ni[i] == pp[i].x[ij[i]]) {
                    siz[i] += pp[i].n[ij[i]++];
                }
                ni[i]++;
            }
            35
        }
        40
    }
    45
}
50
55
static
dumpblock()
{
    register i;
    for (i = 0; i < 2; i++)
        *po[i] = '\0';
    60
}
dumpblock
{
    register i;
    for (i = 0; i < 2; i++)
        *po[i] = '\0';
}
...pr_align

```

**Table 1 (cont')****...dumpblock**

```

5      (void) putc('\n', fx);
for (i = 0; i < 2; i++) {
    if (*out[i] && (*out[i] != ' ' || *(po[i]) != ' ')) {
        if (i == 0)
            nums(i);
        if (i == 0 && *out[1])
            stars();
        putline(i);
        if (i == 0 && *out[1])
            fprintf(fx, star);
        if (i == 1)
            nums(i);
    }
}
20  /*
 * put out a number line: dumpblock()
 */
static
25  nums(ix)
{
    int      ix;      /* index in out[ ] holding seq line */
    char      nline[P_LINE];
    register  i, j;
    register char  *pn, *px, *py;
30
    for (pn = nline, i = 0; i < lmax+P_SPC; i++, pn++)
        *pn = ' ';
    for (i = nc[ix], py = out[ix]; *py; py++, pn++) {
        if (*py == ' ' || *py == '-')
            *pn = ' ';
        else {
            if (i%10 == 0 || (i == 1 && nc[ix] != 1)) {
                j = (i < 0)? -i : i;
                for (px = pn; j; j /= 10, px--)
                    *px = j%10 + '0';
                if (i < 0)
                    *px = '-';
            }
            else
                *pn = ' ';
            i++;
        }
    }
    *pn = '\0';
    nc[ix] = i;
    for (pn = nline; *pn; pn++)
        (void) putc(*pn, fx);
    (void) putc('\n', fx);
}
55  /*
 * put out a line (name, [num], seq, [num]): dumpblock()
 */
static
60  putline(ix)
    int      ix;
    {

```

**putline**

**Table 1 (cont')****...putline**

```

5      int          i;
register char    *px;

10     for (px = namex[ix], i = 0; *px && *px != ':'; px++, i++)
           (void) putc(*px, fx);
     for (; i < lmax+P_SPC; i++)
           (void) putc(' ', fx);

15     /* these count from 1:
     * ni[ ] is current element (from 1)
     * nc[ ] is number at start of current line
     */
     for (px = out[ix]; *px; px++)
           (void) putc(*px&0x7F, fx);
     (void) putc('\n', fx);
}

20

25     /*
     * put a line of stars (seqs always in out[0], out[1]): dumpblock()
     */
static stars()
{
30     stars
{
     int          i;
register char    *p0, *p1, cx, *px;

     if (!*out[0] || (*out[0] == ' ' && *(po[0]) == ' ') ||
         !*out[1] || (*out[1] == ' ' && *(po[1]) == ' '))
         return;
35     px = star;
     for (i = lmax+P_SPC; i; i--)
         *px++ = ' ';

     for (p0 = out[0], p1 = out[1]; *p0 && *p1; p0++, p1++) {
40         if (isalpha(*p0) && isalpha(*p1)) {
             if (xbm[*p0-'A']&xbm[*p1-'A']) {
                 cx = '*';
                 nm++;
45             }
             else if (!dma && _day[*p0-'A'][*p1-'A'] > 0)
                 cx = '!';
             else
                 cx = ' ';
50             }
             else
                 cx = ' ';
             *px++ = cx;
55         }
         *px++ = '\n';
         *px = '\0';
     }
}

```

60

**Table 1 (cont<sup>2</sup>)**

```
/*
 * strip path or prefix from pn, return len: pr_align()
 */
5 static
stripname(pn)
    char    *pn;      /* file name (may be path) */
{
10    register char    *px, *py;

    py = 0;
    for (px = pn; *px; px++)
        if (*px == '/')
15        py = px + 1;
    if (py)
        (void) strcpy(pn, py);
    return(strlen(pn));
20 }
```

25

30 |

35

40

45

50

55

60

**Table 1 (cont')**

```

/*
 * cleanup() -- cleanup any tmp file
 * getseq() -- read in seq, set dna, len, maxlen
5  * g_calloc() -- calloc() with error checkin
 * readjmps() -- get the good jmps, from tmp file if necessary
 * writejmps() -- write a filled array of jmps to a tmp file: nw()
 */
10 #include "nw.h"
#include <sys/file.h>

15 char *jname = "/tmp/homgXXXXXX"; /* tmp file for jmps */
FILE *fj;

20 15 int cleanup(); /* cleanup tmp file */
long lseek();

25 /*
 * remove any tmp file if we blow
 */
cleanup(i)
25 30 {
    int i;
    if (fj)
        (void) unlink(jname);
    exit(i);
}

30 /*
 * read, return ptr to seq, set dna, len, maxlen
 * skip lines starting with ';', '<', or '>'
 * seq in upper or lower case
 */
35 char *
getseq(file, len)
35 40 {
    char *file; /* file name */
    int *len; /* seq len */
    char line[1024], *pseq;
    register char *px, *py;
    int natgc, tlen;
    FILE *fp;

    if ((fp = fopen(file, "r")) == 0) {
        fprintf(stderr, "%s: can't read %s\n", prog, file);
        exit(1);
    }
    tlen = natgc = 0;
    while (fgets(line, 1024, fp)) {
50      if (*line == ';' || *line == '<' || *line == '>')
          continue;
      for (px = line; *px != '\n'; px++)
          if (isupper(*px) || islower(*px))
              tlen++;
    }
    if ((pseq = malloc((unsigned)(tlen+6))) == 0) {
        fprintf(stderr, "%s: malloc() failed to get %d bytes for %s\n", prog, tlen+6, file);
        exit(1);
    }
60    pseq[0] = pseq[1] = pseq[2] = pseq[3] = '\0';
}

```

**Table 1 (cont')****...getseq**

```

5      py = pseq + 4;
*len = tlen;
rewind(fp);

10     while (fgets(line, 1024, fp)) {
        if (*line == ';' || *line == '<' || *line == '>')
            continue;
        for (px = line; *px != '\n'; px++) {
            if (isupper(*px))
                *py++ = *px;
            else if (islower(*px))
                *py++ = toupper(*px);
            if (index("ATGCU", *(py-1)))
                natgc++;
        }
        *py++ = '\0';
        *py = '\0';
        (void) fclose(fp);
        dna = natgc > (tlen/3);
        return(pseq+4);
    }

25     char *
30     g_calloc(msg, nx, sz)
        char *msg;           /* program, calling routine */
        int nx, sz;          /* number and size of elements */
    {
        char *px, *calloc();

        if ((px = calloc((unsigned)nx, (unsigned)sz)) == 0) {
            if (*msg) {
                fprintf(stderr, "%s: g_calloc() failed %s (n=%d, sz=%d)\n", prog, msg, nx, sz);
                exit(1);
            }
        }
        return(px);
    }

40     /*

45     * get final jmps from dx[ ] or tmp file, set pp[ ], reset dmax: main()
46     */
47     readjmps()
48     readjmps
49     {
50         int fd = -1;
51         int siz, i0, i1;
52         register i, j, xx;

53         if (fj) {
54             (void) fclose(fj);
55             if ((fd = open(jname, O_RDONLY, 0)) < 0) {
56                 fprintf(stderr, "%s: can't open() %s\n", prog, jname);
57                 cleanup(1);
58             }
59         }
60         for (i = i0 = i1 = 0, dmax0 = dmax, xx = len0; ; i++) {
61             while (1) {
62                 for (j = dx[dmax].jmp; j >= 0 && dx[dmax].jp.x[j] >= xx; j--)
63                     ;
64             }
65         }
66     }
67 
```

**Table 1 (cont')****...readjmps**

```

    if (j < 0 && dx[dmax].offset && fj) {
        (void) lseek(fd, dx[dmax].offset, 0);
        (void) read(fd, (char *)&dx[dmax].jp, sizeof(struct jmp));
        (void) read(fd, (char *)&dx[dmax].offset, sizeof(dx[dmax].offset));
        dx[dmax].jmp = MAXJMP-1;
    }
    else
        break;
}
if (i >= JMPS) {
    fprintf(stderr, "%s: too many gaps in alignment\n", prog);
    cleanup(1);
}
if (j >= 0) {
    siz = dx[dmax].jp.n[j];
    xx = dx[dmax].jp.x[j];
    dmax += siz;
    if (siz < 0) { /* gap in second seq */
        pp[1].n[i1] = -siz;
        xx += siz;
        /* id = xx - yy + len1 - 1
        */
        pp[1].x[i1] = xx - dmax + len1 - 1;
        gapy++;
        ngapy -= siz;
    /* ignore MAXGAP when doing endgaps */
        siz = (-siz < MAXGAP || endgaps)? -siz : MAXGAP;
        i1++;
    }
    else if (siz > 0) { /* gap in first seq */
        pp[0].n[i0] = siz;
        pp[0].x[i0] = xx;
        gapx++;
        ngapx += siz;
    /* ignore MAXGAP when doing endgaps */
        siz = (siz < MAXGAP || endgaps)? siz : MAXGAP;
        i0++;
    }
}
else
    break;
}
/* reverse the order of jmps
 */
for (j = 0, i0--; j < i0; j++, i0--) {
    i = pp[0].n[j]; pp[0].n[j] = pp[0].n[i0]; pp[0].n[i0] = i;
    i = pp[0].x[j]; pp[0].x[j] = pp[0].x[i0]; pp[0].x[i0] = i;
}
for (j = 0, i1--; j < i1; j++, i1--) {
    i = pp[1].n[j]; pp[1].n[j] = pp[1].n[i1]; pp[1].n[i1] = i;
    i = pp[1].x[j]; pp[1].x[j] = pp[1].x[i1]; pp[1].x[i1] = i;
}
if (fd >= 0)
    (void) close(fd);
if (fj) {
    (void) unlink(jname);
    fj = 0;
    offset = 0;
}

```

**Table 1 (cont')**

```

5      /*
6       * write a filled jmp struct offset of the prev one (if any): nw()
7       */
8      writejmps(ix)
9      {
10         writejmps
11         int ix;
12         {
13             char *mktemp();
14
15             if (!fj) {
16                 if (mktemp(jname) < 0) {
17                     fprintf(stderr, "%s: can't mktemp() %s\n", prog, jname);
18                     cleanup(1);
19                 }
20                 if ((fj = fopen(jname, "w")) == 0) {
21                     fprintf(stderr, "%s: can't write %s\n", prog, jname);
22                     exit(1);
23                 }
24             }
25             (void) fwrite((char *)&dx[ix].jp, sizeof(struct jmp), 1, fj);
26             (void) fwrite((char *)&dx[ix].offset, sizeof(dx[ix].offset), 1, fj);
27         }
28     }

```

**Table 2**

PRO	XXXXXXXXXXXXXXXXXX	(Length = 15 amino acids)
Comparison Protein	XXXXXXXXYYYYYYYY	(Length = 12 amino acids)
% amino acid sequence identity =		

30 (the number of identically matching amino acid residues between the two polypeptide sequences as determined by ALIGN-2) divided by (the total number of amino acid residues of the PRO polypeptide) = 5 divided by 15 = 33.3%

**Table 3**

35 PRO	XXXXXXXXXXXX	(Length = 10 amino acids)
Comparison Protein	XXXXXXXXYYYYYYYYZZYZ	(Length = 15 amino acids)
% amino acid sequence identity =		

40 (the number of identically matching amino acid residues between the two polypeptide sequences as determined by ALIGN-2) divided by (the total number of amino acid residues of the PRO polypeptide) = 5 divided by 10 = 50%

**Table 4**

PRO-DNA	NNNNNNNNNNNNNN	(Length = 14 nucleotides)
Comparison DNA	NNNNNNNLLLLLLLL	(Length = 16 nucleotides)

45 % nucleic acid sequence identity =  
 (the number of identically matching nucleotides between the two nucleic acid sequences as determined by ALIGN-2) divided by (the total number of nucleotides of the PRO-DNA nucleic acid sequence) = 6 divided by 14 = 42.9%

**Table 5**

PRO-DNA	NNNNNNNNNNNN	(Length = 12 nucleotides)
Comparison DNA	NNNNNLLVV	(Length = 9 nucleotides)

5 % nucleic acid sequence identity =

(the number of identically matching nucleotides between the two nucleic acid sequences as determined by ALIGN-2) divided by (the total number of nucleotides of the PRO-DNA nucleic acid sequence) =  
4 divided by 12 = 33.3%

## II. Compositions and Methods of the Invention

### A. Full-Length PRO Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO polypeptides. In particular, cDNAs encoding various PRO polypeptides have been identified and isolated, as disclosed in further detail in the Examples below. However, for sake of simplicity, in the present specification the protein encoded by the full length native nucleic acid molecules disclosed herein as well as all further native homologues and variants included in the foregoing definition of PRO, will be referred to as "PRO/number", regardless of their origin or mode of preparation.

As disclosed in the Examples below, various cDNA clones have been disclosed. The predicted amino acid sequence can be determined from the nucleotide sequence using routine skill. For the PRO polypeptides and encoding nucleic acids described herein, Applicants have identified what is believed to be the reading frame best identifiable with the sequence information available at the time.

### B. PRO Polypeptide Variants

In addition to the full-length native sequence PRO polypeptides described herein, it is contemplated that PRO variants can be prepared. PRO variants can be prepared by introducing appropriate nucleotide changes into the PRO DNA, and/or by synthesis of the desired PRO polypeptide. Those skilled in the art will appreciate that amino acid changes may alter post-translational processes of the PRO, such as changing the number or position of glycosylation sites or altering the membrane anchoring characteristics.

Variations in the native full-length sequence PRO or in various domains of the PRO described herein, can be made, for example, using any of the techniques and guidelines for conservative and non-conservative mutations set forth, for instance, in U.S. Patent No. 5,364,934. Variations may be a substitution, deletion or insertion of one or more codons encoding the PRO that results in a change in the amino acid sequence of the PRO as compared with the native sequence PRO. Optionally, the variation is by substitution of at least one amino acid with any other amino acid in one or more of the domains of the PRO. Guidance in determining which amino acid residue may be inserted, substituted or deleted without adversely affecting the desired activity may be found by comparing the sequence of the PRO with that of homologous known protein molecules and minimizing the number of amino acid sequence changes made in regions of high homology. Amino acid substitutions can be the result of replacing one amino acid with another amino acid having similar structural and/or chemical properties, such as the replacement of a leucine with a serine, i.e., conservative amino acid replacements. Insertions or deletions may optionally be in the range of about 1 to 5 amino acids. The variation allowed may be determined by systematically making insertions, deletions

or substitutions of amino acids in the sequence and testing the resulting variants for activity exhibited by the full-length or mature native sequence.

PRO polypeptide fragments are provided herein. Such fragments may be truncated at the N-terminus or C-terminus, or may lack internal residues, for example, when compared with a full length native protein. Certain fragments lack amino acid residues that are not essential for a desired biological activity of the PRO polypeptide.

PRO fragments may be prepared by any of a number of conventional techniques. Desired peptide fragments may be chemically synthesized. An alternative approach involves generating PRO fragments by enzymatic digestion, e.g., by treating the protein with an enzyme known to cleave proteins at sites defined by particular amino acid residues, or by digesting the DNA with suitable restriction enzymes and isolating the desired fragment. Yet another suitable technique involves isolating and amplifying a DNA fragment encoding a desired polypeptide fragment, by polymerase chain reaction (PCR). Oligonucleotides that define the desired termini of the DNA fragment are employed at the 5' and 3' primers in the PCR. Preferably, PRO polypeptide fragments share at least one biological and/or immunological activity with the native PRO polypeptide disclosed herein.

In particular embodiments, conservative substitutions of interest are shown in Table 6 under the heading of preferred substitutions. If such substitutions result in a change in biological activity, then more substantial changes, denominated exemplary substitutions in Table 6, or as further described below in reference to amino acid classes, are introduced and the products screened.

20 Table 6

	Original Residue	Exemplary Substitutions	Preferred Substitutions
25	Ala (A)	val; leu; ile	val
	Arg (R)	lys; gln; asn	lys
	Asn (N)	gln; his; lys; arg	gln
	Asp (D)	glu	glu
	Cys (C)	ser	ser
	Gln (Q)	asn	asn
30	Glu (E)	asp	asp
	Gly (G)	pro; ala	ala
	His (H)	asn; gln; lys; arg	arg
	Ile (I)	leu; val; met; ala; phe; norleucine	leu
35	Leu (L)	norleucine; ile; val; met; ala; phe	ile
	Lys (K)	arg; gln; asn	arg
	Met (M)	leu; phe; ile	leu
	Phe (F)	leu; val; ile; ala; tyr	leu
40	Pro (P)	ala	ala
	Ser (S)	thr	thr
	Thr (T)	ser	ser
	Trp (W)	tyr; phe	tyr
	Tyr (Y)	trp; phe; thr; ser	phe
45	Val (V)	ile; leu; met; phe; ala; norleucine	leu

Substantial modifications in function or immunological identity of the PRO polypeptide are accomplished by selecting substitutions that differ significantly in their effect on maintaining (a) the

structure of the polypeptide backbone in the area of the substitution, for example, as a sheet or helical conformation, (b) the charge or hydrophobicity of the molecule at the target site, or (c) the bulk of the side chain. Naturally occurring residues are divided into groups based on common side-chain properties:

- (1) hydrophobic: norleucine, met, ala, val, leu, ile;
- 5 (2) neutral hydrophilic: cys, ser, thr;
- (3) acidic: asp, glu;
- (4) basic: asn, gln, his, lys, arg;
- (5) residues that influence chain orientation: gly, pro; and
- (6) aromatic: trp, tyr, phe.

10 Non-conservative substitutions will entail exchanging a member of one of these classes for another class. Such substituted residues also may be introduced into the conservative substitution sites or, more preferably, into the remaining (non-conserved) sites.

15 The variations can be made using methods known in the art such as oligonucleotide-mediated (site-directed) mutagenesis, alanine scanning, and PCR mutagenesis. Site-directed mutagenesis [Carter et al., *Nucl. Acids Res.*, 13:4331 (1986); Zoller et al., *Nucl. Acids Res.*, 10:6487 (1987)], cassette mutagenesis [Wells et al., *Gene*, 34:315 (1985)], restriction selection mutagenesis [Wells et al., *Philos. Trans. R. Soc. London SerA*, 317:415 (1986)] or other known techniques can be performed on the cloned DNA to produce the PRO variant DNA.

20 Scanning amino acid analysis can also be employed to identify one or more amino acids along a contiguous sequence. Among the preferred scanning amino acids are relatively small, neutral amino acids. Such amino acids include alanine, glycine, serine, and cysteine. Alanine is typically a preferred scanning amino acid among this group because it eliminates the side-chain beyond the beta-carbon and is less likely to alter the main-chain conformation of the variant [Cunningham and Wells, *Science*, 244: 1081-1085 (1989)]. Alanine is also typically preferred because it is the most common amino acid. Further, it is frequently found 25 in both buried and exposed positions [Creighton, *The Proteins*, (W.H. Freeman & Co., N.Y.); Chothia, *J. Mol. Biol.*, 150:1 (1976)]. If alanine substitution does not yield adequate amounts of variant, an isoteric amino acid can be used.

### C. Modifications of PRO

30 Covalent modifications of PRO are included within the scope of this invention. One type of covalent modification includes reacting targeted amino acid residues of a PRO polypeptide with an organic derivatizing agent that is capable of reacting with selected side chains or the N- or C-terminal residues of the PRO. Derivatization with bifunctional agents is useful, for instance, for crosslinking PRO to a water-insoluble support matrix or surface for use in the method for purifying anti-PRO antibodies, and vice-versa. Commonly used crosslinking agents include, e.g., 1,1-bis(diazoacetyl)-2-phenylethane, glutaraldehyde, N-35 hydroxysuccinimide esters, for example, esters with 4-azidosalicylic acid, homobifunctional imidoesters, including disuccinimidyl esters such as 3,3'-dithiobis(succinimidylpropionate), bifunctional maleimides such as bis-N-maleimido-1,8-octane and agents such as methyl-3-[(p-azidophenyl)dithio]propioimidate.

40 Other modifications include deamidation of glutaminyl and asparaginyl residues to the corresponding glutamyl and aspartyl residues, respectively, hydroxylation of proline and lysine, phosphorylation of hydroxyl groups of seryl or threonyl residues, methylation of the  $\alpha$ -amino groups of

lysine, arginine, and histidine side chains [T.E. Creighton, Proteins: Structure and Molecular Properties, W.H. Freeman & Co., San Francisco, pp. 79-86 (1983)], acetylation of the N-terminal amine, and amidation of any C-terminal carboxyl group.

Another type of covalent modification of the PRO polypeptide included within the scope of this  
5 invention comprises altering the native glycosylation pattern of the polypeptide. "Altering the native  
glycosylation pattern" is intended for purposes herein to mean deleting one or more carbohydrate moieties  
found in native sequence PRO (either by removing the underlying glycosylation site or by deleting the  
glycosylation by chemical and/or enzymatic means), and/or adding one or more glycosylation sites that are  
not present in the native sequence PRO. In addition, the phrase includes qualitative changes in the  
10 glycosylation of the native proteins, involving a change in the nature and proportions of the various  
carbohydrate moieties present.

Addition of glycosylation sites to the PRO polypeptide may be accomplished by altering the amino  
acid sequence. The alteration may be made, for example, by the addition of, or substitution by, one or more  
15 serine or threonine residues to the native sequence PRO (for O-linked glycosylation sites). The PRO amino  
acid sequence may optionally be altered through changes at the DNA level, particularly by mutating the  
DNA encoding the PRO polypeptide at preselected bases such that codons are generated that will translate  
into the desired amino acids.

Another means of increasing the number of carbohydrate moieties on the PRO polypeptide is by  
chemical or enzymatic coupling of glycosides to the polypeptide. Such methods are described in the art,  
20 e.g., in WO 87/05330 published 11 September 1987, and in Aplin and Wriston, CRC Crit. Rev. Biochem.,  
pp. 259-306 (1981).

Removal of carbohydrate moieties present on the PRO polypeptide may be accomplished  
chemically or enzymatically or by mutational substitution of codons encoding for amino acid residues that  
serve as targets for glycosylation. Chemical deglycosylation techniques are known in the art and described,  
25 for instance, by Hakimuddin, et al., Arch. Biochem. Biophys., 259:52 (1987) and by Edge et al., Anal.  
Biochem., 118:131 (1981). Enzymatic cleavage of carbohydrate moieties on polypeptides can be achieved  
by the use of a variety of endo- and exo-glycosidases as described by Thotakura et al., Meth. Enzymol.,  
138:350 (1987).

Another type of covalent modification of PRO comprises linking the PRO polypeptide to one of a  
30 variety of nonproteinaceous polymers, e.g., polyethylene glycol (PEG), polypropylene glycol, or  
polyoxyalkylenes, in the manner set forth in U.S. Patent Nos. 4,640,835; 4,496,689; 4,301,144; 4,670,417;  
4,791,192 or 4,179,337.

The PRO of the present invention may also be modified in a way to form a chimeric molecule  
comprising PRO fused to another, heterologous polypeptide or amino acid sequence.

35 In one embodiment, such a chimeric molecule comprises a fusion of the PRO with a tag  
polypeptide which provides an epitope to which an anti-tag antibody can selectively bind. The epitope tag is  
generally placed at the amino- or carboxyl- terminus of the PRO. The presence of such epitope-tagged  
forms of the PRO can be detected using an antibody against the tag polypeptide. Also, provision of the  
epitope tag enables the PRO to be readily purified by affinity purification using an anti-tag antibody or  
40 another type of affinity matrix that binds to the epitope tag. Various tag polypeptides and their respective

antibodies are well known in the art. Examples include poly-histidine (poly-his) or poly-histidine-glycine (poly-his-gly) tags; the flu HA tag polypeptide and its antibody 12CA5 [Field et al., *Mol. Cell. Biol.*, 8:2159-2165 (1988)]; the c-myc tag and the 8F9, 3C7, 6E10, G4, B7 and 9E10 antibodies thereto [Evan et al., *Molecular and Cellular Biology*, 5:3610-3616 (1985)]; and the Herpes Simplex virus glycoprotein D (gD) tag and its antibody [Paborsky et al., *Protein Engineering*, 3(6):547-553 (1990)]. Other tag polypeptides include the Flag-peptide [Hopp et al., *BioTechnology*, 6:1204-1210 (1988)]; the KT3 epitope peptide [Martin et al., *Science*, 255:192-194 (1992)]; an alpha-tubulin epitope peptide [Skinner et al., *J. Biol. Chem.*, 266:15163-15166 (1991)]; and the T7 gene 10 protein peptide tag [Lutz-Freyermuth et al., *Proc. Natl. Acad. Sci. USA*, 87:6393-6397 (1990)].

10 In an alternative embodiment, the chimeric molecule may comprise a fusion of the PRO with an immunoglobulin or a particular region of an immunoglobulin. For a bivalent form of the chimeric molecule (also referred to as an "immunoadhesin"), such a fusion could be to the Fc region of an IgG molecule. The Ig fusions preferably include the substitution of a soluble (transmembrane domain deleted or inactivated) form of a PRO polypeptide in place of at least one variable region within an Ig molecule. In a particularly preferred embodiment, the immunoglobulin fusion includes the hinge, CH2 and CH3, or the hinge, CH1, CH2 and CH3 regions of an IgG1 molecule. For the production of immunoglobulin fusions see also US Patent No. 5,428,130 issued June 27, 1995.

#### D. Preparation of PRO

20 The description below relates primarily to production of PRO by culturing cells transformed or transfected with a vector containing PRO nucleic acid. It is, of course, contemplated that alternative methods, which are well known in the art, may be employed to prepare PRO. For instance, the PRO sequence, or portions thereof, may be produced by direct peptide synthesis using solid-phase techniques [see, e.g., Stewart et al., *Solid-Phase Peptide Synthesis*, W.H. Freeman Co., San Francisco, CA (1969); Merrifield, *J. Am. Chem. Soc.*, 85:2149-2154 (1963)]. *In vitro* protein synthesis may be performed using 25 manual techniques or by automation. Automated synthesis may be accomplished, for instance, using an Applied Biosystems Peptide Synthesizer (Foster City, CA) using manufacturer's instructions. Various portions of the PRO may be chemically synthesized separately and combined using chemical or enzymatic methods to produce the full-length PRO.

##### 1. Isolation of DNA Encoding PRO

30 DNA encoding PRO may be obtained from a cDNA library prepared from tissue believed to possess the PRO mRNA and to express it at a detectable level. Accordingly, human PRO DNA can be conveniently obtained from a cDNA library prepared from human tissue, such as described in the Examples. The PRO-encoding gene may also be obtained from a genomic library or by known synthetic procedures (e.g., automated nucleic acid synthesis).

35 Libraries can be screened with probes (such as antibodies to the PRO or oligonucleotides of at least about 20-80 bases) designed to identify the gene of interest or the protein encoded by it. Screening the cDNA or genomic library with the selected probe may be conducted using standard procedures, such as described in Sambrook et al., *Molecular Cloning: A Laboratory Manual* (New York: Cold Spring Harbor Laboratory Press, 1989). An alternative means to isolate the gene encoding PRO is to use PCR methodology

[Sambrook et al., *supra*; Dieffenbach et al., *PCR Primer: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, 1995)].

The Examples below describe techniques for screening a cDNA library. The oligonucleotide sequences selected as probes should be of sufficient length and sufficiently unambiguous that false positives are minimized. The oligonucleotide is preferably labeled such that it can be detected upon hybridization to DNA in the library being screened. Methods of labeling are well known in the art, and include the use of radiolabels like  $^{32}\text{P}$ -labeled ATP, biotinylation or enzyme labeling. Hybridization conditions, including moderate stringency and high stringency, are provided in Sambrook et al., *supra*.

Sequences identified in such library screening methods can be compared and aligned to other known sequences deposited and available in public databases such as GenBank or other private sequence databases. Sequence identity (at either the amino acid or nucleotide level) within defined regions of the molecule or across the full-length sequence can be determined using methods known in the art and as described herein.

Nucleic acid having protein coding sequence may be obtained by screening selected cDNA or genomic libraries using the deduced amino acid sequence disclosed herein for the first time, and, if necessary, using conventional primer extension procedures as described in Sambrook et al., *supra*, to detect precursors and processing intermediates of mRNA that may not have been reverse-transcribed into cDNA.

## 2. Selection and Transformation of Host Cells

Host cells are transfected or transformed with expression or cloning vectors described herein for PRO production and cultured in conventional nutrient media modified as appropriate for inducing promoters, selecting transformants, or amplifying the genes encoding the desired sequences. The culture conditions, such as media, temperature, pH and the like, can be selected by the skilled artisan without undue experimentation. In general, principles, protocols, and practical techniques for maximizing the productivity of cell cultures can be found in *Mammalian Cell Biotechnology: a Practical Approach*, M. Butler, ed. (IRL Press, 1991) and Sambrook et al., *supra*.

Methods of eukaryotic cell transfection and prokaryotic cell transformation are known to the ordinarily skilled artisan, for example,  $\text{CaCl}_2$ ,  $\text{CaPO}_4$ , liposome-mediated and electroporation. Depending on the host cell used, transformation is performed using standard techniques appropriate to such cells. The calcium treatment employing calcium chloride, as described in Sambrook et al., *supra*, or electroporation is generally used for prokaryotes. Infection with *Agrobacterium tumefaciens* is used for transformation of certain plant cells, as described by Shaw et al., *Gene*, 23:315 (1983) and WO 89/05859 published 29 June 1989. For mammalian cells without such cell walls, the calcium phosphate precipitation method of Graham and van der Eb, *Virology*, 52:456-457 (1978) can be employed. General aspects of mammalian cell host system transfections have been described in U.S. Patent No. 4,399,216. Transformations into yeast are typically carried out according to the method of Van Solingen et al., *J. Bact.*, 130:946 (1977) and Hsiao et al., *Proc. Natl. Acad. Sci. (USA)*, 76:3829 (1979). However, other methods for introducing DNA into cells, such as by nuclear microinjection, electroporation, bacterial protoplast fusion with intact cells, or polycations, e.g., polybrene, polyornithine, may also be used. For various techniques for transforming

mammalian cells, see Keown et al., *Methods in Enzymology*, 185:527-537 (1990) and Mansour et al., *Nature*, 336:348-352 (1988).

Suitable host cells for cloning or expressing the DNA in the vectors herein include prokaryote, yeast, or higher eukaryote cells. Suitable prokaryotes include but are not limited to eubacteria, such as 5 Gram-negative or Gram-positive organisms, for example, Enterobacteriaceae such as *E. coli*. Various *E. coli* strains are publicly available, such as *E. coli* K12 strain MM294 (ATCC 31,446); *E. coli* X1776 (ATCC 31,537); *E. coli* strain W3110 (ATCC 27,325) and K5 772 (ATCC 53,635). Other suitable prokaryotic host 10 cells include Enterobacteriaceae such as *Escherichia*, e.g., *E. coli*, *Enterobacter*, *Erwinia*, *Klebsiella*, *Proteus*, *Salmonella*, e.g., *Salmonella typhimurium*, *Serratia*, e.g., *Serratia marcescens*, and *Shigella*, as well 15 as *Bacilli* such as *B. subtilis* and *B. licheniformis* (e.g., *B. licheniformis* 41P disclosed in DD 266,710 published 12 April 1989), *Pseudomonas* such as *P. aeruginosa*, and *Streptomyces*. These examples are 20 illustrative rather than limiting. Strain W3110 is one particularly preferred host or parent host because it is a common host strain for recombinant DNA product fermentations. Preferably, the host cell secretes minimal amounts of proteolytic enzymes. For example, strain W3110 may be modified to effect a genetic mutation 25 in the genes encoding proteins endogenous to the host, with examples of such hosts including *E. coli* W3110 strain 1A2, which has the complete genotype *tonA*; *E. coli* W3110 strain 9E4, which has the complete genotype *tonA ptr3*; *E. coli* W3110 strain 27C7 (ATCC 55,244), which has the complete genotype *tonA ptr3 phoA E15 (argF-lac)169 degP ompT kan*<sup>r</sup>; *E. coli* W3110 strain 37D6, which has the complete genotype 30 *tonA ptr3 phoA E15 (argF-lac)169 degP ompT rbs7 ilvG kan*<sup>r</sup>; *E. coli* W3110 strain 40B4, which is strain 37D6 with a non-kanamycin resistant *degP* deletion mutation; and an *E. coli* strain having mutant periplasmic protease disclosed in U.S. Patent No. 4,946,783 issued 7 August 1990. Alternatively, *in vitro* methods of cloning, e.g., PCR or other nucleic acid polymerase reactions, are suitable.

In addition to prokaryotes, eukaryotic microbes such as filamentous fungi or yeast are suitable 25 cloning or expression hosts for PRO-encoding vectors. *Saccharomyces cerevisiae* is a commonly used lower eukaryotic host microorganism. Others include *Schizosaccharomyces pombe* (Beach and Nurse, *Nature*, 290: 140 [1981]; EP 139,383 published 2 May 1985); *Kluyveromyces* hosts (U.S. Patent No. 4,943,529; Fleer et al., *Bio/Technology*, 9:968-975 (1991)) such as, e.g., *K. lactis* (MW98-8C, CBS683, CBS4574; Louvencourt et al., *J. Bacteriol.*, 154(2):737-742 [1983]), *K. fragilis* (ATCC 12,424), *K. bulgaricus* (ATCC 16,045), *K. wickeramii* (ATCC 24,178), *K. waltii* (ATCC 56,500), *K. drosophilae* (ATCC 36,906; Van 30 den Berg et al., *Bio/Technology*, 8:135 (1990)), *K. thermotolerans*, and *K. marxianus*; *yarrowia* (EP 402,226); *Pichia pastoris* (EP 183,070; Sreekrishna et al., *J. Basic Microbiol.*, 28:265-278 [1988]); *Candida*; *Trichoderma reesia* (EP 244,234); *Neurospora crassa* (Case et al., *Proc. Natl. Acad. Sci. USA*, 76:5259-5263 [1979]); *Schwanniomyces* such as *Schwanniomyces occidentalis* (EP 394,538 published 31 October 1990); and filamentous fungi such as, e.g., *Neurospora*, *Penicillium*, *Tolypocladium* (WO 91/00357 35 published 10 January 1991), and *Aspergillus* hosts such as *A. nidulans* (Ballance et al., *Biochem. Biophys. Res. Commun.*, 112:284-289 [1983]; Tilburn et al., *Gene*, 26:205-221 [1983]; Yelton et al., *Proc. Natl. Acad. Sci. USA*, 81: 1470-1474 [1984]) and *A. niger* (Kelly and Hynes, *EMBO J.*, 4:475-479 [1985]). Methylotropic yeasts are suitable herein and include, but are not limited to, yeast capable of growth on methanol selected from the genera consisting of *Hansenula*, *Candida*, *Kloeckera*, *Pichia*, *Saccharomyces*,

*Torulopsis*, and *Rhodotorula*. A list of specific species that are exemplary of this class of yeasts may be found in C. Anthony, The Biochemistry of Methylotrophs, 269 (1982).

Suitable host cells for the expression of glycosylated PRO are derived from multicellular organisms. Examples of invertebrate cells include insect cells such as *Drosophila* S2 and *Spodoptera* Sf9, as well as plant cells. Examples of useful mammalian host cell lines include Chinese hamster ovary (CHO) and COS cells. More specific examples include monkey kidney CV1 line transformed by SV40 (COS-7, ATCC CRL 1651); human embryonic kidney line (293 or 293 cells subcloned for growth in suspension culture, Graham et al., J. Gen Virol., 36:59 (1977)); Chinese hamster ovary cells/-DHFR (CHO, Urlaub and Chasin, Proc. Natl. Acad. Sci. USA, 77:4216 (1980)); mouse sertoli cells (TM4, Mather, Biol. Reprod., 23:243-251 (1980)); human lung cells (W138, ATCC CCL 75); human liver cells (Hep G2, HB 8065); and mouse mammary tumor (MMT 060562, ATCC CCL51). The selection of the appropriate host cell is deemed to be within the skill in the art.

### 3. Selection and Use of a Replicable Vector

The nucleic acid (e.g., cDNA or genomic DNA) encoding PRO may be inserted into a replicable vector for cloning (amplification of the DNA) or for expression. Various vectors are publicly available. The vector may, for example, be in the form of a plasmid, cosmid, viral particle, or phage. The appropriate nucleic acid sequence may be inserted into the vector by a variety of procedures. In general, DNA is inserted into an appropriate restriction endonuclease site(s) using techniques known in the art. Vector components generally include, but are not limited to, one or more of a signal sequence, an origin of replication, one or more marker genes, an enhancer element, a promoter, and a transcription termination sequence. Construction of suitable vectors containing one or more of these components employs standard ligation techniques which are known to the skilled artisan.

The PRO may be produced recombinantly not only directly, but also as a fusion polypeptide with a heterologous polypeptide, which may be a signal sequence or other polypeptide having a specific cleavage site at the N-terminus of the mature protein or polypeptide. In general, the signal sequence may be a component of the vector, or it may be a part of the PRO-encoding DNA that is inserted into the vector. The signal sequence may be a prokaryotic signal sequence selected, for example, from the group of the alkaline phosphatase, penicillinase, lpp, or heat-stable enterotoxin II leaders. For yeast secretion the signal sequence may be, e.g., the yeast invertase leader, alpha factor leader (including *Saccharomyces* and *Kluyveromyces*  $\alpha$ -factor leaders, the latter described in U.S. Patent No. 5,010,182), or acid phosphatase leader, the *C. albicans* glucoamylase leader (EP 362,179 published 4 April 1990), or the signal described in WO 90/13646 published 15 November 1990. In mammalian cell expression, mammalian signal sequences may be used to direct secretion of the protein, such as signal sequences from secreted polypeptides of the same or related species, as well as viral secretory leaders.

Both expression and cloning vectors contain a nucleic acid sequence that enables the vector to replicate in one or more selected host cells. Such sequences are well known for a variety of bacteria, yeast, and viruses. The origin of replication from the plasmid pBR322 is suitable for most Gram-negative bacteria, the 2 $\mu$  plasmid origin is suitable for yeast, and various viral origins (SV40, polyoma, adenovirus, VSV or BPV) are useful for cloning vectors in mammalian cells.

Expression and cloning vectors will typically contain a selection gene, also termed a selectable marker. Typical selection genes encode proteins that (a) confer resistance to antibiotics or other toxins, e.g., ampicillin, neomycin, methotrexate, or tetracycline, (b) complement auxotrophic deficiencies, or (c) supply critical nutrients not available from complex media, e.g., the gene encoding D-alanine racemase for *Bacilli*.

5 An example of suitable selectable markers for mammalian cells are those that enable the identification of cells competent to take up the PRO-encoding nucleic acid, such as DHFR or thymidine kinase. An appropriate host cell when wild-type DHFR is employed is the CHO cell line deficient in DHFR activity, prepared and propagated as described by Urlaub et al., *Proc. Natl. Acad. Sci. USA*, 77:4216 (1980).  
10 A suitable selection gene for use in yeast is the *trp1* gene present in the yeast plasmid YRp7 [Stinchcomb et al., *Nature*, 282:39 (1979); Kingsman et al., *Gene*, 7:141 (1979); Tschemper et al., *Gene*, 10:157 (1980)]. The *trp1* gene provides a selection marker for a mutant strain of yeast lacking the ability to grow in tryptophan, for example, ATCC No. 44076 or PEP4-1 [Jones, *Genetics*, 85:12 (1977)].

15 Expression and cloning vectors usually contain a promoter operably linked to the PRO-encoding nucleic acid sequence to direct mRNA synthesis. Promoters recognized by a variety of potential host cells are well known. Promoters suitable for use with prokaryotic hosts include the  $\beta$ -lactamase and lactose promoter systems [Chang et al., *Nature*, 275:615 (1978); Goeddel et al., *Nature*, 281:544 (1979)], alkaline phosphatase, a tryptophan (trp) promoter system [Goeddel, *Nucleic Acids Res.*, 8:4057 (1980); EP 36,776], and hybrid promoters such as the tac promoter [deBoer et al., *Proc. Natl. Acad. Sci. USA*, 80:21-25 (1983)].  
20 Promoters for use in bacterial systems also will contain a Shine-Dalgarno (S.D.) sequence operably linked to the DNA encoding PRO.

25 Examples of suitable promoting sequences for use with yeast hosts include the promoters for 3-phosphoglycerate kinase [Hitzeman et al., *J. Biol. Chem.*, 255:2073 (1980)] or other glycolytic enzymes [Hess et al., *J. Adv. Enzyme Reg.*, 7:149 (1968); Holland, *Biochemistry*, 17:4900 (1978)], such as enolase, glyceraldehyde-3-phosphate dehydrogenase, hexokinase, pyruvate decarboxylase, phosphofructokinase, glucose-6-phosphate isomerase, 3-phosphoglycerate mutase, pyruvate kinase, triosephosphate isomerase, phosphoglucose isomerase, and glucokinase.

30 Other yeast promoters, which are inducible promoters having the additional advantage of transcription controlled by growth conditions, are the promoter regions for alcohol dehydrogenase 2, isocytochrome C, acid phosphatase, degradative enzymes associated with nitrogen metabolism, metallothionein, glyceraldehyde-3-phosphate dehydrogenase, and enzymes responsible for maltose and galactose utilization. Suitable vectors and promoters for use in yeast expression are further described in EP 73,657.

35 PRO transcription from vectors in mammalian host cells is controlled, for example, by promoters obtained from the genomes of viruses such as polyoma virus, fowlpox virus (UK 2,211,504 published 5 July 1989), adenovirus (such as Adenovirus 2), bovine papilloma virus, avian sarcoma virus, cytomegalovirus, a retrovirus, hepatitis-B virus and Simian Virus 40 (SV40), from heterologous mammalian promoters, e.g., the actin promoter or an immunoglobulin promoter, and from heat-shock promoters, provided such promoters are compatible with the host cell systems.

40 Transcription of a DNA encoding the PRO by higher eukaryotes may be increased by inserting an enhancer sequence into the vector. Enhancers are cis-acting elements of DNA, usually about from 10 to 300

bp, that act on a promoter to increase its transcription. Many enhancer sequences are now known from mammalian genes (globin, elastase, albumin,  $\alpha$ -fetoprotein, and insulin). Typically, however, one will use an enhancer from a eukaryotic cell virus. Examples include the SV40 enhancer on the late side of the replication origin (bp 100-270), the cytomegalovirus early promoter enhancer, the polyoma enhancer on the 5 late side of the replication origin, and adenovirus enhancers. The enhancer may be spliced into the vector at a position 5' or 3' to the PRO coding sequence, but is preferably located at a site 5' from the promoter.

Expression vectors used in eukaryotic host cells (yeast, fungi, insect, plant, animal, human, or nucleated cells from other multicellular organisms) will also contain sequences necessary for the termination of transcription and for stabilizing the mRNA. Such sequences are commonly available from the 5' and, 10 occasionally 3', untranslated regions of eukaryotic or viral DNAs or cDNAs. These regions contain nucleotide segments transcribed as polyadenylated fragments in the untranslated portion of the mRNA encoding PRO.

Still other methods, vectors, and host cells suitable for adaptation to the synthesis of PRO in recombinant vertebrate cell culture are described in Gething et al., *Nature*, 293:620-625 (1981); Mantei et 15 al., *Nature*, 281:40-46 (1979); EP 117,060; and EP 117,058.

#### 4. Detecting Gene Amplification/Expression

Gene amplification and/or expression may be measured in a sample directly, for example, by conventional Southern blotting, Northern blotting to quantitate the transcription of mRNA [Thomas, *Proc. Natl. Acad. Sci. USA*, 77:5201-5205 (1980)], dot blotting (DNA analysis), or *in situ* hybridization, using an 20 appropriately labeled probe, based on the sequences provided herein. Alternatively, antibodies may be employed that can recognize specific duplexes, including DNA duplexes, RNA duplexes, and DNA-RNA hybrid duplexes or DNA-protein duplexes. The antibodies in turn may be labeled and the assay may be carried out where the duplex is bound to a surface, so that upon the formation of duplex on the surface, the presence of antibody bound to the duplex can be detected.

25 Gene expression, alternatively, may be measured by immunological methods, such as immunohistochemical staining of cells or tissue sections and assay of cell culture or body fluids, to quantitate directly the expression of gene product. Antibodies useful for immunohistochemical staining and/or assay of sample fluids may be either monoclonal or polyclonal, and may be prepared in any mammal. Conveniently, the antibodies may be prepared against a native sequence PRO polypeptide or against a 30 synthetic peptide based on the DNA sequences provided herein or against exogenous sequence fused to PRO DNA and encoding a specific antibody epitope.

#### 5. Purification of Polypeptide

Forms of PRO may be recovered from culture medium or from host cell lysates. If membrane-bound, it can be released from the membrane using a suitable detergent solution (e.g. Triton-X 100) or by 35 enzymatic cleavage. Cells employed in expression of PRO can be disrupted by various physical or chemical means, such as freeze-thaw cycling, sonication, mechanical disruption, or cell lysing agents.

It may be desired to purify PRO from recombinant cell proteins or polypeptides. The following procedures are exemplary of suitable purification procedures: by fractionation on an ion-exchange column; ethanol precipitation; reverse phase HPLC; chromatography on silica or on a cation-exchange resin such as 40 DEAE; chromatofocusing; SDS-PAGE; ammonium sulfate precipitation; gel filtration using, for example,

Sephadex G-75; protein A Sepharose columns to remove contaminants such as IgG; and metal chelating columns to bind epitope-tagged forms of the PRO. Various methods of protein purification may be employed and such methods are known in the art and described for example in Deutscher, Methods in Enzymology, 182 (1990); Scopes, Protein Purification: Principles and Practice, Springer-Verlag, New York 5 (1982). The purification step(s) selected will depend, for example, on the nature of the production process used and the particular PRO produced.

E. Tissue Distribution

The location of tissues expressing the PRO can be identified by determining mRNA expression in various human tissues. The location of such genes provides information about which tissues are most likely 10 to be affected by the stimulating and inhibiting activities of the PRO polypeptides. The location of a gene in a specific tissue also provides sample tissue for the activity blocking assays discussed below.

As noted before, gene expression in various tissues may be measured by conventional Southern blotting, Northern blotting to quantitate the transcription of mRNA (Thomas, *Proc. Natl. Acad. Sci. USA*, 77:5201-5205 [1980]), dot blotting (DNA analysis), or *in situ* hybridization, using an appropriately labeled 15 probe, based on the sequences provided herein. Alternatively, antibodies may be employed that can recognize specific duplexes, including DNA duplexes, RNA duplexes, and DNA-RNA hybrid duplexes or DNA-protein duplexes.

Gene expression in various tissues, alternatively, may be measured by immunological methods, such as immunohistochemical staining of tissue sections and assay of cell culture or body fluids, to 20 quantitate directly the expression of gene product. Antibodies useful for immunohistochemical staining and/or assay of sample fluids may be either monoclonal or polyclonal, and may be prepared in any mammal. Conveniently, the antibodies may be prepared against a native sequence of a PRO polypeptide or against a 25 synthetic peptide based on the DNA sequences encoding the PRO polypeptide or against an exogenous sequence fused to a DNA encoding a PRO polypeptide and encoding a specific antibody epitope. General techniques for generating antibodies, and special protocols for Northern blotting and *in situ* hybridization are provided below.

F. Antibody Binding Studies

The activity of the PRO polypeptides can be further verified by antibody binding studies, in which 30 the ability of anti-PRO antibodies to inhibit the effect of the PRO polypeptides, respectively, on tissue cells is tested. Exemplary antibodies include polyclonal, monoclonal, humanized, bispecific, and heteroconjugate antibodies, the preparation of which will be described hereinbelow.

Antibody binding studies may be carried out in any known assay method, such as competitive binding assays, direct and indirect sandwich assays, and immunoprecipitation assays. Zola, *Monoclonal Antibodies: A Manual of Techniques*, pp.147-158 (CRC Press, Inc., 1987).

Competitive binding assays rely on the ability of a labeled standard to compete with the test sample 35 analyte for binding with a limited amount of antibody. The amount of target protein in the test sample is inversely proportional to the amount of standard that becomes bound to the antibodies. To facilitate determining the amount of standard that becomes bound, the antibodies preferably are insolubilized before or after the competition, so that the standard and analyte that are bound to the antibodies may conveniently 40 be separated from the standard and analyte which remain unbound.

Sandwich assays involve the use of two antibodies, each capable of binding to a different immunogenic portion, or epitope, of the protein to be detected. In a sandwich assay, the test sample analyte is bound by a first antibody which is immobilized on a solid support, and thereafter a second antibody binds to the analyte, thus forming an insoluble three-part complex. See, e.g., US Pat No. 4,376,110. The second 5 antibody may itself be labeled with a detectable moiety (direct sandwich assays) or may be measured using an anti-immunoglobulin antibody that is labeled with a detectable moiety (indirect sandwich assay). For example, one type of sandwich assay is an ELISA assay, in which case the detectable moiety is an enzyme.

For immunohistochemistry, the tissue sample may be fresh or frozen or may be embedded in paraffin and fixed with a preservative such as formalin, for example.

10        G.     Cell-Based Assays

Cell-based assays and animal models for immune related diseases can be used to further understand the relationship between the genes and polypeptides identified herein and the development and pathogenesis of immune related disease.

In a different approach, cells of a cell type known to be involved in a particular immune related 15 disease are transfected with the cDNAs described herein, and the ability of these cDNAs to stimulate or inhibit immune function is analyzed. Suitable cells can be transfected with the desired gene, and monitored for immune function activity. Such transfected cell lines can then be used to test the ability of poly- or monoclonal antibodies or antibody compositions to inhibit or stimulate immune function, for example to modulate T-cell proliferation or inflammatory cell infiltration. Cells transfected with the coding sequences 20 of the genes identified herein can further be used to identify drug candidates for the treatment of immune related diseases.

In addition, primary cultures derived from transgenic animals (as described below) can be used in the cell-based assays herein, although stable cell lines are preferred. Techniques to derive continuous cell 25 lines from transgenic animals are well known in the art (see, e.g., Small *et al.*, *Mol. Cell. Biol.* 5: 642-648 [1985]).

One suitable cell based assay is the mixed lymphocyte reaction (MLR). *Current Protocols in Immunology*, unit 3.12; edited by J E Coligan, A M Kruisbeek, D H Marglies, E M Shevach, W Strober, National Institutes of Health, Published by John Wiley & Sons, Inc. In this assay, the ability of a test compound to stimulate or inhibit the proliferation of activated T cells is assayed. A suspension of responder 30 T cells is cultured with allogeneic stimulator cells and the proliferation of T cells is measured by uptake of tritiated thymidine. This assay is a general measure of T cell reactivity. Since the majority of T cells respond to and produce IL-2 upon activation, differences in responsiveness in this assay in part reflect differences in IL-2 production by the responding cells. The MLR results can be verified by a standard lymphokine (IL-2) detection assay. *Current Protocols in Immunology*, above, 3.15, 6.3.

A proliferative T cell response in an MLR assay may be due to direct mitogenic properties of an assayed molecule or to external antigen induced activation. Additional verification of the T cell stimulatory activity of the PRO polypeptides can be obtained by a costimulation assay. T cell activation requires an antigen specific signal mediated through the T-cell receptor (TCR) and a costimulatory signal mediated through a second ligand binding interaction, for example, the B7 (CD80, CD86)/CD28 binding interaction. 35 CD28 crosslinking increases lymphokine secretion by activated T cells. T cell activation has both negative 40

and positive controls through the binding of ligands which have a negative or positive effect. CD28 and CTLA-4 are related glycoproteins in the Ig superfamily which bind to B7. CD28 binding to B7 has a positive costimulation effect of T cell activation; conversely, CTLA-4 binding to B7 has a T cell deactivating effect. Chambers, C. A. and Allison, J. P., *Curr. Opin. Immunol.* (1997) 9:396. Schwartz, R. 5 H., *Cell* (1992) 71:1065; Linsey, P. S. and Ledbetter, J. A., *Annu. Rev. Immunol.* (1993) 11:191; June, C. H. et al, *Immunol. Today* (1994) 15:321; Jenkins, M. K., *Immunity* (1994) 1:405. In a costimulation assay, the PRO polypeptides are assayed for T cell costimulatory or inhibitory activity.

Direct use of a stimulating compound as in the invention has been validated in experiments with 4-1BB glycoprotein, a member of the tumor necrosis factor receptor family, which binds to a ligand (4-1BBL) 10 expressed on primed T cells and signals T cell activation and growth. Alderson, M. E. et al., *J. Immunol.* (1994) 24:2219.

The use of an agonist stimulating compound has also been validated experimentally. Activation of 4-1BB by treatment with an agonist anti-4-1BB antibody enhances eradication of tumors. Hellstrom, I. and Hellstrom, K. E., *Crit. Rev. Immunol.* (1998) 18:1. Immunoadjuvant therapy for treatment of tumors, 15 described in more detail below, is another example of the use of the stimulating compounds of the invention.

Alternatively, an immune stimulating or enhancing effect can also be achieved by administration of a PRO which has vascular permeability enhancing properties. Enhanced vascular permeability would be beneficial to disorders which can be attenuated by local infiltration of immune cells (e.g., monocytes, 20 eosinophils, PMNs) and inflammation.

On the other hand, PRO polypeptides, as well as other compounds of the invention, which are direct inhibitors of T cell proliferation/activation, lymphokine secretion, and/or vascular permeability can be directly used to suppress the immune response. These compounds are useful to reduce the degree of the immune response and to treat immune related diseases characterized by a hyperactive, superoptimal, or 25 autoimmune response. This use of the compounds of the invention has been validated by the experiments described above in which CTLA-4 binding to receptor B7 deactivates T cells. The direct inhibitory compounds of the invention function in an analogous manner. The use of compound which suppress vascular permeability would be expected to reduce inflammation. Such uses would be beneficial in treating conditions associated with excessive inflammation.

30 Alternatively, compounds, e.g., antibodies, which bind to stimulating PRO polypeptides and block the stimulating effect of these molecules produce a net inhibitory effect and can be used to suppress the T cell mediated immune response by inhibiting T cell proliferation/activation and/or lymphokine secretion. Blocking the stimulating effect of the polypeptides suppresses the immune response of the mammal. This use has been validated in experiments using an anti-IL2 antibody. In these experiments, the antibody binds 35 to IL2 and blocks binding of IL2 to its receptor thereby achieving a T cell inhibitory effect.

#### H. Animal Models

The results of the cell based in vitro assays can be further verified using *in vivo* animal models and assays for T-cell function. A variety of well known animal models can be used to further understand the role 40 of the genes identified herein in the development and pathogenesis of immune related disease, and to test the efficacy of candidate therapeutic agents, including antibodies, and other antagonists of the native

polypeptides, including small molecule antagonists. The *in vivo* nature of such models makes them predictive of responses in human patients. Animal models of immune related diseases include both non-recombinant and recombinant (transgenic) animals. Non-recombinant animal models include, for example, rodent, *e.g.*, murine models. Such models can be generated by introducing cells into syngeneic mice using 5 standard techniques, *e.g.*, subcutaneous injection, tail vein injection, spleen implantation, intraperitoneal implantation, implantation under the renal capsule, *etc.*

Graft-versus-host disease occurs when immunocompetent cells are transplanted into immunosuppressed or tolerant patients. The donor cells recognize and respond to host antigens. The response can vary from life threatening severe inflammation to mild cases of diarrhea and weight loss. 10 Graft-versus-host disease models provide a means of assessing T cell reactivity against MHC antigens and minor transplant antigens. A suitable procedure is described in detail in Current Protocols in Immunology, above, unit 4.3.

An animal model for skin allograft rejection is a means of testing the ability of T cells to mediate *in vivo* tissue destruction and a measure of their role in transplant rejection. The most common and accepted 15 models use murine tail-skin grafts. Repeated experiments have shown that skin allograft rejection is mediated by T cells, helper T cells and killer-effector T cells, and not antibodies. Auchincloss, H. Jr. and Sachs, D. H., *Fundamental Immunology*, 2nd ed., W. E. Paul ed., Raven Press, NY, 1989, 889-992. A suitable procedure is described in detail in Current Protocols in Immunology, above, unit 4.4. Other 20 transplant rejection models which can be used to test the compounds of the invention are the allogeneic heart transplant models described by Tanabe, M. *et al*, *Transplantation* (1994) 58:23 and Tinubu, S. A. *et al*, *J. Immunol.* (1994) 4330-4338.

Animal models for delayed type hypersensitivity provides an assay of cell mediated immune function as well. Delayed type hypersensitivity reactions are a T cell mediated *in vivo* immune response characterized by inflammation which does not reach a peak until after a period of time has elapsed after 25 challenge with an antigen. These reactions also occur in tissue specific autoimmune diseases such as multiple sclerosis (MS) and experimental autoimmune encephalomyelitis (EAE, a model for MS). A suitable procedure is described in detail in Current Protocols in Immunology, above, unit 4.5.

EAE is a T cell mediated autoimmune disease characterized by T cell and mononuclear cell inflammation and subsequent demyelination of axons in the central nervous system. EAE is generally 30 considered to be a relevant animal model for MS in humans. Bolton, C., *Multiple Sclerosis* (1995) 1:143. Both acute and relapsing-remitting models have been developed. The compounds of the invention can be tested for T cell stimulatory or inhibitory activity against immune mediated demyelinating disease using the protocol described in Current Protocols in Immunology, above, units 15.1 and 15.2. See also the models for myelin disease in which oligodendrocytes or Schwann cells are grafted into the central nervous system as 35 described in Duncan, I. D. *et al*, *Molec. Med. Today* (1997) 554-561.

Contact hypersensitivity is a simple delayed type hypersensitivity *in vivo* assay of cell mediated immune function. In this procedure, cutaneous exposure to exogenous haptens which gives rise to a delayed type hypersensitivity reaction which is measured and quantitated. Contact sensitivity involves an initial sensitizing phase followed by an elicitation phase. The elicitation phase occurs when the T lymphocytes 40 encounter an antigen to which they have had previous contact. Swelling and inflammation occur, making

this an excellent model of human allergic contact dermatitis. A suitable procedure is described in detail in *Current Protocols in Immunology*, Eds. J. E. Cologan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach and W. Strober, John Wiley & Sons, Inc., 1994, unit 4.2. See also Grabbe, S. and Schwarz, T, *Immun. Today* 19 (1): 37-44 (1998) .

5 An animal model for arthritis is collagen-induced arthritis. This model shares clinical, histological and immunological characteristics of human autoimmune rheumatoid arthritis and is an acceptable model for human autoimmune arthritis. Mouse and rat models are characterized by synovitis, erosion of cartilage and subchondral bone. The compounds of the invention can be tested for activity against autoimmune arthritis using the protocols described in *Current Protocols in Immunology*, above, units 15.5. See also the model 10 using a monoclonal antibody to CD18 and VLA-4 integrins described in Issekutz, A.C. *et al.*, *Immunology* (1996) 88:569.

15 A model of asthma has been described in which antigen-induced airway hyper-reactivity, pulmonary eosinophilia and inflammation are induced by sensitizing an animal with ovalbumin and then challenging the animal with the same protein delivered by aerosol. Several animal models (guinea pig, rat, non-human primate) show symptoms similar to atopic asthma in humans upon challenge with aerosol 20 antigens. Murine models have many of the features of human asthma. Suitable procedures to test the compounds of the invention for activity and effectiveness in the treatment of asthma are described by Wolyniec, W. W. *et al.*, *Am. J. Respir. Cell Mol. Biol.* (1998) 18:777 and the references cited therein.

25 Additionally, the compounds of the invention can be tested on animal models for psoriasis like diseases. Evidence suggests a T cell pathogenesis for psoriasis. The compounds of the invention can be tested in the scid/scid mouse model described by Schon, M. P. *et al.*, *Nat. Med.* (1997) 3:183, in which the mice demonstrate histopathologic skin lesions resembling psoriasis. Another suitable model is the human skin/scid mouse chimera prepared as described by Nickoloff, B. J. *et al.*, *Am. J. Path.* (1995) 146:580.

30 Recombinant (transgenic) animal models can be engineered by introducing the coding portion of the genes identified herein into the genome of animals of interest, using standard techniques for producing transgenic animals. Animals that can serve as a target for transgenic manipulation include, without limitation, mice, rats, rabbits, guinea pigs, sheep, goats, pigs, and non-human primates, *e.g.*, baboons, chimpanzees and monkeys. Techniques known in the art to introduce a transgene into such animals include pronucleic microinjection (Hoppe and Wanger, U.S. Patent No. 4,873,191); retrovirus-mediated gene transfer into germ lines (*e.g.*, Van der Putten *et al.*, *Proc. Natl. Acad. Sci. USA* 82, 6148-615 [1985]); gene targeting in embryonic stem cells (Thompson *et al.*, *Cell* 56, 313-321 [1989]); electroporation of embryos (*Lo, Mol. Cel. Biol.* 3, 1803-1814 [1983]); sperm-mediated gene transfer (Lavitrano *et al.*, *Cell* 57, 717-73 [1989]). For review, see, for example, U.S. Patent No. 4,736,866.

35 For the purpose of the present invention, transgenic animals include those that carry the transgene only in part of their cells ("mosaic animals"). The transgene can be integrated either as a single transgene, or in concatamers, *e.g.*, head-to-head or head-to-tail tandems. Selective introduction of a transgene into a particular cell type is also possible by following, for example, the technique of Lasko *et al.*, *Proc. Natl. Acad. Sci. USA* 89, 6232-636 (1992).

40 The expression of the transgene in transgenic animals can be monitored by standard techniques. For example, Southern blot analysis or PCR amplification can be used to verify the integration of the

transgene. The level of mRNA expression can then be analyzed using techniques such as *in situ* hybridization, Northern blot analysis, PCR, or immunocytochemistry.

The animals may be further examined for signs of immune disease pathology, for example by histological examination to determine infiltration of immune cells into specific tissues. Blocking 5 experiments can also be performed in which the transgenic animals are treated with the compounds of the invention to determine the extent of the T cell proliferation stimulation or inhibition of the compounds. In these experiments, blocking antibodies which bind to the PRO polypeptide, prepared as described above, are administered to the animal and the effect on immune function is determined.

Alternatively, "knock out" animals can be constructed which have a defective or altered gene 10 encoding a polypeptide identified herein, as a result of homologous recombination between the endogenous gene encoding the polypeptide and altered genomic DNA encoding the same polypeptide introduced into an embryonic cell of the animal. For example, cDNA encoding a particular polypeptide can be used to clone genomic DNA encoding that polypeptide in accordance with established techniques. A portion of the genomic DNA encoding a particular polypeptide can be deleted or replaced with another gene, such as a 15 gene encoding a selectable marker which can be used to monitor integration. Typically, several kilobases of unaltered flanking DNA (both at the 5' and 3' ends) are included in the vector [see e.g., Thomas and Capecchi, *Cell*, 51:503 (1987) for a description of homologous recombination vectors]. The vector is introduced into an embryonic stem cell line (e.g., by electroporation) and cells in which the introduced DNA 20 has homologously recombined with the endogenous DNA are selected [see e.g., Li *et al.*, *Cell*, 69:915 (1992)]. The selected cells are then injected into a blastocyst of an animal (e.g., a mouse or rat) to form aggregation chimeras [see e.g., Bradley, 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 to create a 25 "knock out" animal. Progeny harboring the homologously recombined DNA in their germ cells can be identified by standard techniques and used to breed animals in which all cells of the animal contain the homologously recombined DNA. Knockout animals can be characterized for instance, for their ability to defend against certain pathological conditions and for their development of pathological conditions due to absence of the polypeptide.

#### I. ImmunoAdjuvant Therapy

30 In one embodiment, the immunostimulating compounds of the invention can be used in immunoadjuvant therapy for the treatment of tumors (cancer). It is now well established that T cells recognize human tumor specific antigens. One group of tumor antigens, encoded by the MAGE, BAGE and GAGE families of genes, are silent in all adult normal tissues, but are expressed in significant amounts in tumors, such as melanomas, lung tumors, head and neck tumors, and bladder carcinomas. DeSmet, C. *et al.*, 35 (1996) *Proc. Natl. Acad. Sci. USA*, 93:7149. It has been shown that costimulation of T cells induces tumor regression and an antitumor response both *in vitro* and *in vivo*. Melero, I. *et al.*, *Nature Medicine* (1997) 3:682; Kwon, E. D. *et al.*, *Proc. Natl. Acad. Sci. USA* (1997) 94: 8099; Lynch, D. H. *et al.*, *Nature Medicine* (1997) 3:625; Finn, O. J. and Lotze, M. T., *J. Immunol.* (1998) 21:114. The stimulatory compounds of the invention can be administered as adjuvants, alone or together with a growth regulating agent, cytotoxic agent 40 or chemotherapeutic agent, to stimulate T cell proliferation/activation and an antitumor response to tumor

antigens. The growth regulating, cytotoxic, or chemotherapeutic agent may be administered in conventional amounts using known administration regimes. Immunostimulating activity by the compounds of the invention allows reduced amounts of the growth regulating, cytotoxic, or chemotherapeutic agents thereby potentially lowering the toxicity to the patient.

5        J.        Screening Assays for Drug Candidates

Screening assays for drug candidates are designed to identify compounds that bind to or complex with the polypeptides encoded by the genes identified herein or a biologically active fragment thereof, or otherwise interfere with the interaction of the encoded polypeptides with other cellular proteins. Such screening assays will include assays amenable to high-throughput screening of chemical libraries, making 10 them particularly suitable for identifying small molecule drug candidates. Small molecules contemplated include synthetic organic or inorganic compounds, including peptides, preferably soluble peptides, (poly)peptide-immunoglobulin fusions, and, in particular, antibodies including, without limitation, poly- and monoclonal antibodies and antibody fragments, single-chain antibodies, anti-idiotypic antibodies, and chimeric or humanized versions of such antibodies or fragments, as well as human antibodies and antibody 15 fragments. The assays can be performed in a variety of formats, including protein-protein binding assays, biochemical screening assays, immunoassays and cell based assays, which are well characterized in the art. All assays are common in that they call for contacting the drug candidate with a polypeptide encoded by a nucleic acid identified herein under conditions and for a time sufficient to allow these two components to interact.

20        In binding assays, the interaction is binding and the complex formed can be isolated or detected in the reaction mixture. In a particular embodiment, the polypeptide encoded by the gene identified herein or the drug candidate is immobilized on a solid phase, *e.g.*, on a microtiter plate, by covalent or non-covalent attachments. Non-covalent attachment generally is accomplished by coating the solid surface with a solution 25 of the polypeptide and drying. Alternatively, an immobilized antibody, *e.g.*, a monoclonal antibody, specific for the polypeptide to be immobilized can be used to anchor it to a solid surface. The assay is performed by adding the non-immobilized component, which may be labeled by a detectable label, to the immobilized component, *e.g.*, the coated surface containing the anchored component. When the reaction is complete, the non-reacted components are removed, *e.g.*, by washing, and complexes anchored on the solid surface are detected. When the originally non-immobilized component carries a detectable label, the detection of label 30 immobilized on the surface indicates that complexing occurred. Where the originally non-immobilized component does not carry a label, complexing can be detected, for example, by using a labelled antibody specifically binding the immobilized complex.

35        If the candidate compound interacts with but does not bind to a particular protein encoded by a gene identified herein, its interaction with that protein can be assayed by methods well known for detecting protein-protein interactions. Such assays include traditional approaches, such as, cross-linking, co-immunoprecipitation, and co-purification through gradients or chromatographic columns. In addition, protein-protein interactions can be monitored by using a yeast-based genetic system described by Fields and co-workers [Fields and Song, *Nature (London)* 340, 245-246 (1989); Chien *et al.*, *Proc. Natl. Acad. Sci. USA* 88, 9578-9582 (1991)] as disclosed by Chevray and Nathans, *Proc. Natl. Acad. Sci. USA* 89, 5789-5793 40 (1991). Many transcriptional activators, such as yeast GAL4, consist of two physically discrete modular

domains, one acting as the DNA-binding domain, while the other one functioning as the transcription activation domain. The yeast expression system described in the foregoing publications (generally referred to as the "two-hybrid system") takes advantage of this property, and employs two hybrid proteins, one in which the target protein is fused to the DNA-binding domain of GAL4, and another, in which candidate activating proteins are fused to the activation domain. The expression of a GAL1-*lacZ* reporter gene under control of a GAL4-activated promoter depends on reconstitution of GAL4 activity via protein-protein interaction. Colonies containing interacting polypeptides are detected with a chromogenic substrate for  $\beta$ -galactosidase. A complete kit (MATCHMAKER<sup>TM</sup>) for identifying protein-protein interactions between two specific proteins using the two-hybrid technique is commercially available from Clontech. This system can also be extended to map protein domains involved in specific protein interactions as well as to pinpoint amino acid residues that are crucial for these interactions.

In order to find compounds that interfere with the interaction of a gene identified herein and other intra- or extracellular components can be tested, a reaction mixture is usually prepared containing the product of the gene and the intra- or extracellular component under conditions and for a time allowing for the interaction and binding of the two products. To test the ability of a test compound to inhibit binding, the reaction is run in the absence and in the presence of the test compound. In addition, a placebo may be added to a third reaction mixture, to serve as positive control. The binding (complex formation) between the test compound and the intra- or extracellular component present in the mixture is monitored as described above. The formation of a complex in the control reaction(s) but not in the reaction mixture containing the test compound indicates that the test compound interferes with the interaction of the test compound and its reaction partner.

#### K. Compositions and Methods for the Treatment of Immune Related Diseases

The compositions useful in the treatment of immune related diseases include, without limitation, proteins, antibodies, small organic molecules, peptides, phosphopeptides, antisense and ribozyme molecules, triple helix molecules, *etc.* that inhibit or stimulate immune function, for example, T cell proliferation/activation, lymphokine release, or immune cell infiltration.

For example, antisense RNA and RNA molecules act to directly block the translation of mRNA by hybridizing to targeted mRNA and preventing protein translation. When antisense DNA is used, oligodeoxyribonucleotides derived from the translation initiation site, *e.g.*, between about -10 and +10 positions of the target gene nucleotide sequence, are preferred.

Ribozymes are enzymatic RNA molecules capable of catalyzing the specific cleavage of RNA. Ribozymes act by sequence-specific hybridization to the complementary target RNA, followed by endonucleolytic cleavage. Specific ribozyme cleavage sites within a potential RNA target can be identified by known techniques. For further details see, *e.g.*, Rossi, *Current Biology* 4, 469-471 (1994), and PCT publication No. WO 97/33551 (published September 18, 1997).

Nucleic acid molecules in triple helix formation used to inhibit transcription should be single-stranded and composed of deoxynucleotides. The base composition of these oligonucleotides is designed such that it promotes triple helix formation via Hoogsteen base pairing rules, which generally require sizeable stretches of purines or pyrimidines on one strand of a duplex. For further details see, *e.g.*, PCT publication No. WO 97/33551, *supra*.

These molecules can be identified by any or any combination of the screening assays discussed above and/or by any other screening techniques well known for those skilled in the art.

L. Anti-PRO Antibodies

The present invention further provides anti-PRO antibodies. Exemplary antibodies include 5 polyclonal, monoclonal, humanized, bispecific, and heteroconjugate antibodies.

1. Polyclonal Antibodies

The anti-PRO antibodies may comprise polyclonal antibodies. Methods of preparing polyclonal antibodies are known to the skilled artisan. Polyclonal antibodies can be raised in a mammal, for example, by one or more injections of an immunizing agent and, if desired, an adjuvant. Typically, the immunizing 10 agent and/or adjuvant will be injected in the mammal by multiple subcutaneous or intraperitoneal injections. The immunizing agent may include the PRO polypeptide or a fusion protein thereof. It may be useful to conjugate the immunizing agent to a protein known to be immunogenic in the mammal being immunized. Examples of such immunogenic proteins include but are not limited to keyhole limpet hemocyanin, serum 15 albumin, bovine thyroglobulin, and soybean trypsin inhibitor. Examples of adjuvants which may be employed include Freund's complete adjuvant and MPL-TDM adjuvant (monophosphoryl Lipid A, synthetic trehalose dicorynomycolate). The immunization protocol may be selected by one skilled in the art without undue experimentation.

2. Monoclonal Antibodies

The anti-PRO antibodies may, alternatively, be monoclonal antibodies. Monoclonal antibodies may 20 be prepared using hybridoma methods, such as those described by Kohler and Milstein, Nature, 256:495 (1975). In a hybridoma method, a mouse, hamster, or other appropriate host animal, is typically immunized with an immunizing agent to elicit lymphocytes that produce or are capable of producing antibodies that will specifically bind to the immunizing agent. Alternatively, the lymphocytes may be immunized *in vitro*.

The immunizing agent will typically include the PRO polypeptide or a fusion protein thereof. 25 Generally, either peripheral blood lymphocytes ("PBLs") are used if cells of human origin are desired, or spleen cells or lymph node cells are used if non-human mammalian sources are desired. The lymphocytes are then fused with an immortalized cell line using a suitable fusing agent, such as polyethylene glycol, to form a hybridoma cell [Goding, Monoclonal Antibodies: Principles and Practice, Academic Press, (1986) pp. 59-103]. Immortalized cell lines are usually transformed mammalian cells, particularly myeloma cells of 30 rodent, bovine and human origin. Usually, rat or mouse myeloma cell lines are employed. The hybridoma cells may be cultured in a suitable culture medium that preferably contains one or more substances that inhibit the growth or survival of the unfused, immortalized cells. For example, if the parental cells lack the enzyme hypoxanthine guanine phosphoribosyl transferase (HGPRT or HPRT), the culture medium for the hybridomas typically will include hypoxanthine, aminopterin, and thymidine ("HAT medium"), which 35 substances prevent the growth of HGPRT-deficient cells.

Preferred immortalized cell lines are those that fuse efficiently, support stable high level expression of antibody by the selected antibody-producing cells, and are sensitive to a medium such as HAT medium. More preferred immortalized cell lines are murine myeloma lines, which can be obtained, for instance, from the Salk Institute Cell Distribution Center, San Diego, California and the American Type Culture Collection, 40 Manassas, Virginia. Human myeloma and mouse-human heteromyeloma cell lines also have been described

for the production of human monoclonal antibodies [Kozbor, J. Immunol., 133:3001 (1984); Brodeur et al., Monoclonal Antibody Production Techniques and Applications, Marcel Dekker, Inc., New York, (1987) pp. 51-63].

5 The culture medium in which the hybridoma cells are cultured can then be assayed for the presence of monoclonal antibodies directed against PRO. Preferably, the binding specificity of monoclonal antibodies produced by the hybridoma cells is determined by immunoprecipitation or by an *in vitro* binding assay, such as radioimmunoassay (RIA) or enzyme-linked immunoabsorbent assay (ELISA). Such techniques and assays are known in the art. The binding affinity of the monoclonal antibody can, for example, be determined by the Scatchard analysis of Munson and Pollard, Anal. Biochem., 107:220 (1980).

10 After the desired hybridoma cells are identified, the clones may be subcloned by limiting dilution procedures and grown by standard methods [Goding, supra]. Suitable culture media for this purpose include, for example, Dulbecco's Modified Eagle's Medium and RPMI-1640 medium. Alternatively, the hybridoma cells may be grown *in vivo* as ascites in a mammal.

15 The monoclonal antibodies secreted by the subclones may be isolated or purified from the culture medium or ascites fluid by conventional immunoglobulin purification procedures such as, for example, protein A-Sepharose, hydroxylapatite chromatography, gel electrophoresis, dialysis, or affinity chromatography.

20 The monoclonal antibodies may also be made by recombinant DNA methods, such as those described in U.S. Patent No. 4,816,567. DNA encoding the monoclonal antibodies of the invention can be readily isolated and sequenced using conventional procedures (e.g., by using oligonucleotide probes that are capable of binding specifically to genes encoding the heavy and light chains of murine antibodies). The hybridoma cells of the invention serve as a preferred source of such DNA. Once isolated, the DNA may be placed into expression vectors, which are then transfected into host cells such as simian COS cells, Chinese hamster ovary (CHO) cells, or myeloma cells that do not otherwise produce immunoglobulin protein, to 25 obtain the synthesis of monoclonal antibodies in the recombinant host cells. The DNA also may be modified, for example, by substituting the coding sequence for human heavy and light chain constant domains in place of the homologous murine sequences [U.S. Patent No. 4,816,567; Morrison et al., supra] or by covalently joining to the immunoglobulin coding sequence all or part of the coding sequence for a non-immunoglobulin polypeptide. Such a non-immunoglobulin polypeptide can be substituted for the constant 30 domains of an antibody of the invention, or can be substituted for the variable domains of one antigen-combining site of an antibody of the invention to create a chimeric bivalent antibody.

35 The antibodies may be monovalent antibodies. Methods for preparing monovalent antibodies are well known in the art. For example, one method involves recombinant expression of immunoglobulin light chain and modified heavy chain. The heavy chain is truncated generally at any point in the Fc region so as to prevent heavy chain crosslinking. Alternatively, the relevant cysteine residues are substituted with another amino acid residue or are deleted so as to prevent crosslinking.

*In vitro* methods are also suitable for preparing monovalent antibodies. Digestion of antibodies to produce fragments thereof, particularly, Fab fragments, can be accomplished using routine techniques known in the art.

3. Human and Humanized Antibodies

The anti-PRO antibodies of the invention may further comprise humanized antibodies or human antibodies. Humanized forms of non-human (e.g., murine) antibodies are chimeric immunoglobulins, immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', F(ab')<sub>2</sub> or other antigen-binding subsequences of antibodies) which contain minimal sequence derived from non-human immunoglobulin. Humanized antibodies include human immunoglobulins (recipient antibody) in which residues from a complementary determining region (CDR) of the recipient are replaced by residues from a CDR of a non-human species (donor antibody) such as mouse, rat or rabbit having the desired specificity, affinity and capacity. In some instances, Fv framework residues of the human immunoglobulin are replaced by corresponding non-human residues. Humanized antibodies may also comprise residues which are found neither in the recipient antibody nor in the imported CDR or framework sequences. In general, the humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the FR regions are those of a human immunoglobulin consensus sequence. The humanized antibody optimally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin [Jones et al., Nature, 321:522-525 (1986); Riechmann et al., Nature, 332:323-329 (1988); and Presta, Curr. Op. Struct. Biol., 2:593-596 (1992)].

Methods for humanizing non-human antibodies are well known in the art. Generally, a humanized antibody has one or more amino acid residues introduced into it from a source which is non-human. These non-human amino acid residues are often referred to as "import" residues, which are typically taken from an "import" variable domain. Humanization can be essentially performed following the method of Winter and co-workers [Jones et al., Nature, 321:522-525 (1986); Riechmann et al., Nature, 332:323-327 (1988); Verhoeyen et al., Science, 239:1534-1536 (1988)], by substituting rodent CDRs or CDR sequences for the corresponding sequences of a human antibody. Accordingly, such "humanized" antibodies are chimeric antibodies (U.S. Patent No. 4,816,567), wherein substantially less than an intact human variable domain has been substituted by the corresponding sequence from a non-human species. In practice, humanized antibodies are typically human antibodies in which some CDR residues and possibly some FR residues are substituted by residues from analogous sites in rodent antibodies.

Human antibodies can also be produced using various techniques known in the art, including phage display libraries [Hoogenboom and Winter, J. Mol. Biol., 227:381 (1991); Marks et al., J. Mol. Biol., 222:581 (1991)]. The techniques of Cole et al. and Boerner et al. are also available for the preparation of human monoclonal antibodies (Cole et al., Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, p. 77 (1985) and Boerner et al., J. Immunol., 147(1):86-95 (1991)]. Similarly, human antibodies can be made by introducing of human immunoglobulin loci into transgenic animals, e.g., mice in which the endogenous immunoglobulin genes have been partially or completely inactivated. Upon challenge, human antibody production is observed, which closely resembles that seen in humans in all respects, including gene rearrangement, assembly, and antibody repertoire. This approach is described, for example, in U.S. Patent Nos. 5,545,807; 5,545,806; 5,569,825; 5,625,126; 5,633,425; 5,661,016, and in the following scientific publications: Marks et al., Bio/Technology 10, 779-783 (1992); Lonberg et al., Nature 368 856-859 (1994);

Morrison, Nature 368, 812-13 (1994); Fishwild *et al.*, Nature Biotechnology 14, 845-51 (1996); Neuberger, Nature Biotechnology 14, 826 (1996); Lonberg and Huszar, Intern. Rev. Immunol. 13 65-93 (1995).

The antibodies may also be affinity matured using known selection and/or mutagenesis methods as described above. Preferred affinity matured antibodies have an affinity which is five times, more preferably 5 10 times, even more preferably 20 or 30 times greater than the starting antibody (generally murine, humanized or human) from which the matured antibody is prepared.

#### 4. Bispecific Antibodies

Bispecific antibodies are monoclonal, preferably human or humanized, antibodies that have binding specificities for at least two different antigens. In the present case, one of the binding specificities is for the 10 PRO, the other one is for any other antigen, and preferably for a cell-surface protein or receptor or receptor subunit.

Methods for making bispecific antibodies are known in the art. Traditionally, the recombinant production of bispecific antibodies is based on the co-expression of two immunoglobulin heavy-chain/light-chain pairs, where the two heavy chains have different specificities [Milstein and Cuello, Nature, 305:537-15 539 (1983)]. Because of the random assortment of immunoglobulin heavy and light chains, these hybridomas (quadromas) produce a potential mixture of ten different antibody molecules, of which only one has the correct bispecific structure. The purification of the correct molecule is usually accomplished by affinity chromatography steps. Similar procedures are disclosed in WO 93/08829, published 13 May 1993, and in Traunecker *et al.*, EMBO J., 10:3655-3659 (1991).

20 Antibody variable domains with the desired binding specificities (antibody-antigen combining sites) can be fused to immunoglobulin constant domain sequences. The fusion preferably is with an immunoglobulin heavy-chain constant domain, comprising at least part of the hinge, CH2, and CH3 regions. It is preferred to have the first heavy-chain constant region (CH1) containing the site necessary for light-chain binding present in at least one of the fusions. DNAs encoding the immunoglobulin heavy-chain 25 fusions and, if desired, the immunoglobulin light chain, are inserted into separate expression vectors, and are co-transfected into a suitable host organism. For further details of generating bispecific antibodies see, for example, Suresh *et al.*, Methods in Enzymology, 121:210 (1986).

According to another approach described in WO 96/27011, the interface between a pair of antibody 30 molecules can be engineered to maximize the percentage of heterodimers which are recovered from recombinant cell culture. The preferred interface comprises at least a part of the CH3 region of an antibody constant domain. In this method, one or more small amino acid side chains from the interface of the first antibody molecule are replaced with larger side chains (e.g. tyrosine or tryptophan). Compensatory "cavities" of identical or similar size to the large side chain(s) are created on the interface of the second antibody molecule by replacing large amino acid side chains with smaller ones (e.g. alanine or threonine). 35 This provides a mechanism for increasing the yield of the heterodimer over other unwanted end-products such as homodimers.

Bispecific antibodies can be prepared as full length antibodies or antibody fragments (e.g. F(ab')<sub>2</sub> 40 bispecific antibodies). Techniques for generating bispecific antibodies from antibody fragments have been described in the literature. For example, bispecific antibodies can be prepared can be prepared using chemical linkage. Brennan *et al.*, Science 229:81 (1985) describe a procedure wherein intact antibodies are

proteolytically cleaved to generate  $F(ab')_2$  fragments. These fragments are reduced in the presence of the dithiol complexing agent sodium arsenite to stabilize vicinal dithiols and prevent intermolecular disulfide formation. The  $Fab'$  fragments generated are then converted to thionitrobenzoate (TNB) derivatives. One of the  $Fab'$ -TNB derivatives is then reconverted to the  $Fab'$ -thiol by reduction with mercaptoethylamine and is 5 mixed with an equimolar amount of the other  $Fab'$ -TNB derivative to form the bispecific antibody. The bispecific antibodies produced can be used as agents for the selective immobilization of enzymes.

10  $Fab'$  fragments may be directly recovered from *E. coli* and chemically coupled to form bispecific antibodies. Shalaby *et al.*, *J. Exp. Med.* 175:217-225 (1992) describe the production of a fully humanized bispecific antibody  $F(ab')_2$  molecule. Each  $Fab'$  fragment was separately secreted from *E. coli* and subjected to directed chemical coupling *in vitro* to form the bispecific antibody. The bispecific antibody thus formed was able to bind to cells overexpressing the ErbB2 receptor and normal human T cells, as well as trigger the lytic activity of human cytotoxic lymphocytes against human breast tumor targets.

15 Various technique for making and isolating bispecific antibody fragments directly from recombinant cell culture have also been described. For example, bispecific antibodies have been produced using leucine zippers. Kostelny *et al.*, *J. Immunol.* 148(5):1547-1553 (1992). The leucine zipper peptides from the Fos and Jun proteins were linked to the  $Fab'$  portions of two different antibodies by gene fusion. The antibody homodimers were reduced at the hinge region to form monomers and then re-oxidized to form the antibody heterodimers. This method can also be utilized for the production of antibody homodimers. The "diabody" technology described by Hollinger *et al.*, *Proc. Natl. Acad. Sci. USA* 90:6444-6448 (1993) 20 has provided an alternative mechanism for making bispecific antibody fragments. The fragments comprise a heavy-chain variable domain ( $V_H$ ) connected to a light-chain variable domain ( $V_L$ ) by a linker which is too short to allow pairing between the two domains on the same chain. Accordingly, the  $V_H$  and  $V_L$  domains of one fragment are forced to pair with the complementary  $V_L$  and  $V_H$  domains of another fragment, thereby forming two antigen-binding sites. Another strategy for making bispecific antibody fragments by the use of 25 single-chain Fv (sFv) dimers has also been reported. See, Gruber *et al.*, *J. Immunol.* 152:5368 (1994). Antibodies with more than two valencies are contemplated. For example, trispecific antibodies can be prepared. Tutt *et al.*, *J. Immunol.* 147:60 (1991).

30 Exemplary bispecific antibodies may bind to two different epitopes on a given PRO polypeptide herein. Alternatively, an anti-PRO polypeptide arm may be combined with an arm which binds to a triggering molecule on a leukocyte such as a T-cell receptor molecule (e.g. CD2, CD3, CD28, or B7), or Fc receptors for IgG (Fc $\gamma$ R), such as Fc $\gamma$ RI (CD64), Fc $\gamma$ RII (CD32) and Fc $\gamma$ RIII (CD16) so as to focus cellular defense mechanisms to the cell expressing the particular PRO polypeptide. Bispecific antibodies may also be used to localize cytotoxic agents to cells which express a particular PRO polypeptide. These antibodies possess a PRO-binding arm and an arm which binds a cytotoxic agent or a radionuclide chelator, such as 35 EOTUBE, DPTA, DOTA, or TETA. Another bispecific antibody of interest binds the PRO polypeptide and further binds tissue factor (TF).

##### 5. Heteroconjugate Antibodies

Heteroconjugate antibodies are also within the scope of the present invention. Heteroconjugate antibodies are composed of two covalently joined antibodies. Such antibodies have, for example, been

proposed to target immune system cells to unwanted cells [U.S. Patent No. 4,676,980], and for treatment of HIV infection [WO 91/00360; WO 92/200373; EP 03089]. It is contemplated that the antibodies may be prepared *in vitro* using known methods in synthetic protein chemistry, including those involving crosslinking agents. For example, immunotoxins may be constructed using a disulfide exchange reaction or 5 by forming a thioether bond. Examples of suitable reagents for this purpose include iminothiolate and methyl-4-mercaptopbutyrimidate and those disclosed, for example, in U.S. Patent No. 4,676,980.

#### 6. Effector Function Engineering

It may be desirable to modify the antibody of the invention with respect to effector function, so as to enhance, *e.g.*, the effectiveness of the antibody in treating cancer. For example, cysteine residue(s) may 10 be introduced into the Fc region, thereby allowing interchain disulfide bond formation in this region. The homodimeric antibody thus generated may have improved internalization capability and/or increased complement-mediated cell killing and antibody-dependent cellular cytotoxicity (ADCC). See Caron *et al.*, *J. Exp Med.*, 176: 1191-1195 (1992) and Shope, *J. Immunol.*, 148: 2918-2922 (1992). Homodimeric 15 antibodies with enhanced anti-tumor activity may also be prepared using heterobifunctional cross-linkers as described in Wolff *et al.* *Cancer Research*, 53: 2560-2565 (1993). Alternatively, an antibody can be engineered that has dual Fc regions and may thereby have enhanced complement lysis and ADCC capabilities. See Stevenson *et al.*, *Anti-Cancer Drug Design*, 3: 219-230 (1989).

#### 7. Immunoconjugates

The invention also pertains to immunoconjugates comprising an antibody conjugated to a cytotoxic 20 agent such as a chemotherapeutic agent, toxin (*e.g.*, an enzymatically active toxin of bacterial, fungal, plant, or animal origin, or fragments thereof), or a radioactive isotope (*i.e.*, a radioconjugate).

Chemotherapeutic agents useful in the generation of such immunoconjugates have been described above. Enzymatically active toxins and fragments thereof that can be used include diphtheria A chain, 25 nonbinding active fragments of diphtheria toxin, exotoxin A chain (from *Pseudomonas aeruginosa*), ricin A chain, abrin A chain, modeccin A chain, alpha-sarcin, *Aleurites fordii* proteins, dianthin proteins, *Phytolaca americana* proteins (PAPI, PAPII, and PAP-S), momordica charantia inhibitor, curcin, crotin, sapaonaria officinalis inhibitor, gelonin, mitogellin, restrictocin, phenomycin, enomycin, and the tricothecenes. A variety of radionuclides are available for the production of radioconjugated antibodies. Examples include  $^{212}\text{Bi}$ ,  $^{131}\text{I}$ ,  $^{131}\text{In}$ ,  $^{90}\text{Y}$ , and  $^{186}\text{Re}$ .

Conjugates of the antibody and cytotoxic agent are made using a variety of bifunctional protein-coupling agents such as N-succinimidyl-3-(2-pyridyldithiol) propionate (SPDP), iminothiolane (IT), bifunctional derivatives of imidoesters (such as dimethyl adipimidate HCL), active esters (such as disuccinimidyl suberate), aldehydes (such as glutaraldehyde), bis-azido compounds (such as bis-(*p*-azidobenzoyl) hexanediamine), bis-diazonium derivatives (such as bis-(*p*-diazoniumbenzoyl)-ethylenediamine), diisocyanates (such as tolyene 2,6-diisocyanate), and bis-active fluorine compounds (such as 1,5-difluoro-2,4-dinitrobenzene). For example, a ricin immunotoxin can be prepared as described in Vitetta *et al.*, *Science*, 238: 1098 (1987). Carbon-14-labeled 1-isothiocyanatobenzyl-3-methyldiethylene triaminepentaacetic acid (MX-DTPA) is an exemplary chelating agent for conjugation of radionucleotide to 30 the antibody. See WO94/11026.

In another embodiment, the antibody may be conjugated to a "receptor" (such streptavidin) for utilization in tumor pretargeting wherein the antibody-receptor conjugate is administered to the patient, followed by removal of unbound conjugate from the circulation using a clearing agent and then administration of a "ligand" (e.g., avidin) that is conjugated to a cytotoxic agent (e.g., a radionucleotide).<sup>1</sup>

5                   8.        Immunoliposomes

The antibodies disclosed herein may also be formulated as immunoliposomes. Liposomes containing the antibody are prepared by methods known in the art, such as described in Epstein *et al.*, Proc. Natl. Acad. Sci. USA, 82: 3688 (1985); Hwang *et al.*, Proc. Natl. Acad. Sci. USA, 77: 4030 (1980); and U.S. Pat. Nos. 4,485,045 and 4,544,545. Liposomes with enhanced circulation time are disclosed in U.S. Patent 10 No. 5,013,556.

Particularly useful liposomes can be generated by the reverse-phase evaporation method with a lipid composition comprising phosphatidylcholine, cholesterol, and PEG-derivatized phosphatidylethanolamine (PEG-PE). Liposomes are extruded through filters of defined pore size to yield liposomes with the desired diameter. Fab' fragments of the antibody of the present invention can be 15 conjugated to the liposomes as described in Martin *et al.*, J. Biol. Chem., 257: 286-288 (1982) via a disulfide-interchange reaction. A chemotherapeutic agent (such as Doxorubicin) is optionally contained within the liposome. See Gabizon *et al.*, J. National Cancer Inst., 81(19): 1484 (1989).

M.        Pharmaceutical Compositions

The active PRO molecules of the invention (e.g., PRO polypeptides, anti-PRO antibodies, and/or variants of each) as well as other molecules identified by the screening assays disclosed above, can be administered for the treatment of immune related diseases, in the form of pharmaceutical compositions.

Therapeutic formulations of the active PRO molecule, preferably a polypeptide or antibody of the invention, are prepared for storage by mixing the active molecule having the desired degree of purity with optional pharmaceutically acceptable carriers, excipients or stabilizers (*Remington's Pharmaceutical Sciences* 16th edition, Osol, A. Ed. [1980]), in the form of lyophilized formulations or aqueous solutions. Acceptable carriers, excipients, or stabilizers are nontoxic to recipients at the dosages and concentrations employed, and include buffers such as phosphate, citrate, and other organic acids; antioxidants including ascorbic acid and methionine; preservatives (such as octadecylmethylbenzyl ammonium chloride; hexamethonium chloride; benzalkonium chloride, benzethonium chloride; phenol, butyl or benzyl alcohol; 30 alkyl parabens such as methyl or propyl paraben; catechol; resorcinol; cyclohexanol; 3-pentanol; and m-cresol); low molecular weight (less than about 10 residues) polypeptides; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino acids such as glycine, glutamine, asparagine, histidine, arginine, or lysine; monosaccharides, disaccharides, and other carbohydrates including glucose, mannose, or dextrans; chelating agents such as EDTA; sugars such as 35 sucrose, mannitol, trehalose or sorbitol; salt-forming counter-ions such as sodium; metal complexes (e.g., Zn-protein complexes); and/or non-ionic surfactants such as TWEEN<sup>TM</sup>, PLURONICS<sup>TM</sup> or polyethylene glycol (PEG).

Compounds identified by the screening assays disclosed herein can be formulated in an analogous manner, using standard techniques well known in the art.

40               Lipofections or liposomes can also be used to deliver the PRO molecule into cells. Where antibody

fragments are used, the smallest inhibitory fragment which specifically binds to the binding domain of the target protein is preferred. For example, based upon the variable region sequences of an antibody, peptide molecules can be designed which retain the ability to bind the target protein sequence. Such peptides can be synthesized chemically and/or produced by recombinant DNA technology (see, *e.g.*, Marasco *et al.*, *Proc. 5 Natl. Acad. Sci. USA* **90**, 7889-7893 [1993]).

The formulation herein may also contain more than one active compound as necessary for the particular indication being treated, preferably those with complementary activities that do not adversely affect each other. Alternatively, or in addition, the composition may comprise a cytotoxic agent, cytokine or growth inhibitory agent. Such molecules are suitably present in combination in amounts that are effective 10 for the purpose intended.

The active PRO molecules may also be entrapped in microcapsules prepared, for example, by coacervation techniques or by interfacial polymerization, for example, hydroxymethylcellulose or gelatin-microcapsules and poly-(methylmethacrylate) microcapsules, respectively, in colloidal drug delivery systems (for example, liposomes, albumin microspheres, microemulsions, nano-particles and nanocapsules) or in 15 macroemulsions. Such techniques are disclosed in *Remington's Pharmaceutical Sciences* 16th edition, Osol, A. Ed. (1980).

The formulations to be used for *in vivo* administration must be sterile. This is readily accomplished by filtration through sterile filtration membranes.

Sustained-release preparations or the PRO molecules may be prepared. Suitable examples of 20 sustained-release preparations include semipermeable matrices of solid hydrophobic polymers containing the antibody, which matrices are in the form of shaped articles, *e.g.*, films, or microcapsules. Examples of sustained-release matrices include polyesters, hydrogels (for example, poly(2-hydroxyethyl-methacrylate), or poly(vinylalcohol)), polylactides (U.S. Pat. No. 3,773,919), copolymers of L-glutamic acid and  $\gamma$ -ethyl-L-glutamate, non-degradable ethylene-vinyl acetate, degradable lactic acid-glycolic acid copolymers such as 25 the LUPRON DEPOT<sup>TM</sup> (injectable microspheres composed of lactic acid-glycolic acid copolymer and leuprolide acetate), and poly-D-( $\text{--}$ )-3-hydroxybutyric acid. While polymers such as ethylene-vinyl acetate and lactic acid-glycolic acid enable release of molecules for over 100 days, certain hydrogels release proteins for shorter time periods. When encapsulated antibodies remain in the body for a long time, they may 30 denature or aggregate as a result of exposure to moisture at 37°C, resulting in a loss of biological activity and possible changes in immunogenicity. Rational strategies can be devised for stabilization depending on the mechanism involved. For example, if the aggregation mechanism is discovered to be intermolecular S-S bond formation through thio-disulfide interchange, stabilization may be achieved by modifying sulfhydryl residues, lyophilizing from acidic solutions, controlling moisture content, using appropriate additives, and developing specific polymer matrix compositions.

35 N. Methods of Treatment

It is contemplated that the polypeptides, antibodies and other active compounds of the present invention may be used to treat various immune related diseases and conditions, such as T cell mediated diseases, including those characterized by infiltration of inflammatory cells into a tissue, stimulation of T-cell proliferation, inhibition of T-cell proliferation, increased or decreased vascular permeability or the 40 inhibition thereof.

Exemplary conditions or disorders to be treated with the polypeptides, antibodies and other compounds of the invention, include, but are not limited to systemic lupus erythematosis, rheumatoid arthritis, juvenile chronic arthritis, osteoarthritis, spondyloarthropathies, systemic sclerosis (scleroderma), idiopathic inflammatory myopathies (dermatomyositis, polymyositis), Sjögren's syndrome, systemic 5 vasculitis, sarcoidosis, autoimmune hemolytic anemia (immune pancytopenia, paroxysmal nocturnal hemoglobinuria), autoimmune thrombocytopenia (idiopathic thrombocytopenic purpura, immune-mediated thrombocytopenia), thyroiditis (Grave's disease, Hashimoto's thyroiditis, juvenile lymphocytic thyroiditis, atrophic thyroiditis), diabetes mellitus, immune-mediated renal disease (glomerulonephritis, tubulointerstitial 10 nephritis), demyelinating diseases of the central and peripheral nervous systems such as multiple sclerosis, idiopathic demyelinating polyneuropathy or Guillain-Barré syndrome, and chronic inflammatory demyelinating polyneuropathy, hepatobiliary diseases such as infectious hepatitis (hepatitis A, B, C, D, E and other non-hepatotropic viruses), autoimmune chronic active hepatitis, primary biliary cirrhosis, 15 granulomatous hepatitis, and sclerosing cholangitis, inflammatory bowel disease (ulcerative colitis: Crohn's disease), gluten-sensitive enteropathy, and Whipple's disease, autoimmune or immune-mediated skin diseases including bullous skin diseases, erythema multiforme and contact dermatitis, psoriasis, allergic diseases such as asthma, allergic rhinitis, atopic dermatitis, food hypersensitivity and urticaria, immunologic diseases of the lung such as eosinophilic pneumonias, idiopathic pulmonary fibrosis and hypersensitivity pneumonitis, transplantation associated diseases including graft rejection and graft -versus-host-disease.

In systemic lupus erythematosus, the central mediator of disease is the production of auto-reactive 20 antibodies to self proteins/tissues and the subsequent generation of immune-mediated inflammation. Antibodies either directly or indirectly mediate tissue injury. Though T lymphocytes have not been shown to be directly involved in tissue damage, T lymphocytes are required for the development of auto-reactive antibodies. The genesis of the disease is thus T lymphocyte dependent. Multiple organs and systems are 25 affected clinically including kidney, lung, musculoskeletal system, mucocutaneous, eye, central nervous system, cardiovascular system, gastrointestinal tract, bone marrow and blood.

Rheumatoid arthritis (RA) is a chronic systemic autoimmune inflammatory disease that mainly involves the synovial membrane of multiple joints with resultant injury to the articular cartilage. The pathogenesis is T lymphocyte dependent and is associated with the production of rheumatoid factors, auto-antibodies directed against self IgG, with the resultant formation of immune complexes that attain high 30 levels in joint fluid and blood. These complexes in the joint may induce the marked infiltrate of lymphocytes and monocytes into the synovium and subsequent marked synovial changes; the joint space/fluid if infiltrated by similar cells with the addition of numerous neutrophils. Tissues affected are primarily the joints, often in symmetrical pattern. However, extra-articular disease also occurs in two major forms. One form is the development of extra-articular lesions with ongoing progressive joint disease and 35 typical lesions of pulmonary fibrosis, vasculitis, and cutaneous ulcers. The second form of extra-articular disease is the so called Felty's syndrome which occurs late in the RA disease course, sometimes after joint disease has become quiescent, and involves the presence of neutropenia, thrombocytopenia and splenomegaly. This can be accompanied by vasculitis in multiple organs with formations of infarcts, skin ulcers and gangrene. Patients often also develop rheumatoid nodules in the subcutis tissue overlying 40 affected joints; the nodules late stage have necrotic centers surrounded by a mixed inflammatory cell

infiltrate. Other manifestations which can occur in RA include: pericarditis, pleuritis, coronary arteritis, intestinal pneumonitis with pulmonary fibrosis, keratoconjunctivitis sicca, and rheumatoid nodules.

5 Juvenile chronic arthritis is a chronic idiopathic inflammatory disease which begins often at less than 16 years of age. Its phenotype has some similarities to RA; some patients which are rheumatoid factor positive are classified as juvenile rheumatoid arthritis. The disease is sub-classified into three major categories: pauciarticular, polyarticular, and systemic. The arthritis can be severe and is typically destructive and leads to joint ankylosis and retarded growth. Other manifestations can include chronic anterior uveitis and systemic amyloidosis.

10 Spondyloarthropathies are a group of disorders with some common clinical features and the common association with the expression of HLA-B27 gene product. The disorders include: ankylosing spondylitis, Reiter's syndrome (reactive arthritis), arthritis associated with inflammatory bowel disease, spondylitis associated with psoriasis, juvenile onset spondyloarthropathy and undifferentiated spondyloarthropathy. Distinguishing features include sacroileitis with or without spondylitis; inflammatory asymmetric arthritis; association with HLA-B27 (a serologically defined allele of the HLA-B locus of class I 15 MHC); ocular inflammation, and absence of autoantibodies associated with other rheumatoid disease. The cell most implicated as key to induction of the disease is the CD8+ T lymphocyte, a cell which targets antigen presented by class I MHC molecules. CD8+ T cells may react against the class I MHC allele HLA-B27 as if it were a foreign peptide expressed by MHC class I molecules. It has been hypothesized that an epitope of HLA-B27 may mimic a bacterial or other microbial antigenic epitope and thus induce a CD8+ T 20 cells response.

25 Systemic sclerosis (scleroderma) has an unknown etiology. A hallmark of the disease is induration of the skin; likely this is induced by an active inflammatory process. Scleroderma can be localized or systemic; vascular lesions are common and endothelial cell injury in the microvasculature is an early and important event in the development of systemic sclerosis; the vascular injury may be immune mediated. An immunologic basis is implied by the presence of mononuclear cell infiltrates in the cutaneous lesions and the presence of anti-nuclear antibodies in many patients. ICAM-1 is often upregulated on the cell surface of fibroblasts in skin lesions suggesting that T cell interaction with these cells may have a role in the pathogenesis of the disease. Other organs involved include: the gastrointestinal tract: smooth muscle atrophy and fibrosis resulting in abnormal peristalsis/motility; kidney: concentric subendothelial intimal 30 proliferation affecting small arcuate and interlobular arteries with resultant reduced renal cortical blood flow, results in proteinuria, azotemia and hypertension; skeletal muscle: atrophy, interstitial fibrosis; inflammation; lung: interstitial pneumonitis and interstitial fibrosis; and heart: contraction band necrosis, scarring/fibrosis.

35 Idiopathic inflammatory myopathies including dermatomyositis, polymyositis and others are disorders of chronic muscle inflammation of unknown etiology resulting in muscle weakness. Muscle injury/inflammation is often symmetric and progressive. Autoantibodies are associated with most forms. These myositis-specific autoantibodies are directed against and inhibit the function of components, proteins and RNA's, involved in protein synthesis.

40 Sjögren's syndrome is due to immune-mediated inflammation and subsequent functional destruction of the tear glands and salivary glands. The disease can be associated with or accompanied by inflammatory

connective tissue diseases. The disease is associated with autoantibody production against Ro and La antigens, both of which are small RNA-protein complexes. Lesions result in keratoconjunctivitis sicca, xerostomia, with other manifestations or associations including biliary cirrhosis, peripheral or sensory neuropathy, and palpable purpura.

5 Systemic vasculitis are diseases in which the primary lesion is inflammation and subsequent damage to blood vessels which results in ischemia/necrosis/degeneration to tissues supplied by the affected vessels and eventual end-organ dysfunction in some cases. Vasculitides can also occur as a secondary lesion or sequelae to other immune-inflammatory mediated diseases such as rheumatoid arthritis, systemic sclerosis, *etc.*, particularly in diseases also associated with the formation of immune complexes. Diseases in  
10 the primary systemic vasculitis group include: systemic necrotizing vasculitis: polyarteritis nodosa, allergic angiitis and granulomatosis, polyangiitis; Wegener's granulomatosis; lymphomatoid granulomatosis; and giant cell arteritis. Miscellaneous vasculitides include: mucocutaneous lymph node syndrome (MLNS or Kawasaki's disease), isolated CNS vasculitis, Behet's disease, thromboangiitis obliterans (Buerger's disease) and cutaneous necrotizing venulitis. The pathogenic mechanism of most of the types of vasculitis listed is  
15 believed to be primarily due to the deposition of immunoglobulin complexes in the vessel wall and subsequent induction of an inflammatory response either via ADCC, complement activation, or both.

20 Sarcoidosis is a condition of unknown etiology which is characterized by the presence of epithelioid granulomas in nearly any tissue in the body; involvement of the lung is most common. The pathogenesis involves the persistence of activated macrophages and lymphoid cells at sites of the disease with subsequent chronic sequelae resultant from the release of locally and systemically active products released by these cell types.

25 Autoimmune hemolytic anemia including autoimmune hemolytic anemia, immune pancytopenia, and paroxysmal nocturnal hemoglobinuria is a result of production of antibodies that react with antigens expressed on the surface of red blood cells (and in some cases other blood cells including platelets as well) and is a reflection of the removal of those antibody coated cells via complement mediated lysis and/or  
30 ADCC/Fc-receptor-mediated mechanisms.

In autoimmune thrombocytopenia including thrombocytopenic purpura, and immune-mediated thrombocytopenia in other clinical settings, platelet destruction/removal occurs as a result of either antibody or complement attaching to platelets and subsequent removal by complement lysis, ADCC or FC-receptor mediated mechanisms.

35 Thyroiditis including Grave's disease, Hashimoto's thyroiditis, juvenile lymphocytic thyroiditis, and atrophic thyroiditis, are the result of an autoimmune response against thyroid antigens with production of antibodies that react with proteins present in and often specific for the thyroid gland. Experimental models exist including spontaneous models: rats (BUF and BB rats) and chickens (obese chicken strain); inducible models: immunization of animals with either thyroglobulin, thyroid microsomal antigen (thyroid peroxidase).

Type I diabetes mellitus or insulin-dependent diabetes is the autoimmune destruction of pancreatic islet  $\beta$  cells; this destruction is mediated by auto-antibodies and auto-reactive T cells. Antibodies to insulin or the insulin receptor can also produce the phenotype of insulin-non-responsiveness.

Immune mediated renal diseases, including glomerulonephritis and tubulointerstitial nephritis, are the result of antibody or T lymphocyte mediated injury to renal tissue either directly as a result of the production of autoreactive antibodies or T cells against renal antigens or indirectly as a result of the deposition of antibodies and/or immune complexes in the kidney that are reactive against other, non-renal antigens. Thus other immune-mediated diseases that result in the formation of immune-complexes can also induce immune mediated renal disease as an indirect sequela. Both direct and indirect immune mechanisms result in inflammatory response that produces/induces lesion development in renal tissues with resultant organ function impairment and in some cases progression to renal failure. Both humoral and cellular immune mechanisms can be involved in the pathogenesis of lesions.

Demyelinating diseases of the central and peripheral nervous systems, including Multiple Sclerosis; idiopathic demyelinating polyneuropathy or Guillain-Barré syndrome; and Chronic Inflammatory Demyelinating Polyneuropathy, are believed to have an autoimmune basis and result in nerve demyelination as a result of damage caused to oligodendrocytes or to myelin directly. In MS there is evidence to suggest that disease induction and progression is dependent on T lymphocytes. Multiple Sclerosis is a demyelinating disease that is T lymphocyte-dependent and has either a relapsing-remitting course or a chronic progressive course. The etiology is unknown; however, viral infections, genetic predisposition, environment, and autoimmunity all contribute. Lesions contain infiltrates of predominantly T lymphocyte mediated, microglial cells and infiltrating macrophages; CD4+ T lymphocytes are the predominant cell type at lesions. The mechanism of oligodendrocyte cell death and subsequent demyelination is not known but is likely T lymphocyte driven.

Inflammatory and Fibrotic Lung Disease, including Eosinophilic Pneumonias; Idiopathic Pulmonary Fibrosis, and Hypersensitivity Pneumonitis may involve a disregulated immune-inflammatory response. Inhibition of that response would be of therapeutic benefit.

Autoimmune or Immune-mediated Skin Disease including Bullous Skin Diseases, Erythema Multiforme, and Contact Dermatitis are mediated by auto-antibodies, the genesis of which is T lymphocyte-dependent.

Psoriasis is a T lymphocyte-mediated inflammatory disease. Lesions contain infiltrates of T lymphocytes, macrophages and antigen processing cells, and some neutrophils.

Allergic diseases, including asthma; allergic rhinitis; atopic dermatitis; food hypersensitivity; and urticaria are T lymphocyte dependent. These diseases are predominantly mediated by T lymphocyte induced inflammation, IgE mediated-inflammation or a combination of both.

Transplantation associated diseases, including Graft rejection and Graft-Versus-Host-Disease (GVHD) are T lymphocyte-dependent; inhibition of T lymphocyte function is ameliorative. Other diseases in which intervention of the immune and/or inflammatory response have benefit are infectious disease including but not limited to viral infection (including but not limited to AIDS, hepatitis A, B, C, D, E and herpes) bacterial infection, fungal infections, and protozoal and parasitic infections (molecules (or derivatives/agonists) which stimulate the MLR can be utilized therapeutically to enhance the immune response to infectious agents), diseases of immunodeficiency (molecules/derivatives/agonists) which stimulate the MLR can be utilized therapeutically to enhance the immune response for conditions of

inherited, acquired, infectious induced (as in HIV infection), or iatrogenic (*i.e.*, as from chemotherapy) immunodeficiency, and neoplasia.

It has been demonstrated that some human cancer patients develop an antibody and/or T lymphocyte response to antigens on neoplastic cells. It has also been shown in animal models of neoplasia 5 that enhancement of the immune response can result in rejection or regression of that particular neoplasm. Molecules that enhance the T lymphocyte response in the MLR have utility *in vivo* in enhancing the immune response against neoplasia. Molecules which enhance the T lymphocyte proliferative response in the MLR (or small molecule agonists or antibodies that affected the same receptor in an agonistic fashion) can be used therapeutically to treat cancer. Molecules that inhibit the lymphocyte response in the MLR also function *in* 10 *vivo* during neoplasia to suppress the immune response to a neoplasm; such molecules can either be expressed by the neoplastic cells themselves or their expression can be induced by the neoplasm in other cells. Antagonism of such inhibitory molecules (either with antibody, small molecule antagonists or other means) enhances immune-mediated tumor rejection.

Additionally, inhibition of molecules with proinflammatory properties may have therapeutic benefit 15 in reperfusion injury; stroke; myocardial infarction; atherosclerosis; acute lung injury; hemorrhagic shock; burn; sepsis/septic shock; acute tubular necrosis; endometriosis; degenerative joint disease and pancreatitis.

The compounds of the present invention, *e.g.*, polypeptides or antibodies, are administered to a mammal, preferably a human, in accord with known methods, such as intravenous administration as a bolus or by continuous infusion over a period of time, by intramuscular, intraperitoneal, intracerebrospinal, 20 subcutaneous, intra-articular, intrasynovial, intrathecal, oral, topical, or inhalation (intranasal, intrapulmonary) routes. Intravenous or inhaled administration of polypeptides and antibodies is preferred.

In immunoadjuvant therapy, other therapeutic regimens, such administration of an anti-cancer agent, may be combined with the administration of the proteins, antibodies or compounds of the instant invention. For example, the patient to be treated with the immunoadjuvant of the invention may also 25 receive an anti-cancer agent (chemotherapeutic agent) or radiation therapy. Preparation and dosing schedules for such chemotherapeutic agents may be used according to manufacturers' instructions or as determined empirically by the skilled practitioner. Preparation and dosing schedules for such chemotherapy are also described in *Chemotherapy Service* Ed., M.C. Perry, Williams & Wilkins, Baltimore, MD (1992). The chemotherapeutic agent may precede, or follow administration of the immunoadjuvant or may be given 30 simultaneously therewith. Additionally, an anti-estrogen compound such as tamoxifen or an anti-progesterone such as onapristone (see, EP 616812) may be given in dosages known for such molecules.

It may be desirable to also administer antibodies against other immune disease associated or tumor associated antigens, such as antibodies which bind to CD20, CD11a, CD18, ErbB2, EGFR, ErbB3, ErbB4, or vascular endothelial factor (VEGF). Alternatively, or in addition, two or more antibodies binding the 35 same or two or more different antigens disclosed herein may be coadministered to the patient. Sometimes, it may be beneficial to also administer one or more cytokines to the patient. In one embodiment, the PRO polypeptides are coadministered with a growth inhibitory agent. For example, the growth inhibitory agent may be administered first, followed by a PRO polypeptide. However, simultaneous administration or administration first is also contemplated. Suitable dosages for the growth inhibitory agent are those 40 presently used and may be lowered due to the combined action (synergy) of the growth inhibitory agent and

the PRO polypeptide.

For the treatment or reduction in the severity of immune related disease, the appropriate dosage of an a compound of the invention will depend on the type of disease to be treated, as defined above, the severity and course of the disease, whether the agent is administered for preventive or therapeutic purposes, 5 previous therapy, the patient's clinical history and response to the compound, and the discretion of the attending physician. The compound is suitably administered to the patient at one time or over a series of treatments.

For example, depending on the type and severity of the disease, about 1  $\mu$ g/kg to 15 mg/kg (e.g., 0.1-20 mg/kg) of polypeptide or antibody is an initial candidate dosage for administration to the patient, 10 whether, for example, by one or more separate administrations, or by continuous infusion. A typical daily dosage might range from about 1  $\mu$ g/kg to 100 mg/kg or more, depending on the factors mentioned above. For repeated administrations over several days or longer, depending on the condition, the treatment is sustained until a desired suppression of disease symptoms occurs. However, other dosage regimens may be useful. The progress of this therapy is easily monitored by conventional techniques and assays.

15        O.     Articles of Manufacture

In another embodiment of the invention, an article of manufacture containing materials (e.g., comprising a PRO molecule) useful for the diagnosis or treatment of the disorders described above is provided. The article of manufacture comprises a container and an instruction. Suitable containers include, for example, bottles, vials, syringes, and test tubes. The containers may be formed from a variety of 20 materials such as glass or plastic. The container holds a composition which is effective for diagnosing or treating the condition and may have a sterile access port (for example the container may be an intravenous solution bag or a vial having a stopper pierceable by a hypodermic injection needle). The active agent in the composition is usually a polypeptide or an antibody of the invention. An instruction or label on, or associated with, the container indicates that the composition is used for diagnosing or treating the condition 25 of choice. The article of manufacture may further comprise a second container comprising a pharmaceutically-acceptable buffer, such as phosphate-buffered saline, Ringer's solution and dextrose solution. It may further include other materials desirable from a commercial and user standpoint, including other buffers, diluents, filters, needles, syringes, and package inserts with instructions for use.

30        P.     Diagnosis and Prognosis of Immune Related Disease

Cell surface proteins, such as proteins which are overexpressed in certain immune related diseases, are excellent targets for drug candidates or disease treatment. The same proteins along with secreted 35 proteins encoded by the genes amplified in immune related disease states find additional use in the diagnosis and prognosis of these diseases. For example, antibodies directed against the protein products of genes amplified in multiple sclerosis, rheumatoid arthritis, or another immune related disease, can be used as diagnostics or prognostics.

For example, antibodies, including antibody fragments, can be used to qualitatively or quantitatively detect the expression of proteins encoded by amplified or overexpressed genes ("marker gene products"). The antibody preferably is equipped with a detectable, e.g., fluorescent label, and binding can be monitored by light microscopy, flow cytometry, fluorimetry, or other techniques known in the art. These 40 techniques are particularly suitable, if the overexpressed gene encodes a cell surface protein. Such binding

assays are performed essentially as described above.

*In situ* detection of antibody binding to the marker gene products can be performed, for example, by immunofluorescence or immunoelectron microscopy. For this purpose, a histological specimen is removed from the patient, and a labeled antibody is applied to it, preferably by overlaying the antibody on a biological sample. This procedure also allows for determining the distribution of the marker gene product in the tissue examined. It will be apparent for those skilled in the art that a wide variety of histological methods are readily available for *in situ* detection.

The following examples are offered for illustrative purposes only, and are not intended to limit the scope of the present invention in any way.

10 All patent and literature references cited in the present specification are hereby incorporated by reference in their entirety.

#### EXAMPLES

15 Commercially available reagents referred to in the examples were used according to manufacturer's instructions unless otherwise indicated. The source of those cells identified in the following examples, and throughout the specification, by ATCC accession numbers is the American Type Culture Collection, Manassas, VA.

##### EXAMPLE 1: Microarray analysis of stimulated T-cells

20 Nucleic acid microarrays, often containing thousands of gene sequences, are useful for identifying differentially expressed genes in diseased tissues as compared to their normal counterparts. Using nucleic acid microarrays, test and control mRNA samples from test and control tissue samples are reverse transcribed and labeled to generate cDNA probes. The cDNA probes are then hybridized to an array of nucleic acids immobilized on a solid support. The array is configured such that the sequence and position of each member of the array is known. For example, a selection of genes known to be expressed in certain 25 disease states may be arrayed on a solid support. Hybridization of a labeled probe with a particular array member indicates that the sample from which the probe was derived expresses that gene. If the hybridization signal of a probe from a test (in this instance, activated CD4+ T cells) sample is greater than hybridization signal of a probe from a control (in this instance, non-stimulated CD4 + T cells) sample, the gene or genes overexpressed in the test tissue are identified. The implication of this result is that an 30 overexpressed protein in a test tissue is useful not only as a diagnostic marker for the presence of the disease condition, but also as a therapeutic target for treatment of the disease condition.

35 The methodology of hybridization of nucleic acids and microarray technology is well known in the art. In one example, the specific preparation of nucleic acids for hybridization and probes, slides, and hybridization conditions are all detailed in PCT Patent Application Serial No. PCT/US01/10482, filed on March 30, 2001 and which is herein incorporated by reference.

40 In this experiment, CD4+ T cells were purified from a single donor using the RosetteSep™ protocol from (Stem Cell Technologies, Vancouver BC) which contains anti-CD8, anti-CD16, anti-CD19, anti-CD36 and anti-CD56 antibodies used to produce a population of isolated CD4 + T cells. Isolated CD4+ T cells were activated with an anti-CD3 antibody (used at a concentration that does not stimulate proliferation) together with either ICAM-1 or anti-CD28 antibody. At 24 or 72 hours cells were harvested,

RNA extracted and analysis run on Affimax (Affymetrix Inc. Santa Clara, CA) microarrays. Non-stimulated (resting) cells were harvested immediately after purification, and subjected to the same analysis. Genes were compared whose expression was upregulated at either of the two timepoints in activated vs. resting cells.

Below are the results of these experiments, demonstrating that various PRO polypeptides of the present invention are significantly overexpressed in isolated CD4 + T cells activated by anti-CD3/ICAM-1 or anti-CD3/anti-CD28 as compared to isolated resting CD4+ T cells. As described above, these data demonstrate that the PRO polypeptides of the present invention are useful not only as diagnostic markers for the presence of one or more immune disorders, but also serve as therapeutic targets for the treatment of those immune disorders.

The nucleic acids of Figure 1 (SEQ ID NO:1), Figure 3 (SEQ ID NO:3), Figure 5 (SEQ ID NO:5), Figure 7 (SEQ ID NO:7), Figure 9 (SEQ ID NO:9), Figure 11 (SEQ ID NO:11), Figure 13 (SEQ ID NO:13), Figure 15A-B (SEQ ID NO:15), Figure 17 (SEQ ID NO:17), Figure 19 (SEQ ID NO:19), Figure 21 (SEQ ID NO:21), Figure 23 (SEQ ID NO:23), Figure 25 (SEQ ID NO:25), Figure 27 (SEQ ID NO:27), Figure 29 (SEQ ID NO:29), Figure 31 (SEQ ID NO:31), Figure 33 (SEQ ID NO:33), Figure 35 (SEQ ID NO:35), Figure 37 (SEQ ID NO:37), Figure 39 (SEQ ID NO:39), Figure 41 (SEQ ID NO:41), Figure 43 (SEQ ID NO:43), Figure 45 (SEQ ID NO:45), Figure 47 (SEQ ID NO:47), Figure 49 (SEQ ID NO:49), Figure 51 (SEQ ID NO:51), Figure 53 (SEQ ID NO:53), Figure 55 (SEQ ID NO:55), Figure 57 (SEQ ID NO:57), Figure 59 (SEQ ID NO:59), Figure 61 (SEQ ID NO:61), Figure 63 (SEQ ID NO:63), Figure 65 (SEQ ID NO:65), Figure 67 (SEQ ID NO:67), Figure 69 (SEQ ID NO:69), Figure 71 (SEQ ID NO:71), Figure 73 (SEQ ID NO:73), Figure 75 (SEQ ID NO:75), Figure 77 (SEQ ID NO:77), Figure 79 (SEQ ID NO:79), Figure 81 (SEQ ID NO:81), Figure 83 (SEQ ID NO:83), Figure 85 (SEQ ID NO:85), Figure 87 (SEQ ID NO:87), Figure 89 (SEQ ID NO:89), Figure 91 (SEQ ID NO:91), Figure 93 (SEQ ID NO:93), Figure 95 (SEQ ID NO:95), Figure 97 (SEQ ID NO:97), Figure 99 (SEQ ID NO:99), Figure 101 (SEQ ID NO:101) and Figure 103 (SEQ ID NO:103) show increase in expression upon stimulation with anti-CD3/ICAM1 and also show increase in expression upon stimulation with anti-CD3/anti-CD28.

#### EXAMPLE 2: Use of PRO as a hybridization probe

The following method describes use of a nucleotide sequence encoding PRO as a hybridization probe.

DNA comprising the coding sequence of full-length or mature PRO as disclosed herein is employed as a probe to screen for homologous DNAs (such as those encoding naturally-occurring variants of PRO) in human tissue cDNA libraries or human tissue genomic libraries.

Hybridization and washing of filters containing either library DNAs is performed under the following high stringency conditions. Hybridization of radiolabeled PRO-derived probe to the filters is performed in a solution of 50% formamide, 5x SSC, 0.1% SDS, 0.1% sodium pyrophosphate, 50 mM sodium phosphate, pH 6.8, 2x Denhardt's solution, and 10% dextran sulfate at 42°C for 20 hours. Washing of the filters is performed in an aqueous solution of 0.1x SSC and 0.1% SDS at 42°C.

DNAs having a desired sequence identity with the DNA encoding full-length native sequence PRO can then be identified using standard techniques known in the art.

EXAMPLE 3: Expression of PRO in *E. coli*

This example illustrates preparation of an unglycosylated form of PRO by recombinant expression in *E. coli*.

The DNA sequence encoding PRO is initially amplified using selected PCR primers. The primers 5 should contain restriction enzyme sites which correspond to the restriction enzyme sites on the selected expression vector. A variety of expression vectors may be employed. An example of a suitable vector is pBR322 (derived from *E. coli*; see Bolivar et al., *Gene*, 2:95 (1977)) which contains genes for ampicillin and tetracycline resistance. The vector is digested with restriction enzyme and dephosphorylated. The PCR amplified sequences are then ligated into the vector. The vector will preferably include sequences which 10 encode for an antibiotic resistance gene, a trp promoter, a polyhis leader (including the first six STII codons, polyhis sequence, and enterokinase cleavage site), the PRO coding region, lambda transcriptional terminator, and an argU gene.

The ligation mixture is then used to transform a selected *E. coli* strain using the methods described 15 in Sambrook et al., *supra*. Transformants are identified by their ability to grow on LB plates and antibiotic resistant colonies are then selected. Plasmid DNA can be isolated and confirmed by restriction analysis and DNA sequencing.

Selected clones can be grown overnight in liquid culture medium such as LB broth supplemented with antibiotics. The overnight culture may subsequently be used to inoculate a larger scale culture. The 20 cells are then grown to a desired optical density, during which the expression promoter is turned on.

After culturing the cells for several more hours, the cells can be harvested by centrifugation. The cell pellet obtained by the centrifugation can be solubilized using various agents known in the art, and the solubilized PRO protein can then be purified using a metal chelating column under conditions that allow tight binding of the protein.

PRO may be expressed in *E. coli* in a poly-His tagged form, using the following procedure. The 25 DNA encoding PRO is initially amplified using selected PCR primers. The primers will contain restriction enzyme sites which correspond to the restriction enzyme sites on the selected expression vector, and other useful sequences providing for efficient and reliable translation initiation, rapid purification on a metal chelation column, and proteolytic removal with enterokinase. The PCR-amplified, poly-His tagged sequences are then ligated into an expression vector, which is used to transform an *E. coli* host based on 30 strain 52 (W3110 fuhA(tonA) lon galE rpoHts(htpRts) clpP(lacIq)). Transformants are first grown in LB containing 50 mg/ml carbenicillin at 30°C with shaking until an O.D.600 of 3-5 is reached. Cultures are then diluted 50-100 fold into CRAP media (prepared by mixing 3.57 g (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 0.71 g sodium citrate•2H<sub>2</sub>O, 1.07 g KCl, 5.36 g Difco yeast extract, 5.36 g Sheffield hycase SF in 500 mL water, as well as 35 110 mM MPOS, pH 7.3, 0.55% (w/v) glucose and 7 mM MgSO<sub>4</sub>) and grown for approximately 20-30 hours at 30°C with shaking. Samples are removed to verify expression by SDS-PAGE analysis, and the bulk culture is centrifuged to pellet the cells. Cell pellets are frozen until purification and refolding.

*E. coli* paste from 0.5 to 1 L fermentations (6-10 g pellets) is resuspended in 10 volumes (w/v) in 7 M guanidine, 20 mM Tris, pH 8 buffer. Solid sodium sulfite and sodium tetrathionate is added to make final concentrations of 0.1M and 0.02 M, respectively, and the solution is stirred overnight at 4°C. This step 40 results in a denatured protein with all cysteine residues blocked by sulfitolization. The solution is centrifuged

at 40,000 rpm in a Beckman Ultracentrifuge for 30 min. The supernatant is diluted with 3-5 volumes of metal chelate column buffer (6 M guanidine, 20 mM Tris, pH 7.4) and filtered through 0.22 micron filters to clarify. The clarified extract is loaded onto a 5 ml Qiagen Ni-NTA metal chelate column equilibrated in the metal chelate column buffer. The column is washed with additional buffer containing 50 mM imidazole

5 (Calbiochem, Utrol grade), pH 7.4. The protein is eluted with buffer containing 250 mM imidazole. Fractions containing the desired protein are pooled and stored at 4°C. Protein concentration is estimated by its absorbance at 280 nm using the calculated extinction coefficient based on its amino acid sequence.

The proteins are refolded by diluting the sample slowly into freshly prepared refolding buffer consisting of: 20 mM Tris, pH 8.6, 0.3 M NaCl, 2.5 M urea, 5 mM cysteine, 20 mM glycine and 1 mM 10 EDTA. Refolding volumes are chosen so that the final protein concentration is between 50 to 100 micrograms/ml. The refolding solution is stirred gently at 4°C for 12-36 hours. The refolding reaction is quenched by the addition of TFA to a final concentration of 0.4% (pH of approximately 3). Before further 15 purification of the protein, the solution is filtered through a 0.22 micron filter and acetonitrile is added to 2-10% final concentration. The refolded protein is chromatographed on a Poros R1/H reversed phase column using a mobile buffer of 0.1% TFA with elution with a gradient of acetonitrile from 10 to 80%. Aliquots of fractions with A280 absorbance are analyzed on SDS polyacrylamide gels and fractions containing homogeneous refolded protein are pooled. Generally, the properly refolded species of most 20 proteins are eluted at the lowest concentrations of acetonitrile since those species are the most compact with their hydrophobic interiors shielded from interaction with the reversed phase resin. Aggregated species are usually eluted at higher acetonitrile concentrations. In addition to resolving misfolded forms of proteins from the desired form, the reversed phase step also removes endotoxin from the samples.

Fractions containing the desired folded PRO polypeptide are pooled and the acetonitrile removed using a gentle stream of nitrogen directed at the solution. Proteins are formulated into 20 mM Hepes, pH 6.8 with 0.14 M sodium chloride and 4% mannitol by dialysis or by gel filtration using G25 Superfine 25 (Pharmacia) resins equilibrated in the formulation buffer and sterile filtered.

Many of the PRO polypeptides disclosed herein were successfully expressed as described above.

#### EXAMPLE 4: Expression of PRO in mammalian cells

This example illustrates preparation of a potentially glycosylated form of PRO by recombinant 30 expression in mammalian cells.

The vector, pRK5 (see EP 307,247, published March 15, 1989), is employed as the expression vector. Optionally, the PRO DNA is ligated into pRK5 with selected restriction enzymes to allow insertion of the PRO DNA using ligation methods such as described in Sambrook et al., *supra*. The resulting vector is called pRK5-PRO.

35 In one embodiment, the selected host cells may be 293 cells. Human 293 cells (ATCC CCL 1573) are grown to confluence in tissue culture plates in medium such as DMEM supplemented with fetal calf serum and optionally, nutrient components and/or antibiotics. About 10 µg pRK5-PRO DNA is mixed with about 1 µg DNA encoding the VA RNA gene [Thimmappaya et al., *Cell*, 31:543 (1982)] and dissolved in 500 µl of 1 mM Tris-HCl, 0.1 mM EDTA, 0.227 M CaCl<sub>2</sub>. To this mixture is added, dropwise, 500 µl of 50 mM HEPES (pH 7.35), 280 mM NaCl, 1.5 mM NaPO<sub>4</sub>, and a precipitate is allowed to form for 10 minutes 40

at 25°C. The precipitate is suspended and added to the 293 cells and allowed to settle for about four hours at 37°C. The culture medium is aspirated off and 2 ml of 20% glycerol in PBS is added for 30 seconds. The 293 cells are then washed with serum free medium, fresh medium is added and the cells are incubated for about 5 days.

5        Approximately 24 hours after the transfections, the culture medium is removed and replaced with culture medium (alone) or culture medium containing 200  $\mu$ Ci/ml  $^{35}$ S-cysteine and 200  $\mu$ Ci/ml  $^{35}$ S-methionine. After a 12 hour incubation, the conditioned medium is collected, concentrated on a spin filter, and loaded onto a 15% SDS gel. The processed gel may be dried and exposed to film for a selected period of time to reveal the presence of PRO polypeptide. The cultures containing transfected cells may undergo 10 further incubation (in serum free medium) and the medium is tested in selected bioassays.

In an alternative technique, PRO may be introduced into 293 cells transiently using the dextran sulfate method described by Somparyrac et al., *Proc. Natl. Acad. Sci.*, 78:7575 (1981). 293 cells are grown to maximal density in a spinner flask and 700  $\mu$ g pRK5-PRO DNA is added. The cells are first concentrated from the spinner flask by centrifugation and washed with PBS. The DNA-dextran precipitate is incubated 15 on the cell pellet for four hours. The cells are treated with 20% glycerol for 90 seconds, washed with tissue culture medium, and re-introduced into the spinner flask containing tissue culture medium, 5  $\mu$ g/ml bovine insulin and 0.1  $\mu$ g/ml bovine transferrin. After about four days, the conditioned media is centrifuged and filtered to remove cells and debris. The sample containing expressed PRO can then be concentrated and purified by any selected method, such as dialysis and/or column chromatography.

20        In another embodiment, PRO can be expressed in CHO cells. The pRK5-PRO can be transfected into CHO cells using known reagents such as CaPO<sub>4</sub> or DEAE-dextran. As described above, the cell cultures can be incubated, and the medium replaced with culture medium (alone) or medium containing a radiolabel such as  $^{35}$ S-methionine. After determining the presence of PRO polypeptide, the culture medium may be replaced with serum free medium. Preferably, the cultures are incubated for about 6 days, and then 25 the conditioned medium is harvested. The medium containing the expressed PRO can then be concentrated and purified by any selected method.

Epitope-tagged PRO may also be expressed in host CHO cells. The PRO may be subcloned out of the pRK5 vector. The subclone insert can undergo PCR to fuse in frame with a selected epitope tag such as a poly-his tag into a Baculovirus expression vector. The poly-his tagged PRO insert can then be subcloned 30 into a SV40 promoter/enhancer containing vector containing a selection marker such as DHFR for selection of stable clones. Finally, the CHO cells can be transfected (as described above) with the SV40 promoter/enhancer containing vector. Labeling may be performed, as described above, to verify expression. The culture medium containing the expressed poly-His tagged PRO can then be concentrated and purified by any selected method, such as by Ni<sup>2+</sup>-chelate affinity chromatography.

35        PRO may also be expressed in CHO and/or COS cells by a transient expression procedure or in CHO cells by another stable expression procedure.

Stable expression in CHO cells is performed using the following procedure. The proteins are expressed as an IgG construct (immunoadhesin), in which the coding sequences for the soluble forms (e.g.

extracellular domains) of the respective proteins are fused to an IgG1 constant region sequence containing the hinge, CH2 and CH2 domains and/or is a poly-His tagged form.

Following PCR amplification, the respective DNAs are subcloned in a CHO expression vector using standard techniques as described in Ausubel et al., *Current Protocols of Molecular Biology*, Unit 3.16, 5 John Wiley and Sons (1997). CHO expression vectors are constructed to have compatible restriction sites 5' and 3' of the DNA of interest to allow the convenient shuttling of cDNA's. The vector used expression in CHO cells is as described in Lucas et al., *Nucl. Acids Res.* 24:9 (1774-1779 (1996), and uses the SV40 early promoter/enhancer to drive expression of the cDNA of interest and dihydrofolate reductase (DHFR). DHFR expression permits selection for stable maintenance of the plasmid following transfection.

10 Twelve micrograms of the desired plasmid DNA is introduced into approximately 10 million CHO cells using commercially available transfection reagents Superfect® (Qiagen), Dasurer® or Fugene® (Boehringer Mannheim). The cells are grown as described in Lucas et al., *supra*. Approximately  $3 \times 10^7$  cells are frozen in an ampule for further growth and production as described below.

15 The ampules containing the plasmid DNA are thawed by placement into water bath and mixed by vortexing. The contents are pipetted into a centrifuge tube containing 10 mL of media and centrifuged at 1000 rpm for 5 minutes. The supernatant is aspirated and the cells are resuspended in 10 mL of selective media (0.2 µm filtered PS20 with 5% 0.2 µm diafiltered fetal bovine serum). The cells are then aliquoted into a 100 mL spinner containing 90 mL of selective media. After 1-2 days, the cells are transferred into a 20 250 mL spinner filled with 150 mL selective growth medium and incubated at 37°C. After another 2-3 days, 250 mL, 500 mL and 2000 mL spinners are seeded with  $3 \times 10^5$  cells/mL. The cell media is exchanged with 20 fresh media by centrifugation and resuspension in production medium. Although any suitable CHO media may be employed, a production medium described in U.S. Patent No. 5,122,469, issued June 16, 1992 may actually be used. A 3L production spinner is seeded at  $1.2 \times 10^6$  cells/mL. On day 0, pH is determined. On day 1, the spinner is sampled and sparging with filtered air is commenced. On day 2, the spinner is sampled, 25 the temperature shifted to 33°C, and 30 mL of 500 g/L glucose and 0.6 mL of 10% antifoam (e.g., 35% polydimethylsiloxane emulsion, Dow Corning 365 Medical Grade Emulsion) taken. Throughout the production, the pH is adjusted as necessary to keep it at around 7.2. After 10 days, or until the viability dropped below 70%, the cell culture is harvested by centrifugation and filtering through a 0.22 µm filter. The filtrate was either stored at 4°C or immediately loaded onto columns for purification.

30 For the poly-His tagged constructs, the proteins are purified using a Ni-NTA column (Qiagen). Before purification, imidazole is added to the conditioned media to a concentration of 5 mM. The conditioned media is pumped onto a 6 ml Ni-NTA column equilibrated in 20 mM Hepes, pH 7.4, buffer containing 0.3 M NaCl and 5 mM imidazole at a flow rate of 4-5 ml/min. at 4°C. After loading, the column is washed with additional equilibration buffer and the protein eluted with equilibration buffer containing 35 0.25 M imidazole. The highly purified protein is subsequently desalting into a storage buffer containing 10 mM Hepes, 0.14 M NaCl and 4% mannitol, pH 6.8, with a 25 ml G25 Superfine (Pharmacia) column and stored at -80°C.

40 Immunoadhesin (Fc-containing) constructs are purified from the conditioned media as follows. The conditioned medium is pumped onto a 5 ml Protein A column (Pharmacia) which had been equilibrated in 20 mM Na phosphate buffer, pH 6.8. After loading, the column is washed extensively with equilibration

buffer before elution with 100 mM citric acid, pH 3.5. The eluted protein is immediately neutralized by collecting 1 ml fractions into tubes containing 275  $\mu$ l of 1 M Tris buffer, pH 9. The highly purified protein is subsequently desalted into storage buffer as described above for the poly-His tagged proteins. The homogeneity is assessed by SDS polyacrylamide gels and by N-terminal amino acid sequencing by Edman degradation.

5 Many of the PRO polypeptides disclosed herein were successfully expressed as described above.

**EXAMPLE 5: Expression of PRO in Yeast**

The following method describes recombinant expression of PRO in yeast.

10 First, yeast expression vectors are constructed for intracellular production or secretion of PRO from the ADH2/GAPDH promoter. DNA encoding PRO and the promoter is inserted into suitable restriction enzyme sites in the selected plasmid to direct intracellular expression of PRO. For secretion, DNA encoding PRO can be cloned into the selected plasmid, together with DNA encoding the ADH2/GAPDH promoter, a native PRO signal peptide or other mammalian signal peptide, or, for example, a yeast alpha-factor or 15 invertase secretory signal/leader sequence, and linker sequences (if needed) for expression of PRO.

Yeast cells, such as yeast strain AB110, can then be transformed with the expression plasmids described above and cultured in selected fermentation media. The transformed yeast supernatants can be analyzed by precipitation with 10% trichloroacetic acid and separation by SDS-PAGE, followed by staining of the gels with Coomassie Blue stain.

20 Recombinant PRO can subsequently be isolated and purified by removing the yeast cells from the fermentation medium by centrifugation and then concentrating the medium using selected cartridge filters. The concentrate containing PRO may further be purified using selected column chromatography resins.

Many of the PRO polypeptides disclosed herein were successfully expressed as described above.

25 **EXAMPLE 6: Expression of PRO in Baculovirus-Infected Insect Cells**

The following method describes recombinant expression of PRO in Baculovirus-infected insect cells.

30 The sequence coding for PRO is fused upstream of an epitope tag contained within a baculovirus expression vector. Such epitope tags include poly-his tags and immunoglobulin tags (like Fc regions of IgG). A variety of plasmids may be employed, including plasmids derived from commercially available plasmids such as pVL1393 (Novagen). Briefly, the sequence encoding PRO or the desired portion of the coding sequence of PRO such as the sequence encoding the extracellular domain of a transmembrane protein or the sequence encoding the mature protein if the protein is extracellular is amplified by PCR with primers complementary to the 5' and 3' regions. The 5' primer may incorporate flanking (selected) restriction 35 enzyme sites. The product is then digested with those selected restriction enzymes and subcloned into the expression vector.

40 Recombinant baculovirus is generated by co-transfected the above plasmid and BaculoGold<sup>TM</sup> virus DNA (Pharmingen) into *Spodoptera frugiperda* ("Sf9") cells (ATCC CRL 1711) using lipofectin (commercially available from GIBCO-BRL). After 4 - 5 days of incubation at 28°C, the released viruses are harvested and used for further amplifications. Viral infection and protein expression are performed as

described by O'Reilley et al., Baculovirus expression vectors: A Laboratory Manual, Oxford: Oxford University Press (1994).

Expressed poly-his tagged PRO can then be purified, for example, by  $\text{Ni}^{2+}$ -chelate affinity chromatography as follows. Extracts are prepared from recombinant virus-infected Sf9 cells as described by 5 Rupert et al., Nature, 362:175-179 (1993). Briefly, Sf9 cells are washed, resuspended in sonication buffer (25 mL Hepes, pH 7.9; 12.5 mM  $\text{MgCl}_2$ ; 0.1 mM EDTA; 10% glycerol; 0.1% NP-40; 0.4 M KCl), and sonicated twice for 20 seconds on ice. The sonicates are cleared by centrifugation, and the supernatant is diluted 50-fold in loading buffer (50 mM phosphate, 300 mM NaCl, 10% glycerol, pH 7.8) and filtered through a 0.45  $\mu\text{m}$  filter. A  $\text{Ni}^{2+}$ -NTA agarose column (commercially available from Qiagen) is prepared 10 with a bed volume of 5 mL, washed with 25 mL of water and equilibrated with 25 mL of loading buffer. The filtered cell extract is loaded onto the column at 0.5 mL per minute. The column is washed to baseline  $A_{280}$  with loading buffer, at which point fraction collection is started. Next, the column is washed with a secondary wash buffer (50 mM phosphate; 300 mM NaCl, 10% glycerol, pH 6.0), which elutes 15 nonspecifically bound protein. After reaching  $A_{280}$  baseline again, the column is developed with a 0 to 500 mM Imidazole gradient in the secondary wash buffer. One mL fractions are collected and analyzed by SDS-PAGE and silver staining or Western blot with  $\text{Ni}^{2+}$ -NTA-conjugated to alkaline phosphatase (Qiagen). Fractions containing the eluted His<sub>10</sub>-tagged PRO are pooled and dialyzed against loading buffer.

Alternatively, purification of the IgG tagged (or Fc tagged) PRO can be performed using known chromatography techniques, including for instance, Protein A or protein G column chromatography.

20 Many of the PRO polypeptides disclosed herein were successfully expressed as described above.

#### EXAMPLE 7: Preparation of Antibodies that Bind PRO

This example illustrates preparation of monoclonal antibodies which can specifically bind PRO.

Techniques for producing the monoclonal antibodies are known in the art and are described, for 25 instance, in Goding, supra. Immunogens that may be employed include purified PRO, fusion proteins containing PRO, and cells expressing recombinant PRO on the cell surface. Selection of the immunogen can be made by the skilled artisan without undue experimentation.

Mice, such as Balb/c, are immunized with the PRO immunogen emulsified in complete Freund's adjuvant and injected subcutaneously or intraperitoneally in an amount from 1-100 micrograms.

30 Alternatively, the immunogen is emulsified in MPL-TDM adjuvant (Ribi Immunochemical Research, Hamilton, MT) and injected into the animal's hind foot pads. The immunized mice are then boosted 10 to 12 days later with additional immunogen emulsified in the selected adjuvant. Thereafter, for several weeks, the mice may also be boosted with additional immunization injections. Serum samples may be periodically obtained from the mice by retro-orbital bleeding for testing in ELISA assays to detect anti-PRO antibodies.

35 After a suitable antibody titer has been detected, the animals "positive" for antibodies can be injected with a final intravenous injection of PRO. Three to four days later, the mice are sacrificed and the spleen cells are harvested. The spleen cells are then fused (using 35% polyethylene glycol) to a selected murine myeloma cell line such as P3X63AgU.1, available from ATCC, No. CRL 1597. The fusions generate hybridoma cells which can then be plated in 96 well tissue culture plates containing HAT

(hypoxanthine, aminopterin, and thymidine) medium to inhibit proliferation of non-fused cells, myeloma hybrids, and spleen cell hybrids.

The hybridoma cells will be screened in an ELISA for reactivity against PRO. Determination of 1 "positive" hybridoma cells secreting the desired monoclonal antibodies against PRO is within the skill in the 5 art.

The positive hybridoma cells can be injected intraperitoneally into syngeneic Balb/c mice to produce ascites containing the anti-PRO monoclonal antibodies. Alternatively, the hybridoma cells can be grown in tissue culture flasks or roller bottles. Purification of the monoclonal antibodies produced in the ascites can be accomplished using ammonium sulfate precipitation, followed by gel exclusion 10 chromatography. Alternatively, affinity chromatography based upon binding of antibody to protein A or protein G can be employed.

#### EXAMPLE 8: Purification of PRO Polypeptides Using Specific Antibodies

Native or recombinant PRO polypeptides may be purified by a variety of standard techniques in the 15 art of protein purification. For example, pro-PRO polypeptide, mature PRO polypeptide, or pre-PRO polypeptide is purified by immunoaffinity chromatography using antibodies specific for the PRO polypeptide of interest. In general, an immunoaffinity column is constructed by covalently coupling the anti-PRO polypeptide antibody to an activated chromatographic resin.

Polyclonal immunoglobulins are prepared from immune sera either by precipitation with 20 ammonium sulfate or by purification on immobilized Protein A (Pharmacia LKB Biotechnology, Piscataway, N.J.). Likewise, monoclonal antibodies are prepared from mouse ascites fluid by ammonium sulfate precipitation or chromatography on immobilized Protein A. Partially purified immunoglobulin is covalently attached to a chromatographic resin such as CnBr-activated SEPHAROSE<sup>TM</sup> (Pharmacia LKB Biotechnology). The antibody is coupled to the resin, the resin is blocked, and the derivative resin is washed 25 according to the manufacturer's instructions.

Such an immunoaffinity column is utilized in the purification of PRO polypeptide by preparing a fraction from cells containing PRO polypeptide in a soluble form. This preparation is derived by solubilization of the whole cell or of a subcellular fraction obtained via differential centrifugation by the addition of detergent or by other methods well known in the art. Alternatively, soluble PRO polypeptide 30 containing a signal sequence may be secreted in useful quantity into the medium in which the cells are grown.

A soluble PRO polypeptide-containing preparation is passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of PRO polypeptide (e.g., high 35 ionic strength buffers in the presence of detergent). Then, the column is eluted under conditions that disrupt antibody/PRO polypeptide binding (e.g., a low pH buffer such as approximately pH 2-3, or a high concentration of a chaotrope such as urea or thiocyanate ion), and PRO polypeptide is collected.

#### EXAMPLE 9: Drug Screening

This invention is particularly useful for screening compounds by using PRO polypeptides or 40 binding fragment thereof in any of a variety of drug screening techniques. The PRO polypeptide or

fragment employed in such a test may either be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. One method of drug screening utilizes eukaryotic or prokaryotic host cells which are stably transformed with recombinant nucleic acids expressing the PRO polypeptide or fragment. Drugs are screened against such transformed cells in competitive binding assays. Such cells, either in viable 5 or fixed form, can be used for standard binding assays. One may measure, for example, the formation of complexes between PRO polypeptide or a fragment and the agent being tested. Alternatively, one can examine the diminution in complex formation between the PRO polypeptide and its target cell or target receptors caused by the agent being tested.

Thus, the present invention provides methods of screening for drugs or any other agents which can 10 affect a PRO polypeptide-associated disease or disorder. These methods comprise contacting such an agent with an PRO polypeptide or fragment thereof and assaying (I) for the presence of a complex between the agent and the PRO polypeptide or fragment, or (ii) for the presence of a complex between the PRO polypeptide or fragment and the cell, by methods well known in the art. In such competitive binding assays, the PRO polypeptide or fragment is typically labeled. After suitable incubation, free PRO polypeptide or 15 fragment is separated from that present in bound form, and the amount of free or uncomplexed label is a measure of the ability of the particular agent to bind to PRO polypeptide or to interfere with the PRO polypeptide/cell complex.

Another technique for drug screening provides high throughput screening for compounds having 20 suitable binding affinity to a polypeptide and is described in detail in WO 84/03564, published on September 13, 1984. Briefly stated, large numbers of different small peptide test compounds are synthesized on a solid substrate, such as plastic pins or some other surface. As applied to a PRO polypeptide, the peptide test compounds are reacted with PRO polypeptide and washed. Bound PRO polypeptide is detected by methods well known in the art. Purified PRO polypeptide can also be coated directly onto plates for use in the aforementioned drug screening techniques. In addition, non-neutralizing antibodies can be used to capture 25 the peptide and immobilize it on the solid support.

This invention also contemplates the use of competitive drug screening assays in which neutralizing 30 antibodies capable of binding PRO polypeptide specifically compete with a test compound for binding to PRO polypeptide or fragments thereof. In this manner, the antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with PRO polypeptide.

#### EXAMPLE 10: Rational Drug Design

The goal of rational drug design is to produce structural analogs of biologically active polypeptide 35 of interest (*i.e.*, a PRO polypeptide) or of small molecules with which they interact, *e.g.*, agonists, antagonists, or inhibitors. Any of these examples can be used to fashion drugs which are more active or stable forms of the PRO polypeptide or which enhance or interfere with the function of the PRO polypeptide *in vivo* (*c.f.*, Hodgson, *Bio/Technology*, 9: 19-21 (1991)).

In one approach, the three-dimensional structure of the PRO polypeptide, or of a PRO 40 polypeptide-inhibitor complex, is determined by x-ray crystallography, by computer modeling or, most typically, by a combination of the two approaches. Both the shape and charges of the PRO polypeptide must be ascertained to elucidate the structure and to determine active site(s) of the molecule. Less often, useful

information regarding the structure of the PRO polypeptide may be gained by modeling based on the structure of homologous proteins. In both cases, relevant structural information is used to design analogous PRO polypeptide-like molecules or to identify efficient inhibitors. Useful examples of rational drug design may include molecules which have improved activity or stability as shown by Braxton and Wells,

5 Biochemistry, 31:7796-7801 (1992) or which act as inhibitors, agonists, or antagonists of native peptides as shown by Athauda *et al.*, J. Biochem., 113:742-746 (1993).

It is also possible to isolate a target-specific antibody, selected by functional assay, as described above, and then to solve its crystal structure. This approach, in principle, yields a pharmacore upon which subsequent drug design can be based. It is possible to bypass protein crystallography altogether by 10 generating anti-idiotypic antibodies (anti-ids) to a functional, pharmacologically active antibody. As a mirror image of a mirror image, the binding site of the anti-ids would be expected to be an analog of the original receptor. The anti-id could then be used to identify and isolate peptides from banks of chemically or biologically produced peptides. The isolated peptides would then act as the pharmacore.

By virtue of the present invention, sufficient amounts of the PRO polypeptide may be made 15 available to perform such analytical studies as X-ray crystallography. In addition, knowledge of the PRO polypeptide amino acid sequence provided herein will provide guidance to those employing computer modeling techniques in place of or in addition to x-ray crystallography.

The foregoing written specification is considered to be sufficient to enable one skilled in the art to practice the invention. The present invention is not to be limited in scope by the construct deposited, since 20 the deposited embodiment is intended as a single illustration of certain aspects of the invention and any constructs that are functionally equivalent are within the scope of this invention. The deposit of material herein does not constitute an admission that the written description herein contained is inadequate to enable the practice of any aspect of the invention, including the best mode thereof, nor is it to be construed as limiting the scope of the claims to the specific illustrations that it represents. Indeed, various modifications 25 of the invention in addition to those shown and described herein will become apparent to those skilled in the art from the foregoing description and fall within the scope of the appended claims.

What is claimed:

1. Isolated nucleic acid having at least 80% nucleic acid sequence identity to:

5 (a) a nucleotide sequence encoding the polypeptide shown in Figure 2 (SEQ ID NO:2), Figure 4 (SEQ ID NO:4), Figure 6 (SEQ ID NO:6), Figure 8 (SEQ ID NO:8), Figure 10 (SEQ ID NO:10), Figure 12 (SEQ ID NO:12), Figure 14 (SEQ ID NO:14), Figure 16 (SEQ ID NO:16), Figure 18 (SEQ ID NO:18), Figure 20 (SEQ ID NO:20), Figure 22 (SEQ ID NO:22), Figure 24 (SEQ ID NO:24), Figure 26 (SEQ ID NO:26), Figure 28 (SEQ ID NO:28), Figure 30 (SEQ ID NO:30), Figure 32 (SEQ ID NO:32), Figure 34 (SEQ ID NO:34), Figure 36 (SEQ ID NO:36), Figure 38 (SEQ ID NO:38), Figure 40 (SEQ ID NO:40), Figure 42 (SEQ ID NO:42), Figure 44 (SEQ ID NO:44), Figure 46 (SEQ ID NO:46), Figure 48 (SEQ ID NO:48), Figure 50 (SEQ ID NO:50), Figure 52 (SEQ ID NO:52), Figure 54 (SEQ ID NO:54), Figure 56 (SEQ ID NO:56), Figure 58 (SEQ ID NO:58), Figure 60 (SEQ ID NO:60), Figure 62 (SEQ ID NO:62), Figure 64 (SEQ ID NO:64), Figure 66 (SEQ ID NO:66), Figure 68 (SEQ ID NO:68), Figure 70 (SEQ ID NO:70), Figure 72 (SEQ ID NO:72), Figure 74 (SEQ ID NO:74), Figure 76 (SEQ ID NO:76), Figure 78 (SEQ ID NO:78), Figure 80 (SEQ ID NO:80), Figure 82 (SEQ ID NO:82), Figure 84 (SEQ ID NO:84), Figure 86 (SEQ ID NO:86), Figure 88 (SEQ ID NO:88), Figure 90 (SEQ ID NO:90), Figure 92 (SEQ ID NO:92), Figure 94 (SEQ ID NO:94), Figure 96 (SEQ ID NO:96), Figure 98 (SEQ ID NO:98), Figure 100 (SEQ ID NO:100), Figure 102 (SEQ ID NO:102) or Figure 104 (SEQ ID NO:104).

20

2. Isolated nucleic acid having at least 80% nucleic acid sequence identity to a nucleotide sequence selected from the group consisting of the nucleotide sequence shown in Figure 1 (SEQ ID NO:1), Figure 3 (SEQ ID NO:3), Figure 5 (SEQ ID NO:5), Figure 7 (SEQ ID NO:7), Figure 9 (SEQ ID NO:9), Figure 11 (SEQ ID NO:11), Figure 13 (SEQ ID NO:13), Figure 15A-B (SEQ ID NO:15), Figure 17 (SEQ ID NO:17), Figure 19 (SEQ ID NO:19), Figure 21 (SEQ ID NO:21), Figure 23 (SEQ ID NO:23), Figure 25 (SEQ ID NO:25), Figure 27 (SEQ ID NO:27), Figure 29 (SEQ ID NO:29), Figure 31 (SEQ ID NO:31), Figure 33 (SEQ ID NO:33), Figure 35 (SEQ ID NO:35), Figure 37 (SEQ ID NO:37), Figure 39 (SEQ ID NO:39), Figure 41 (SEQ ID NO:41), Figure 43 (SEQ ID NO:43), Figure 45 (SEQ ID NO:45), Figure 47 (SEQ ID NO:47), Figure 49 (SEQ ID NO:49), Figure 51 (SEQ ID NO:51), Figure 53 (SEQ ID NO:53), Figure 55 (SEQ ID NO:55), Figure 57 (SEQ ID NO:57), Figure 59 (SEQ ID NO:59), Figure 61 (SEQ ID NO:61), Figure 63 (SEQ ID NO:63), Figure 65 (SEQ ID NO:65), Figure 67 (SEQ ID NO:67), Figure 69 (SEQ ID NO:69), Figure 71 (SEQ ID NO:71), Figure 73 (SEQ ID NO:73), Figure 75 (SEQ ID NO:75), Figure 77 (SEQ ID NO:77), Figure 79 (SEQ ID NO:79), Figure 81 (SEQ ID NO:81), Figure 83 (SEQ ID NO:83), Figure 85 (SEQ ID NO:85), Figure 87 (SEQ ID NO:87), Figure 89 (SEQ ID NO:89), Figure 91 (SEQ ID NO:91), Figure 93 (SEQ ID NO:93), Figure 95 (SEQ ID NO:95), Figure 97 (SEQ ID NO:97), Figure 99 (SEQ ID NO:99), Figure 101 (SEQ ID NO:101) and Figure 103 (SEQ ID NO:103).

40 3. Isolated nucleic acid having at least 80% nucleic acid sequence identity to a nucleotide sequence selected from the group consisting of the full-length coding sequence of the nucleotide sequence shown in Figure 1 (SEQ ID NO:1), Figure 3 (SEQ ID NO:3), Figure 5 (SEQ ID NO:5), Figure 7 (SEQ ID

NO:7), Figure 9 (SEQ ID NO:9), Figure 11 (SEQ ID NO:11), Figure 13 (SEQ ID NO:13), Figure 15A-B (SEQ ID NO:15), Figure 17 (SEQ ID NO:17), Figure 19 (SEQ ID NO:19), Figure 21 (SEQ ID NO:21), Figure 23 (SEQ ID NO:23), Figure 25 (SEQ ID NO:25), Figure 27 (SEQ ID NO:27), Figure 29 (SEQ ID NO:29), Figure 31 (SEQ ID NO:31), Figure 33 (SEQ ID NO:33), Figure 35 (SEQ ID NO:35), Figure 37 (SEQ ID NO:37), Figure 39 (SEQ ID NO:39), Figure 41 (SEQ ID NO:41), Figure 43 (SEQ ID NO:43), Figure 45 (SEQ ID NO:45), Figure 47 (SEQ ID NO:47), Figure 49 (SEQ ID NO:49), Figure 51 (SEQ ID NO:51), Figure 53 (SEQ ID NO:53), Figure 55 (SEQ ID NO:55), Figure 57 (SEQ ID NO:57), Figure 59 (SEQ ID NO:59), Figure 61 (SEQ ID NO:61), Figure 63 (SEQ ID NO:63), Figure 65 (SEQ ID NO:65), Figure 67 (SEQ ID NO:67), Figure 69 (SEQ ID NO:69), Figure 71 (SEQ ID NO:71), Figure 73 (SEQ ID NO:73), Figure 75 (SEQ ID NO:75), Figure 77 (SEQ ID NO:77), Figure 79 (SEQ ID NO:79), Figure 81 (SEQ ID NO:81), Figure 83 (SEQ ID NO:83), Figure 85 (SEQ ID NO:85), Figure 87 (SEQ ID NO:87), Figure 89 (SEQ ID NO:89), Figure 91 (SEQ ID NO:91), Figure 93 (SEQ ID NO:93), Figure 95 (SEQ ID NO:95), Figure 97 (SEQ ID NO:97), Figure 99 (SEQ ID NO:99), Figure 101 (SEQ ID NO:101), and Figure 103 (SEQ ID NO:103).

15

5. A vector comprising the nucleic acid of Claim 1.

6. The vector of Claim 5 operably linked to control sequences recognized by a host cell transformed with the vector.

20

7. A host cell comprising the vector of Claim 5.

8. The host cell of Claim 7, wherein said cell is a CHO cell, an *E.coli* cell or a yeast cell.

25

9. A process for producing a PRO polypeptide comprising culturing the host cell of Claim 7 under conditions suitable for expression of said PRO polypeptide and recovering said PRO polypeptide from the cell culture.

10. An isolated polypeptide having at least 80% amino acid sequence identity to:

30

(a) an amino acid sequence of the polypeptide shown in Figure 2 (SEQ ID NO:2), Figure 4 (SEQ ID NO:4), Figure 6 (SEQ ID NO:6), Figure 8 (SEQ ID NO:8), Figure 10 (SEQ ID NO:10), Figure 12 (SEQ ID NO:12), Figure 14 (SEQ ID NO:14), Figure 16 (SEQ ID NO:16), Figure 18 (SEQ ID NO:18), Figure 20 (SEQ ID NO:20), Figure 22 (SEQ ID NO:22), Figure 24 (SEQ ID NO:24), Figure 26 (SEQ ID NO:26), Figure 28 (SEQ ID NO:28), Figure 30 (SEQ ID NO:30), Figure 32 (SEQ ID NO:32), Figure 34 (SEQ ID NO:34), Figure 36 (SEQ ID NO:36), Figure 38 (SEQ ID NO:38), Figure 40 (SEQ ID NO:40), Figure 42 (SEQ ID NO:42), Figure 44 (SEQ ID NO:44), Figure 46 (SEQ ID NO:46), Figure 48 (SEQ ID NO:48), Figure 50 (SEQ ID NO:50), Figure 52 (SEQ ID NO:52), Figure 54 (SEQ ID NO:54), Figure 56 (SEQ ID NO:56), Figure 58 (SEQ ID NO:58), Figure 60 (SEQ ID NO:60), Figure 62 (SEQ ID NO:62), Figure 64 (SEQ ID NO:64), Figure 66 (SEQ ID NO:66), Figure 68 (SEQ ID NO:68), Figure 70 (SEQ ID NO:70), Figure 72 (SEQ ID NO:72), Figure 74 (SEQ ID NO:74), Figure 76 (SEQ ID NO:76), Figure 78

(SEQ ID NO:78), Figure 80 (SEQ ID NO:80), Figure 82 (SEQ ID NO:82), Figure 84 (SEQ ID NO:84), Figure 86 (SEQ ID NO:86), Figure 88 (SEQ ID NO:88), Figure 90 (SEQ ID NO:90), Figure 92 (SEQ ID NO:92), Figure 94 (SEQ ID NO:94), Figure 96 (SEQ ID NO:96), Figure 98 (SEQ ID NO:98), Figure 100 (SEQ ID NO:100), Figure 102 (SEQ ID NO:102), or Figure 104 (SEQ ID NO:104).

5

12. A chimeric molecule comprising a polypeptide according to Claim 10 fused to a heterologous amino acid sequence.

10 13. The chimeric molecule of Claim 12, wherein said heterologous amino acid sequence is an epitope tag sequence or an Fc region of an immunoglobulin.

14. An antibody which specifically binds to a polypeptide according to Claim 10.

15 15. The antibody of Claim 14, wherein said antibody is a monoclonal antibody, a humanized antibody or a single-chain antibody.

16. A composition of matter comprising (a) a polypeptide of Claim 10, (b) an agonist of said polypeptide, (c) an antagonist of said polypeptide, or (d) an antibody that binds to said polypeptide, in combination with a carrier.

20

17. The composition of matter of Claim 16, wherein said carrier is a pharmaceutically acceptable carrier.

25 18. The composition of matter of Claim 16 comprising a therapeutically effective amount of (a), (b), (c) or (d).

19. An article of manufacture, comprising:

a container;

a label on said container; and

30 a composition of matter comprising (a) a polypeptide of Claim 10, (b) an agonist of said polypeptide, (c) an antagonist of said polypeptide, or (d) an antibody that binds to said polypeptide, contained within said container, wherein label on said container indicates that said composition of matter can be used for treating an immune related disease.

35 20. A method of treating an immune related disorder in a mammal in need thereof comprising administering to said mammal a therapeutically effective amount of (a) a polypeptide of Claim 10, (b) an agonist of said polypeptide, (c) an antagonist of said polypeptide, or (d) an antibody that binds to said polypeptide.

21. The method of Claim 20, wherein the immune related disorder is systemic lupus erythematosis, rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an idiopathic inflammatory myopathy, Sjögren's syndrome, systemic vasculitis, sarcoidosis, autoimmune hemolytic anemia, autoimmune thrombocytopenia, thyroiditis, diabetes mellitus, 5 immune-mediated renal disease, a demyelinating disease of the central or peripheral nervous system, idiopathic demyelinating polyneuropathy, Guillain-Barré syndrome, a chronic inflammatory demyelinating polyneuropathy, a hepatobiliary disease, infectious or autoimmune chronic active hepatitis, primary biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis, inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's disease, an autoimmune or immune-mediated skin disease, a bullous skin disease, 10 erythema multiforme, contact dermatitis, psoriasis, an allergic disease, asthma, allergic rhinitis, atopic dermatitis, food hypersensitivity, urticaria, an immunologic disease of the lung, eosinophilic pneumonias, idiopathic pulmonary fibrosis, hypersensitivity pneumonitis, a transplantation associated disease, graft rejection or graft-versus-host-disease.

15 22. A method for determining the presence of a PRO polypeptide in a sample suspected of containing said polypeptide, said method comprising exposing said sample to an anti-PRO83478 anti-PRO69889, anti-PRO37073, anti-PRO4984, anti-PRO71039, anti-PRO71191, anti-PRO271, anti-PRO71042, anti-PRO71242, anti-PRO6015, anti-PRO34336, anti-PRO71043, anti-PRO71044, anti-PRO71045, anti-PRO71046, anti-PRO83479, anti-PRO535, anti-PRO4426, anti-PRO34447, anti-PRO2023, 20 anti-PRO25349, anti-PRO1265, anti-PRO34298, anti-PRO738, anti-PRO71049, anti-PRO21341, anti-PRO1125, anti-PRO4369, anti-PRO177, anti-PRO83480, anti-PRO71052, anti-PRO37125, anti-PRO83481, anti-PRO71054, anti-PRO69462, anti-PRO268, anti-PRO615, anti-PRO25402, anti-PRO71055, anti-PRO71056, anti-PRO34332, anti-PRO71057, anti-PRO71058, anti-PRO83482, anti-PRO34454, anti-PRO6243, anti-PRO1864, anti-PRO71060, anti-PRO71061, anti-PRO793, anti-PRO83483 or anti-25 PRO60929 antibody and determining binding of said antibody to a component of said sample.

30 23. A method of diagnosing an immune related disease in a mammal, said method comprising detecting the level of expression of a gene encoding PRO83478, PRO69889, PRO37073, PRO4984, PRO71039, PRO71191, PRO271, PRO71042, PRO71242, PRO6015, PRO34336, PRO71043, PRO71044, PRO71045, PRO71046, PRO83479, PRO535, PRO4426, PRO34447, PRO2023, PRO25349, PRO1265, PRO34298, PRO738, PRO71049, PRO21341, PRO1125, PRO4369, PRO177, PRO83480, PRO71052, PRO37125, PRO83481, PRO71054, PRO69462, PRO268, PRO615, PRO25402, PRO71055, PRO71056, PRO34332, PRO71057, PRO71058, PRO83482, PRO34454, PRO6243, PRO1864, PRO71060, PRO71061, PRO793, PRO83483, or PRO60929 polypeptide (a) in a test sample of tissue cells obtained from the 35 mammal, and (b) in a control sample of known normal tissue cells of the same cell type, wherein a higher or lower level of expression of said gene in the test sample as compared to the control sample is indicative of the presence of an immune related disease in the mammal from which the test tissue cells were obtained.

40 24. A method of diagnosing an immune related disease in a mammal, said method comprising (a) contacting an anti-PRO83478 anti-PRO69889, anti-PRO37073, anti-PRO4984, anti-PRO71039, anti-

PRO71191, anti-PRO271, anti-PRO71042, anti-PRO71242, anti-PRO6015, anti-PRO34336, anti-PRO71043, anti-PRO71044, anti-PRO71045, anti-PRO71046, anti-PRO83479, anti-PRO535, anti-PRO4426, anti-PRO34447, anti-PRO2023, anti-PRO25349, anti-PRO1265, anti-PRO34298, anti-PRO738, anti-PRO71049, anti-PRO21341, anti-PRO1125, anti-PRO4369, anti-PRO177, anti-PRO83480, anti-PRO71052, anti-PRO37125, anti-PRO83481, anti-PRO71054, anti-PRO69462, anti-PRO268, anti-PRO615, anti-PRO25402, anti-PRO71055, anti-PRO71056, anti-PRO34332, anti-PRO71057, anti-PRO71058, anti-PRO83482, anti-PRO34454, anti-PRO6243, anti-PRO1864, anti-PRO71060, anti-PRO71061, anti-PRO793, anti-PRO83483 or anti-PRO60929 antibody with a test sample of tissue cells obtained from said mammal and (b) detecting the formation of a complex between the antibody and the polypeptide in the test sample, wherein formation of said complex is indicative of the presence of an immune related disease in the mammal from which the test tissue cells were obtained.

25. A method of identifying a compound that inhibits the activity of PRO83478, PRO69889, PRO37073, PRO4984, PRO71039, PRO71191, PRO271, PRO71042, PRO6015, PRO34336, PRO71043, PRO71044, PRO71045, PRO71046, PRO83479, PRO535, PRO4426, PRO34447, PRO2023, PRO25349, PRO1265, PRO34298, PRO738, PRO71049, PRO21341, PRO1125, PRO4369, PRO177, PRO83480, PRO71052, PRO37125, PRO83481, PRO71054, PRO69462, PRO268, PRO615, PRO25402, PRO71055, PRO71056, PRO34332, PRO71057, PRO71058, PRO83482, PRO34454, PRO6243, PRO1864, PRO71060, PRO71061, PRO793, PRO83483, or PRO60929 polypeptide, said method comprising contacting cells which normally respond to said polypeptide with (a) said polypeptide and (b) a candidate compound, and determining the lack responsiveness by said cell to (a).

26. A method of identifying a compound that inhibits the expression of a gene encoding a PRO83478, PRO69889, PRO37073, PRO4984, PRO71039, PRO71191, PRO271, PRO71042, PRO71242, PRO6015, PRO34336, PRO71043, PRO71044, PRO71045, PRO71046, PRO83479, PRO535, PRO4426, PRO34447, PRO2023, PRO25349, PRO1265, PRO34298, PRO738, PRO71049, PRO21341, PRO1125, PRO4369, PRO177, PRO83480, PRO71052, PRO37125, PRO83481, PRO71054, PRO69462, PRO268, PRO615, PRO25402, PRO71055, PRO71056, PRO34332, PRO71057, PRO71058, PRO83482, PRO34454, PRO6243, PRO1864, PRO71060, PRO71061, PRO793, PRO83483, or PRO60929 polypeptide, said method comprising contacting cells which normally express said polypeptide with a candidate compound, and determining the lack of expression said gene.

27. The method of Claim 26, wherein said candidate compound is an antisense nucleic acid.

35 28. A method of identifying a compound that mimics the activity of a PRO83478, PRO69889, PRO37073, PRO4984, PRO71039, PRO71191, PRO271, PRO71042, PRO71242, PRO6015, PRO34336, PRO71043, PRO71044, PRO71045, PRO71046, PRO83479, PRO535, PRO4426, PRO34447, PRO2023, PRO25349, PRO1265, PRO34298, PRO738, PRO71049, PRO21341, PRO1125, PRO4369, PRO177, PRO83480, PRO71052, PRO37125, PRO83481, PRO71054, PRO69462, PRO268, PRO615, PRO25402, 40 PRO71055, PRO71056, PRO34332, PRO71057, PRO71058, PRO83482, PRO34454, PRO6243, PRO1864,

PRO71060, PRO71061, PRO793, PRO83483, or PRO60929 polypeptide, said method comprising contacting cells which normally respond to said polypeptide with a candidate compound, and determining the responsiveness by said cell to said candidate compound.

5        31.        A method of stimulating the immune response in a mammal, said method comprising administering to said mammal an effective amount of a PRO83478, PRO69889, PRO37073, PRO4984, PRO71039, PRO71191, PRO271, PRO71042, PRO71242, PRO6015, PRO34336, PRO71043, PRO71044, PRO71045, PRO71046, PRO83479, PRO535, PRO4426, PRO34447, PRO2023, PRO25349, PRO1265, PRO34298, PRO738, PRO71049, PRO21341, PRO1125, PRO4369, PRO177, PRO83480, PRO71052, 10 PRO37125, PRO83481, PRO71054, PRO69462, PRO268, PRO615, PRO25402, PRO71055, PRO71056, PRO34332, PRO71057, PRO71058, PRO83482, PRO34454, PRO6243, PRO1864, PRO71060, PRO71061, PRO793, PRO83483, or PRO60929 polypeptide antagonist, wherein said immune response is stimulated.

15        32.        A method of diagnosing an inflammatory immune response in a mammal, said method comprising detecting the level of expression of a gene encoding PRO83478, PRO69889, PRO37073, PRO4984, PRO71039, PRO71191, PRO271, PRO71042, PRO71242, PRO6015, PRO34336, PRO71043, PRO71044, PRO71045, PRO71046, PRO83479, PRO535, PRO4426, PRO34447, PRO2023, PRO25349, PRO1265, PRO34298, PRO738, PRO71049, PRO21341, PRO1125, PRO4369, PRO177, PRO83480, PRO71052, PRO37125, PRO83481, PRO71054, PRO69462, PRO268, PRO615, PRO25402, PRO71055, 20 PRO71056, PRO34332, PRO71057, PRO71058, PRO83482, PRO34454, PRO6243, PRO1864, PRO71060, PRO71061, PRO793, PRO83483, or PRO60929 polypeptide (a) in a test sample of tissue cells obtained from the mammal, and (b) in a control sample of known normal tissue cells of the same cell type, wherein a higher or lower level of expression of said gene in the test sample as compared to the control sample is indicative of the presence of an inflammatory immune response in the mammal from which the test tissue 25 cells were obtained.

1/105  
**FIGURE 1**

**ATGGCTT**AGAGATCCACATGTCAGACCCCAGTGCCTCATCGAGAACTTAAATGAGCAGCTGAAGGTTAACAG  
GAAGCTTGGAGATCCGTCTGCCATTACGCAACCTGTAGTTGGTAGCGATTGTGGGCTCTATCGCACTGGC  
AAATCCTACCTGATGAACAAGCTGGCTGGGAAGAACAGGGCTCTCTGTCATCTACGGTGCAGTCTCACACC  
AAGGGAAATTGGATATGGTGTGCCTCATCCAACTGGCAAATCACACATTAGTTCTGCTTGACACCGAGGGC  
CTGGGAGATGTAGAGAAGGCTGACAACAAGAATGATATCCAGATCTTGCACTGGCACTTACTGAGCAGCACC  
TTTGTGTACAATACTGTGAACAAAATTGATCAGGGTGTATCGACCTACTGCACAATGTGACAGAACTGACAGAT  
CTGCTCAAGGCAAGAACTCACCGACCTTGACAGGGTTGAAGATCCTGCTGACTCTGCGAGCTTCCAGAC  
TTAGTGTGGACTCTGAGAGATTCTGCTTAGGCCTGGAAATAGATGGCAACTGTGACACCAGATGAATACTG  
GAGAATTCCCTAACGGCAAAGCAAGGTAGTGTCAAAGAGTTCAAAATTGCTCGTGTGTATAACAG  
AAGTTCTTCCAAAAAAGAAATGCTTATCTTGACTTACCTGCTCACCAAAAGCTGCCAACTTGAAACA  
CTGCCTGATGAGCTAGAGCCTGAATTGTCAACAAGTGACAGAATTCTGTCCTACATCTTAGCCATTCT  
ATGACCAAGACTCTTCAGGTGGCATATGGTCAATGGATCTCGTCTAAAGAACCTGGTGTGACCTATGTCAAT  
GCCATCAGCAGTGGGGATCTGCCCTGCATAGAGAAAGCAGTCTGGCCTGGCTCAGAGAGAGAACTCAGCTGCA  
GTGCAAAAGGCCATTGCCCACTATGACCAGCAAATGGCCAGAAAGTGCAGCTGCCATGGAAACCTCCAGGAG  
CTGCTGGACCTGCACAGGACCAAGTGAGAGGGAGGCCATTGAAGTCTTCATGAAAACACTCTTCAAGGATGTAGAC  
CAAAGTTCCAGAAAGAATTGGAGACTCTACTAGATGCAAAACAGAATGACATTGTAACGGAACCTGGAAGCA  
TCCTCGGATTATTGCTCGGCTTACTTAAGGATATTGGCTCTAGAAGAAGCAGTGAAGCAGGGATTAT  
TCTAAGCCAGGAGGCCATAATCTCTCATTGAGAAACTGAAGGCAAAGTACTATCGGGAGCCTCGG  
AAAGGAATAAGGCTGAAGAAGTCTGCAGAAATTTAAAGTCCAAGGAGTCTGTGAGTCATGCAATTACAG  
ACTGACCAGGCTCTCACAGAGACGGAAAAAAGAAGAAAGAGGCACAAGTGAAGCAGAAGCTGAAAAGGCTGAA  
GCGCAAAGGTGGCGCGATTCAAAGGCAGAACGAGCAAATGATGCGAGGAGAGGAGACTCCATCAGGAACAA  
GTGAGACAAATGGAGATAGCCAAACAAATTGGCTGGCAGAGCAACAGAAAATGCAGGAACACAGATGCAGGAA  
CAGGCTGCACAGCTCAGCACAACATTCAAGCTAAAATAGAAGCCTCTCAGTGAGCTCCAGCACGCCAGAGG  
ACTGTTAATAACGATGATCCATGTGTTTACTCTAA

2/105  
**FIGURE 2**

MALEIHMSDPMCLIENFNEQLKVNQEALEILSAITQPVVVVAIVGLYRTGKSYLMNKLAGKNKGSVASTVQSHT  
KGIWIWCVPHPNWPNTLVLLDTEGLGDVEKADNKNDIQIFALALLSSTFVYNTVNKIDQGAIDLHNVTELTD  
LLKARNSPDLDLDRVEDPADSASFDPDVWTLRDFCLGLEIDGQLVTPDEYLENSLRPKQGSDQRVQNFNLPRLCIQ  
KFFPKKKCFIFDLPAHQKKLAQLETLPDDELEPEFVQQVTEFCSYIFSHSMTKTLPGGIMVNGSRLKNLVLTYVN  
AISSGDLPCIENAVLALAQRENSAAVQKAIAYDQQMGQKQVQLPMETLQELLDLHRTSEREAIEVFMKNSFKDWD  
QSFQKELETLLDAKQNDICKRNLEASSDYCSCALLKDIFGPLEEAVQGIVSKPGGHNLFIQKTEELKAKYYREPR  
KGIQAEEVLQKYLKSKEVSVAIIQTDQALTEKEKKKEAQVKAEEAKAEQRLAAIQRQNEQMMQERERLHQEQ  
VRQMEIAKQNWLAEQQKMQEQQMQEAAQLSTTFQAQNRSLLSELQHAQRTVNNDDPCVLL

3/105  
**FIGURE 3**

AGCAGGTTCGAATGCTCTTACTCCCTTGTGGAGCAAAGAAAAAGCAGGAGTATTGAACAAATCACTAAG  
ACTCATGGAACAATTATTGGCATTACTTCAGGGATTGTCTGGCTTCTCATTATTCTATTAGTACAAGTG  
AAACAGCCTCGAAAAAAGGTATGGCTGAAAACCGCTTTAATAAAACCGGGTCCAAGAAGTGTGATCC  
CCTCATTATGAACGTTCACTAAGGGACAAAGAGATTCTGCAGACCTGGCAGACTGTGCGAAGAATTGGAC  
AACTACCAGAGGATGCGGCGCTCCACCGCTCCGCTGCATCCACGACCACACTGTGGTGCAGGCCTCC  
AGCGTCAAACAAAGCAGGACCAACCTCAGTCCATGGAGCTCCCTCCGAAATGACTTGCACAACCACAGCCA  
ATGAAAACATTAATAGCACCTCAAGAAAAGTAGTTACACTTCAAACAGGGACATGAGTGCCTGAGCAGGCC  
CTGGAAGACCGAGTAATGGAGGAGATCCCTGTGAAATTATGTCAGGGGGGAGAAGATTCTGCACAAGCATTCC  
ATATCCATTGACTTCTAATCTCTGCTAATGGTATGTGAATTCTTAGGGTGTGTACGTACGCAGCCTCCAGGGC  
ACCATACTGTTCCAGCAGCCAACCCCTTCTCCATCACAACATACGAAGACCTGATTACCGTTAACCTATTG  
TATGGTATGTTTATTCTCAGGCAGTCTATATATGTTAAACCAATCAAGGAACCTACTCTATTCACTGGAA  
ACAATAATCATCTCTATTGCTTGGTGTCTTATAGGAAGCACTGCCAGTTAAAGAGCATTAGAAGAGGTGGTTG  
GATGGAGCCAGGCTCAGGCTGCCCTTCGTTAGCAACAAGAAGACTGCTTGTACTGATAACAGCTCTGCAA  
TATTTGATGCCACAATAAACTGATTTCTTACATTCTTTATTCTCTTCTAAATTAAATTGTTT  
TATAAGCCTATCGTTTACCATTCATTCTTACATAAGTACAAGTGGTTAATGTACCACTACACTTCAGTATAG  
GCATTGTTCTTGAGTGTGTCAAACATACAGCTAGTTACTGTGCCATTAAAGACCCAGTGTATTCAACCCATCTG  
TTCTCTGGCTAATCTGTACTTCTGCCATTAAATTACTGGGCCATTCTTCTTCTGAGAAATAA  
TAGATGATGATTATTACCTTCAATTATATTCTCAGTTACTAGAAAATTCTATACTCTGGGATATA  
TGTACCATGTCAGCTATGACTAAAATTGAAAAAGATAAAATTCTAGCAAGCCTTGAAGTTACCAAGTA  
TAGTCACATTCACTGACAGCCCATTCACTTCAATTCTGACTTCACTTGGAGAGGCCTATAATTAC  
ATTATTGCAATGTTCTTCGCTAGATTGTACATAGCTCCATTCTGGTTTGTCTACAGCATATGGT  
AACCAAGGTTAGATGCCAGTTAAATTCTTAGAAATTGGATGAGCCTTGAGATTGCTCTTAACGGACATG  
CATTTCTAGCTTTATCAAGAATAACAATTCCACTTTTAAACTGCACATTGACTTTTATGGTAT  
AAAAACAATAATTATAAACATAAAAGCTATTGTGTTTTAGACTTTGATATTATTGATACTGTACAAACT  
TTATTAAATCAAGATGAAAGACCTACAGGACAGATTCTTCAGTGTACATCAGTGGCTTGTATGCAAATAT  
GCTGTGTTGGACCTGGACGCTATAACTTATTGTAAGACCTGGAAATGTGGACATAAGCTCTTCTTCTT  
GTTACTGTATTAGTTGTGATAAAATTCTACTGTGATATTATGCTCTAAACTACACAAATCCCATA  
TAAAATATACTTGTACCTGACCCCTTAATCATGTTATTATGCCACCAAGGTTGTGGATCTTAAGGTATGT  
GAAAGGAACCTATTATCAAATTGTAAGTAATACAGACATGCCATTAAAAGAGGTAAATTCTTGTGTTCTATAT  
TTTGTGATAAAATTCTCAATGAAATAAGTGAAGTTCACTGGATTCTAACCTTAAATTACATATATG  
GTTTCTCAGATTAGTGAAATTGTGACCTTAAATTAAATACACATATACTGCCTCAG

4/105  
**FIGURE 4**

MLFTSFVEQKKKAGVFEQITKTHGTIIGITSGIVLVLLIISILVQVKQPRKKVMACKTAFNKTGFQEVDPPHYELFSLRDKEISADLADLSEELDNYQRMRRSSTASRCIHDDHCGSQASSVKQSRTNLSSMELPLRNDFAQPQPMKTFNSTFKKSSYTFKQGHECPEQALEDRVMEEIPCEIYVRGREDSAQASISIDF

5/105  
**FIGURE 5**

TCTGGGCGCGCGGACGTCAAGTTGAGTTCTGTGTTCTCCCCGCCC GTGTCCCGCCCCGACCCGCGCCCGCG **ATGC**  
TGGCGCTGCGTGC GGCTCCCGCTGGCTGGCCTGCTCTCCGTCCCGCCTCCGTGCCGTGCCCTGCCCTCCCGCG  
CCCGCGCCTGCA GCAAGGGCTCCGGCGACCCGCTCTCTCCCTCCGGAAACCCGCTCGTGTACCTGGACG  
TGGACGCCAACGGGAAGCCGCTCGGCCGCTGGTGCTGGAGCTGAAGGCAGATGTCGTCCCAAAGACAGCTGAGA  
ACTTCAGAGCCCTGTGCACTGGTGAGAAGGGCTCGGCTACAAAGGCTCCACCTCCACAGGGTGTACCCCTCCT  
TCATGTGCCAGGCGGGGCACTTCACCAACCACAATGGCACAGGCGGGAAAGTCCATCTACGGAAGCCGTTCC  
ACGAGAACCTTACACTGAAGCACGTGGGCCAGGTGTCCATGGCTAATGCTGGCTAACACCAACGGCT  
CCCAGTTCTCATCTGCACCAATAAGACAGACTGGTTGGATGGCAAGCAGTGTGTGTTGGTCACGTCAAAGAGG  
GCATGGACGTCGTGAAGAAAATAGAATCTTCGGCTCTAAGAGTGGGAGGACATCCAAGAAGATTGTCATCACAG  
ACTGTGCCAGTTGAGC **TAAT**CTGTGCCAGGGTGCTGGCATGGTGGCAGCTGCAAATGTCCATGCACCCAGGTG  
GCCGCGTGGGCTGTCA GCCAAGGTGCCTGAAACGATA CGTGTGCCCACTCCACTGTACAGTGTGCCTGAGGAA  
GGCTGCTAGGGATGTTAGACCTCGGCCAGGACCCACCACATTGCTTCTAATACCCACCCCTCCTACGACCTCA  
TTTCTGGGATCTTGTGGACATGATGTCACCCACCCCTGTCAAGCATTGCCGTGATTGCCAGCCCAGATTCA  
ATCTGTGCCCTGGACATGGTGTGGATGGGTTGCCATCCAAGTGAAGGCTTTGAGATCTTCCTGGTTACTG  
TCAGTTTGCAAAAGGACTCTAATACCTGTTAATATTGCTTCTAATTGGGATAATTAAACAGATTGA  
CTAGAAGTGAACACTCAACTTCCCCGTGCTGTGGTGTGACCTGAGTTGGTACACAGGCCACAGACCCC  
AGAGCTTGGCTTTGAAACACAACACTCAGGGCTTTGTGAAGGTTCCCCGCTGAGATCTTCCTGGTTACTG  
TGAAGCCTGTTGGTTGCTGTCGTTTGAGGAGGGCCATGGGGTAGGAGCAGTTGAACCTGGAAACAAA  
CCTCACTTGAGCTGTGCC TAGACAATGTGAATTCTGTGCTAACAGAAGTGGCCTGTAAGCTCTGTGCTCC  
GGAGGGAAAGCATTCTGGTAGGCTTGATTTCTGTGTAAAGAAATTCAATCTACTCATGATGTGTTATG  
CATAAAACATTCTGGAACATGGATTGTGTTCACCTAAATGTGAAAATAATCCTATTTCTATGGAAAAAAA  
AAAAAAAAAAAAAAA

6/105  
**FIGURE 6**

MLALRCGSRWLGLLSVPRSVPLRLPAARACSKGSGDPSSSSSGNPLVYLDVDANGKPLGRVVLELKADVVPKTA  
ENFRALCTGEKGFGYKGSTFHHRVIPSFMQAGDFTNHNGTGGKSIYGSRFPDENFTLKHVGPGVLSMANAGPNTN  
GSQFFICTIKTDWLDGKHVVFGHVKEGMDVVKKIESFGSKSGRTSKKIVITDCGQLS

7/105  
**FIGURE 7**

TGTTCTTGAGCCCAGCTCTTCTCGTCTCCACCCAGCTTCCCGCATTGGAAGAAGGGACCGTCCCTTCCTT  
GTCTTGGCCACCCAAATCCTGGTATCGAAAGGGTGAACGGACCGGAAGTGTGCAGCAGCGACGGTCCCCAGCT  
AATCGACGCCCGAAGTAGCAATTACTAGACAAGCATTCCGCCGCCGCTCGCTATGGCGGCAATTCCCCCAGAT  
TCCTGGCAGCCACCCAACGTTACTTGGAGACCAAGCATGGGAATCATTGTGCTGGAGCTGTACTGGAAGCATGCT  
CCAAAGACCTGTAAGAACCTTGCTGAGTTGGCTCGAGGTTACTACAATGGCACAAATTCCACAGAATTATC  
AAAGACTTCATGATCCAAGGAGGTGACCCAACAGGGACAGGTGAGGTGGTGCATCTATCTATGGCAAACAGTT  
GAAGATGAACCTCATCCAGACTGAAATTACGGGGGCTGGAATTCTCGCAATGGCAATGCGGGGCCAGATACC  
AATGGCAGCCAGTTCTTGACCCCTGCCACCCAGTGGCTGACGGCAAACACACCATTGGCCGAGTG  
TGTCAAGGCATAGGAATGGTGAATCGCTGGGAATGGTAGAAACAAACTCCAGGACGCCGTGGACGACGTG  
AAGATCATTAAAGGCATACCCCTCTGGGTAGACTTGCTACCCCTCTGAGCAGCTCTGAGATGGCCCCAGTGA  
CCAGCTCTAGATGACATAGAACATGTAATGCTAAATTCTATTTGGCTTGCAAGTCATGAAGCTTAGGAG  
GCCCTGGCATCTGGGTGAGTTAGAGATGGAAGTACATTAAATAGGATGCTCTTCTCTCCAGTCCTAA  
GGTTGCCAGAGCATTGACAAAATGCCCTGTTATCAATAGGTGACTACTACACATGAACCATATACTGCTG  
CTTCTGTGATGTCATGCTGATATACTGCAACATGTCAGCTTCTGCTGACGGTTGGCTAA  
CCAAACTCTCCTAAGGAGATTATATTCTGGCTACACAGCAGCTTGATGGCTGACAGCCACAGAATTCCA  
AACCAAGTAGTGTCTGTCAGCCCTTTAACTCTGTGACGCCCTATTCTAGCTTTACATTGTTCTCTAGGG  
AATGTATGCATCTCTATATATTCTCCCTCTAAACACAGAACATCAACAGTGTGTTCTGACACTCAGACA  
TCCCACGCAAAGCCACATTGAATTGGCCAAATGAAAAACACATCCAACATCAAGTTCTAAGAAGGTGTCAA  
GTGGGGAATAATAATAATGTATAATAATCAAGAAATTAGTTATTAAAGGAAGCAGAACGATTGACCATT  
CCCAGAGAAGAGGAGAAATCTGTAGTGAGCAAAGGACAGACCATGAATCCTCTTGAGAAGTAGTACTCTCAGAA  
AGGAGAACGCCACTCAAGTTCTTAACCCAAAGACTTTAGAGAAATTAGGTCCAAGATTATATGTTAGTT  
GTTTATGTATAAAATAACTTCTGGATTTGTGGGGAGGAGCAGGAGAGGAAGGAGTTAACCTATGTAATA  
CATAGAAACTCCACAATAAAATGCCATTGATGGTTGAAAAA

8/105  
**FIGURE 8**

MAAIPPD SWQPPNVY LETSMGI IIVLELYWKHAPKTCKNFAELARRGYYNGTKFHRIIKDFMIQGGDPTGTGRGGA  
SIYGKQF EDELHPDLKFTGAGILAMANAGPDTNGSQFFVTLAPTQWLDGKHTIFGRVCQGIGMVNRVGMVETNSQ  
DRPVDDVKIIKAYPSG

9/105  
**FIGURE 9**

GTAACGGATGGTGCGCCAACGTGAGAGGAAACCCGTGCGCGCTGCCTTCTGTCCCCAAGCCGTTAGACG  
CGGATGAAGTGCAAAACAAACTTCCATAGAGGAGTTGCAAAGTCCAGTTATACCAAACAGTAATCAGA  
TTCCATTGGAAGCTAAAGATTTGAGAGCCTTTGACTATATGCAACTAATTGATTTCAAGCTGGAACTTT  
TAAAAAAAACATTAAGCAAAATGAAAATGCTTCTGAAAGCAGCTCCTTTGAAAGGTGTGATGCTTGGAAAG  
CCATTTCTGTGCTTGATCCACTAATGCTAAGGACACATTAGGATTGGTCATGAAATAGAATGCACCACCATG  
AGCATCATCACCTACAAGCTCTAACAAAGAAGATATCTGAAAATTCAGAGGATGAGCGCATGGAGCTCAGTA  
AGAGCTTCGAGTATACTGTATTATCCTGAAAACCAAAGATGTGAGTCTTGGCTGAGTAAAGGAGACTT  
GGACCAAACACTGTGACAAAGCAGAGTCTCAGTTCTGAAAATGTTAAAGAGTTGAGTCATTAATATGGACA  
CAAATGACATGTGGTTAATGATGAGAAAAGCTTACAAATACGCCTTGATAAGTATAGAGACCAATACAACGTG  
TCTTCTTGACGCCACTACGTTGCTATCATTGAAAACCTAAAGTATTTTGTAAAAAGGATCCATCAC  
AGCCTTCTATCTAGGCCACACTATAAAATCTGGAGACCTGAAATATGTGGGTATGAAAGGAGGAATTGCTTAA  
GTGTTAGAATCAATGAAAAGACTAACGCCTCTCAATATCCCAGAAAAGTGTCTGAAACAGGGAGGGATGATT  
GGAAGATATCCGAAGATAAACAGCTAGCAGTTGCCTGAAATATGCTGGAGTATTGCAAGAAAATGCAAGATG  
CTGATGGAAAAGATGTATTAATACCAATCTGTTGGCTTCTATTAAAGAGGCAATGACTTATCACCCCAACC  
AGGTAGTAGAAGGCTGTTGTCAGATATGGCTGTTACTTTAATGGACTGACTCCAAATCAGATGCATGTGATGA  
TGTATGGGTATACCGCTTAGGCATTGGCATATTCATGATGCATTGGTTTCTACCTCAAATGGTT  
CTGACAATGACTTGAGAAGTGGTAGAAAAGCGTGAATATGATCTTGTATAAGACGTGTGTCATTATTTGTAG  
TAGTAACTACATATCCAATACAGCTGTATGTTCTTTCTAATTGGTGGCACTGGTATAACCACCCAT  
TAAAGTCAGTAGTACATTTAAATGAGGGTGGTTTTTCTTAAACACATGAACATTGAAATGTGAAATTGTTGGAA  
AAAAGTGTGTTAAGAATAATAATTGCAAATAACTATTAATAATATTATGTGATAAAATTCTAAC

10/105  
**FIGURE 10**

MHHHEHHHLQAPNKEDILKISEDERMELSKSFRVYCIILVKPKDVSLLWAAVKEWTWKHCDKAEFFSSENVKEFES  
INMDTNDMWLMMRKAYKYAFDKYRDQYNWFFLARP TTFAI IENLKYFLLKKDPSPQFYLGHTIKSGDLEYVGMEG  
GIVLSVESMKRLNSLLNIPEKCPEQGGMIWKISEDKQLAVCLKYAGVFAENAEDADGKDVFNTKSVGLSIKEAMT  
YHPNQVVEGCCSDMAVTFNGLTPNQMHVMMYGVYRLRAFGHIFNDALVFLPPNGSDND

11/105  
**FIGURE 11**

TCAGTGGCGTCGCGCGAAGGCTAAGGGAGTGTGGCGGGCGGCTCCGGGAGCCAAC **ATGCCTCGGTATGCGCAGC**  
TGGTCATGGCCCCCGCGGGCAGCGGAAGAGCACC TACTGTGCCACCATGGTCCAGCACTGTGAAGCCCTCAACC  
GGTCTGTCCAAGTTGTAACCTGGATCCAGCAGCAGAACACTTCAACTACTCCGTATGGCTGACATCCGGGAAC  
TGATCGAGGTGGATGATGTAATGGAGGATGATTCTCGCATTGGTCCCAACGGAGGATTGGTATTTGCATGG  
AGTACTTTGCCAATAATTTGACTGGCTGGAGAAC TGTCTGGCCATGTAGAGGACGACTATATCCTTTGATT  
GTCAGGTCAAGATTGAGTTGTACACTCACCTGCCTGTGATGAAACAGCTGGTCCAGCAGCTCGAGCAGTGGAGT  
TCCGAGTCTGTGGAGTTCTTGATTCTCAGTTCATGGTGGAGTCATTCAAGTTATTTCTGGCATCTGG  
CAGCCCTGAGTGCCATGATCTCTAGAAATTCCGCAAGTCACATCATGACAAAAATGGATCTGCTGAGTAAAA  
AAGCAAAAAGGAAATTGAGAAATTTAGATCCAGACATGTATTCTTATTAGAAGATTCTACAAGTGACTTAA  
GAAGCAAAAATTCAAGAAACTGACTAAAGCTATATGTGGACTGATTGATGACTACAGCATGGTCGATTTTAC  
CTTACGATCAGTCAGATGAAAGAACATTGCATTGCAGCATATTGATTTCGCATTCAATATGGAGAAG  
ACCTAGAATTAAAGAACCAAAGGAACGTGAAGATGAGTCTCCCTATGTTGACGAATATTTCAAGAACATGCC  
AGGATGAA **TGAAGAGTT**ACTAAAGTAACCATCTAAAGAGCTTGTGGCCAAACCAGCAGAACATTCTCTTC  
AAAGGATGCAATAGTAGAAAGCTACTTATTAAATGAAAAAAAGTAAAACCTCGTTCTTATCAGCCTCATGCCT  
GAATCAAATTAAATTATTCTGAAACTGCTGCTTTAAAGTGAATCTTTAGTATTATAACAGCATCACTTT  
AGATTTGTAAGTCAAAATTGAAATGAATGCACATAGATTATATAAAATTAGCACCTGAGCTAAGGTTAAGGC  
TGGCTAAACTTATTTCACTTTGTATTATTGAGATGCAGGAATTACTGTAACAAAATATGTATGTCCGA  
AGGGAAAAAGCTGCAAGGATATATAAGACCACTGCTATCTGATCTTCCCATTTCCTATATTGAAAATGTA  
TATTATTATATAACTAAAAAGAAAATAACTATGTTGAGATATGTATGTATATAAAAGAAACAAAG  
GTTTTAATGATTCTGGACCTAGATAACAAGTAAAAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAA

12/105  
**FIGURE 12**

MPRYAQLVMGPAGSGKSTYCATMVQHCEALNRSVQVNLDPAAEHFNYSVMA  
GLVFCMEYFANNFDWLENCLGHVEDDYILFDCPGQIELYTHLPVMKQLVQQLEQWEFRVCGVFLVDSQFMVESFK  
FISGILAALSAMISLEIPQVNIMTKMDLLSKAKKEIEKFLDPDMYSLL  
EDSTSDLRSKKFKLTKAICGLIDDY  
SMVRFLPYDQSDEESMNIALQHIDFAIQYGEDLEFKEPKEREDESSSMFDEYFQECQDE

13/105  
**FIGURE 13**

CCGGGACAGCTCGCGCCCCGAGAGCTAGCCGTCGAGGAGCTGCCTGGGACGTTGCCTGGGCCAGC  
CTGGCCGGGTACCCCTGGCATGAGGAGATGGGCTGTTGCTCTGGTCCCCTGCTCCCTGCTGCCGGCTCCTA  
CGGACTGCCCTCTACAAACGGCTTACTACTCCAACAGGCCAACGACCAGAACCTAGGCAACGGTCA  
AGACCTCCTTAATGGAGTGAAGCTGGTGGAGACACCCGAGGAACCCCTGTTACCTACCAAGGGCCAGTGT  
GATCCTGCCCTGCCGCTACCGCTACGAGCCGCCCTGGTCTCCCCGCCGGCGTGTGCGTGTCAAATGGTGGAA  
GCTCGGAGAACGGGGCCCCAGAGAAGGACGTGCTGGTGGCATCGGGCTGAGGCACCGCTCCTTGGGACTACCA  
AGGCCGCGTGCACCTGCCGAGGACAAAGACATGACGTCTCGCTGGAGATCCAGGATCTGCCGCTGGAGGACTA  
TGGCGTTACCGCTGTGAGGTATTGACGGCTGGAGGATGAAAGCGGTCTGGTGGAGCTGGAGCTGCCGGTGT  
GGTCTTCCTTACAGTCCCCAACGGCGCTACCAGTTCAACTTCCACGAGGGCCAGCAGGTCTGCA  
GGCTGCCGTGGTGGCCTCCTTGAGCAGCTCTCCGGGCTGGAGGAGGGCTGGACTGGTGCAACGCC  
GCTGCAGGATGCCACGGTCAGTACCCCCATCATGTTGCCCGGCAGCCCTGCCGGTGGCCGGCTGGCACCTGG  
CGTGCAGGATGCCACGGCCCTGCACCGCTATGATGTATTCTGCTTCGCTACTGCCCTCAAGGG  
GCCGGGTGACTACCTGGAGCACCCCTGAGAAGCTGACGCTGACAGAGGCAAGGGAGGCCTGCCAGGAAG  
CACGATGCCAAGGTGGGACAGCTTTGCCCTGGAAGTTCCATGGCCTGGACCGCTGCCAGCTGGCTGG  
GGCAGATGGTAGCGTCCGCTACCGCTGGTCACTGCACTGTTACTGCTACCGCCAGCATAGACTGGGCC  
CTTGCTTCCCGACCCGCAAGGCCCTGTTACGGTGTATTATTGAGTGGTCTGGGCTGGAGGCATTAA  
CCCTGCCGCATTCCCTACTGGCTGTATTATTGAGTGGTCTGGGCTGGAGGCATTAA  
TGTTTTATACTTCTCAATTAAATTCTTAAACATTTTACTATTGGTAAAGCAAACAGAACCCAA  
GCCCTCCCTTGCTCTGGATGCCCACTCCAGGAATCATGCTTGCCTCCCTGGGCCATTGCGGTTTG  
TCTGGAGGGTCCCCGCCATCCAGGCTGGTCTCCCTCCCTAAGGAGGTTGGTGCCAGAGTGGCGGTGGC  
TCTAGAATGCCGCCGGAGTCCGGCATGGTGGGACAGTTCCCTGCCCTCAGCCTGGGGAAGAAGAGGG  
CTCGGGGGCCTCCGGAGCTGGCTTGGGCCCTCTCTGCCCACCTACTTCTCTGTGAAGCCGCTGACTG  
CCAGTTCTAGGCTCCAGGCAGGAAAGCTGAGGGAAGGAAGAAACTCCCCCTCCCTCCCTCGGT  
CAAAGAATCTGTTGTCATTGTTCTCTGTTCCCTGTGTGGGAGGGCCCTCAGGTGTGTACTTT  
GGACAATAAATGGTGTATGACTGCCCTCCGCCAAAAAA

14/105  
**FIGURE 14**

MGLLLLVPLLLPGSYGLPFYNGFYYNSNSANDQNLGNGHGKDLLNGVKLVVETPEETLFTYQGASVILPCRYRYE  
PALVSPRRVRVKWWKLSENGAPEKDVLVAIGLRHRSFGDYQGRVHLRQDKEHDVSLEIQDLRLEDYGRYRCEVID  
GLEDESGLVELELRGVVFQSPNGRYQFNFHEGQQVCAEQAAVVASFEQLFRAWEGLDWNCAGWLQDATVQYP  
IMLPRQPCGGPGLAPGVRSYGRHRLHRYDVFCFATALKGRVYYLEHPEKLTLEAREACQEDDATIAKVGQLF  
AAWKFHGLDRCDAGWLADGSVRYPVVHPHPNCPPPEPGVRSFGFPDPQSRLYGVYCYRQH

15/105  
**FIGURE 15A**

16/105  
**FIGURE 15B**

CCCGAGTAGCTGGATTACAGGCACACACCACGCCAGCTAGTTTTGTATTTAGTAGAGACGGGTT  
TCACCATGCAAGCCCAGCTGCCACGTAGGTTAAAGCAAGGGCGTGAAGAAGGCACAGTGAGGTATGTGGCT  
GTTCTCGTGGTAGTTCATCGGCCCTAAATAGACCTGGCATTAAATTCAAGAAGGATTGGCATTCTCTTCTT  
GACCCCTCTCTTAAAGGGTAAAGATTAATGTTAGAATGACAAAGATGAATTATTACAATAATCTGATGTAC  
ACAGACTGAAACACACACACATAACCCCTAATCAAAACGTTGGGGAAAAATGTATTGGTTTGTCCTTCATC  
CTGTCTGTGTTATGTGGTGGAGATGGTTTCATTCTTCATTACTGTTGTTATCCTTGATCTGAAATA  
CCTTAATTATTAATATCTGTTGTCAGAGCTCTGCCATTGAGTACCTGTTAGTTAGTATTATTTATGT  
GTATCGGGAGTGTGTTAGTCTGTTATTGCAAAACCGATCTCAAAGATTCCCTTGAAACGCTTTT  
CCCCCTCTTAATTTTATATTCTTACTGTTTACTAAATATTAAGTGTCTTGACAATTGGTGCTCATGTG  
TTTGGGGACAAAGTGAATGAAATCTGTCATTATACAGAAAGTTAAATTCTCAGATCAAATGTGCCTTAATAA  
ATTGTTTCATTAGATTCAACAGTGATAGACTTGCCATTAAATACACGTCAATTGGAGGGCTCGTATTG  
TAAATAGCCTGATGCTATTGGAAAATAAACCAAGTGAACAATATTCTATTGTACTTTCAGAACCAATT  
GTCTCATTATTCTGTTAGCTGAAGAATTGTATTACATTGGAGAGTAAAAACTAAACACG

17/105  
**FIGURE 16**

MAARGRRAWSVLLGLVLFVLASRLVLPASELKRAGPRRRASPEGCRSGQAAASQAGGARGDARGAQLWPPGS  
DPGGPRDRNFLFVGVMTAQKYLQTRAVAAYRTWSKTPGKVQFFSSEGSDTSPVIPVPLRGVDDSYPPQKKSF  
MMLKYMHDHYLDKYEWFMRAADDVYIKGDRLENFLRSILNSSEPLFLGQTGLGTEEMGKLALEPGENFCMGGPGV  
IMSREVLRRMVPHIGKCLREMYTTEDVEVGRCVRRFAGVQCVWSYEMQQLFYENYEQNKKGYIRDLHNSKIHQA  
ITLHPNKNPPYQYRLHSYMLSRKISELRHRTIQLHREIVLMSKYSNTEIHKEDLQLGIPPSFMRFQPRQREEILE  
WEFLTGKYLSSAVDGQPPRRGMDSAQREALDIVMQVEMINANAKTRGRIIDFKEIQGYRRVNPMYGAEYILD  
LLLLYKKHKGKKMTPVRRHAYLQQTFSKIQFVEHEELDAQELAKRINQESGSSLFLSNSLKKLVFQLPGSKSE  
HKEPKDKKINILIPSLSGRFDMFVRFMGNFEKTCLIPNQNVKLVVLNFNSDSNPDKAKQVELMTDYRIKYPKADMQ  
ILPVSGEFSRALALEVGSSQFNNESSLFFCDVDLVFTTEFLQRCRANTVLGQQIYFPIIFSQYDPKIVYSGKVP  
DNHFAFTQKTGFWRNYGFGITCIYKGDLVRVGGFDVSIQGWGLEDVDFNKVVQAGLKTFRSQEVGVVHHPVF  
CDPNLDPKQYKMCGLGSKASTYGSTQQLAEMWLEKNDPSYSKSSNNNGSVRTA

18/105  
**FIGURE 17**

GCTGGAGCCGGGCCGGGCGATGTGGAGCGCGGGCCGGCGGGCTGCCTGGCCGGTGCCTGGGGCTGCTGC  
TGGCGCTTTAGTGCCTGGCGGTGGTGCCTGCCAAGACCGGTGCAGGAGCTCGTGCACCTGCCTGGGTGCTGAAGC  
TGCTCAATACGCACCACCGCGTGCCTGCACTCGCACGACATCAAATACGGATCCGGCAGCGGCCAGCAATCGG  
TGACCGGCGTAGAGGGCGTGGACGACCGAATAGCTACTGGCGGATCCGGCGGCTCGGAGGGCGGGTGCCCGT  
GCGGGTCCCCGGTGCCTGCCTGGCGAGGCGGTGAGGCTCACGCATGTGCTTACGGGCAAGAACCTGCACACGCACC  
ACTTCCCCTGCCTGCAACAAACCAGGAGGTGAGTGCCTTGGGAAGACGGCGAGGGCGACGGACCTGGACC  
TATGGACAGTGCGCTGCTCTGGACAGCACTGGAGCGTGAGGCTGCTGTGCCTTACAGCATGTGGGCACCTCTG  
TGTTCCCTGTCAGTCACGGTGAGCAGTATGGAAGCCCCATCCGTGGGCAGCATGAGGTCCACGGCATGCCAGTG  
CCAACACGCACAATACGTGGAAGGCCATGGAAGGCATCTCATCAAGCCTAGTGTGGAGCCCTGCAGGTACG  
ATGAACTTGAGTGTGGATGGATGGGTGGCAGGTGGGGCTGCAGGGCCACTCTGGCAG  
AGACTTGGTTGTAGGGTCCTCAAGTGCCTTGTGATTAAAGAATGTTGGTCTATGA

19/105  
**FIGURE 18**

MWSAGRGGAAWPVLLGLLLALLVPGGAAKTAGELVTCGSVLKLLNTHHRVRLHSHDIKYGSGSGQQSVTGV  
EAS  
DDANSYWRIRGGSEGGCPCGSPVRCGQAVRLTHVLTGKNLHTHHFPSPLSNNQEVSAGFEGDGE  
GDDLDLWTVRC  
S  
GQHWEREAAVRLQHVGTSVFLSVTGEQYGSPIRGQHEVHGMP  
SANTHNTWKAMEGIFIKPSVEPSAGHDEL

20/105  
**FIGURE 19**

TCTCAGGGCTCATAACAGGAAACTATTGCTGTCAAGTTCCAGAGAAAAGCTTCTGTCGTCCAAGTTACTAA  
CCAGGCTAAACCACATAGACGTGAAGGAAGGGGCTAGAAGGAAGGGAGTGCCCCACTGTTGATGGGTAAGAGGA  
TCCGTACTGAGAAGTTGACCAAGAGAGGGTCTCACCATGCGCACAGTTCTCTGTACCTGTGTGGAGGAAAAGT  
ACTGAGTGAAGGGCAGAAAAAGAGAAAACAGAAATGCTGCCCTGGAGAACTGCTAACCTAGGCTACTGTG  
ATTTGACTATCTCTAGTGGCGAAGCGGAGGGTGCCTGCTCAACCAAACAACCTATTAATGCTGCAAACACTAGC  
AAGGAGAACATGCTTAGCTTCAGCAGTTATGTATGGATGAAAACAGATTACACAGAACTACTCGAAAGTA  
CTCGCAGAAGTTAACACTCATGCCGTAAAGATGGCTACAAATGCTGTGCTTGTGCCCTCATCGCATTAA  
AGAAATTGATCATAATAACATGGAAATAATCCTGAGAGGCCAGCCTTCCTGCACAAAAGCCTACAGGAAAGAA  
ACAAATGAGACCAAGGAAACCAACTGTACTGATGAGAGAATAACCTGGCTCCAGACCTGATCAGAATTGGAC  
CTTCAGATTGTCAGTGGCCATCACTCATGACGGTATTACAGATGCATAATGTTAACACCTGATGGAAATTTC  
CATCGTGGATATCACCTCAAGTGTAGTTACACCTGAACACTGACCTGTTCAAAACAGGAATAGAACTGCAGTA  
TGCAGGGCAGTTGCAAGGGAAAGCCAGCTGCCAGATCTCTGGATCCCAGAGGGCGATTGTGCCACTAACAGAA  
TACTGGAGCAATGGCACAGTGACTGTTAAGAGTACATGCCACTGGAGGTCCACAATGTTCTACCGTGACCTGC  
CACGTCTCCATTGACTGGCAACAAGAGTCTGTACATAGAGCTACTTCCTGTTCCAGGTGCCAAAAATCAGCA  
AAATTATATATTCCATATATCATCCTTACTATTATTTTGACCATCGTGGATTCTTGGTTGAAAGTC  
AATGGCTGCAGAAAATATAAATTGAATAAAACAGAATCTACTCCAGTTGTTGAGGAGGATGAAATGCAAGCCAT  
GCCAGCTACACAGAGAAGAACATCCTCTATGATACAAACAAAGGTGAAGGCATCTCAGGCATTACAAAGT  
GAAGTTGACACAGACCTCCATACTTTTAAGTTGTTGACTCTAGTACCAAGAAACAACAAACAGAGATA  
TATAATTACTGCTGATTTCTACAGTTAGAATGAAGACTTATATTGAAATTAGGTTCCAAGGTTCTTAG  
AAGACATTAAATGGATTCTCATACCCCTGTATAATTGAAATTGGATTCTAGCTGCTACCAGCTAGTT  
CTCTGAAGAACTGATGTTATTACAAAGAAAATACATGCCCATGCCAAATATTCAAATTGTCAGGACAGTAAT  
AATGAAAACCAATTCTCAAGAAATAACTGAAGAAGGAGCAAGTGTGAACAGTTCTGTTGATCCTTCAGA  
ATATTAAATGTACATATGACATGTATATGCCTATGGTATATGTCATAATTGTCATATGTCATACATATA  
TGCACATATCTTGTCAAGGCACCAAGGGAAACAATACACTGCATTACTGTTCTATACATATGAAAACCTAATAA  
TATAAGTCTTAGAGATCATTATCATGACAAGTAGAGCTACCTCATTCTTTAATGGTTATATAAAATTCC  
ATTGTATAGTTATATCATTATTAAATTAAAAACACCTTAATGTTGATATTAGATTCTTTAAGTTGTTA  
TTCTTTAAGTTGTTGTTGTTGATAAAACAAACACATAGAAATGTTCTGTTCATATCTCTTGTGTTG  
AGTATATCTGTTAGGATAACTTCTTGAGTGAATTGTCAGGTCAAAGGGTTGTCATTTACTATTGATATATA  
TGTTAAATTGTCATAATATATGTCATAATTCCCTCAACATTGTTAAATGTCATTTCTTAAATTCTATT  
TTAATAACTGTTACTATTCTGCTTACAGTTGCCACTTCTTTAATCAACCAGATTAAATGATGAG  
ATTATAATAAGAATTACTATTAAATAAAATGGATTATA

21/105  
**FIGURE 20**

MLCPWRTANLGLLILITIFLVAEAEGAAQPNNSLMLQTSKENHALASSSLCMDEKQITQNYSKVLAEVNTSWPVK  
MATNAVLCCPPIALRNLIITWEIILRGQPSCTKAYRKETNETKETNCTDERITWVSRPDQNSDLQIRPVAITHD  
GYYRCIMVTPDGNFHRYGHQVLVTPELTLFQNRNRTAVCKAVAGKPAAQISWIPEGDCATKQEYWSNGTVTVKS  
TCHWEVHNVSTVTCHVSHLTGNKSLYIELLPVPGAKKSAKLYIPTYIILTIIILTIVGFIWLLKVNGCRKYKLNKT  
ESTPVVEEDEMQPYASYTEKNNPLYDTTNKVKASQALQSEVDTDLHTL

22/105  
**FIGURE 21**

AAGGCTGTGGACCCAGAGAAGGTGGCAGGTGGCCCCCTAGGAGAGCTCTGGGCACATTGAATCTCCAAAC  
TCCAATAATAAAATTGAAGACTTGGCAGAGAGTGTGTGTGTATGGTGTGGCGTAGGACAGGTT  
TCGGGGATGCGCGGTACCGCGTACCCCTCGAGGGCCCCACCCAGACGCCAGGCCCTCCCACTCCC  
CCTCAAGCAGCCCAGCCGGGACTTCCGTCGCGGGAAAGGGCGGGGACCCCTGAGCAGAAGGTGCGGAGCGG  
CCTGCCGGGTGGTCTGGCTTCCGTCGGCGCTGTACCCAGAGCTGAAGAGGAGCAGCGCGCC  
CGCGGACCCGGCAAGGCTGGCGGACTCGGGCTCCGAGGGACGCCATGGGGGAGGCAGGGCGCCCTTC  
TGGCTGTGGCCGCTGCCAAGCTGGCGCTGCTGCCCTGTGTGGTGGCTTTCCAGCGGACGCGTCCCCAGGGC  
AGCGCCGGGCCACTGCAGTGTACGGAGTGGACCCTGGCGACTTGAACCTGCTCGTGGAGCCCTTGGGAC  
CTGGGAGCCCCCTCGAGTTACACCTCAGAGCCAAGTACCGTTCCAACAAACCCAGACTGTGGCAGTGGCA  
GCCGGACGGAGCTGGGTGGCATTCTCGGGAACAGCTCACCATGTCTGACAAACTCCTGTCTGGGCACTAAG  
GCAGGCCAGCCTCTGGCCCCCGTCTCGTAACCTAGAAACCCAAATGAAGCCAACGCCCGGCTGGGC  
CCTGACGTGGACTTTCCGAGGATGACCCCTGGAGGCCACTGTCCATTGGGCCCCACCTACATGCCATCTCAT  
AAAGTTCTGATCTGCCAGTTCAACTACCGAAGATGTCAGGAGGGCCTGGACCCCTGCTGGAACGGAGCTGAAG  
ACCATACCCCTGACCCCTGTTGAGATCCAAGATTGGAGCTAGCCACTGGTACAAAGTGTATGGCGCTGCCGG  
ATGGAGAAAGAAGAGGATTGTGGGGCGAGTGGAGGCCATTITGTCTTCCAGACACGCCCTTGCTCCAAA  
GATGTGTGGGTATCAGGGAACCTCTGTGGGACGCCCTGGAGGAGAGGAACCTTGCTTCTATGGAAGGCCAGGG  
CCCTGTGTGCAGGTGAGCTACAAAGTCTGGTTCTGGGTTGGAGGTGAGCTGAGTCCAGAAGGAATTACCTGC  
TGCTGCTCCCTAATTCCAGTGGGGCGAGTGGCCAGGGTGTCCGCTGTCAACGCCACAAGCTGGAGCCTCTC  
ACCAACCTCTTGGCTGCTGGATTGCCCTTGCCCCCGTAGCGTGGCAGTCAGCAGCATCGCTGGAGC  
ACGGAGCTACTGGTGACCTGGCAACCGGGCCTGGGAACCACCTGGAGCATGTAGTGGACTGGCTCGAGATGGG  
GACCCCTGGAGAAACTCAACTGGTCCGGCTCCCCCTGGAACCTCAGTGTCTGTTACAGGGAAATTCACT  
GTCGGGGTCCCTATGAATCACTGTGACCGCAGTCTCTGCTTCAGGCTGGCCTCTGCATCCTCGTCTGGGG  
TTCAGGGAGGAATTAGCACCCCTAGTGGGCCACGCTTGGCAGCTCAAGATGCCCTCCAGGGACCCCGCC  
ATAGCGTGGGAGAGGTCCAAGGCACCAGCTCGAGGCCACCTCACCCACTACACCTGTGTGCACAGAGTGG  
ACCAGCCCCCTCGTCTGCATGAATGTGAGTGGCAACACACAGAGTGTACCCCTGCCCTGACCTCCGGTCCC  
TGTGAGCTGTGGGTGACAGCATCTACCATCGCTGGACAGGGCCCTCTGGTCCCCTCCGGTCTCATCTACCA  
GATAACACCCCTGAGGTGGAAAGTTCTGCCGGCATCTTATTCTGTGGGGCTTCTGTGGGGTGTGGCTG  
AGCCTGGCACCTCTGAAGGTGCTACCACTAAGGACAAAGTGTGCCCCGCTGGGCTGGAGAAAGTCC  
GATCCTGCCAACAGCAGTTCAGGCCAGCCCCACATGGAGCAAGTACCTGAGGCCAGGCCCTGGGACTTGGCC  
ATCCTGGAAAGTGGAGGGAGATGGAGCCCCGCCGGTTATGGAGTCCCTCCAGCCGCCAGGCCACCGCCCGCT  
GACTCTGGGTATGAGAAGCACTCCTGCCACACCTGAGGAGCTGGGCTTCTGGGGCCCCCAGGCCACAGGTT  
CTGGCCTGAACCACACGTCTGGCTGGGCTGCCAGCCAGGCTAGAGGGATGCTATGCAGGTTGCACCCAGTC  
CTGGATTAGCCCTCTGATGGATGAAGACACTGAGGACTCAGAGAGGCTGAGTCACCTACCTGAGGACACCCAGC  
CAGGCAGAGCTGGATTGAAGGACCCCTATAGAGAAGGGCTTGGCCCCATGGGAAGACACGGATGGAAGGTGG  
AGCAAAGGAAATACATGAAATTGAGAGTGGCAGCTGCCCTGCAAATCTGTTCCGCTGTAACAGAACTGAATT  
GGACCCCGACAGTGGCTCACGCCCTGTAATCCAGCACTTGGCAGGCCAGGTGGAAGGGATCACTAGAGCTA  
GGAGTTGAGACCAGGCTGGCAATATAGCAAGACCCCTACTACAAAAATAAAACATCAAAACAAAAACATT  
AGCTGGGATGATGGCACACACCTGTAGTCCGAGCAGCTTGGGAGGCTGAGGTGGAGGATGGTGAAGGCCAGG  
AGTCGAAGCTGCAGGGACCTCTGATTGCAACACTGCACTCCAGGCTGGTAACAGAAATGAGACCTTATCTCAA  
ATAAAACAAACTAATAAAAGCAAAAAAAAAAAAGAAAAGAAAAACACTGCAATTGGGACCATCTCAGC  
TCCCTTGATCCAGGTGCAGCATGGACTGAGTTCTGACAACAGAAATGTGGTCAGAAGTGCACATATGCCAACACG  
GGGTCTGGGTGGGGCTCCCCACATCCTTGCCTATGAGCTGGAACATAACACATGCCTATGATCCAGCT  
TTGGTCATACCAAGGGAGGTGGAGCAAGAAATGAAAGGAACCTGAATCCCTGAATGACTGCATGGATAGAA  
CCACTAAGAAAATAAACTTTATATTTTATA

23/105  
**FIGURE 22**

MRGGRGAPFWLWPLPKLALLPLLLWVLFQRTRPQGSAGPLQCYGVGPLGDLNCSWEPLGDLGAPSELHLQSQKYRS  
NKTQTVAVAAGRSWVAIPREQLTMSDKLLVWGTKAGQPLWPPVFNLETQMKPNAPRLGPDVDFSEDDPLEATVH  
WAPPTWPSHKVLICQFHYRRCQEAAWTLLEPELKTIPLTPEVIQDLELATGYKVYGRCRMEKEEDIWGEWSPILS  
FQTPPSAPKDWWVSGNLCGTPGGEPLLWKAPEGPCVQVSYKVWFVGGRELSPEGITCCCSLIPSGAEWARVSA  
VNATSWEPLTNLSLVCLDSASAPRSVAVSSIAGSTELLVTWQPGPGEPLEHVVWDWARDGDPLEKLNWVRLPPGNL  
SALLPGNFTVGVPYRITVTAVSASGLASASSVWGFREELAPLVGPTLWRLQDAPPGTPAIAWGEVPRHQLRGHLT  
HYTLCAQSGTSPSVCMNVSGNTQSVTLPDLPWGPCELVWTASTIAGQGPPGPILRLHLPDNTLRWKVLPGILFLW  
GLFLLGCGLSLATSGRCYHLRKVLPRVVWEKVPDPANSSSGQPHMEQVPEAQPLGDLPILEVEEMEPPVMESS  
QPAQATAPLDGYEKHFLPTPEELGLLGPPRPQVLA

24/105  
**FIGURE 23**

GGCACGAGGGCTGCCCTGGCGCTGCCGGCGGGCCATGGTGGTTGGATTGAGCCGGCCGCCGGGGCGCCG  
AGTCGGAGGGGGTGGCAGTGAGCGGCAGAGGC TACGGGGCTCGGTTGGCTGACTGGGAGTCGGCAGCGG  
CAGGAACCATGCGAGGCCAGCGGAGCCTGCTGGGCCCGCCCTGCCTCCGCCCTCTGCTGCTGG  
GTTACAGGCCGCGCTGTCACCTACTCCGGGTCTAGTACAGCGCTGGCCTACGGAAGGTCTGCCCTGCC  
CCCTGCTCTACAACCTCTGGGGCAGTGACACCGCTGTTGATGCTGCCCTTGAGCCTGTCTACTGGCTGGTAG  
ACAACGTGATCCGCTGGTTGGAGTGGGAGGAATGATATGCCACCGTCCATCTGTAAGAAGTGCATTAC  
CCAAGCCAGCCCACACACCAGCATCTGCAACAGGTGTGCTGAAGATGGATCACCACTGCCCTGGC  
TAAACAATTGTGTGGGCCACTATAACCATCGGTACTTCTCTTTCTGCTTTCATGACTCTGGCTGTGCT  
ACTGCAGCTATGGAAGTTGGACCTTCCGGGAGGCTATGCTGCCATTGAGACTTACCAAGACCCAC  
CCACCTTCTCCTTCGAGAAAGGATGACTCACAAGAGTCTTGTCTACCTCTGGTCTGTGAGTTCTGTGGC  
TTGCCCTGGGTGCCCTAAGTGTATGGCATGCTGTTCTCATCAGTCGAGGTGAGACTAGACATCGAAAGGCACATCA  
ACAAGAAGGAGAGACGTCGGCTACAGGCCAAGGGCAGAGTATTAGGAATCCTACAACACTACGGCTGCTGGACA  
ACTGGAAGGTATTCTGGGTGTGGATACAGGAAGGCACTGGCTACTCGGGTGTCTTACCTCTAGTCACTGC  
CCCATGGGAATGGAATGAGCTGGGAGCCCCCTCCCTGGGTGACTGCTCACTCAGCCTCTGTGATGGCAGTGTGAG  
CTGGACTGTGTCAGCCACGACTCGAGCACTCATTCTGCTCCCTATGTTATTCAAGGGCCTCCAAGGGCAGCTT  
TCTCAGAACCTTGATCAAAAAGAGCCAGTGGCCTGCCCTAGGGTACCATGCAGGACAATTCAAGGACCAGCT  
TTTACCAACTGCAGAACAGAACACAATGGGAGAAATCTTAGGACTGACATCCCTTACTCAGGCAAACAGAAC  
TTCCAACCCCAGACTAGGGTCAGGCAGCTAGCTACCTACCTGCCAGTGTGACCCGGACCTCCAGGATA  
CAGCACTGGAGTTGGCCACCACCTCTACTTGCTGTGAAAAAAACACCTGACTAGTACAGCTGAGATCTGG  
CTTCTCAACAGGGCAAAGATAACCAGGCCTGCTGAGGTACTGCCACTTCTCACATGCTGCTTAAGGGAGCAC  
AAATAAAAGGTATTGATTTAAAGATAAAAAAAAAAAAAAA

25/105  
**FIGURE 24**

MRGQRSLLLGPARLCLRLLLGYRRRCPPLRGLVQRWRYGKVCLRSLLYNSFGGSDTAVDAAFEPVYWLVDNV  
IRWFGVGRNDIATVSICKCIYPKKPARTHHCSICNRCVLMKMDHHCPWLNNCVGHYNHRYFFSFCFFMTLGCVYCS  
YGSWDLFREAYAAIETYHQTPPPTFSFRERMTHKSLVYLWFLCSSVALALGALTWHAVLISRGETSIERHINKK  
ERRRLQAKGRVFRNPYNYGCLDNWKVFLGVDIGRHWLTRVLLPSSHLPHGNNGMSWEPPPWTVAHSASVMAV

26/105  
**FIGURE 25**

ACGAGGGGAGCTCCGGTGCCTTCCCGCAGCGCTACCCGCC**ATG**CGCCTGCCGCCGGCCGCCTGGGGCT  
CCTGCCGCTCTGCTGCTGCCGCCGCCGGAGGCCGCAAGAAGCCGACGCCCTGCCACCGGTGCCGGGG  
GCTGGTGGACAAGTTAACCAAGGGATGGTGGACACCGCAAAGAAGAACTTGGCGGGAACACGGCTTGGGA  
GGAAAAGACGCTGTCCAAGTACGAGTCCAGCGAGATTGCCCTGCTGGAGATCCTGGAGGGCTGTGCGAGAGCAG  
CGACTTCGAATGCAATCAGATGCTAGAGGCCAGGAGGACACCTGGAGGCCCTGGTGGCTGCAGCTGAAGAGCGA  
ATATCCTGACTTATTGAGTGGTTTGTGAAGACACTGAAAGTGTGCTGCTCTCCAGGAACCTACGGTCCCGA  
CTGTCTCGCATGCCAGGGCGGATCCCAGAGGCCCTGCAGCGGGAAATGCCACTGCAGCGGAGATGGGAGCAGACA  
GGCGACGGGTCTGCCGGTGCCACATGGGGTACCAAGGGCCGCTGTGCACTGACTGCATGGACGGCTACTCAG  
CTCGCTCCGGAACGAGACCCACAGCATCTGCACAGCCTGTGACGAGTCCTGCAAGACGTGCTCGGCCTGACCAA  
CAGAGACTGCGCGAGTGTGAAGTGGGCTGGGTCTGGACGAGGGCCTGTGGATGTGGACGAGTGTGCGGC  
CGAGCCGCTCCCTGCAGCGCTGCCAGTTCTGTAAGAACGCAACGGCTCCTACACGTGCGAAGATGTGGACGA  
GTGCTCACTAGCAGAAAAAACCTGTGTGAGGAAAACGAAACTGCTACAATACTCCAGGGAGCTACGTCTGTGT  
GTGTCTGACGGCTTCGAAGAACGGAAGATGCCGTGTGCGCCGCCAGAGGCTGAAGCCACAGAAGGAGAAAG  
CCCGACACAGCTGCCCTCCCGCAAGACCTG**TAAT**GTGCCGGACTTACCCCTTAAATTATTACAGAAGGATGTCCC  
GTGGAAAATGTGCCCTGAGGATGCCGTCTCTGCAGTGGACAGCGGCCGGAGAGGCTGCCTGCTCTAACGG  
TTGATTCTCATTTGTCCCTAAACAGCTGCATTTCTGGTTGTTAAACAGACTTGTATATTTGATACAGTT  
CTTGTAATAAATTGACCATTGTAGGTAATCAGGAAAAAAAAAAAAAA

27/105  
**FIGURE 26**

MRLPRRAALGLPLLLLPPAPEAAKKPTPCHRRCRGLVDKFNQGMVDTAKKNFGGGNTAEEKTLSKYESSEIRL  
LEILEGLCESSDFECNQMLEAQEEHLEAWWLQLKSEYPDLFEWFCVKTLKVCCSPGTYGPDCLACQGGSQRPCSG  
NGHCSGDGSRQGDGSCRCHMGYQGPLCTDCMDGYFSSLRNEHTSICTACDESCKTCSGLTNRDCGECEVGWLDE  
GACVDVDECAAEPPCSAAQFCKNANGSYTCEDVDECSLAEKTCVRKNENCYNTPGSYVCVCPDGFETEDACVP  
PAEAEATEGESPTQLPSREDL

28/105  
**FIGURE 27**

GTGCAGTTGCGGCTCCAGGGCC **ATG**CGGGAGGAGCAGGGCAGGGAAACGGGACTCGGTTCCCAAGCCGTCGGTGCT  
GTTCCCTCCACCCAGACCTGGCGTGGCGGGCGCTGAGCGGCTGGTGTGGACGCGCGCTGGCGCTGCAGGCAGCG  
CGGGTGTAGCGTGAAGATCTGGACAGCGCACTACGACCCGGGCCACTGTTCGCCAGAGAGCCGAGCTACCGGT  
GCGCTGTGCCGGGGACTGGCTGCCCGAGGGCTGGGCTGGGGCGCCGCCGCGCCGCTGCGCCTACGTGCG  
CATGGTTTCTGGCGCTACGTGCTGTTCTGCCGACGAGGAGTTGACGTGGTAGTGTGCGACCAGGTGTC  
TGCCTGTATCCCAGTGGCTCAGGCTGGTAGACGGCGGAAGAAGATCCTATTTACTGTCACTTCCAGATCTGCT  
TCTCACCAAGAGAGATTCTTTCTAAACGACTATACAGGGCCCCAATTGACTGGATAGAGGAATACACCACAGG  
CATGGCAGACTGCATCTAGTCACAGCCAGTTACAGCTGCTGTTAAGGAAACATTCAAGTCCCTGTCTCA  
CATAGACCCCTGATGTCCCTATCCATCTAAATGTCACCAGCTTGACTCAGTTGTTCTGAAAGCTGGATGA  
CCTAGTCCCCAAGGGGAAAAAATCCTGCTGCTCCATCAACAGATACGAAAGGAAGAAAATGACTTTGGC  
ACTGGAAGCCCTAGTACAGCTGCGTGGAGATTGACATCCCAAGATTGGAGAGGGTTATCTGATCGTGGCAGG  
TGGTTATGACGAGAGAGTCCCTGGAGAATGTGGAACATTATCAGGAATTGAAGAAAATGGTCAACAGTCCGACCT  
TGGCCAGTATGTGACCTTCTTGAGGTCTTCAGACAAACAGAAAATCTCCCTCTCACAGCTGACGTGTC  
GCTTACACCCAAGCAATGAGCACTTGGCATTGTCCTCTGGAGCCATGTACATGCACTGCCCAGTCATTGC  
TGTTAATTGGGTGGACCCCTGGAGTCATTGACCACAGTGTACAGGGTTCTGTGAGCCTGACCGGTGCA  
CTTCTCAGAAGCAATAGAAAAGTCATCCGTGAACCTCCCTAAAGCCACCATGGGCCTGGCTGGAAGAGCCAG  
AGTGAAGGAAAATTTCCTGAAGCATTACAGAACAGCTACCGATATGTTACCAAATGCTGGTA**TATC**  
AGATTGTTTAAGATCTCCATTAATGTCATTTATGGATTGTAGACCCAGTTGAAACCAAAAAAGAAACCT  
AGAATCTAATGCAGAAGAGATCTTAAAAAAACTTGAGTCTGAATGTGAGCCACTTCCTATATACCACA  
CTCCCTGTCCACTTTCAGAAAAACCATGCTTTATGCTATAATCATTCAAATTGCCAGTGTAAAGTTAC  
AAATGTGGTGTCAATTCCATGTTCAAGCAGAGTATTTAATTATATTCTCGGGATTATTGCTCTGTCTATAA  
ATTTGAATGATACTGTGCCTTAATTGGTTTCATAGTTAAGTGTATCATTCAAAAGTTGATTAATTGGC  
TTCATAGTATAATGAGAGCAGGGCTATTGAGTTCCAGATTCAATCCACCGAAGTGTCACTGTCATCTGTTAG  
GGAATTGGTTGTCTGTCTTGCGCTGGATCCATAGCGAGAGTGTCTGTATTAAAGATAATTGTAT  
TTTGCACACTGAGATATAATAAGGTGTTATC

29/105  
**FIGURE 28**

MAEEQGRERDSVPKPSVLFLHPDLGVGGAERLVLDAAALALQARGCSVKIWTAAHYDPGHCFAESRELPVRCAGDWL  
PRGLGWGGRGAACVAYVRMVFLALYVLFLADEEFDVVVCDQVSACIPVFRLLARRKKILFYCHFPDLLLTKRDSF  
LKRLYRAPIDWIEEYTTGMADCILVNSQFTAASFKETFKSLSHIDPDVLYPSLNVTSDSVVPEKLDLVPKGKK  
FLLLSINRYERKKNLTLALEALVQLRGRLLSQDWERVHLIVAGGYDERVLENVEHYQELKKMVQQSDLGQYVTFL  
RSFSDKQKISLLHSCTVLYTPSNEHFGIVPLEAMYMQCPVIAVNSGGPLESIDHSVTGFLCEPDPVHFSEAIEK  
FIREPSLKATMGLAGRARVKEKF SPEAFTEQLYRYVTKLLV

30/105  
**FIGURE 29**

GTATTTATTGACTTTGCCAAGGCTTGGTCACAACAATCATATTACGTAATTTCCCCCTTGGTGGCAGAAC  
TGTAGCAATAGGGGGAGAAGACAAGCAGCGGATGAAGCGTTTCTCAGCTTGGATTGCTTCGACCTGACATC  
CGTTGTAACCGTTGCCACTTCTCAGATATTTATAAAAAAGTACCACTGAGTCAGTGAGGGCCACAGATTGG  
TATTAATGAGATAACGAGGGTTGCTGGGTGTTCTGAGCTAAGTGTCAAGACTGTAGTGGAGTTGCA  
GCTAACATGGGTTAGGTTAAACCGTGGGGGATGCAACCCCTTGCCTTGCATATGTAGGCCTACTGGCTTGTG  
TAGCTGGAGTAGTTGGGTGCTTGTGTTAGGAGGATCCAGATCATGTTGGCTACAGGGAGATGCTCTTTGAG  
AGGCTCCTGGGCATTGATTCAATCTCATTCTGGATATGTTCATTGAGTAAAGGAGGAGACCTCA  
TACGCTATTAAATGTCACTTTGCCTATGCCCGTTTGGTCATGTTCAATTAAATTGTGAGGAAGGCGC  
AGCTCCTCTGCACGTAGATCATTAAAGCTAATGTAAGCACATCTAAGGAATAACATGATTAAAGGTTG  
AAATGGCTTAAATGAGATCATTGGGTTGAGGGTGTGTTATTTGAGTCATGAATGTACAAGCTCTGTGAATCAGAC  
CAGCTAAATACCCACACCTTTTCTGGCTTACACAGTTATTTATTTGAGCTTACAGAGCTGGCTCATACCAAATAAGTT  
TTTGAGGAGGCCATGGCTTTCACACAGTTATTTATTTGAGCTTACAGAGCTGGCTCATACCAAATAAGTT  
TTGAGTGGCTGTACACTTGAGGCAACTAAAAGGCTCAACGTTGATCAGTTCTTCAGGAAACATTG  
TGCTCTAACAGTATGACTATTCTTCCCCACTCTAAACAGTGTGATGTGTTAT

31/105  
**FIGURE 30**

MCSLSKGGETLIRYLNVTFLPMPRFVLVMFQLI  
VRKAQLLSARRSFFKANVSTSKGIT

32/105  
**FIGURE 31**

AAAAGCGAGTGAAGAGAGCGCGACGGCGGCGGGCGCAGCTATTGCTGGACGCCAGTGGAGAGCGA  
GGCCTGAGCCTCTCGCTAGGATCAAATGGTTCAATCCCAGAATACTATGAAGGCAAGAACGTCCCTCAC  
AGGAGCTACCGTTCTAGGAAGGTGCTCTGGAAAAGTTGCTGAGGTCTGTGCTTAAGGTGAATTCAAGTATA  
TGTGTTGGTGAAGGAGAAAGCTGGACAGACACCACAAGAGCGAGTGGAAAGAAGTCTTAGTGGCAAGCTTTGA  
CAGATTGAGAGATGAAAATCCAGATTTAGAGAGAAAATTATAGCAATCAACAGCGAAGTCAACCCAACTAACT  
GGCTCTCAGTGAAGAAGATAAAGAGGTGATCATAGATTCTACCAATATTATATTCCACTGTGCAGCTACAGTAAG  
GTTTAATGAAAATTAAAGGGATGCTGTTCAAGTTAAATGTGATTGCAACGCGACAGCTTATTCTCTTGACAAACA  
AATGAAGAATCTGGAAGTGTTCATGCATGTCAACAGCATATGCCTACTGTAATCGCAAGCATATTGATGAAGT  
AGTCTATCCACCACCTGTGGATCCAAGAAGCTGATTGATTCTTAGAGTGGATGGATGGCTAGTAAATGA  
TATCAGGCCAAATTGATAGGAGACAGACCTAATACATACATACACAAAAGCATTGGCAGAATATGTTGTACA  
ACAAGAAGGAGCAAACATAATGTGGCAATTGTAAGGCCATCGATTGTTGGTGCCAGTTGGAAAGAACCTTTCC  
AGGATGGATTGATAACTTAATGGACCAAGTGGTCCTTTATTGCGGGCAGGGAAAGGAATTCTCGAACAAATACG  
TGCCTCCAACAATGCCCTGCAGATCTGTTCTGTAGATGTAGTTGCAACATGAGTCTGCGGCAGCCTGGTA  
TTCCGGAGTTAATAGACCAAGAACATCATGGTGTATAATTGTACAACAGGCAGCACTAATCCTTCCACTGGGG  
TGAAGTTGGTATGATTACCTGTGTTTGAATGTTAGAATAATCTAAAGAACCAAAAAAAAAAAAAAAA  
AAAAAA

33/105  
**FIGURE 32**

MVSIPEYYEGKNVLLTGTGFLGVILLEKLLRSCP KVNSVYVLVRQKAGQTPQERVEEVLSGKLFDRRLRDENPDF  
REKIIAINSELTQPKLALSEEDKEVIIIDSTNI IFHCAATVRFNENLRDAVQLNVIATRQLILLAQQMKNLEVFMH  
VSTAYAYCNRKHIDEVVYPPPVDPKKLIDSLEWMD DGLVNDITPKLIGDRPNTYIYTKALAELYVQQEGAKLNVA  
IVRPSIVGASWKEPFPGWIDNFNGPSGLFIAAGKGILRTIRASNALADLVPVDVVVNMSLAAAWYSGVNRPRNI  
MVYNCTTGSTNPFWGEVGMILPVFLNVRINLKEPKKKKK

34/105  
**FIGURE 33**

GAACGAGGGTCTAGCTGCCGCCACCGAACAGCCTGTCCTGGTCCCCGGCTCCCTGCCCGCGCCAGTCATG

ACCCCTGCGCCCCCTCACTCCTCCCGCTCCATCTGCTGCTGCTGCTGCTCAGTGCAGGGGTGTGCCGGGCTGAG  
GCTGGGCTCGAAACCGAAAGTCCCGTCCGGACCCCTCAAGTGGAGACCGCTTCACATACACTACACGGGAAGCTTGGTAGATGGACGTATTATTGACACC  
GAGCCCCGCTGCTTTGGAGACACGCTTCACATACACTACACGGGAAGCTTGGTAGATGGACGTATTATTGACACC  
TCCCTGACCAGAGACCCTCTGGTTATAGAACTTGGCAAAAGCAGGTGATTCCAGGTCTGGAGCAGAGTCTTCTC  
GACATGTGTGGAGAGAAGCGAAGGGCAATCATCCTTCTCACTTGGCCTATGGAAAACGGGGATTCCACCA  
TCTGTCCCAGCGGATGCAGTGGTGCAGTATGACGTGGAGCTGATTGCACTAATCGAGCCAACACTGGCTAAAG  
CTGGTGAAGGGCATTTGCCTCTGGTAGGGATGGCCATGGTGCCAGCCCTCCTGGCCTCATTGGGTATCACCTA  
TACAGAAAGCCAATAGACCCAAAGTCTCCAAAAGAAGCTCAAGGAAGAGAAACAGAGCAAAAGAAA  
TAATAAATAATAAATTTAAAAACTTAAAAAAAAAAAAAAAAAAAAAAAAAAA

35/105  
**FIGURE 34**

MTLRPSLLPLHLLLLLSSAAVCRAEAGLETESPVRTLQVETLVEPPEPCAEPAAFGDTLHIHYTGSLVDGRIID  
TSLTRDPLVIELGQKQVIPGLEQSLLDMCVGEKRRAIIPSHLAYKRGFPPSVPADAVVQYDVELIALIRANYWL  
KLVKGILPLVGMAMVPALLGLIGYHLYRKANRPKVSKKKLKEEKRNKSKKK

36/105  
**FIGURE 35**

CCTATTCTACGGCTGACCCCTGGTGGCACGTGGACTGTCGCCACGCAAGTCTGGTCCTCGGCATTGACC  
GGGGCCTTGTGTTGGGAGCCTCTCTTAAGCTGCCCTGTCGCCAGAGAGTTGGAGGGCGGGTTGGGTGCG  
GTGTCTGATTGGGGCTCGCACCGCAGCACGCTGGAGTCCCGCTAGGTACCAAGTTAGCGTCAGGGAGCTGGTC  
AGCGGTGCCGGGACACCCGTGTGGCAGGCCAGCGCTCTGGAGAATCCCAGCAGCCCTGCTCCCTG  
CAGCCAGGTGTAGTTCGGGAGCCACTGGGCCAAAGTGAGAGTCCAGCGGTCTTCCAGCGCTTGGCCACGGCG  
GCGGCCCTGGGAGCAGAGGTGGAGCGACCCATTACGCTAAAGATTGAAGGCTGGGTTGGCTGGCCCTGCTTCT  
GGGGGCCCTGCTGGGAACGCCCTGGCTGGAGGAGCCAGGATCTCCACTGTGGAGCATGCAGGGCTCTGGTGA  
TGAACATAGAATGGGAAATTGCCAGGTGGACCCCAAGAAGACCATTAGATGGGATCTTCCGGATCAATCCAGA  
TGGCAGCCAGTCAGTGGTGGAGGTGCCTATGCCGCTCAGAGGCCACCTCACAGAGCTGCTGGAGGAGATATG  
TGACCGGATGAAGGAGTATGGGAACAGATTGATCCTCCACCCATCGCAAGAACTACGTACGTAGTGGCCG  
GAATGGAGAATCCAGTGAACCTGGACCTACAAGGCATCCGAATCGACTCAGATATTAGCGGCACCCCTCAAGTTGC  
GTGTGAGAGCATTGTGGAGGAATACGAGGATGAACTCATTGAATTCTTCCGAGAGGCTGACAATGTTAAAGA  
CAAACTTGCAGTAAGCAACAGATCTTGACCATGCCCTGCACATATCGCATGAGCTTGAACCACTGG  
AGCAGCCACACTGGCTTGATGGATCACCCCCAGGAGGGAAAATGGTGGCAATGCCTTTATATATTATGTTT  
TACTGAAATTAACTGAAAAATATGAAACC

37/105  
**FIGURE 36**

MKGWGLALLGALLGTAWARRSQDLHCGACRALVDELEWEIAQVDPKKTIQMGSFRINPDGSQSVVEVPYARSE  
AHLTELEEICDRMKEYGEQIDPSTHRKNYVRVGRNGESELDLQGIRIDS DISGTLKFA CESIVEEYEDELIE  
FFSREADNVKDKLCSKRTDLCDHALHISHDEL

38/105  
**FIGURE 37**

GAGGGTTGAAGGACCCAGGC GTGTAGCCCTGCTCCAGAGACCTGGGCATGGAGGAGAGTGTCGTACGGCCCTCA  
GTGTTGTGGATGGACAGACCGACATCCCATT CACGAGGCTGGGACGAAGCCACCGGAGACAGTCGTGCAGT  
GTGGCCCGGGTGGGTCTGGGTCTTGCTGTTGCTGATGGGGCTGGGCTGCCGTCAAGGCTGGTCCCTCCTG  
CAGCTGCACTGGCGTCTAGGAGAGATGGTCACCCGCCCTGACGGACCTGCAGGCTCTGGGAGCAGCTGATA  
CAAGAGCGAAGGTCTCACGAGGTCAACCCAGCAGCGCATCTCACAGGGCCAACCTCCAGCTTGACCGGCAGCGG  
GGGCGCTGTTATGGGAGACTCAGCTGGGCTGGCTTCTGAGGGCCTCAGCTACCACGATGGGCCCTGTG  
GTCACCAAAGCTGGCTACTACTACATCTACTCCAAGGTGCAGCTGGCGGTGTGGCTGCCGCTGGCCTGGC  
AGCACCATCACCCACGGCCTCTACAAGCGCACACCCCGTACCCGAGGAGCTGGAGCTGTTGGTACACCTG  
TCACCCCTGCGGACGGGCCACCAGCAGCTCCGGGTCTGGTGGGACAGCAGCTCCTGGTGGTACACCTG  
GAGGCTGGGGAGGAGGTGGTGTGCTGGATGAAACGCCTGGTTCGACTCGTGATGGTACCCGGTCTTAC  
TTCGGGGCTTCATGGTGTTGAAGGAAGGAGCGTGGTGCATTGGACATGGGTCTGACACGTGGAGAACTCAGAGGG  
TGCCTCAGGGAAAGAAAACCTCAGAAGCAGGGCTGGCGTGGCTCTGCCTGTAATCCCAGCACTTGGG  
AGGCCAAGGCAGGC GGATCACCTGAGGT CAGGAGTT CGAGACCA GCCTGGCTAACATGGCAAACCCATCTCTA  
CTAAAAATACAAAAATTAGCCGGACGTGGTGGTGCCTGTAATCCAGCTACTCAGGAGGCTGAGGCAGGATA  
ATTTGCTTAAACCCGGGAGGC GGAGGTTGCAGTGAGCCGAGATCACACCACTGCACTCCAACCTGGAAACGCA  
GTGAGACTGTGCCTCAAAAAAAAAAAAAAAA

39/105  
**FIGURE 38**

MEESVVRPSVFVVDGQTDIPFTRLGRSHRRQCSVARVGLGLLLLMGAGLAVQGWFLLQLHWRLGEMVTRLPDG  
PAGSWEQLIQERRSHEVNPAAHLTGANSSLTGSGGPLLWETQLGLAFLRGLSYHDGALVVTKAGYYYIYSKVQLG  
GVGCPLGLASTITHGLYKRTPRYPEELELLVSQQSPCGRATSSSRVWWDSFLGGVVHLEAGEEVVVRVLDRLV  
RLRDGTRSYFGAFMV

40/105  
**FIGURE 39**

GGCACGAGGGGAGTGGAAAGTTCTCCGGCAGCCCTGAGATCTCAAGAGTGACATTGTGAGACCAAGCTAATTGA  
TTAAAATTCTCTTGGAAATCAGCTTGCTAGTATCATACCTGTGCCAGATTCATCATGGGAAACAGCTGTTACAA  
CATAGTAGCCACTCTGTTGCTGGCCTCAACTTGAGAGGACAAGATCATTGCAGGATCCTGTAGTAACGCC  
AGCTGGTACATTCTGTGATAATAACAGGAATCAGATTGCAGTCCTGTCCAAATAGTTCTCCAGCGCAGG  
TGGACAAAGGACCTGTGACATATGCAGGCAGTGAAAGGTGTTTCAGGACCAGGAAGGAGTGTCCCTCCACAG  
CAATGCAGAGTGTGACTGCACTCCAGGGTTCACTGCCTGGGGCAGGATGCAGCATGTTGAAACAGGATTGAA  
ACAAGGTCAAGAACTGACAAAAAAAGGTTGAAAGACTGTTGTTGGACATTAAACGATCAGAAACGTGGCAT  
CTGTCGACCCCTGGACAAACTGTTCTTGGATGGAAAGTCTGTGCTTGTGAATGGGACGAAGGAGAGGGACGTGGT  
CTGTGGACCACATCTCCAGCCGACCTCTCTCCGGGAGCATCCTCTGTGACCCCGCCTGCCCTGCGAGAGAGCCAGG  
ACACTCTCCGCAGATCATCTCCTCTTCTGCGCTGACGTCAGTCGTTGCTCTCCTGCTGTTCTCCTCAC  
GCTCCGTTCTCTGTTAAACGGGGCAGAAAGAAACTCCTGTATATATTCAAACAAACCATTATGAGACCAGT  
ACAAACTACTCAAGAGGAAGATGGCTGTAGCTGCCGATTCCAGAAGAAGAAGAAGGAGGATGTGAACGTTGAAA  
TGGAAAGTCAATAGGGCTGTTGGACTTTCTGAAAAGAAGCAAGGAAATATGAGTCATCCGCTATCACAGCTTC  
AAAAGCAAGAACACCATCCTACATAATACCCAGGATTCACACACACGTCTTTCTAAATGCCAATGAGTTG  
GCCTTAAAAATGCACCACTTTTTTTTGACAGGGTCTCACTCTGTCACTGTCACCCAGGCTGGAGTGCAGTGGC  
ACCACCATGGCTCTGCAAGCCTGACCTCTGGGAGCTCAAGTGATCCTCTGCCTCAGTCTCTGAGTAGCTGG  
AACTACAAGGAAGGGCCACCACACCTGACTAACCTTTGTTTTGTTGGTAAAGATGGCATTACCATGTT  
GTACAGGCTGGCTCAAACTCCTAGGTCACCTTGGCCTCCAAAGTGCTGGGATTACAGACATGAACTGCCAGG  
CCCGGCCAAAATAATGCACCACTTTAACAGAACAGACAGATGAGGACAGAGCTGGTATAAAAAAAAAAAAA  
AAAGCATTCTAGATACCACTAACAGGTTGAGCTAGTTTTGAAATCCAAAGAAAATTATAGTTAAATT  
CAATTACATAGTCCAGTGGTCCAACATAATTATAATCAAATCAATGCAGTTGTTGGTGTCAATATGA  
CATATGACAATAAGCCACGAGGTGCAGTAAGTACCCGACTAAAGTTCCGTGGTTCTGTATGTAACACGACAT  
GCTCCACCGTCAGGGGGAGTATGAGCAGAGTGCCTGAGTTAGGGTCAAGGACAAAAACCTCAGGCCTGGAGG  
AAGTTTGGAAAGAGTCAAGTGTCTGATATCCTATGGCTTCCATCCTCACACCTCTGCCTTGTCCIGC  
TCCCTTTAACGCCAGGTTACATTCTAAACCTAACATAATTTATACCAAGCCAATAATGA  
ACTGCAATGAA

41/105  
**FIGURE 40**

MGNSCYNIVATLLLVLNFERTRSIQLDPCSNCPAGTFCDNNRNQICSPCPPNSFSSAGGQRTCDICRQCKGVFRTR  
KECSSTSNAECDCPTGFHCLGAGCSMCEQDCKQGQELTKKGCKDCCFGTFNDQKRGICRPWTNCSDLGKSVLVNG  
TKERDVVCGPSPADLSPGASSVTPPAPAREPGHSPQIISFFLALTSTALLFLFFLTLRFSVVKRGRKKLLYIFK  
QPFMRPVQTTQEEEDGCSCRFPEEEEGGCEL

42/105  
**FIGURE 41**

GGAAACAAAGCTGGAGCTCCACCGCGGTGGCGGCCGCTCTAGAACTAGTGGATCCCCCGGGCTGCAGGAATTCGG  
CACGAGCAGAAGAGGGGGCTAGCTAGCTGTCTCTGCAGGACAGGGGAGACCCCGCGCCCCCCCCTGGTGTGAGGCGG  
CCTCACAGGGCGGGTGGCTGGCGAGCCGACGCAGGGGGAGAGGCTGTGAGGAGTGTGTGAAACAGGACCC  
GGGACAGAGGAACCATGGCTCCCGAGAACCTGAGCACCTTTGCCTGTGCTATAACCTCATCGGGCGGTGA  
TTGCCGGACGAGATTCTATAAGATCTGGGGGTGCCTCGAAGTGCCTCTATAAAGGATATTAAAAGGCCTATA  
GGAAACTAGCCCTGCAGCTTCATCCGACCGGAACCTGATGATCCACAAGGCCAGGAGAAATTCCAGGATCTGG  
GTGCTGTTATGAGGTCTGTCAAGATAGTGAGAACCGAAACAGTACGATACTTATGGTGAAGAAGGATTAAAAG  
ATGGTCATCAGAGCTCCATGGAGACATTTCACACTCTTGGGGATTTGGTTATGTTGGAGGAACCC  
CTCGTCAGCAAGACAGAAATATCCAAGAGGAAGTGATATTATTGTAGATCTAGAAGTCACTTGGAAGAAGTAT  
ATGCAGGAAATTGTGGAAGTAGTTAGAAACAAACCTGTGGCAAGGCAGGCTCTGGAAACGGAAGTGCAATT  
GTCGGCAAGAGATGCGGACCAACCGCTGGCCCTGGCGCTTCAAATGACCCAGGAGGTGGTCTGCACGAAT  
GCCCTAATGTCAAACTAGTGAATGAAGAACGACGCTGGAAAGTAGAAATAGAGCCTGGGTGAGAGACGGCATGG  
AGTACCCCTTATTGGAGAAGGTGAGCCTCACGTGGATGGGAGCCTGGAGATTACGGTCCGAATCAAAGTTG  
TCAAGCACCAATATTGAAAGGAGAGGAGATGATTGTACACAAATGTGACAATCTCATTAGTTGAGTCACGG  
TTGGCTTGAGATGGATATTACTCACTGGATGGTACAAGGTACATATTTCCGGGATAAGATCACCAGGCCAG  
GAGCGAAGCTATGAAAGAAAGGGGAAGGGCTCCCCAACCTTGACAACAACAAATATCAAGGGCTTTGATAATCA  
CTTTGATGTGGATTTCAAAAAGAACAGTTAACAGAGGAAGCGAGAGAACGGTATCAAACAGCTACTGAAACAAAG  
GGTCAGTGCAGAAGGTATACAATGGACTGCAAGGATATTGAGAGTGAATAAAATTGGACTTGTAAATAAGT  
GAATAAGCGATATTATCTGCAAGGTTTTGTGTGTTTGTGTTTATTTCAATATGCAAGTTAGGC  
TTAATTTTATCTAATGATCATGAAATGAATAAGAGGGCTTAAGAATTGTCCATTGCATTGGAAAAG  
AATGACCAGCAAAAGGTTACTAATACGTCCCTTGGGGATTAAATGTCGGTGTGCGCCTGAGTTCAAG  
AATTAAAGCTGCAAGAGGACTCCAGGAGCAAAAGAACACAATATAGAGGGTTGGAGTTAGCAATTTCATTC  
AAAATGCCACTGGAGAAGTCTGTTAAATACATTGTGTTATTTT

43/105  
**FIGURE 42**

MAPQNLSTFCLLLLYLIGAVIAGRDFYKILGVPRSASIKDIKKAYRKLALQLHPDRNPDDPQAQEKFQDLGAAYE  
VLSDSEKRKQYDTYGEEGLKDGHQSSHGDIFSHFFGDFGFMFGGTPRQQDRNIPRGSDIIVDLEVYAGNF  
VEVVRNKPVARQAPGKRKCNCRQEMRTTQLGPGRFQMTQEVCDECPNVKLVNEERTLEVEIEPGVRDGMEYPFI  
GEGEPHVDGEPGDLRFRRIKVVKHPIFERRGDDLYTNVTISLVESLVGFEDEMITHLDGHKVHISRDKITRPGAKLW  
KKGEGLPNFDNNNIKGSLLIITFDVDFPKEQLTEEARREGIKQLLKQGSVQKVYNGLQGY

44/105  
**FIGURE 43**

GACAGTGGAGGGCAGTGGAGAGGACCGCGCTGTCCGTGTCACCAAGAGCTGGAGACACCATCTCCCACCGAGA  
**GTCATGGCCCCATTGGCCCTGCACCTCCTCGTCCCCATCCTCCTCAGCCTGGTGGCCTCCAGGACTGG**  
AAGGCTGAACGCAGCCAAGACCCCTCGAGAAATGCATGCAGGATCCTGACTATGAGCAGCTGCTCAAGGTGGT  
ACCTGGGGCTCAATCGGACCCCTGAAGCCCCAGAGGGTGATTGGTGGCCTGGTGGCCGGCTGGTGGCC  
GCCAAGGTGCTCAGCGATGCTGGACACAAGGTCACCATCCTGGAGGCAGATAACAGGATCGGGGCCGATCTC  
ACCTACCGGGGACAGAACACGGGCTGGATTGGGAGCTGGAGCCATGCGCATGCCAGCTCACAGGATCCTC  
ACAAGCTCTGCCAGGGCCTGGGCTCAACCTGACCAAGTTCACCCAGTACGACAAGAACACGTGGACGGAGGTG  
CACGAAGTGAAGCTGCGCAACTATGTTGAGAAGGTGCCGAGAACAGTGGCTACGCCCTGCGTCCCCAGGAA  
AAGGGCCACTCGCCCGAAGACATCTACAGATGGCTCTCAACCAGGCCCTAAAGACCTCAAGGCAGTGGCTGC  
AGAAAGGCAGTGAAGAAGTTGAAAGGCACACGCTCTTGAATATCTCTGGGAGGGAACCTGAGCCGGCC  
GCCGTGCAGCTCTGGGAGACGTGATGTCGAGGATGGCTTCTTCTATCTCAGCTCGCCGAGGCCCTCCGGGCC  
CACAGCTGCCTCAGCGACAGACTCCAGTACAGCCGATCGTGGTGGCTGGACCTGCTGCCGCGCGCTGCTG  
AGCTCGCTGCCGGCTTGTGCTGTTGAACCGCCCGTGGTGGCATGACCCAGGGACCGACGATGTGACGTG  
CAGATCGAGACCTCTCCCCCGCGGAATCTGAAGGTGCTGAAGGCCGACGTGGTGTGCTGACGGCGAGCGGA  
CCGGCGGTGAAGCGCATCACCTCTGCCGCCGCTGCCCGCCACATGCGAGGCCGCTGCCGAGGCTGCACTAC  
GTGCCGCCACCAAGGTGTTCTAAAGCTTCCGAGGCCCTCTGGCGCGAGGGCGCTGCTGCTGGCCTCGTACAC  
AACACCGATGCCCGCGCATGATTTCTACCCGCCGCCGCGAGGGCGCTGCTGCTGGCCTCGTACCGTGGCG  
TGGTCGGACGCCGGCAGCGTTCGCCGGCTTGAGCCGGAAAGAGGCCTGCTGCTGCCGCTGACGACGTGGCG  
GCATTGCACGGCCTGCGCCAGCTCTGGACGGCACCGCGCTGCTCAAGCCTGGCGAGGACAGCAC  
AGCCAGGGTGGCTTGTGGTACAGCCGCCGCGCTCTGGCAAACCGAAAAGGATGACTGGACGGTCCCTATGGC  
CGCATCTACTTGCCTGGCGAGCACCCGCCACGGCTGGGTGGAGACGGCGGTCAAGTCCGCGCTGCG  
GCCGCATCAAGATCAACAGCCGAAGGGGCCGATCGGACACGGCCAGCCCCGAGGGCACGCATCTGACATG  
GAGGGCGAGGGCATGTGATGGGTGGCCAGCAGCCCTCGCATGACCTGGCAAAGGAAGAAGGCAGCCACCC  
CCAGTCCAAGGCCAGTTATCTCCAAAACACGACCCACACGAGGACCTCGCAT**TAA**AGTATTTCGGAAAAAA

45/105  
**FIGURE 44**

MAPLALHLLVLVPILLSLVASQDWKAERSQDPFEKCMQDPDYEQLLKVVTLKPQRVIVVGAGVAGLVAA  
KVLSDAGHKV TILEADNRIGGRIFTYRDQNTGWIGELGAMRMPSSHRIILHKLCQGLGLNLTKFTQYDKNTWTEVH  
EVKLRNYVVEKVPEKLGYALRPQEKGHSPEDIYQMALNQALKDLKALGCRKAMKKFERHTLLEYLLGEGNLSRPA  
VQLLGDVMSEDGFFYLSFAEALRAHSCLSDRLQYSRIVGGWDLLPRALLSSLSGLVLLNAPVVAMTQGPHDVHQ  
IETSPPARNLKVLKADVVLLTASGPAVKRITFSPPLPRHMQEALRRLHYVPATKVFLSFRRPFWREEHIEGGHSN  
TDRPSRMIFYPPPREGALLLASYTWSDAAAFAGLSREEALRLALDDVAALHGPVVRQLWDGTGVVKRWAEDQHS  
QGGFVVQPPALWQTEKDDWTVPYGRIYFAGEHTAYPHGWVETAVKSALRAAIKINSRKGPASDTASPEGHASDME  
GQGHVHGVASSPSHDLAKEEGSHPPVQGQLSLQNTTHTRTSH

46/105  
**FIGURE 45**

TGCACAAGCAGAACAGGTTCTCCTCCCCAGTCACCAGTTGCTGAGTTAGAATTGTCTGCAATGG  
CCGCCCTGCAGAAATCTGTGAGCTCTTCCTTATGGGGACCCCTGCCACCAGCTGCCTCCTCTGGCCCTCT  
TGGTACAGGGAGGAGCAGCTGCCCATCAGCTCCACTGCAGGCTTGACAAGTCCAACCTCCAGCAGCCCTATA  
TCACCAACCGCACCTCATGCTGGCTAAGGAGGCTAGCTTGGCTGATAACAACACAGACGTTCTCATTGGGG  
AGAAAATGTTCCACGGAGTCAGTATGAGTGAGCGCTGCTATCTGATGAAGCAGGTGCTGAACCTCACCCCTGAAG  
AAGTGCTGTCCCTCAATCTGATAGGTTCCAGCCTTATATGCAGGAGGTGGGCCCTCCTGGCCAGGCTCAGCA  
ACAGGCTAACGACATGTCATATTGAAGGTGATGACCTGCATATCCAGAGGAATGTGCAAAAGCTGAAGGACACAG  
TGAAAAAGCTTGGAGAGAGTGAGGAGATCAAAGCAATTGGAGAACTGGATTGCTGTTATGTCTGAGAAATG  
CCTGCATTTGACCAGAGCAAAGCTGAAAATGAATAACTAACCCCCCTTCCCTGCTAGAAATAACAATTAGATGC  
CCCAAAGCGATTTT

<sup>47/105</sup>  
**FIGURE 46**

MAALQKSVSSFLMGTLATSCLLLALLVQGGAAAPISSHCRLDKSNFQQPYITNRTFMLAKEASLADNNNTDVRLI  
GEKLFHGVSMSERCYLMKQVLNFTLEEVLFQSDRFQPYMQEVVPFALARLSNRLSTCHIEGDDLHIQRNVQKLKD  
TVKKLGESGEIKAIGELDLLFMSLRNACI

48/105  
**FIGURE 47**

GAGGGGTAGAGATGCAGAAAGGCAGAAAGGAGAAAATTCAAGGATAACTCTCTGAGGGGTGAGCCAAGCCCTGCC  
ATGTAGTGCACGCAGGACATCAACAAACACAGATAACAGGAAATGATCCATTCCCTGTGGTCACTTATTCTAAAG  
GCCCAACCTCAAAGTTCAAGTAGTGATATGGATGACTCCACAGAAAGGGAGCAGTCACGCCCTACTTCTTGCC  
TTAAGAAAAGAGAAGAAAATGAAACTGAAGGAGTGTGTTCCATCCTCCCACCGAAGGAAAGGCCCTCTGTCCGAT  
CCTCCAAAGACGGAAAGCTGCTGGCTGCAACCTTGCTGCTGGCAGTGCTGTCTGCTGCCACGGTGGTGTCTT  
TCTACCAGGTGGCCGCCCTGCAAGGGGACCTGGCCAGCCTCCGGCAGAGCTGCAGGGCACCACGGAGAAGC  
TGCCAGCAGGAGCAGGAGCCCCAAGGCCGCTGGAGGAAGCTCCAGCTGTACCGCGGGACTGAAAATTTG  
AACCACAGCTCCAGGAGAAGGAACTCCAGTCAGAACAGCAGAAATAAGCGTGCGTTCAGGGTCCAGAAGAAA  
CAGTCACTCAAGACTGCTTGCAACTGATTGAGACAGTGAACACCAACTATACAAAAGGATCTACACATTG  
TTCCATGGCTCTCAGCTTAAAAGGGGAAGTGCCCTAGAAGAAAAGAGAATAAAATATTGGTCAAAGAAACTG  
GTTACTTTTTATATATGGTCAGGTTTATATACTGATAAGACCTACGCCATGGGACATCTAATTAGAGGAAGA  
AGGTCCATGTCTTGGGATGAATTGAGTCGGTACATTGTTGATGTATTCAAATATGCCTGAAACACTAC  
CCAATAATTCTGCTATTCACTGGCATTGCAAAACTGGAAGAAGGAGATGAACCTCAACTTGCAATACCAAGAG  
AAAATGCACAAATATCACTGGATGGAGATGTCACATTGGTGCATTGAAACTGCTGTGAACTTACACCCA  
TGTCTGTAGCTATTCCTCCCTTCTGTACCTCTAAGAAGAAGAATCTAACAGAAAAAAAAAAAAAAA  
AAAAAAAAAAA

49/105  
**FIGURE 48**

MDDSTEREQSRLTSCLKKREEMKLKECVSILPRKESPSVRSSKGKLLAATLLLALLSCCLTVVSFYQVAALQGD  
LASLRAELQGHHAEKILPAGAGAPKAGLEEAPAVTAGLKIFEPPAPGEGNSSQNSRNKRAVQGPEETVTQDCLQLI  
ADSETPTIQKGSYTFVPWLLSFKRGSALEEKENKILVKETGYFFIYGQVLYTDKTYAMGHLIQRKKVHVGDELS  
LVTLFRCIQNMPETLPNNSCYSAGIAKLEEGDELQLAIPRENAQISLDGDTFFGALKLL

50/105  
**FIGURE 49**

51/105  
**FIGURE 50**

MYRLLSAVTARAAPGLASSCGRGVHQAGLPPGHGVGGGLGLALGVLAGGLSGAAPAQSPAAPDPE  
ASPLAEPPEQSLAPWSPQTPAPPSCSRCFARAIESSRDLLHRIKDEVGAPGIVVGVSVDGKEVWSEGLGYADVEN  
RVPCKPETVMRIASISKSLSLTVALAKLWEAGKLDLDIPVQHYVPEFPEKEYEGEKVSVTRLLISHLSGIRHYEK  
DIKKVKEEKAYKALKMMKENVAFEQEKEGKSNEKNDFTKFKTEQENEAKCRNSKPGKKNDFEQGELYLREKfen  
SIESLRLFKNDPLFFKPGSQFLYSTFGYTLAAIVERASGCKYLDYMQKIFHLDMLTTVQEENEPVIYNRAR

52/105  
**FIGURE 51**

GGAGCCCAGATTCTGGAAGAGCCCTAGAGCTTGCCTTTCTCCTGCAGCACTAACGAAACCAGTTT  
GCAATCAATTCTGTCAAAGGCSACCCTACTCTCCATCCGTCTTCTCCAGCCCAGACACTCACAGCCCCCT  
GCCAGACCAGGGGACCTCGGAGAGGAAGGACAGAGGTCAGGATCTCCTCCCTCGGGACCCAAGGSCACAA  
AGGAGAGCTCCGTGGAGAGAAGAAAATCATTTGACTCTGGGGACACAGATTGCTGCCACAGAGGCTGATGGAC  
AACCAAGGGCGAGAGAGAAAGTGAAGGCTGGTGTGGTTGCAAAGGGATGAGGATGACGCTCTCTGTGTGAAGAC  
GTGGAGCTACAAGACGGAGATCTGCCCCGAAGAAAAAAATTTTGAGAGAATTCCCAGATTGAAAGAAGAT  
CTGAAAGGGAACATTGACAAGCTCCGTGCCCTCGCAGACGATATTGACAAAACCCACAAGAAATTACCAAGGCT  
AACATGGTGGCACCTACTGCTGTCATCTCTGGAGTGATGAGCCTCTGGGTTAGCCCTGCCCAAGCAACA  
GGAGGAGGAAGCCTGCTGCTCCACCGCTGGTCAAGGTTGGCAACAGCAGCTGGGTCAACAGCATCGTGA  
GGTACGTTGGAACGCTCCAAAATAAGAACGCCCAGCACGGCGGAAGACATACTGCCACCTACGACCAAGAG  
GACAGGGAGGATGAGGAAGAGAACGGCAGACTATGTCACAGCTGGAAAGATTATCTATAATCTTAGAACACC  
TTGAAGTATGCCAAGAAAACGTCGTGCACTTGGAAACTCAGAGCCAACCCACGCTGGCAATGCTACCAAG  
CGTCTCTGACCACTGGCCAAGTCTCCCTCCGGAGCCGCGTGCAGGTGCAAAGGCCTTGCGGGAAACAACACTG  
GCGATGACCAAAATGTCGCGTGTGGAGGTGTGATGTCGCCTCTCCCTGGCTATGACTGGCCACTCTC  
TCAAAGGAATGGAAGCACCTGAAGGAAGGAGCAAGGACAAAGTTGCGGAAGAGTTGAGAGCCAAGGCCTGGAG  
CTGGAGAGGAAACTCACAGAACTCACCCAGCTCTACAAGAGCTTGCAAGGAGGTTAAGGGAGCATGTTGGATG  
GTGGGGAGGATTAACTGGACCTGCAACCGAGGCTACTGGAAGGAGTTAAGGGAGCATGTTGGATG  
CTGTGGCTGT  
ATGGCCTGGAGGTCAAATAATATCAAGTACATCTGGAGATGAGGGTGCTGTCTGGACAGACCTCGGCATG  
CCTCTGTTCTCCTCAATGCTCTTAAGGCTATGTCTGGAAAAGGGCTTCCCTGTTGTTGTTGTT  
GTTTGTGTTGTTGTTGAGACTCCAGTCTGGGTGTCAGAATGAGACCCATCTCAAAAAAAAAAAAAAAA  
AAAAGAAGAAGAATAACAGTCATGTATCTCTGGTACAGGGACGCATTCTGATAAATGTGTCATTAGGCAATTGC  
ATTGTAGTGTGATTATCACAGATTGTACTTATAACAAACTTAGATGGCATAGCCTACTGCATACTAGGCTATAT  
GGGAGGCCTATTGCTCCAGGCACGCACCTGTACAGCATGTGACTACTGAATACTATAGGCAATTGCAGCACA  
ATGGGAAATATTGTGTATCTAAACATATGTAAACAGAGAAAAGTAAAAATATGGCATAAAAGATAAAGA  
ATTGGCTCTCTGTACAGGGACTTACTACGAATGGAGCTGCAAGGGCTGAGAGTTGCTCCAGATGAGTCAGTGA  
GTGGTGAATGAATGTGAAGGCCTAGGCATTACTGTATACTACTGTAGGCTTATAAACACAGCACACTTAGGGT  
ACACAAAATGCATATTAAACATTCTCCTTCAGTATATTAGGCAATAGGAATTTCAGTCAGTCCACTATAAT  
CTTATCAAACCATGGTGTATATGCAGTTGACCGAACATTGTTATTGGACACATAACTATAGTTGAAAGAATAA  
GCAAAAAGTCTATCTAGGTGTGCTGTGAGCAACTTTAATTATTCTCCTGCTGCAATATGAGTTAATCTT  
CTCTGATCGATGTAGATTCCAGGAAGGGGTGTCAGGACAATTACCTTCCTCTGGAGAAACTCCCTTAATCAA  
ATAAGAGAACTCAAAGAAAATCCCTCCCTGCTTGGAGGGAGGTGGCAGCAGTGGTCAGAGATA  
GACCTTGTCTTATTCTGAGGCCCTCAGTCCTTATTCAAAGCACTCAGCATGCCAAAGCACCTATT  
TTAGGGTATCTTTCTGAGGCCCTAAACACTGTGTTGGGATGTCACACTGTGACAGGAAAATATCTGGGCC  
AGAATCACTAAGGAAAACCAAGCTTAGGAAACTCTTAGGGCAAACCCACCTCCACTCTATTCAAAGTTATC  
TCTCTGCTCACTGAGATAGATACTATCTGATTGCTCTTGGAAAGGCTAATCAGAAACTCAAAGAATGCAA  
CTGTTTGTGTCACCTATCTGACCTGGAAAGCTCCCTCCCCACTGAACCAATGTTCTTACATATATTGAT  
TAATGTCTATGTCTCCCTAAATGTATAAAACCAAGGTATGCCCAACCACATTGGCACATGTCATCAGGACT  
TCCTGAGTGTGTCACAGTGTGCTCAACCTTGGAAAATAACTTCTAAATTAACTGAGACAAAAAAA  
AAAAA

53/105  
**FIGURE 52**

MDNQAERESEAGVGLQRDEDDAFLCEDVELQDGDL SPEEKIFLREFPRLKEDLKGNIDKLRALADDIDKTHKKFT  
KANMVATSTAVISGVMSLLGLALAPATGGGSLLLSTAGQGLATAAGVTSIVSGTLERSKNKEAQARAEDILPTYD  
QEDREDEEEKADYVTAAGKIIYNLRNLTKYAKKNVRAFWKLRLANATKRLTTGQVSSRSRVQVQKAFAGT  
TLAMTKNARVLGGVMSAFSLGYDLATLSKEWKHLKEGARTKFAELRAKALELERKLTELTQLYKSLQQKVRSA  
RGVGKDLTGTCTEAYWKELEHVWMWLWLCVCLCVYVQFT

54/105  
**FIGURE 53**

GGGCCAGGCCGCCCCCGCGTGCCTGCAGCGGCCAGAGCGTGCAGGGCGCCCGCTACTCACTAGCTGAGG  
TGGCAGTGGTCCACCAACATGGAGCTCTCGCAGATGTCGGAGCTCATGGGCTGTCGGTGTGCTTGGGCTGCT  
GGCCCTGATGGCGACGGCGGGTAGCGCAGGGGTGGCTGCGCAGGGGAGGGAGAGGAGCGGCGCCCGCTG  
CAAAAAAGCAAATGGATTCCACCTGACAAATCTTGGGATCCAAGAAGCAGAAACAATATCAGGGATTGGAA  
GGAGAAGCCTCAACAACACAACCTCACCCACCGCTCCTGGCTGAGCTGAAGAGCACAGCAGGAACATATC  
TTGCATGGACTTTAGCAGCAATGGAAATACCTGGCTACCTGTGAGATGATCGCACCATCCGATCTGGAGCAC  
CAAGGACTTCCTGCAGCGAGAGCAACCGCAGCATGAGAGCAGGGCTGGACCACGCCACCCCTGGTGCCT  
CAGCCCTGACTGCAGAGCCTTCATCGTCTGGCTGGCAACGGGACACCCCTCGTGTCTCAAGATGACCAAGCG  
GGAGGATGGGGCTACACCTCACGCCACCCAGAGGACTTCCCTAAAAGCACAAGGCGCTGATCGACAT  
TGGCATTGCTAACACAGGAAAGTTATCATGACTGCCTCCAGTGACACCAGTCCTCATCTGGAGCCTGAAGGG  
TCAAGTGTGTCTACCATCAACACCAACCAGATGAACAACACACACGCTGCTGTATCTCCCTGTGGCAGATTGT  
AGCCTCGTGTGGCTCACCCAGATGTGAAGGTTGGAAAGTCTGCTTGGAAAGAAGGGGAGTTCCAGGAGGT  
GGTGCAGGCCCTCGAACTAAAGGCCACTCCGGCTGTGACTCGTTGCTTCTCAACGACTCACGGAGGAT  
GGCTTCTGCTCCAAGGATGGTACATGAAACTGTGGACACAGATGTGGAATACAAGAAGAAGCAGGACCCCTA  
CTTGCTGAAGACAGGCCCTTGAGGCGGCGGGTGCCTGGCCCTGGCCCTCTCCCCAACGCCA  
GGTCTGGCCTTGCCAGTGGCAGTAGTATTGATCTCTACAATACCCGGCGGGCGAGAAGGAGGAGTGCTTGA  
GCGGGTCCATGGCGAGTGTATGCCAATTGTCCTTGACATCAGTGGCCGTTCTGGCCTCTGGGACCG  
GGCGGTGCGCTGTTACAACACTCTGGCCACCGAGCCATGGTGGAGGAGATGCAGGGCACCTGAAGCGGGC  
CTCCAACGAGAGCACCGCCAGAGGCTGCAGCAGCAGCTGACCCAGGCCAAGAGACCTGAAGAGCCTGGTGC  
CCTGAAGAAGTGAACCTGGGAGGGCCGGCGAGAGGATTGAGGAGGAGGGATCTGGCCTCTCATGGCCTGCT  
GCCATTTCTCCCAGGTGAAAGCCTTCAGAAGGAGTCTCTGGTTTCTACTGGTGGCCCTGCTTCTTCCC  
ATTGAAACTACTCTGTCTACTTAGGTCTCTCTTCTTGCTGGCTGTGACTCCTCCCTGACTAGTGGCAAGGT  
GCTTTCTCTCCCAGGCCAGGGTGGAAATCTGCTCCACCTGGCACTGAGGAGAATGGTAGAGGAGGAGG  
AGAGAGAGAGAGAATGTGATTTGGCCTGTGGCAGCACATCCTCACACCCAAAGAAGTTGAAATGTTCCAG  
AACAAACCTAGAGAACACCTGAGTACTAAGCAGCAGTTGCAAGGATGGGAGACTGGGATAGCTCCCATCACAG  
AACTGTGTTCCATCAAAAGACACTAAGGATTCTCTGGCCTCAGTTCTATTTGTAAGATGGAGAATAATC  
CTCTCTGTAACCTGCAAAGATGATATGAGGCTAAGAGAATATCAAGTCCCAGGTCTGGAAAGAAAAGTGA  
AAAGAGTAGTACTATTGTCCAATGTCAAGGAGTGGTAAAGTGGAAACCAAGTGTGCTTGAACCAAATTAGAA  
ACACATTCTGGAGGCCAAAGTTCTGGGACTTGATCATACTTATATGGTGGACTCTCTCTTCCGG  
AGATGATATCTGTTAAGGAGACCTCTTCAAGTCAAGTCAAGATATTGAGTGCCACTCTGTGCC  
CAAATAAATATGAGCTGGGATTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
AAAAAAA

55/105  
**FIGURE 54**

MELSQMSELMGLSVLLGLLALMATAAVARGWLRA GEERSGRPACQKANGFPDKSSGSKQKQYQRIRKEKPQQH  
NFTHRLLAAALKSHSGNISCMDFSSNGKYLATCADRTIRIWSTKDFLQREHRSMRANVELDHATLVRFSPDCRA  
FIVWL LANGDTLRVFKMTKREDGGYTFTATPEDFPKKHAKPVIDIGIANTGKFIMTASSDTVLIWSLKGQVLSTI  
NTNQMNNTHAAVSPCGRFVASC GFTPDVKVWEVCFGKKGEFQEVVRAFELKGHSAAVHSFAFSNDSRRMASVSKD  
GTWKLWDTDVEYKKQDPYLLKTGRFEEAAGAACRCLALSPNAQVLALASGSSIHLYNTRGEKEECFERVHGE  
IANLSFDITGRFLASCGDRAVRLFHNTPGH RAMVEEMQGHLKRASNESTRQLQQQLTQAQETLKSLGALKK

56/105  
**FIGURE 55**

GGCACGAGGGTAAGCGCTAGGGCGCTCGCGCGCAGCAGAAAATGGCGGCTTCCAGGTGGCGCGAAGGC  
CGTGGCTCTGTTGTGCCTCTGACCTGCTGCTGCTGCTACTGCTACCACCGCTGGTCCTGCGCGGCCGA  
AGGCTGCCGGGACGCCGACGAGTCAACCCACCTCCCCGGAAGAAGAAGGATATTGCGATTACAATGA  
TGCAGACATGGCGCTTCTGGAGCAATGGGAGAAAGATGATGACATTGAAGAAGGAGATCTTCCAGAGCACAA  
GAGACCTCAGCACCTGCACTCTCAAAGATAGACCAAGCAAGCCTGAAAGCATATTGAAAATGACGAAAAA  
AGGGAAAGACTCTCATGATGTTGCACTGTATCAGGAAGCCCTACTGAGAAGGAGACAGAGGAAATTACGAGCCT  
CTGGCAGGGCAGCCTTCAATGCCAATGACGTCAGAGGTCATTGGGATCAGACCGTGTATCTCAT  
GCTTCGCGATGGGAGCTACGCCCTGGGAGATCAAGGACTTTGGTCGGTCAAGACAGGTGTGCTGATGTAACCT  
GGAGGGCCAGGTGTACCCGGAAAGGAGGAGGAAGCAAAGAGAAAAATAAAACAAAGCAAGACAAGGGCAAAA  
AAAGAAGGAAGGAGATCTGAAATCTGGTCTCCAAGGAAGAAATCGAGCTGGAAATAAGAGAAGACCTGTG  
ATGGGGCAGCAGTGACGCGCTGGGGGGACAGGTGGACGTGGAGAGCTTTGCCAGCTCCTGGGTGGAGT  
GGTCTCAGGCAACTGCACACCGGATGACATTCTAGTGTCTCTAGAAAGGTCTGCCACATGACCAGTTGTGGT  
CAAAGAATTACTGCTTAATAGGCTCAAGTAAGAAGACAGATGTTCTAATTAATACGGACACTGACAAATT  
ATGTTACTATAAAATCTCCTACATGAAATGTGACTGTGTTGCTTTCCATTACACTGACCCCTGAGCC  
CCTTCAGTGCCTGAGGTGCCCTCACCTGTGCCCTCGCTCGCCAGTTGATTGGACCTGTTTC  
CCACCTCAGCCCCATGCTTGTCAAACGTGTTGCCCTCAGCCAAACAGGGCTGGGAGGAAAGAGAGTCC  
TGCTTCTGCCCTGGCTTCCCACGGGAGGTGAAGTGACAGCCTTGACAGCTAGAAATGTGGAGGAC  
GCAGCAAGTCTGGACATGCCAGCCATATGAATGATATCTCTTCTTGAAACCACCTCCACCTGCTCTTAT  
GGACACCCCGCTGGGCAAGGCAAGGGATGGGACAAACTTGAAGAGATTGGCTTGTGATCCTCAGTGTCCA  
GTGTCAACCCAGATGGGCCGGAAAGCTGCCCTGAAGCCAACGAATGCCATTGCTCTGAAAAGCAATGATGAGTCT  
CCTGAGCAAAGCCCTATGAGAGGACGGCTGCCACCCACCCGTTCCATGGCTCCCACGCCAGCCCCCACGCC  
ACCAGGTACCGCCTAGCCACACTGCACTGTGGCCCTGGGAAGAGCTCTGCTCAGCTACTGATGGTGCCTGA  
CATGTGTTGTATGGGCAGGCAGGGTCTCACCTGAGCCTCACTGCCACCTAGGCAGGGTATGGCTTGTGTA  
GACAGGGAGCCATGCACAGCAGGAAATGGGCAAACGTGAAGGGAGGAGACCTCCATGGAACCTGCCTACCTC  
GAGCTGCCCTAGGACAACCAGCAAGTCCCTGCCCTCTGAAGGCCAGTGTCCCACCATCAAATGATGGCA  
TTGATCTACAAGATCTCCTCCCCACTCTACACTGTGCTCTAGAAAGAGTTCTGAGAGGTGAAGAAATCTG  
ACCAAGAGTGTGGAGTAAATGGTGCACCCAGATTAGAAACTCTAAACCCCTGAAACCCACACAGCACTCTCCAC  
TTGCAAGTTCACAGCTTGAGCATCCGAGGGCATGGGAATCCACGTCTACACTCAGCCCATCTGGCAGGCAACAT  
GCATTATCACCAACCGGGAAAGTGCCTACACTAGGTACCTACATACATCAGCTTGGCAACCGGGATCT  
TAATTAAAAGACAGAGGCTCCAAATGATTCAGGACTGGCTGGAAAGGATACTCTCTTCTTCTTATGTCT  
CCTTGGCTGGTTGAAGGGTTGCATTGCTCTGATTTCTCATGCCCATATGATGGTTCTGATTTGA  
TGTTTCCATAGTGAGGATGGTGTATAACTAAGAAGCCCAGCTTGAGTTAGGAAGAGATTAGGATCCAAGTCCTAA  
CCTTAAGCAAGTGGCTGGCTGTTAAATTCTGGTGCCTCATCTTAAATGAAATAATAGTACCTGCTGCACAG  
GTTATTAAATGAGTAAGTGAGATAATTATAACATGTTGCAAGGCTGAACATTGAATAGCAACCACAA  
TGATCATTATCATTATTGCCGTGGAGCTTCTCTGCCCTCACAGCCTCTGAGGACTCTCCCTGCACCTAG  
TTTTGTGATGATAATGGGTGTCATGCTGACATGTCCTCCTGGATGGCTTCTCCAGCACGCCAGCAAGCCTG  
TGCTAGGCTCACCACCGAAGCTGCCCTGGACAAGCCATGTAGGTTCTCTGGCTTAGGACACTTGTGAAAC  
AAATAAGAGTTGCAAGCTGCAAAGGACTGTACAAACGCTAGGCATCAGCTAGAATGTTGAAGAAGATGCAGGA  
CCCCCTGGCCAGTCTGAATAGCAGGAATTGGTAGGAGGGTGAGACTAAGCCACTTTCTTATGATACTTTGTTG  
ATAAAAGGCCACTTCCAAGCAAAAAAAAAAAAAAA

57/105  
**FIGURE 56**

MAASRWARKAVVLLCASDLLLLLLPPPGSCAAEGSPGTPDESTPPPRKKKDIRDYNDADMARLLEQWEKDDD  
IEEGDLPEHKRPSAPVDFS KIDPSKPE S I L K M T K K G K T L M M F V T V G S P T E K E T E E I T S L W Q G S L F N A N Y D V Q R F  
IVGSDRAIFMLRDGSY AWEIKDFLVGQDRCADVTLEGQVYPGKGGSKEKNKTQDKGKKKEGDLKSRSK EEN  
RAGNKREDL

58/105  
**FIGURE 57**

GGCACGAGGCCGGAGTCACGGAAGCCGTTTCGCTTCACCTTCCTGGCTGTAGAGCGCTTCCCCCTGGCGGGT  
GAGAGTGCAGAGACGAAGGTGCGAGATGAGCACTATGTTCGCGGACACTCTCCTCATCGTTTATCTCTGTGTG  
CACGGCTCTGCTCGCAGAGGGCATAACCTGGGTCTGGTTACAGGACAGACAAGTACAAGAGACTGAAGGCAGA  
AGTGGAAAACAGAGTAAAAAAATTGGAAAAGAAGAAGGAAACAATAACAGAGTCAGCTGGTCGACAACAGAAAAA  
GAAAATAGAGAGACAAGAAGAGAAACTGAAGAATAACAACAGAGATCTATCAATGGTCGAATGAAATCCATGTT  
TGCTATTGGCTTTGTTACTGCCCTAATGGGAATGTTCAATTCCATATTGATGGTAGAGTGGTGGCAAAGCT  
TCCTTTACCCCTCTTCTTACATCCAAGGACTGTCATCGAAATCTGCTGGGAGATGACACCACAGACTGTT  
CTTCATTTCTGTATATTCTCTGTACTATGTCGATTGACAGAACATTAGAAGATTCTGGCCTTGGCCCTTC  
ACGAGCCGCCACCAAGCAGGCAGGTGGATTCTTGGCCACCACCTCTGGGAAGTTCTTGAACTCAAGA  
ACTCTTATTTCTATCATTCTCTAGACACACACATCAGACTGGCAACTGTTGTAGCAAGAGCCATAGG  
TAGCCTTACTACTTGGGCCTCTTCTAGTTGAATTATTTTAAGCCTTGGGTATGATTAGAGTGAAATGG  
CAGCCAGCAAACCTGATAGTGCTTTGGTCTAGATGATTTTATCAAATAAGTGGATTGATTAGTTAAGTCAG  
GTAATGTTATGTAATGAAAAACAAATAGCATCCTCTTGTTCATTACATAAGTATTTCTGTGGGACCGACT  
CTCAAGGCACTGTGTATGCCCTGCAAGTTGGCTGTCTATGAGCATTAGAGATTAGAAGAAAATTAGTTGT  
TTAACCCCTTGTAACTGTTGTTGTTGTTGTTTCAAGCCAAATACATGACATAAGATCAATAAGAGG  
CCAAATTTAGCTTTATGTACAAGGAAAAAAAAAAAAAA

59/105  
**FIGURE 58**

MSTMFADTILLIVFISVCTALLAEGITWVLVYRTDKYKRLKAEVEKQSKKLEKKETITESAGRQQKKIERQEEK  
LKNNNRDLSMVRMKSMAIGFCFTALMGMFNSIFDGRVVAKLPFTPLSYIQGLSHRNLLGDDTTDCSFIFLYILC  
TMSIRQNIQKILGLAPSRAATKQAGGFLGPPPSGKFS

60/105  
**FIGURE 59**

GTTCGCCGCCGCCGCCGCCGGCACCTGGAGTTTCAGACTCCAGATTCCCTGTCAACCACGAGGAGTCCAGA  
GAGGAAACGCGGAGCAGGAGACAAACAGTACCTGACGCCCTTTCAGCCCCGGATGCCAGCAGGG**AT**GGCGAC  
AAGATCTGGCTGCCCTTCCCGTGCCTCTGGCCGCTCTGCCCTCGGTGCTGCTGCCCTGGGGCGGCCGGCTC  
ACACCTCCCTCGATAGCGACTTACCTTACCCCTCCCGCCGGCAGAAGGAGTGCTCTACCAAGCCCATGCC  
CTGAAGGCCTCGCTGGAGATCGAGTACCAAGTTAGATGGAGCAGGATTAGATATTGATTCCATCTGCCCT  
CCAGAAGGAAACCTAGTTTGAACAAAGAAAATCAGATGGAGTTCACACTGTAGAGACTGAAGTTGGTGA  
TACATGTTCTGCTTGTACAATACTACATTCAAGCACCATTCTGAGAAGGTGATTTCCTTGAATTAATCCTGGATAAT  
ATGGGAGAACAGGCACAAGAACAGAACAGATGGAAGAAAATATACTGGCACAGATATGGATATGAAACTG  
GAAGACATCCTGGAATCCATCAACAGCATCAAGTCCAGACTAACAGAAAAGTGGGCACATACAAACTCTGCTTAGA  
GCATTGAGCTGTGATCGAACATACAAGAAAGCAACTTTGATAGAGTCATTTCTGGTCTATGGTTAATT  
GTGGTCATGGTGGTGGTGCAGCCATTCAAGTTATGCTGAAGAGTCTGTTGAAGATAAGAGGAAAAGTAGA  
ACT**AAA**ACTCCAAACTAGAGTACGTAACATTGAAAAATGAGGCATAAAATGCAATAAAACTGTTACAGTCAGA  
CCATTAAATGGTCTTCTCCAAATATTTGAGATATAAAAGTAGGAAACAGGTATAATTAAATGTGAAAATTAAG  
TCTTCACTTCTGTGCAAGTAATCCTGCTGATCCAGTTGACTTAAGTGTGTAACAGGAATATTTGCAGAATAT  
AGGTTTAAGTGAATGAAGCCATATTAATAACTGCATTTCTAACATTGAAAAATTGCAAATGTCTAGGTGA  
TTAAATAATGAGTATTGGGCCTAAA

61/105  
**FIGURE 60**

MGDKIWLPFPVLLAALPPVLLPAGFTPSLDSDFTFTLPAGQKECFYQPMPLKASLEIEYQVLDGAGLDIDFH  
LASPEGKTLVFEQRKSDGVHTVETEVGDYMFCDNTFSTISEKVIFFELILDNMGEQAQEQQEDWKYITGTDILD  
MKLEDILESINSIKSRLSKSGHIQTLRAFEARDRNQESNFDRVNFWSMVNLVVMVVSAIQVYMLKSLFEDKR  
KSRT

62/105  
**FIGURE 61**

AACTGATCGCGCCTAGTCCCACCGCGTGTGCTAGTGAGCCGGAGCCGGCACGGCGCAGTGGCGGCCCGGC  
CTGCAGGAGCCCGACGGGGCTCTGCCATGGGGAGTGACCGCGCTGCACCGCTTCCCGGCAGCGGCAGA  
CATGAGGAGACCCCGCAGGGCAGCGCGCGCTCGTGAGCCCCGGATGGAGGAAATACGGCGGGAC  
GTGCTGGCGGCCCGCGCGCGCGCTTGGGCCGGTGGACGTACCCAGCGCTCGATTAACAAAATATT  
GTGTTACTATGTTCACTAAATTGAAAGGCTGTGGACTTTCGAATCATATGATCTCTAAAAGCTGTTCAC  
ATTGTTCAAGTCATTTATATTAAAACGGGACTGCACTTTATGGTTCAAAAGCCATTTCTCT  
GGGAAAACATTACCAAACACCAGTGGATCAAAATTAAACATGCAGTGCTGGGTGATTATTCACTCTG  
TGGTTTTGGCCTCACTTTGTTGACCACTAAGGACTTGCTGCTATTGAGCACAGTGATATTGTTGTCATT  
TCACTACTCAGTGTGTTGTCACCAGTCTGGAGGAGGACAGCAAAGACAAGGGAGCTGCTTTTCAATT  
GCTGTGATCTGTTATTGCTTTGACAATGATGATCTCATGGCTAAAATGGCTGAACACCCTGAAGGACATCAT  
GACAGTGCTCTAACTCATATGCTTACACAGCCATTGCCCTCTAGGTGTGGCAGATCACAAGGGTGGAGTATT  
TTGCTAGTACTGGCTTGTGTTAAAGTGGTTTCATACAGCTCCAGAAAGCTCTGTCGACGTTGGTGA  
GCTAAACGCTCTCAAGCTTATCTCATCTGTTCTGTGCTCTTGCCATGGTCATTGTTCTGTG  
ACAACGTGAGTAAAGTGGAGTCTGGTTCTCATTATGCCCTTGCAACGGTATCTTTTGTGATGAC  
CTGGATTCTACGTGGATTCCATTGTCAGTCAAATGGAAGTCCAAATGTGCTCGTTATGGATCCTTCCC  
ATTTTATTAGTGTCTCCTTTGGAAATTGGACACATCCAATAACAGACCAGCTCGGGCTATGAACAAA  
GCAGCACACCAGGAGAGCACTGAACACGTCTGTCTGGAGGAGTGGTAGTGAGTGCTATATTCTCATTTGCT  
CCAATATCTTATCATCTCCCTTAAGAGAGGACAAAAGGTACCCATTGGATATTCTCTGAAGGAACACCT  
CTTATAACTCATGGGTGATGCTTTCAGCATAGCTCTCAATCGATCCCTAGGTTATTAAGGAATCACTAAA  
CAAATTCTGAGGAGAGTGACTCTAGGCAGATCTTACTTCTTGCTGCTGAATCTGCTTTACCTTGTGAA  
TTATTCTATGGCGTGTGACCAATAGTCTGGCCTGATCTGGATGGATTCCACATGCTTTGACTGCTCTGCT  
TTAGTCATGGACTTTGCTGCCCTGATGAGTAGGTGAAAGCCACTCGGATTTCTCTATGGTACGCCGA  
ATAGAAATTCTGCTGGATTATTAATGGACTTTCTAATAGTAATAGCCTTGTGTTATGGAGTCAGTG  
GCTAGATTGATTGATCCTCCAGAATTAGACACTCACATGTTAACACCAGTCTCAGTTGGAGGGCTGATAGTAAAC  
CTTATTGGTATCTGTGCCCTTAGCCATGCCATGCCATGGAGCTCTCAAGGAAGCTGCACTCATCT  
GATCACAGCCATTCACCCATATGCATGGACACAGTGACCATGGCATGGTCACAGCCACGGATCTGCGGGTGG  
GGCATGAATGCTAACATGAGGGGTGATTCTACATGTTGGCAGATACTCTGGCAGCATTGGTGTGATCGTA  
TCCACAGTCTTATAGAGCAGTTGGATGGTCATCGCTGACCCACTCTGTTCTCTTTATTGCTATATTAATA  
TTCTCAGTGTGTTCCACTGATTAAGATGCCAGGAAACATTACAGGTGACATCTGATGTGCTAGAACAAAGAATAGTACAG  
CAGGTTACAGGAATACTTAAAGATGCTGGAGTAAACAATTAAACATTCAAGTGGAAAGGAGGCATACTTCAA  
CATATGTCGGCTAACAGTACTGGATTCTGATGTTCTGGCTATGACAAAACAAATGGAATCCATGAAATACTGC  
AAAGATGGTACTTACATCATGTGAATAACTCAAGAATTACCCCTGGAGAATAACAATGAAGATTAAATGACTC  
AGTATTGTAATATTGCCAGAAGGATAAAATTACACATTAACATGATGCTACAGAAACAGAGTCCCTACTACTGGATC  
AAGGAATCTTCTGAGGAAATTAAACAGAATGAAACATTAATGGTAAAGTGGAGTAATTATTAAATTA  
TGTGTATAAGGAATCAAATTGAGTAAACATGATGTTACATCATCTCGAAAATAGATATGATGGATTCT  
AGTGAAGACCAAAATTACTTCTGTTACTTCTATCAGGAAGCATCTCCATTGTAATATGATTTACATGTTA  
TTACAAAGACCCAAATGAAAATTAGTCCATTGGCATAGCCTAAAGATAAAATAGGAATAAAAGTTCTA  
TATTTATGGAAAAA

63/105  
**FIGURE 62**

MAKMAEHPEGHHDSALTHMLYTAIAFLGVADHKGGVLLLVLALCCKVGFHTASRKLSVDGGAKRLQALSHLVSV  
LLLCPWVIVLVSVTTESKVESWFSLIMPATVIFFVMIIDFYVDSICSVKMEVSKCARYGSFPIFISALLFGNFWT  
HPITDQLRAMNKAHQESTEHVLSGGVVVSAIFFILSANILSSPSKRGQKGTIGYSPEGTPLYNFMGDAFQHSS  
QSIPRFIKESLKQILEESDSRQIFYFLCLNLLFTFVELFYGVLTNSLGLISDGFHMLFDCSALVMGLFAALMSRW  
KATRIFSYGYGRIEILSGFINGLFLIVIAFFVFME SVARLIDPPELDTHMLTPVSVGGLIVNLIGICAFSHAHSH  
AHGASQGSCHSSDHSHSHHMHGHSDHGHGHGSAGGMNANMRGVFLHVLADTLGSIGVIVSTVIEQFGWFIA  
DPLCSLFIAILIFLSVVVPLIKDACQVLLRLPPEYEKELHIALEKIQKIEGLISYRDPHFWRHSASIVAGTIHQ  
VTSDVLEQRIVQQVTGILKDAGVNNLTIQVEKEAYFQHMSGLSTGFHDVLAMTKQMESMKYCKDGTYIM

64/105  
**FIGURE 63**

GGGGCGAGACCTACGACGCCGGCGAGCAGTGGCCGTACGCCCTAAAAGATGCGGGTCTGGCACCTCTAATTGC  
TCTCGTGTATCGGTGCCCGACTTACGATGGCTCGCCCAACCTACTACCTTCTGTCGGCCCTGCTCTGCG  
TGCCTCCCTACTCGTAGGAAACTGCCGCCCTCTGCCACGGCTGCCACCCAACGCGAAGACGTAACCCGTG  
TGACTTTGACTGGAGAGAAGTGGAGATCCTGATGTTCTCAGTGCATTGTGATGATGAAGAACCGCAGATCCAT  
CACTGTGGAGCAACATATAGGCAACATTTCTATGTTAGTAAAGTGGCCAACACAATTCTTCTCCGTTGGA  
TATTCGATGGGCCTACTTACATCACACTCTGCAAGTGTCCCTGATGACGTGCAAACCCCCCTATATATGGG  
CCCTGAGTATATCAAGTACTTCAATGATAAAACCATGGTAGAGGAACCTAGAACGGGACAAGAGGGTCACTTGGAT  
TGTGGAGTTCTTGCCAATTGGCTAATGACTGCCAATCATTTGCCCTATCTATGCTGACCTCCCTTAAATA  
CAACTGTACAGGGCTAAATTGGAGGTGGATGTTGACGCTACTGATGTTAGTACGCGGTACAAAGTGAG  
CACATCACCCCTACCAAGCAACTCCCTACCCCTGATCTGTTCAAGGTGGCAAGGAGGCAATGCGCGGCCACA  
GATTGACAAGAAAGGACGGGCTGTCATGGACCTCTCTGAGGAGAATGTGATCCGAGAATTAACTTAAATGA  
GCTATACCAGCGGGCCAAGAAACTATCAAAGGCTGGAGACAATATCCCTGAGGAGCAGCCTGTCGGCTCAACCCC  
CACACAGTGTCAAGATGGGAAAACAAGAAGGATAAATAAGATCCTCACTTGGCAGTGCTTCCCTGTCAA  
TTCCAGGCTCTTCCATAACCACAAGCCTGAGGTGCAGCTTTATTTATGTTTCCCTTGGCTGTGACTGGGTG  
GGGCAGCATGCAGCTCTGATTTAAAGAGGCATCTAGGGAATTGTCAGGCACCCCTACAGGAAGGCTGCCATGC  
TTGTGGCCAACGTGTTCACTGGAGCAAAGAAAGAGATCTCATAGGACGGAGGGGAAATGGTTTCCCTCCAAG  
CTTGGGTCACTGTGTTAAGTGCCTATCAGCTATTCAAGACATCTCCATGGTTCTCATGAAACTCTGTGGTTCA  
TCATTCCCTTCTAGTTGACCTGCACAGCTGGTTAGACCTAGATTAACCTAAGGTAAGATGCTGGGTATAGA  
ACGCTAAGAATTTCCTCCAAGGACTCTGCTTCTCAAGCCCTCTGGCTCGTTATGGCTTCATTAAAGT  
ATAAGCTAACTTGTGCTAGTCCTAAGGAGAAACCTTAACCACAAAGTTTATCATTGAAGACAATATTGA  
ACAACCCCTATTTGTGGGATTGAGAAGGGGTGAATAGAGGCTTGAGACTTCCCTTGTGGTAGGACTTGG  
AGGAGAAATCCCCTGGACTTCACTAACCTCTGACATACTCCCCACACCCAGTTGATGGCTTCCGTAAATAAA  
AGATTGGGATTCCTTTG

65/105  
**FIGURE 64**

MAVLAPLIALVYSVPRLSRWLAQPYYLLSALLSAAFLLVRKLPPCHGLPTQREDGNPCDFDWREVEILMFLSAI  
VMMKNRRSITVEQHIGNIFMFSKVANTILFFRLDIRMGLLYITILCIVFLMTCKPPLYMGPEYIKYFNDKTIDEEL  
ERDKRVTWIVEFFANWSNDCQSFAPIYADLSLKYNCTGLNFGKVDVGRYTDVSTRYKVSTSPLTKQLPTLILFQG  
GKEAMRRPQIDKKGRAVSWTFSEENVIREFNLNELYQRACKLSKAGDNIPEEQPVASTPTVSDGENKKDK

66/105  
**FIGURE 65**

TGCAGTCTGCTGAGGGCGGCCGAAGTGGCTGGCTATTAGATGAGGCTCTGCTGCTCTCCTAGTGGCGGC  
GTCTGCGATGGTCCGGAGCGAGGCCTCGGCCAATCTGGGCGGCCTGCCAGCAAGAGATTAAGATGCAGTACGCC  
ACGGGGCCGCTGCTCAAGTCCAGATTGTGTTCTGAGGTTATAGGCAGGGTGTGAGGAGTACATGCGGGTT  
ATTAGCCAGCGGTACCCAGACATCCGCATTGAAGGAGAGAATTACCTCCCTCAACCAATATATAGACACATAGCA  
TCTTCCTGTCAGTCTCAAACACTAGTATTAAATAGGCTTAATAATTGTTGGCAAGGATCCTTGCTTCTTGCG  
ATGCAAGCTCCTAGCATCTGGCAGTGGGGCCAAGAAAATAAGGTTATGCATGTATGATGGTTCTTCTTGAGC  
AACATGATTGAGAACCACTAGTGTATGTCAAACAGGTGCATTGAGATAACTTTAAATGATGTACCTGTTGGTCTAAG  
CTGGAATCTGGTCACCTTCATGCCAACACTGTTCAAATTCTGACAATGAAAGCTCAATGTGCAT  
ATGGATTCAATCCCACACCATCGATCATAGCACCCACCTATCAGCACTGAAAACCTTTGCATTAAGGGATCATT  
GCAAGAGCAGCGTGACTGACATTATGAAGGCCTGACTGAAGACAGCAAGCTGTTAGTACAGACCAGATGCTTC  
TTGGCAGGCTCGTTGACCTCTGGAAAACCTCAATGCAAGATAGTGGCTGCTGGCATATTTGGAATTCT  
GCACATTGATGGAGTGCAATAATACTGTATAGCTTCCCCACCTCCCACAAAATCACCAGTTAATGTGTGT  
GTGTGTTTTTTAAGGTAAACATTACTACTTGTAACTTTCTTAGTCATATTGGAAAAAGTAGAAAATT  
GGAGTTACATTGGATTTTTTCCAA

67/105  
**FIGURE 66**

MQYATGPLLKQICVSXGYRRVFEYMRVISQRYPDIRIEGENYLPQPIYRHIASFLSVFKLVLIGLIIVGKDFF  
AFFGMQAPSIWQWGQENKVYACMMVFFLSNMIENQCMSTGAFEITLNDVPVWSKLESGHLPSMQQLVQILDNEMK  
LNVHMDSIPHRS

68/105  
**FIGURE 67**

GGCGGGGGTGGCCGGTGATACCCGGGCCTTATAGTCCCGCCGCTCTCCACCTCCCTCTCC  
TCTCCTCTGGAGCAGAGGAGGTGTGGCGGTGGCTGGAGAAAGCGCGCGGAGGATGGAGGAAGGAGGCG  
GCGTACGGAGTCTGGTCCCGGGCGGGCGGTGTTACTGGCCTCTGCGGCCTCTGGAGGCGTCCGGCGGCC  
GAGCCCTCCTCAACTCAGCGATGACATCCCTTCGAGTCAGTCACCTGGCCGGCACCGAGTCTCTGCCACAA  
CTGGAGTTTATATAAAGAAGATAATTATGTCATCATGACAAC TGACATAAAAGAAAAATATAAATGCATACTTC  
CCCTTGACAGTGGGATGAGGAAGAAGAAAAGGATTATAAAGGCCCTAATCCAAGAGAGCTTTGGAGGCC  
TATTAAACAAAGCAGTTGCTACAGAATTGAGTCTTATTGACTACGAAGTATGTCATGAAAACACATT  
GGCAGTACCATGAAGAGAAAGAAACTGGTCAGAAAATAATTCACGAGTACTACCTGGAAATATGGCCA  
AGAACCTCTATTGAAAAGAACGAGAACGAGAAAAGGAAAATCAAATGAGATTCCCACAAAAATATCG  
AAGGTAGATGACACCATACTATCCTGTGGGAATGGGAATGGTACACCTTGAGTTGAAACAGAACCGGCC  
GATCAAGTACTGTGATGTACATATGTCATCCTGAACTAAGCATGAAATTCTTCAGTAGCTGAAGTTACAACCT  
GTGAATATGAAGTTGTCATTGACACCACTCTGTGCAGTCATCCTAAATATAGGTCAGAGCATCCTGTGA  
ATGACATATTGCAATCACTGCCAGGATCTCCATTAGCCCTCACCTGAGGAGCTGGAGCAGCAGGAAG  
AAATACTAAGGTGCCTTATTAGGAGAAATAAGAGGAAGATTGCAATCAACTAAAGAAGAGAGATTCCAGCGA  
TCCACAAGTCGATTGCTATTGGCTCTCAGCCAGTGCCTACTGTTGGACAAACCCACATATCCAAATTGACAGATG  
ACCAACTCATAAAAGAGTTCTTAGTGGTTCTACTGCTTCTGTGGGGTGTGGTGGAAATATGAATTCT  
GCTATGCCAAACATGTACATCAATACCATGAGGACAAGGATAGTGGAAAACCTCTGTGGTGTGGGACATGGA  
ACCAAGAAGAGCATATTGAATGGCTAAGAAGAATACTGCTAGAGCTTACATCTCAAGACGATGGTACCCAGA  
CAGTCAGGATGGTGTACATTATGGAAATGGAGATATTGTGATATAACTGACAAACCAAGACAGGTGACTG  
TAAAACCTAAAGTGCAGAACGAAATCAGATTCACCTCATGCTTACTGTATATATGCTAGAGCCTCACTCCTGTCAAT  
ATATTCTGGGGTGAATCTCAGTGTAAATCTTAGATAACAGCAGATGAAAATGGACTCTTCTCTCC  
CCAACTAAAGGATATTAAAGTTAGGGAAAGAAAAGATCATGAAAGTCACTGATAATTCTGTCCCCTGTGTCT  
CATTATAGAGTCTCAGCCATTGGACCTCTCTAAAGGATGGTATAAAATGACTCTCAACCACATTGTGAATACA  
TATGTGTATATAAGAGGTTATTGATAAAACTCTGAGGCAGACATTGTCGCTTTTCATTGGTGTGTC  
TTATAAACTGACTGTTTCTTGCTGGATACTGTGATTCCAAAATAATCTCATCCAAGCAAGTTAGAGTCCA  
GCCTAACATCAAATGTCATAATTGTGACCTATTGAAAGTTAAATAATAGATTATTGTAAATTATAGTAT  
ATGTAAGTAGCTAATGAAGTAAAGATCATGAAGAAAGAAATTGATAGGTGAAATGAGAGACCATGAAAATATG  
TAAATTCTAGTACCTGAAATCCTTCAACAGATTATATAGCACTGCTCTGCAAGTAGTTAAACTAGAAA  
CTGGGCACATGGTAGAGGGCTCACATGGGAGTTGCTCTCACCCCTGTTAATCTCAAGAAACTCTTATTATAATAG  
GTTGCTCTCTCAGAACCTTTATCTATTACTTTCTTCTTATGAGTATGTTACTCTCAGAGTATCTATCT  
GATGTAGACAGTGGTGTGATGCTCTGAGACTCAGAATGGTTACTCTAACAAAACACTGTGCTGTCTATCCCTTG  
TACTTGCTACTGTAATATGGATTCACTCTGAAACAGTTACAGCACAATATTAAAGTGAATAAAATG  
TCCACAAGCAAA

69/105  
**FIGURE 68**

MEEGGGGVRSVPAGPVLVLCGLLEASGGGRALPQLSDDIPFRVNWPGEFLSLPTTGVLKEDNYVIMTTAHKE  
KYKICILPLVTSGDEEEEKDYKGPNPRELLEPLFKQSSCSYRIESYWTYEVCHGKHIRQYHEEKETGQKINIHEY  
LGNMLAKNLLFEKEREAEKEKSNEIPTKNIEGQMTPYYPVGMGNTPCSLKQNRPSSSTVMYICHPESKHEILS  
VAEVTTCEYEVVILTPILCSPKRYFRASPVNDIFCQSLPGSPFKPLTLRQLEQQEEILRVPFRRNKEEDLQSTK  
EERFPAIHKSTAIIGSQPVLTVGTTISKLTDDQLIKEFLSGSYCFRGGVGWWKYEFCYGKHHQYHEDKDSGKTS  
VVVGTWNQEEHIEWAKKNTARAYHLQDDGTQTVRMVSHFYGNQDICDITDKPRQTVKLCKESDSPHAVTVYML  
EPHSCQYILGVESPVICKILDADENGLLSPN

70/105  
**FIGURE 69**

AACAATAGGAAACGTCAAATTGGGATAGTCGGCAGTTCTGGCCCTGCAGCTGGAGGTACCCCTGAGTTCTGAGG  
GTCGTAGTGTGTTCTGGTATTCTCATCGCGTCACCTACCGGTGTGGACAAGTAAAGTTGAATCAGCTTC  
TCCATGGCCTGGGCACAGTCCCGCTGAGCCATTTCTGGCTAAAAGTCCCCGCCAGAGGCCAATTG  
TCGCGGGCGGGTGGAGATCGCAGGTCGCTCAGGCTTGAGATGGGTCAAGGGTTGTGGAGAGTGGTCAGAAACC  
AGCAGCTGCAACAAGAAGGCTACAGTGAGCAAGGCTACCTCACCAGAGCAGAGCAGGAGAATGGCTGCGAGCA  
ACATTCTAACACCAATCATGTAACAAAGTCCAAGGAGGATTGACATATATCATCTTGAGCTAACGTTACCATCTTG  
CGAAAGAACAGGAAGGATTCAATTGGAAATGTTGCCTCCTGAGCTAACGTTACCATCTTG  
ATGCAACTGACCTTGCTGGCTCATGTGTTGGCAGGACCTTGCATGAACTCTCTGGCAAGGGTTGT  
GCAAATCCACTGGGTCACTGGCCATATACAATAAGAACCCACCTTAGGATTCTTTAGAAAAGTGTATA  
TGCAGCTGGATGAAGGCAGCCTCACCTTAATGCCAACCCAGATGAGGGAGTGAACACTTATGTCCAAGGGTA  
TCCTGGATGATTGCCAAAGGAAATAGCAAAGTTATCTCTGTACAAGAACACTAAATTGGAAAAACTGAGAA  
TCTATCTGATGAAAGGAGAGATGTTGGATGACCTTGTAAACATTGCATAATTAGAAAATCAGTTCTGCCAA  
ATGCACTGAGAGAAATTTCGTATATCCATGCCCTGAAGAGCTGGAGAGTATCTGAAACTCTTATAACAA  
AGTTCTCACATAGATTCTGTGCTGCAACCTGATTAAATGCGAGAACTTGGCCTTAGTCTGATGCTGTCTATG  
TACTGTGCTACTCTTGATTCTACTTCCATTGACCTCACTAGCCCTCATGTGAAGAATAAAATGTCAAAAGGG  
AATTATTGCAAATACCGTCGCGCTGCTCAAATATTAGTGAAGATTGTAGGGCATCTTATGACAATATCT  
ACCTTATTGGCATGGCTGCATAAAAAGCACAATTGCTAGGACTTCAGTTTACTTCAGACTAAAGCTACCC  
AAGGACTTAGCAGATGGGGTACATCAGTGCTGGCATTGAGCTGAGTATAACATCAAGCTCAGTGTG  
AACCTTTCTTGCATTTCTATTTCTATTAGTAATTCCCTGGGAACCAAATAATTGAGCTTACGAAATTTCT  
AATTGTTATCACGTTGCACAAAGCAGAGCCACTGTCTAACACAGCTTAAACGAAATGATAACTGACATT  
ATACTCTAAAGATGGGTATTGTGCTTAGATTGCTGAAAAACTTTATCCATTCCATTCTTATAACAAAT  
ACCATGTAATGTGTACATATTAACAAAGAGATTATAGTCATAATTATTGAAAGATTAAACTAAAG  
TTTTCTTCTCTCAAACGTGAAATTATTGATTCTGATCTGAAACTATTGTCTCGTAAAGTAA  
GATCTGACTCAGACAGAAACCAATACCAGCTCCTTAAACTTGAAGAGTGTGATTGTTACTATA  
TTACTATGCAAACGTGCTTAAACCATTTAGGATTAAAACATCAAAATTATGATTACATTGATAGGAAT  
TCAAGTCTCGTAAGTCCTTAAACCATTTAGGATTAAAACATCAAAATTATGATTACATTGATAGGAAT  
AAAATAAAATTATTAGAACTCTGGTAAAAAAAAAAAAAA

71/105  
**FIGURE 70**

MGQGLWRVVRNQQLQQEGYSEQGYLTREQSRRMAASNIISNTNHRKQVQGGIDIYHLLKARKSKEQEGFINLEMLP  
PELSFTILSYLNATDLCLASCVWQDLANDELLWQGLCKSTWGHWSIYNKNPPLGFSFRKVYMQLDEGSLTFNANP  
DEGVNYFMSKGILDDSPKEIAKFIFCTRLNWKKLRIYLDERRDVLDDLVTLNFRNQFLPNALREFFRHIAPE  
ERGEYLETLITKFSHRCACNPDIIMRELGLSPDAVYVLCYSLILLSIDLSPHVKNKMSKREFIRNTRRAAQNIS  
EDFVGHLYDNIYLIGHVAA

72/105  
**FIGURE 71**

TGAGAGTCCTCTAGACAGCGCTCCCTCGCAGCACCCTGAGCAGCTGCGCTGAGCAGCCCCGGAGGGCGGAAGTG  
GGAGCTGCGACCGCGCTCCCTGTGAGGTGGGCAAGCGCGAAATGCGCCCTCCGGAGTCAGTCCCTG  
GCAGTCCTGGTGTGCTTGCTTGGGGTGTCCCTGGACGCACGGCGGGAGCAACGTCGCGTCATCACGGAC  
GAGAACTGGAGAGAACTGCTGGAAGGAGACTGGATGATAGAATTATGCCCCGTGGTGCCCTGCTGTCAAAAT  
CTTCAACCGGAATGGGAAAGTTTGCTGAATGGGAGAAGATCTTGAGGTTAATATTGCGAAAGTAGATGTCACA  
GAGCAGCCAGGACTGAGTGGACGGTTATCATAACTGCTCTCCACTATTATCATTGTAAGATGGTGAATT  
AGGCCTATCAGGGTCCAAGGACTAAGAAGGACTTCATAAAACTTTATAAGTGATAAAGAGTGGAAAGAGTATTGAG  
CCCCTTCATCATGGTTGGTCCAGGTTCTGATGAGTAGTATGTCAGCAGCTTCAGCTATCTATGTGG  
ATCAGGACTTGCCATAACTACTTTATTGAAGACCTGGATTGCCAGTGTGGGATCATATACTGTTTGCTTA  
GCAACTCTGTTCCGGACTGTTATTAGGACTCTGTATGATATTGTGGCAGATTGCCCTTGTCCCTCAAAAGG  
CGCAGACACAGCGTACCCATACCCTCAAAAAAATTATTACAGAATCTGCACAAACCTTGAAAAAGTGGAG  
GAGGAACAAGAGGCAGGATGAAGAAGATGTTCAGAAGAAGCTGAAAGTAAAGAAGGAACAAACAAAGACTT  
CCACAGAATGCCATAAGACAAACGCTCTGGTCCATCATTGCCACAGATAAATCCTAGTTAAATTATAGTT  
ATCTTAATATTGATTTGATAAAACAGAAGATTGATCATTGTTGGTTGAAGTGAACTGTGACTTTTT  
GAATATTGCAGGGTTCACTAGATTGTCATTAAATTGAAGAGTCTACATTAGAACATAAA

73/105  
**FIGURE 72**

MAPSGSLAVPLAVLVLLWGAPWTHGRRSNVRVITDENWRELLEGDWMIEFYAPWCPACQNLQPEWESFAEWGED  
LEVNIAKVDVTEQPGLSGRFIITALPTIYHCKDGEFRRYQGPRTKKDFINFISDKEWKSIEPVSSWFGPGSVLMS  
SMSALFQLSMWIRTCHNYFIEDLGLPVWGSYTVFALATLFSGLLLGLCMIFVADCLCPSKRRRPQPYPYPSKKLL  
SESAQPLKKVEEEQEADEEDVSEEEAESKEGTNKDFPQNAIRQRSLGPSLATDKS

74/105  
**FIGURE 73**

GC GG CG GG CAG CG GG GAC GG GAC AT GG AG AG CG GG CCT AC GG CG GG CA AG GC GG CG CT CCT CG  
AC CT CG GG CG CCT CCT GAC G CAG CG CAG GT GG TG CG CG CG CG TG CT TG CT TG CT CG CCT TG AT CG TG TT CT  
CCT GC AT CT AT GG TG AG GG CT AC AG CA AT GC CC AC GAG T CT AG CAG AT GT ACT CG TG TT CA AC CG CA AC GAG G  
AT GC CT GC CG CT AT GG CAG TG CC AT CG GG TG CT GG CCT CCT CG CG CCT TT CT TG CT TG CG AC CG CG  
AT TT CC CC CA GAT CAG CA AC GC CA CT GAC CG CA AG T AC CT GG TG AT TG AC CT TG CT TT CT CAG CT CT TG GA  
CCT CCT GT GG TT GT GG TT CT G CT CCT CAC CA ACC AG TG GG CG AG TG AC CA AC CC GA AG GG AC GT G CT GG TG  
GG CG CA CT TG GAG GG CAG CC AT CA CCT CAG CT CCT TT CCT AC TT CT CT GG TG TG CT GG CCT CCT CG  
CCT ACC AG CG CT AC AAG G CT GG CG TG GAC GACT C AT CC CAG A AT AC G TT GAC CC AC CCT CAC CC AG A AC G CG GAG ACC ACC G  
CCT AC G C CT CCT AC CC CAG GT G C AT CT GT GG AC A AC T AC CA AC AG CC ACC CCT CAC CC AG A AC G CG GAG ACC ACC G  
AG GG CT ACC AG CC G C C C C T GT G T A C T TG AG CG GG TT AG CG TG GG A AG GG GG AC AG AG AG GG CC C T C C C T CT G  
CC CT GG ACT TT CC CAT GAG C CT CCT GG A ACT G CC AG CC C T CT TT CAC CT GT TG CAG CT GAC AC  
AC AG CT AAG G AG C CT CAT AG C CT GG CG GG CT GG CAG AG CC AC ACC CA AG T G C CT GT G CC CAG AG GG G CT T CAG  
TC AG CG CT C ACT CCT CC CAG GG C ACT TT TAG G A A AG G T T T AG C T AG T G T T T C CT CG CTT A AT G AC CT C  
AG CC CC G C CT TG CAG TG G C T A G A AG CC AG CAG GT G C C AT GT G C T ACT G AC A AG T G C CT CAG CT C C C C C G G C C C  
GG GT CAG G C C G T G G G AG C C G T ATT AT CT G C GT T C T G C C A A AG A C T C G T G G G G C C A T C A C A C C T G C C C T G T G  
CAG CG GAG C C G G ACC AG G C T C T G T G C C T C A C T CAG G T T G C T T C C C T G T G C C C A C T G C T G T A T G A T C T G G G G  
GCC ACC ACC CT GT G C C G G T G G C C T C T G G C T G C C T C C C G T G G T G A G G G C G G G G C T G G T G C T A T G G C A C T T C C  
T C C T G C T C C C A C C C T G G C A G C A G G G A A G G G C T T T G C C T G A C A A C A C C C A G C T T T A T G T A A A T A T C T G C A G T T  
G T T A C T T A G G A A G C C T G G G A G G G C A G G G G T G C C C A T G G C T C C A G A C T C T G T C T G T G C C G A G T G T A T T A T A A A  
A T C G T G G G G A G A T G C C C G G C T G G G A T G C T G T T T G G A G A C G G A A T A A A T G T T T C C T A T T C A G T A

75/105  
**FIGURE 74**

MESGAYGAAKAGGSFDLRRFLTQPQVVARAVCLVFALIVFSCIYGE<sup>Y</sup>GE<sup>Y</sup>SNAHESKQMYCVFNRNEDACRYGSAIG  
VLAFLASAFFLVVDAYFPQISNATDRKYLVIGDLLFSALWTFLWFVGFCFLTNQWAVTNPKDVLVGADSVRAAIT  
FSFFSIFS<sup>W</sup>GVLASLAYQRYKAGVDDFIQNYVDPTPDPNTAYASYPGASVDNYQQPPFTQNAETTEGYQPPP<sup>V</sup>Y

76/105  
**FIGURE 75**

CAGAGCTGCTGTCATGGCGGCCGCTCTGTGGGGCTTCTTCCGCTGCTGCTGCTGCTATCGGGGGATGT  
CCAGAGCTCGGAGGTGCCCGGGCTGCTGCTGAGGGATCAGGGAGTGGGAGTGGGTCGGCATAGGAGATCGCTCAA  
GATTGAGGGCGTGCAGTTGTTCCAGGGTGAAGCCTCAGGACTGGATCTCGGCGGCCGAGTGCCTGGTAGACGG  
AGAAGAGCACGTCGGTTCTTAAGACAGATGGGAGTTTGTTGTTCATGATATACTCTGGATCTTGTAGT  
GGAAGTTGATCTCCAGTTACAGATTGATCCCCTCGAGTGGATATCAGTCGAAAGGAAAATGAGAGCAAG  
ATATGTGAATTACATCAAAACATCAGAGGTTGTCAGACTGCCCTATCCTCTCCAAATGAAATCTCAGGTCCACC  
TTCTTACTTTATTAAAAGGAATCGTGGGCTGGACAGACTTCTAATGAACCCAATGTTATGATGATGGTCT  
TCCTTATTGATATTGTGCTCTGCCTAAAGTGGTCAACACAAAGTGATCCTGACATGAGACGGAAATGGAGCA  
GTCAATGAATATGCTGAATTCCAACCATGAGTTGCCTGATGTTCTGAGTTCATGACAAGACTCTCTTCAA  
ATCATCTGCCAAATCTAGCAGCGGCAGCAGTAAAACAGGCAAAAGTGGGCTGGCAAAAGGAGGTAGTAGGCCG  
TCCAGAGCTGGCATTGCACAAACACGGCAACACTGGTGGCATCCAAGTCTGGAAACCGTGTGAAGCAACTA  
CTATAAAACTTGAGTCATCCGACGTTGATCTTACAACGTGTATGTTAACCTTTAGCACATGTTGTACTT  
GGTACACGAGAAAACCCAGTTCATCTTGCTGTATGAGGTCAATATTGATGTCACTGAATTAAATTACAGTG  
TCCTATAGAAAATGCCATTAATAAATTATGAACACTACATACATTATGTATATTAATTAATTAACATCTTAATCCA  
GAAAAAAA

77/105  
**FIGURE 76**

MAAALWGFFPVLLLLLSDVQSEVPGAAEGSGGSVGIGDRFKIEGRAVPGVKPQDWISAARVLVDGEEHV  
GFLKTGDSFVVHDIPSGSYVVEVSPAYRFDPVRVDIRSKGKMRARYVNYIKTSEVVRLPYPLQMKSSGPPSYFI  
KRESWGWTDFLMNPMVMMVLPLLIFVLLPKVVNTSDPDMRREMEQSMNMLNSNHELPDVSEFMTRLFSSKSSGK  
SSSGSSKTGKSGAGKRR

78/105  
**FIGURE 77**

TGGGACTTATAGAAGGGAGAGGAGCGAACATGGCAGCGCGTGGCGGTTGGTGTCTCTGTGACCATGGTGG  
TGGCGCTGCTCATGTTGCGACGTTCCCTCAGCCTCTGCCAAAGAAAGAAGGAGATGGTGTATCAGAAAAGG  
TTAGTCAGCTGATGGAATGGACTAACAAAGACCTGTAATAAGAATGAATGGAGACAAGTCCGTCGCCCTGTGA  
AAGCCCCACCGAGAAATTACTCGTTATCGTCATGTTACTGCTCTCCAACGTGCATAGACAGTGTGCGTTGCA  
AGCAAGCTGATGAAGAATTCCAGATCCTGGCAAACCTCTGGCGATACTCCAGTCATTCAACCAACAGGATATTT  
TTGCCATGGTGGATTGATGAAGGCTCTGATGTATTTCAGATGCTAACATGAATTCACTGCTCCAACTTCATCA  
ACTTCTGCAAAGGGAAACCAAACGGGTGATACATATGAGTTACAGGTGCGGGGTTTCAGCTGAGCAGA  
TTGCCCGGTGGATGCCGACAGAACTGATGTCATATTAGACTGATTAGACCCCCAAATTATGCTGGTCCCTTA  
TGTGGGGATTGCTTGGCTTATTGGTGGACTTGTGTATCTCGAAGAAGTAATATGGAATTCTCTTTAATA  
AAACTGGATGGGCTTTGCAGCTTGTGTTGTGCTATGACATCTGGCAAATGTGGAACCATATAAGAG  
GACCACCATATGCCATAAGAACCCCCACACGGGACATGTGAATTATATCCATGGAAGCAGTCAGCCCAGTTG  
TAGCTGAAACACACATTGTTCTGTTAATGGTGAGTTACCTAGGAATGGTGCTTGTGAAAGCTGCTA  
CCTCTGACATGGATATTGAAAGCGAAAGATAATGTTGTTGCTGGTATTGGACTTGTGATTATTCTCAGTT  
GGATGCTCTCATTTAGATCTAAATATCATGGCTACCCATACAGCTTCTGATGAGTAAAAAGGTCCCAGAG  
ATATATAGACACTGGAGTACTGAAATTGAAAACGAAAATCGTGTGTTGAAAAGAAGAATGCAACTTGTAT  
ATTCTGTATTACCTCTTCAAGTGATTAAATAGTTAATCATTTAACCAAAGAAGATGTGTTGCTAA  
CAAGCAATCCTCTGTCAAAATCTGAGGTATTGAAAATAATTACCTCTTAACCTCTTCCCAGTGAACCTTA  
TGGAACATTAAATTAGTACAATTAAGTATATTAAAATTGAAAACTACTACTTTGTTTAGTTAGAACAAA  
GCTCAAAACTACTTAGTTAACTGGTCATCTGATCTTATATTGCCTTATCCAAGATGGGAAAGTAAGTCTG  
ACCAGGTGTTCCCACATATGCCTGTTACAGATAACTACATTAGGAATTCTTAGCTTCTCATCTTGTG  
GATGTGTATACCTTACGCATCTTCTTTGAGTAGAGAAATTATGTTGTCATGTGGTCTCTGAAAATGGAAC  
ACCATTCTCAGAGCACACGTCTAGCCCTCAGCAAGACAGTTGTTCTCCTCCTGCATATTCTACTGCG  
CTCCAGCCTGAGTGATAGACTCTGTCCTAAACAGGATTATAATTCTG  
TGAGTATGGTGTAACTACCTGTATTAGAAAGATTTCAGATTCCATCTCTTAGTTCTTTAAGGTG  
ACCCATCTGTGATAAAAATATAGCTTAGTCTAAACGTTGAGGCTCAGTTAGAAAGGACTCCCTGCCAGGC  
TTAGAGTTGTCATTCCATTGTACCTAAGAGAAAATAGGCTCAGTTAGAAAGGACTCCCTGCCAGGC  
GCAGTGACTTACGCCGTAAATCTCAGCACTTGGGAGGCCAGGCAGGCAGATCAGGAGTCAGGAGTTGAGAC  
CATCCTGGCCAACATGGTAAACCCCGTCTACTAAAATATAAAATTAGCTGGGTGTTGAGGAGCCTG  
TAATCCCAGCTGCACAGGAGGCTGAGGCACGAGAACACTGAACAGGAGATGGAGGTTCACTGAGGCCAGA  
TCACGCCACTGCACTCCAGCCTGGCAACAGAGCGAGACTCCATCTAAAAAAAAAAAAAA

79/105  
**FIGURE 78**

MAARWRFWCVSVTMVALLIVCDVPSASAQRKKEMVLSEKVSQLMETNKRPVIRMNGDKFRRLVKAPPRNYSVI  
VMFTALQLHRQCVVCKQADEEFQILANSWRYSSAFTNRIFFAMVDFDEGSDVFQMLNMNSAPTFINFPAKGKPKR  
GDTYELQVRGFSAEQIARWIADRTDVNIRVIRPPNYAGPLMLGLLLAVIGGLVYLRRSNMEFLFNKTGWAFAAALC  
FVLAMTSGQMWNHIRGPPYAHKNPHTGHVNYIHGSQAQFVAETHIVLLFNGGVTLMVLLCEAATSDMDIGKRK  
IMCVAGIGLVVLFFSWMLSIFRSKYHGPYSFLMS

80/105  
**FIGURE 79**

GGGGGAGGCCCGTCGATCCTGGGTTGGAGGAGGTGGCGCGCTGAGGCTGCGCGTGAAGACGGCGGGATG  
GTGGGGCGGGAGAAAGAGCTCTATACACTTGTCCCCGGAGCTGTCGGCTGGTGGAGGAGGAAGTTAACATC  
CCTAATAGGAGGGTTCTGGTTACTGGTGCCACTGGCTCTGGCAGAGCTGTACACAAAGAATTCA  
AATTGGCATGCAGTTGGCTGTGGTTTCAAAGAGCAAGACCAAAATTGAACAGGTTAATCTGGGATTCTAAT  
GCAGTTCATCACATCATGATTTCAGCCCCATGTTATAGTACATTGTGCAGCAGAGAGAAGACAGATGTT  
GTAGAAAATCAGCCAGATGCTGCCCTCAACTTAATGTGGATGCTCTGGGAATTAGCAAAGGAAGCAGCTGCT  
GTTGGAGCATTCTCATCTACATTAGCTCAGATTATGTTGATGGAACAAATCCACCTTACAGAGAGGAAGAC  
ATACCAAGCTCCCCTAAATTGTATGGCAAAACAAAATTAGATGGAGAAAAGGCTGCTGGAGAACAAATCTAGGA  
GCTGCTGTTGAGGATCCTATTCTGTATGGGAAGTTGAAAGCTGAAGAACAGTGCAGTGA  
GATAAAAGTGCAGTCAGCAACAAGTCAGCAAACATGGATCACTGGCAGCAGAGGTTCCCACACATGCAAAGAT  
GTGGCCACTGTGCGCCAGCTAGCAGAGAAGAGAATGCTGGATCCATCAATTAGGAAACCTTCACTGGTCT  
GGCAATGAACAGATGACTAAGTATGAAATGGCATGTGCAATTGCAAGATGCCTCAACCTCCCAGCAGTC  
AGACCTATTACTGACAGCCCTGTCTAGGAGCACACGTCAGGAAATGCTCAGCTGACTGCTCAAATTGGAG  
ACCTTGGGCAATTGGCAACGAACACCATTGCAATTGGAATCAAAGAACACTTGGCCTTCCTCATTGACAAG  
AGATGGAGACAAACGGTCTTCATTAGTTTATTGTGTTGGGTTCTTTTTAAATGAAAAGTATAGTATG  
TGGCCCTTTAAAGAACAAAGGAATAGTTGTATGAGTACTTTAATTGACTCTAGGATCTCAGGTAA  
ATGATGCTCTGCACTAGTGAAATTGCTAAAGAAACTAAAGGGCAGTCATGCCCTGTTGCAAGTAATTTC  
TTTATCATTATGTTGCTGGCTAAACTTGGAGTTGAGTATAGTAAATTGATCCTAAATATTGAGGGTC  
AGGATGAAGCAGATCTGCTGTAGACTTTCAAGATGAAATTGTCATTCTCGTAACCTCCATATTCAGGATTT  
TGAAGCTGTTGACCATTGATTATTAAATTGTGGAATAGTATAAAATCATTGGTGTTCATTATT  
TGCTTGCCCTGAGCTCAGATCAAATGTTGAAGAACAGGAACTTATTGCAAGTTACGTACAGTTTATGC  
TTGAGATATTCAACATGTTATGTATATTGAACTTCTACAGCTGATGCCTCTGCTTTATAGCAGTTATGG  
GGAGCACTGAAAGAGCGTGTACATGTTTATTGCAAGTGTGATATTGATATGTTGAGGTTATACTCATAAT  
TATAAAATATAACTGTCCTTTCATCCCAGTTGCGCTAAGTGTGATATTGATATGTTGAGGTTATACTCATAAT  
AATGGGCCTGTAAGTCTTCACTGAAATATAAAATATGACTGCTGGCATGTAATGCTTAGTTT  
CTGTATTACTCTTTAAATGTAAGGACCAACTCTAAACTAATTGTTCTTGTGCTTAATT  
AAAAATTACATTCTGTGATGTAACATGTGATACATACAAAAGAACATAGTTAATATGATTGAAATAAACAC  
AATAAAATTAAAAAAAAAAAAAA

81/105  
**FIGURE 80**

MVGREKELSIHFVPGSCLRVEEEVNIPNRRVLVTGATGLLGRAVHKEFQQNNWHAVGCGFRRARPKEQVNLLDS  
NAVHHIIHDFQPHVIVHCAAERRPDVVENQPDAAQLNVDASGNLAKEAAAVGAFLIYISSDYVFDGTNPPYREE  
DIPAPLNLYGKTKLGEKAVLENNLGAAVLRIPILYGEVEKLEESAVTVMFDKVQFSNKSANMDHWQQRFPTHVK  
DVATVCRQLAEKRMLDPSIKGTFHWSGNEQMTKYEMACAIADAFNLPSSHLRPITDSPVLGAQRPRNAQLDCSKL  
ETLGIGQRTPFTRIGIKESSLWPFLIDKRWQTVFH

82/105  
**FIGURE 81**

GAAGCAGCAGGTACCCCTCCACATCCCTAGGGCTCTGTGATGTAGGCAGAGGCCGTGGGAGTCAGCATGCCGC  
GTGGCTGGGCCGCCCTGCTCTGCTGCTCCAGGGAGGCTGGGCTGCCCGACCTCGCTGCTACACCG  
ATTACCTCCAGACGGTCATCTGCATCCTGAAATGTGGAACCTCCACCCAGCAGCCTCACCCCTACCTGGCAAG  
ACCAGTATGAAGAGCTGAAGGACGAGGCCACCTCCTGAGCCTCCACAGGTGGCCACAATGCCACGCATGCCA  
CCTACACCTGCCACATGGATGTATTCACTTCATGGCGACGACATTTCACTGGCTGAGAGCATCAAGCCGGCTCCCTTCACACGACTG  
GCAACTACTCCCAGGAGTGTGGCAGCTTCCTGGCTGAGAGCATCAAGCCGGCTCCCTTCACACGACTG  
TGACCTTCTCAGGACAGTATAATATCTCCTGGCGCTCAGATTAGAAGACCCCTGCCTTACATGCTGAAGGGCA  
AGCTTCAGTATGAGCTGCAGTACAGGAACCGGGAGACCCCTGGGCTGTGAGTCCGAGGAGAAAGCTGATCTCAG  
TGGACTCAAGAAGTGTCTCCCTCCTCCCCCTGGAGTTCCGAAAGACTCGAGCTATGAGCTGCAGGTGCGGGCAG  
GGCCCATGCCCTGGCTCCTCCTACAGGGGACCTGGAGTGAATGGAGTGACCCGGTACATTTCAAGACCCAGTCAG  
AGGAGTTAAAGGAAGGCTGGAACCTCACCTGCTGCTCTCCTCCTGCTGTACAGTCTCATTCTGCCCT  
GGAGCCTGAAGACCCATCATTGTGGAGGCTATGGAAGAAGATATGGGCCGCCCCAGCCCTGAGGGTTCTCA  
TGCCCCCTGTACAAGGGCTGCAGGGAGACTCAAGAAATGGGTGGGTGACCCCTCACTGGCTCCAGCCTGGAGC  
TGGGACCCCTGGAGCCCAGAGGTGCCCTCCACCCCTGGAGGTGACAGCTGCCACCCACCGAGCCGGCAAGA  
GGCTGCAGCTCACGGAGCTACAAGAACAGCAGAGCTGGTGGAGTGTGACGGTGTGCCAAGCCCAGCTCTGGC  
CGACAGCCCAGAACTCGGGGGCTCAGCTTACAGTGAGGAGAGGGATGGCCATACGCCCTGGTGTCCATTGACA  
CACTGACTGTGCTAGATGCGAGAGGGCCATGCACTGGCCCTGCAGCTGTGAGGATGACGGCTACCCAGCCCTGG  
ACCTGGATGCTGGCCTGGAGCCCAGCCAGGCCTAGAGGACCCACTCTGGATGCAAGGGACCACAGTCTGCTC  
GTGGCTGTCTCAGCTGGCAGCCCTGGCTAGGAGGGCCCTGGAGCAGACTAAAGCCACCC  
TTGCAGATGGGAGGACTGGCTGGGGACTGCCCTGGGTGGCCGGTACCTGGAGGGTCTCAGAGAGTGAGG  
CGGGCTACCCCTGGCCGCTGGATATGGACACGTTGACAGTGGCTTGTGGCTCTGACTGCAGCAGCCCTG  
TGGAGTGTGACTTCACCAGCCCCGGGACGAAGGACCCCCCGGAGCTACCTCCGCCAGTGGTGGTCAATTCTC  
CGCCACTTCGAGCCCTGGACCCAGGCCAGTAATGAGGCTGACTGGATGTCAGAGCTGCCAGGCCACTGG  
CCCTGAGCCAGAGACAAGGTACCTGGCTGTGATGTGAAGACACCTGCAGCCTTGGTCTGATGGCT  
TGAGCCTGATTTACAGTGTCTGTGATGTGATGTGATGTGATGTGATGTGATGTGATGTGATGTG  
GTGTCTTAGGTGCGCAGTGGCATGTCCACGTGTGATGTGATTGACAGTGGCCTGTGGGCTGGATAATGCCCAT  
GGTACTCCATGCATTACCTGCCCTGTGCATGTGACTCACGGAGCTACCCATGTGCACAAGTGTGCACAGT  
AAACGTGTTGTGGTCAACAGATGACAACAGCGTCTCCCTCAGGGCTTGTGTTGCAAGTGGTCCACAGC  
ATCTCCGGGGTTGTGGGATCAGGGCATTCGCTGTGACTGAGGGGGAGCCAGCCCTCAGCGTCTGCCCTCAG  
GAGCTGCAAGAAGTCCATATTGTTCTTACCTGCCAACAGGAAGCGAAAGGGATGGAGTGAGCCCATGGT  
ACCTGGGAATGGCAATTGGCGGCCCTGGACGAAGGTCTGAATCCGACTCTGATACCTCTGGCTGTG  
CTACCTGAGCCAAGTCGCCCTCCCTCTGGCTAGAGTTCTTATCCAGACAGTGGGAAGGCATGACACACC  
TGGGGAAATTGGCGATGTACCCGTGTACGGTACGCAGCCAGAGCAGACCCCTCAATAAACGTCACTGCCCT  
CTTCTGCCAGAGCCAGGGCGGGGGTGGAGAACATCAATCGTACGCAGCAGCCTGGCACCAGGGCC  
GTCCCGCCTGCAGAGGGCACTCGGGGGTTCCAGGCTAAATCAGTCCGTTCTGCTCTGGAAACAGCTC  
CCCACCAACCAAGATTCTTCTAACTCTGCTACTAAGTTTAAATCCCTTATGCACCCAAGAGATA  
TTTATTAAACACCAATTACGTACGCCATGGCTCATGGGACCCACCCCGTGGCACTCATGGAGGGGCTGC  
AGGTTGGAACATATGCAGTGTGCTCCGGCCACACATCCTGCTGGGCCCTACCCCTGCCCAATTCAATCCTGCCA  
ATAAACTCTGTCTTATTGTTCATCCTGGAGAATTGA

83/105  
**FIGURE 82**

MPRGWAAPLLLLLQGGWGCPDLVCYTDYLQTVICILEMWNLHPSTLTLWQDQYEELKDEATSCSLHRSAHNAT  
HATYTCHMDVFHFMADDIFSVNITDQSGNYSQECGSFLLAESIKPAPPFNVTVTFSQQYNISWRSDYEDPAFYML  
KGKLQYELQYRNRGDPWAVSPRRKLISVDSRSVSLPLEFRKDSSYELQVRAGPMPGSSYQGTWSEWSDPVIFQT  
QSEELKEGWNPHLLLLLIVFIPAFWSLKTHPLWRLWKKIWAVPSPERFFMPLYKGCSGDFKKWVGAPFTGSS  
LELGPWSPEVPSTLEVY SCHPPRSPA KRLQLTELQEP AELVESDGVPKPSFWPTAQNSGG SAYSEERDRPYGLVS  
IDTVTVLDAEGPCTWPCSCEDDGYPALDLDAGLEPSPGLEDPLLAGTTVLSCGCVSAGSPGLGGPLGSLLDRLK  
PPLADGEDWAGGLPWGGRSPGGVSESEAGSPLAGLDMDTFDSGFVGSDCSPVECDFTSPGDEGP PRSYLRQWVV  
IPPLSSPGPQAS

84/105  
**FIGURE 83**

GGCACGAGGAGGTGTGGACGCTGTGTATGAAATGTCTTCCAGGACCCAAGTTCTCACCATGGGGATGT  
GTCCATTGGTGCAGGAGCCCTGGGGCTGCTGCCCTGGCATTGCTGCTGCAACACAGACGTGTTCTGTCAA  
GCCCCAGAAAGCGGCCCTGGAGTACCTGGAGGATATAGACCTGAAAACACTGGAGAAGGAACCAAGGACTTCAA  
AGCAAAGGAGCTATGGAAAAAAATGGAGCTGTATTATGGCGTGCAGGCCAGGCTGTTCTGTGAGA  
GGAAGCTGCGGATCTGCTCTCCCTGAAAAGCATGTTGGACCAGCTGGCGTCCCCCTATGCAGTGGTAAAGGA  
GCACATCAGGACTGAAGTGAAGGATTCCAGCCTTATTCAAAGGAGAAATCTCCTGGATGAAAAGAAAAAGTT  
CTATGGTCCACAAAGGCCAGATGATGTTATGGGATTATCCGTCTGGAGTGTGGTACAACCTCTCCGAGC  
CTGGAACGGAGGCTCTGGAAACCTGGAAGGAGAAGGCTTCATCCTGGGGAGTTTGTGGTGGGATCAGG  
AAAGCAGGGCATTCTTCTTGAGCACCGAGAAAAAGAATTGGAGACAAAGTAAACCTACTTCTGTTCTGGAAGC  
TGCTAAGATGATCAAACCACAGACTTGGCCTCAGAGAAAAAAATGATTGTGTGAAACTGCCAGCTCAGGGATAA  
CCAGGGACATTCACCTGTTCATGGGATGTTATTGTTCCACTCGTGTCCCTAAGGAGTGAGAAACCCATTATA  
CTCTACTCTCAGTATGGATTATTAATGTATTAAATATTCTGTTAGGCCACTAAGGAAAATAGCCCCAAAC  
AAGACTGACAAAAATCTGAAAAACTAATGAGGATTATTAAGCTAAAACCTGGGAAATAGGAGGCTAAATTGAC  
TGCCAGGCTGGTGCAGTGGCTCACACCTGTAATCCCAGCACTTGGGAGGCCAAGGTGAGCAAGTCACTGAGG  
TCGGGAGTTCGAGACCAGCCTGAGCAACATGGCGAACCCCCGTCTACTAAAATACAAAATCACCAGGTGT  
GGTGGCAGGCACCTGTAGTCCCAGCTACCCGGAGGCTGAGGCAGGAGAATCAGTGAACCTGGGAGGTGGAGGT  
TGCAGTGAACCTGAGACTCACACCACTGTATTCCAGCTGGGTGACTGAGACTCTAACTAAAAAAAAAAAAAAA  
AAA

85/105  
**FIGURE 84**

MWSTIGAGALGAAALALLANTDVFLSKPQKALEYLEDIDLKTLEKEPRTFKAKELWEKNGAVIMAVRRPGCFLC  
REEAADLSSLKSMLDQLGVPLYAVVKEHIRTEVKDFQPYFKGEIFLDEKKKFYGPQRRKMMFMGFIRLGWYNFF  
RAWNGGFSGNLEGEgefILGGVFVVSGKQGILLEHREKEFGDKVNLLSVLEAAKMIKPQTLASEKK

86/105  
**FIGURE 85**

GGCACGAGGCCTCGTGCAGCCGGCTTGGTACCTCAGCGAGCGCCAGCGTCCGGCCGGCTGGCTATGTT  
CGTGCCTCGAAAGAGTTACGAGGTGGTCCAGAGCCAGGGCTCTCTCGTGGCCTCGGACGT  
GGATGCTCTGTGCGTGCAGATCCTCAGGCCTTGGTACCTGACCGTGAATATACGCTGGTTCCAGT  
TTCTGGGTGGCAAGAACCTGAAACTGCATTCTGAGCATAAAGAACAGTTTCAATTATTCTCATAAACTG  
TGGAGCTAATGTAGACCTATTGGATATTCTCAACCTGATGAAGACACTATATTCTTGATGACACCCATAG  
GCCAGTCAATGTCGTCATGTATACAACGATACCCAGATCAAATTACTCATTAAACAAGATGATGACCTGAGT  
TCCCGCCTATGAAGACATCTCAGGGATGAAGAGGAGGATGAAGAGCATTAGGAAATGACAGTGTGGTCAGA  
GCCTCTGAGAAGCGCACCGTTAGAAGAGGAGATAGTGGAGCAAACCATGCGGAGGAGGCAGCGCAGAGTG  
GGAGGCCGGAGAAGAGACATCCTCTTGACTACGAGCAGTATGAATATCATGGGACATCGTCAGCCATGGTGT  
GTTTGAGCTGGCTGGATGCTGTCAGGACCTGAATGACATGCTGTGGTGGCCATCGTGGACTAACAGACCA  
GTGGGTGCAAGACAAGACTCAAATGAAATACGTGACTGATGTTGGTGTCTGCAGCGCCACGTTCCGCCA  
CAACCACCGAACGAGGATGAGGAGAACACACTCTCCGTGACTGCACCGATCTCCTTGAGTATGACCTCCG  
CCTGGTGTCTACCAGCACTGGCCCTCATGACAGCGTGCACACCCAGCTATACCGCAGCCAGGTTCAAGCT  
GTGGTGTGCAATGGACAGAACGGCTCCAGGAGTTCCCTGCAGACATGGGCTTCCCTGAAGCAGGTGAAGCA  
GAAGTCCAGGCCATGGACATCTCCTGAAGGAGATTGCGGGAAATGATTGAAGAGTCTGCAAATAATTGG  
GATGAAGGACATGCGCGTGCAGACTTCAGCATTGTTGGCTCAAGCACAAGTTCTGGCAGCGACGTGGT  
CTTGCCACCATGTTGAGAGCCCCGAGAAGGATGGCTCAGGGACAGATCACTCATCCAGGCTCTGGA  
CAGCCTCTCAGGAGTAACCTGGACAAGCTGACCATGGCTGGAACCTGCCAAGAAGCAGCTGCGAGCCACCCA  
GCAGACCATGCCAGCTGCTTGCACCAACCTGTCATCTCCAGGGCCTTCTGTACTGCTCTCATGGA  
GGCACTCCAGATGTCATGCTGTTCTAGGCCGGCATCCCTAAGCCTGTCAGCAAACACCTGCTCAAGTCCTT  
TGTGTGTTGACAAAGAACCGGCCGTGCAAACCTGTCGCCCCCTGGTGTGGCTGCCCCCTGAGCATGGAGCATGG  
CACAGTGACCGTGGTGGCATCCCCCAGAGACCGACAGCTGGACAGGAAGAACTTTTGGGAGGGCGTTGA  
GAAGGCAGGGAAAGCACCAGCTCCGGATGCTGCACAACCATTGACCTCTCAGTAATTGAGCTGAAAGCTGA  
GGATCGGAGCAAGTTCTGGACGCACCTATTCCCTCCTGTCTTAGGAATTGATTCTCCAGAATTGACCTTCTT  
ATTATGTAACTGGCTTCATTTAGATTGTAAGTTATGGACATGATTGAGATGAGTAAAGCCATTTTTATTAAA  
TAAAATGCTTATTAGGCTCCGCCCCAAAAAAAAAAAAAAAAAAAAAA

87/105  
**FIGURE 86**

MFVSDFRKEFYEVVQSQRVLLFVASDVLACACKILQALFQCDHVQYTLVPVSGWQELETAFLEHKEQFHYFILI  
NCGANVDLLDILQPDEDTIFFVCDTHRPNVVNVYNDTQIKLLIKQDDDLEVPAYEDIFRDEEDEHSGNDSDG  
SEPESEKRTRLEEEIVEQTMRRQRREWEARRDILFDYEQYEYHGTSSAMVMFELAWMLSKDLNDMLWWAIVGLT  
DQWVQDKITQMKYVTDVGVLQRHVSRRHNEDEENTLSVDCTRISFEYDLRIVLYQHWSLHDSDLNTSYTAARF  
KLWSVHGQKRLQEFLADMGLPLKQVKQKFQAMDISLKENLREMIEESANKFGMKDMRVQTFSIHFGFKHKFLASD  
VVFATMSLMEESPEKDGSQTDHFIQALDSLRSRNLDKLYHGLELAKKQLRATQQTIASCLCTNLVISQGPFLYCSL  
MEGTPDVMLFSRPASLSSLKHLKSFVCSTKNRRCKLLPLVMAAPLSMEHGTVTVVGIPPETDSSDRKNFFGRA  
FEKAAESTSSRMLHNHFDLSVIELKAEDRSKFLDALISLLS

88/105  
**FIGURE 87**

ACCGGGCACCGGACGGCTGGGTACTTCGTCTTAATTAGTCATGCCGTGAGCCAGGAAGGGCTGTGTT  
TATGGGAAGCCAGTAACACTGTGCCACTATCTCTTCCGTGGTGCACATCTACATTTGGACTCGGAATTAT  
GAGGTAGAGGTGGAGGCGGAGCCGGATGTCAGAGGTCTGAAATAGTCACCATGGGGAAAATGATCCGCCTGCT  
GTGAAAGCCCCCTCTCATTCCGATCGCTTTTGGCCTGATGATTGAAAATAAGTCCTGTTGCACCAGATGCA  
GATGCTGTTGCTGCACAGATCCTGCACTGCTGCCATTGAAGTTTCCAATCATCGTCATTGGATCATTGCA  
TTGATATTAGCACTGGCCATTGGTCTGGGCATCCACTCGACTGCTCAGGGAAAGTACAGATGTCGCTCATCCTT  
AAGTGTATCGAGCTGATAGCTCGATGTGACGGAGTCTCGGATTGCAAAGACGGGGAGGACGAGTACCGCTGTGTC  
CGGGTGGGGGTGAGAATGCCGTGCTCCAGGTGTTCACAGCTGCTCGGAAGACCATGTGCTCCGATGACTGG  
AAGGGTCACTACGCAAATGTTGCTGTGCCAACTGGGTTCCAAGCTATGTGAGTTAGATAACCTCAGAGTG  
AGCTCGCTGGAGGGGCAGTCCGGGAGGAGTTGTGTCATCGATCACCTTGCAGATGACAAGGTGACTGCA  
TTACACCACTCAGTATATGTGAGGGAGGGATGTGCCTCTGCCACGTGGTTACCTTGCAGTGCACAGCCTGTGGT  
CATAGAAGGGCTACAGCTACGCATCGTGGGTGAAACATGTCCTGCTCGCAGTGGCCCTGGCAGGCCAGC  
CTTCAGTTCCAGGGCTACCCACCTGTGCGGGGCTCTGTCATCACGCCCTGTGGATCATCACTGCTGCACACTGT  
GTTTATGACTTGTACCTCCCCAAGTCATGGACCATCCAGGTGGGTCTAGTTCCCTGTTGACAATCCAGCCCCA  
TCCCACTTGGGAGAAGATTGTCACACAGCAAGTACAAGCCAAAGAGGCTGGCAATGACATGCCCTTATG  
AAGCTGGCCGGGCCACTCACGTTCAATGGTACATCTGGGTCTCTATGTGGTCTGCAAGCTTCTGGTCAA  
GAGGATTTGCAATTGCTCATTGAAGCATTCTTATGATGGCTGCTTATAATCCTGTCAGATATTAAATAATTCA  
ACTCCTGATTGATGTTGGTGTGGCATCAGTGATTCTTCTCATTAATGATGCTCCTAAAAAA  
AAAAAA

89/105  
**FIGURE 88**

MGENDPPAVEAPFSFRSLFGLDDLKISPVAPDADAAQILSLLPLKFFPIIVIGIIIALILALAIGLGIHFDCSG  
KYRCRSSFKCIELIARCDGVSDCKDGEDEYRCVRVGGQNAVLOVFTAASWKTMCSDDWKGHYANVACAQLGFPSY  
VSSDNLRVSSLEQFQREEFVSIDHLLPDDKVTALHHSVYVREGCASGHVVTIQCACGHRRGYSSRIVGGNMSSL  
SQWPWQASLQFQGYHLCGGSVITPLWIITAHCVYDLYLPKSWTIQVGLVSLLDNPAPSHLVEKIVYHSKYKPKR  
LGNDIALMKLAGPLTFNGTSGSLCGSAALPLFQEDLQLLIEAFL

90/105  
**FIGURE 89**

**ATGGCACAGCACGGGCGATGGGCGCGTTCGGGCCCTGTGCGGCCTGGCGCTGCTGTGCGCGCTCAGCCTGGGT**  
CAGCGCCCCACCGGGGGTCCCGGGTGC GG CCTGGGCCTCCTGCTTGGGACGGAACGGACGCCGCTGCTGC  
CGGGTTCACACGACGCCGCTGCTGCCCGGATTACCCGGCGAGGAGTGTGCTCCGAGTGGGACTGCATGTGTGTC  
CAGCCTGAATTCCACTGCGGAGACCCCTTGCTGCACGACCTGCCGCCACCACCTTGTCCCCCAGGCCAGGGGTA  
CAGTCCCAGGGAAATTCAAGTTGGCTTCCAGTGTATCGACTGTGCCTCGGGACCTCTCCGGGGCCACGAA  
GGCCACTGCAAACCTTGGACAGACTGCACCCAGTCGGGTTCTCACTGTGTTCCCTGGGAACAAGACCCACAAC  
GCTGTGTGCGTCCCAGGGTCCCGCCGGCAGAGCCGCTTGGGTGGCTGACCGTCGTCCCTGCCGTGGCCGCC  
TGCCTCCTCCTGACCTCGGCCAGCTGGACTGCACATCTGGCAGCTGAGGAGTCAGTGCATGTGGCCCCGA  
GAGACCCAGCTGCTGGAGGTGCCCGCTGACCGAAGACGCCAGAAGCTGCCAGTCCCCGAGGAAGAGCGG  
GGCGAGCGATCGGCAGAGGAGAAGGGCGGCTGGAGACCTGTGGGTTGA

91/105  
**FIGURE 90**

MAQHGAMGAFRALCGIALLCALS LGQRPTGGPGCGPGRLLGTGTDARCCR VHTTRCCR DYPGEECCSEWDCMCV  
QPEFHCGDPCCTTCRHHPCPPGQGVQSQGKFSFGFQCIDCASGTFSGGHEGHCKPWTDCTQFGFLTVFPGNKTHN  
AVCVPGSPPAEPPLGWLTVVLLAVAACVLLLTSAQLGLHIWQLRSQCMWPRETQLLLEVPPSTEDARSCQFPEEER  
GERSAEEKGRLGDLWV

92/105  
**FIGURE 91**

AGTGCAGAAGATGCGAAAGGTGTTGATCACCGGGCTAGCAGTGGCATGGCTGGCCCTGCAAGCGGC  
TGCTGGCGGAAGATGAGCTTCATCTGTGTTGGCGTGCAGGAACATGAGCAAGGCAGAACAGCTGTCTGTGCTG  
CTCTGCTGGCCTCTCACCCCCACTGCTGAGGTACCCATTGTCAGGTCAGCAACCTGCAGTCGGTCTTC  
GGGCCTCCAAGGAACCTAACGAAAGGTTAGAGATTAGACTGTATATCTAAATGCTGGGATCATGCCTAAC  
CACAACTAAATATCAAAGCACTTTCTTGCCCTTTCAAGAAAAGTGAATTGATATGTTCTCCACAGCTGAAG  
GCCTGCTGACCCAGGGTATAAGATCACTGCTGATGGACTTCAGGAGGTGTTGAGACCAATGTCTTGCCATT  
TTATCCTGATTGGAACTGGAGCCTCTCCTCTGTACAGTGACAATCCATCTCAGCTCATCTGGACATCATCTC  
GCAGTGCAAGGAAATCTAATTCAGCCTCGAGGACTTCCAGCACAGCAAAGGAAGAACCTACAGCTTCCA  
AATATGCCACTGACCTTTGAGTGTGGCTTGAACAGGAACCTCAACCAGCAGGGTCTTATCTTCAATGTGGCT  
GTCCAGGTACAGCATTGACCATATGAAATTCTGCCTCCGTTATATGGACGCTGTTGATGCCGGCAA  
TATTGCTACTTCGCTTTTGCAAATGCAATTGACACCATAATGAAACAGAAGCTCTGGTATGGCTT  
TCCACCAAAAGCCTGAATCTCTCAATCCTCTGATCAAATATCTGAGTGCCACCACTGGCTTGGAAAGAAATTATA  
TTATGACCCAGAAGATGGACCTAGATGAAGACACTGCTGAAAAATTATCAAAGTTACTGGAACGGAAAGC  
ACATTAGGGTCACTATTCAAAAACAGATAATCAGGCCAGGCTCAGTGGCTCATGCCCTAATTCCAGCACTTG  
GGAGGCCAAGGCAGAAGGATCACCTGAGGACCAAGGAGTCAAGACCAGCCTGAGAAACATAGTGAGCCCTTGTCTC  
TACAAAAGAAATAAAATAAGCTGGGTGTGGCATGCGCATGTAGTCCCAGCTACTCAGAAGGATGAGGT  
GGGAGGATCTTGAGGCTGGAGGCAGAGGTTGCAGTGAGCTGAGATTGTGCCACTGCACTCCAGCCTGGGTA  
CAGCGAGACCTGCTCAAAATATGTATATTTAATATATAAAAACCAGAGCTGACAATGACACTCTGGAA  
CATTGCATACCTCTGTACATTCTGGGTACATGGATTCTACTGAGTTGATAATATGCATTGTAATAAACTA  
TGAACATGAAAAA

93/105  
**FIGURE 92**

MRKVVLITGASSGIGLALCKRLLAEDDELHLCACRNMSKAEAVCAALLASHPTAEVTIVQVDVSNLQSVFRASK  
ELKQRFQRLDCIYLNAGIMPNPQLNIKALFFGLFSRKVIHMFSTAEGLLTQGDKITADGLQEVFETNVFGHFILI  
RELEPLLCHSDNPSQLIWTSSRSARKSNFSLEDFQHSKGKEPYSSSKYATDLLSVALNRNFNQQGLYSNVACPGT  
ALTNLTYGILPPFIWTLMPAILLLRFFANAFTLTPYNGTEALVWLFHQKPESLNPLIKYLSATTGFGRNYIMTQ  
KMDLDEDTAEKFYQKLLELEKHIRVTIQKTDNQARLSGSCL

94/105  
**FIGURE 93**

AGCCTGGGGCGGCCAGGAACCACCGTTAAGGTGTCTCTTCTGGATGGTGGAGGTTGGAAAAAGGCTC  
CTGTAACCCCTCCAGGATGAACCACCTGCCAGAAGACATGGAGAACGCTCTCACCGGGAGCCAGAGCTCCCAT  
GCTTCTCTGCGCAATATCCATTCCATCAACCCCACACAACATGGCCAGGATTGAGTCCTATGAAGGAAGGAA  
AAGAAAGGCATATCTGATGTCAGGAGGACTTCTGTTGTCACCTTGACCTCTTATTGTAACATTACTG  
TGGATAATAGAGTTAAATGTGAATGGAGGCATTGAGAACACATTAGAGAAGGAGGTGATGCAGTATGACTACTAT  
TCTTCATATTTGATATATTCTCTGGCAGTTTCGATTAAAGTGTAAACTTGCAATATGCTGTGAGA  
CTGCGCCATTGGTGGCAATAGCGTTGACAACGGCAGTGACCAGTGCCTTTACTAGCAAAAGTGATCCTTCG  
AAGCTTTCTCAAGGGCTTTGGCTATGTGCTGCCATCATTCACTCCTGCCTGGATTGAGACGTGG  
TTCCCTGGATTTCAAAGTGTACCTCAAGAACGAGAACAGACTCCTGATAGTTAGGATGCTCAGAG  
AGGGCAGCACTTACCTGGTGGCTTCTGATGGTCAGTTTATTCCCTCTGAATCCGAAGCAGGATCTGAA  
GAAGCTGAAGAAAACAGGACAGTGAGAACCACTTTAGAAACTATGAGTACTACTTTGTTAAATGTGAAAAC  
CCTCACAGAAAAGTCATCGAGGCAAAAGAGGCAGGAGTGGAGTCCTCGACAGTAAAGTTGAAATGGTGA  
CGTCCACTGCTGGCTTATTGAACAGCTAATAAAGATTATTGTAATACCTCACAGCAGTGCACCATATC  
CATGCACATTAGTGCCTGCCTGGCTGTAAGGTAATGTCATGATTCACTCCTCTTCAGTGAGACTGAGCC  
TGATGTGTTAACAAATAGGTGAAGAAAGTCTGTGCTGTATTCTTAATCAAAGACTTAATATATTGAAGTAACA  
CTTTTTAGTAAGCAAGATACTTTTATTCAATTACAGAATGGAATTTTGTTCATGTCAGATTAT  
TTTGTATTCTTTAACACTCTACATTCCCTGTTTTACTCATGCACATGTGCTTTGACAGTTA  
AAAAGTGAATAAAATCTGACATGTCAATGTGGCTAGTTTATTCTGTTGCATTATGTGTATGGCCTGA  
AGTGTGGACTGCAAAAGGGAGAAAGGAATTGCGAATACATGAAATGTCACCAGACATTGTATTATT  
TATCATGAAATCATGTTCTGATTGTTCTGAAATGTTCTAAACTCTTATTGAATGCACAAAATGACT  
TAAACCATTCATCATGTTCTGCGCAATTCAATTAAATGAACTAAATTAAAAAA  
AAAAAAAAAAAAAA

95/105  
**FIGURE 94**

MNHLPEDMENALTGSQSSHASLRNIHSINPTQLMARIESYEGREKKGISDVRRTFCLFVTFDLLFVTLLWIELNVNGGIENTLEKEVMQYDYYSSYFDIFLLAVFRFKVLILAYAVCRLRHWWAIALTTAVTSAFLLAKVILSKLFSQGAFGYVLPIISFILAWIETWFLDFKVLPOEAEEENRLLIVQDASERAALIPGGLSDGQFYSPPESEAGSEEAEKQDSEKPLLEL

96/105  
**FIGURE 95**

GTCTGTATTTCAAGGACTCCAAAGCGAGGCCGGGGACTGAAGGTGTGGGTGTCGAGCCCTCTGGCAGAGGGTTAA  
CCTGGGTCAAATGCACGGATTCTCACCTCGTACAGTTACGCTCTCCCGGGCACGTCCGAAGGATTGGAAGTCC  
TGAGCGCTCAAGTTGTCCGTAGTCGAGAGAAGGCCATGGGAGGTGCCGCCACCGGACGCXGGAGCTTCTGT  
AGAGCATTGTGCCTATTCCCCGAGCTTGTGACTGCCGATTCCAAGTCCCTGAGCTGTGACTGCCGATTCCAAGTCCCTGAGCTGT  
CAGAAGCGGCTTCCCTACGTCCAGAGCCATTACCCGAATCTGGATGGGACCGCCTCCGGAGCTGTTGGC  
AAAGATGAACAGCAGAGAATTCAAAGGACCTGCTAATATCTGTAAGACGGCAGCTACAGCAGGCATTGGC  
TGGGTGTATGGGGAATACCAGCTTTATTGATGCTAAACAACATACATTGAGCAGGCCAGGAGAAATTAT  
CATAACCGGTTGATGCTGTGCAATCTGCACATCGTGTGCCACAGGCTTCATTGCTTATGGCTGGCGCTGG  
GGTGGAGAACTGCAGTGTGACTATATTCAACACAGTGAACACTAGTCTGAATGTATACCGAAATAAGAT  
GCCTTAAGCCATTGTAATTGCAAGGACTGTCACGGGAAGTCTTTAGGATAAACGTAGGCCTGCGTGGCTG  
GTGGCTGGTGGCATATTGGAGCCTGCTGGGACTCCTGTAGGAGGCCTGCTGATGCCATTCAAGTACTCT  
GGTGAGACTGTTCAAGGAAAGAAAACAGAAGGATCGAAAGGCCTCCATGAGCTAAACTGGAAGAGTGGAAAGGC  
AGACTACAAGTTACTGAGCACCTCCCTGAGAAAATTGAAAGTAGTTACAGGAAGATGAACCTGAGAATGATGCT  
AAGAAAATTGAAGCACTGCTAACCTCCTAGAAACCCCTCAGTAATAGATAAACAAAGACAAGGACTTGAAAGTGC  
TCTGAACTGAAACTCACTGGAGAGCTGAAGGGAGCTGCCATGTCGATGAATGCCAACAGACAGGCCACTCTT  
GGTCAGCCTGCTGACAAATTAAAGTGTGCTGGTACCTGTGGTGGCAGTGGCTTGCTCTGTCTTTCTTTCTTT  
TAACTAAGAATGGGCTGTTGACTCTCACTTACCTTACCTAAATTAAACATAACTTATGTTGTATTAAAT  
CTATCAATATATGCATACATGAATATATCCACCCACCTAGATTTAAGCAGTAAATAAACATTCGAAAAGAT  
TAAAGTTGAATTTACAGTTAAAAAAAAAAAAAAAAAAAAAA

97/105  
**FIGURE 96**

MEVPPPDAGSFLCRALCLFPRVFAAEAVTADSEVLEERQKRLPVVPEPYYPESGWDRRLRELFKDEQQRISKDLA  
NICKTAATAGIIGWVYGGIPAFIHKQQYIEQSQAIEIYHNRFDAVQSAHRAATRGFIRYGWRWGWRTAVFVTIFN  
TVNTSLNVYRNKDALSHFVIAGAVTGSFLRINVGLRGLVAGGIIGALLGTPVGGLLMAFQKYSGETVQERKQKDR  
KALHELKLEEWKGRLQVTEHLPEKIESSLQEDEPENDAKKIEALLNLPNPSVIDKQDKD

98/105  
**FIGURE 97**

GCACGAGCG**A**TGTCGCTCGTGTGCTAAGCTGGCCGCCGTGTGCAAGGAGGCCGTACCCCGAGGCCGACCGTT  
CAATGTGGCTCTGAAACTGGGCCATCTCCAGAGTGGATGCTACAACATGATCTAATCCCCGGAGACTTGAGGGAC  
CTCCGAGTAGAACCTGTTACAACACTAGTGTGCAACAGGGGACTATTCAATTGATGAATGTAAGCTGGGTACTC  
CGGGCAGATGCCAGCATCCGTTGTGAAGGCCACCAAGATTGTGTGACGGGAAAAGCAACTTCCAGTCCTAC  
AGCTGTGTGAGGTGCAATTACACAGAGCCTCCAGACTCAGACCCAGACCCCTCTGGTGGTAAATGGACATTTCC  
TACATCGGCTTCCCTGTAGAGCTGAACACAGTCTATTCAATTGGGCCATAATATTCTAATGCAAATATGAAT  
GAAGATGGCCCTTCCATGTCTGTGAATTCACCTCACAGGCTGCTAGACCACATAATGAAATATAAAAAAAAG  
TGTGTCAAGGCCGAAGCCTGTGGGATCCGAACATCACTGCTTGTAAAGAAGAATGAGGAGACAGTAGAAGTGAAC  
TTCAACAACCCTCCCTGGAAACAGATACTGGCTTTATCAACACAGCACTATCATGGGTTTCTCAGGTG  
TTTGAGCCACACCAGAAGAAACAAACGCGAGCTTCAGTGGTACTGGGATAGTGAAGGTGCTACG  
GTGCAGCTGACTCCATATTCTACTTGTGGCAGCGACTGCATCCGACATAAAGGAACAGTTGTGCTCTGCCCA  
CAAACAGGCCCTTCCCTGGATAACAAACAAAGCAAGCAGGCCAGGCTGGCTGCCCTCCTCTGCTGTCT  
CTGCTGGTGCCACATGGGTGCTGGCAGGGATCTATCTAATGTGGAGGCACGAAAGGATCAAGAAGACTTCC  
TTTCTACCACCAACTACTGCCCTTCAAGGTCTTGTGGTTACCCATCTGAAATATGTTCCATCACACA  
ATTGTTACTTCACTGAATTCTCAAAACCATTGCAAGAAGTGAAGGTCACTCCTGAAAGTGGCAGAAAAGAAA  
ATAGCAGAGATGGGTCCAGTGCAGTGGCTGCCACTCAAAGAAGGCAGCAGACAAAGTCGTTCCCTTCC  
AATGACGTCAACAGTGTGCGATGGTACCTGTGGCAAGAGCGAGGGCAGTCCAGTGAGAACTCTCAAGACTCT  
TCCCTTGCTT**TAAC**CTTCTGCACTGATCTAAGAAGCCAGATTCTGCACAAATACGTGGTGGTCTACTT  
TAGAGAGATTGATAACAAAGACGATTACAATGCTCTCAGTGTCTGCCCAAGTACCACTCATGAAGGATGCCAC  
TGCTTCTGCAACTCTCATGTCAAGTAGCAGGTGTCAGCAGGAAAAGATCACAGCCTGCCACGATGG  
CTGCTGCTCTGTAGCCCACCCATGAGAAGCAAGAGACCTTAAAGGCTTCTATCCCACCAATTACAGGGAAAA  
AACGTGTGATGATCTGAAGCTTACTATGCAGCCTACAAACAGCCTTAGTAATTAAACATTTATACCAATAAA  
ATTTCAAATATTGCTAACTAATGCACTAACTAACGATTGGAAACTACATTTACAACCTCAAAGCTGTTTA  
TACATAGAAATCAATTACAGTTTAATTGAAAATATAACCATTGATAATGCAACAATAAGCATCTCAGCC  
AAAAAA

99/105  
**FIGURE 98**

MSLVLLSLAALCRSAVPREPTVQCGSETGP SPEWMLQHDLIPGDLRDLRVEPVTTSVATGDYSILMNVSWVLRAD  
ASIRLLKATKICVTGKSNFQSYSCVRCNYTEAFQTQTRPSGGKWTFSYIGFPVELNTVYFIGAHNIPNANMNEDG  
PSMSVNFTSPGCLDHIMKYKKKCVKAGSLWDPNITACKNEETVEVNFTTPLGNRYMALIQHSTIIGFSQVFEP  
HQKKQTRASVVIPTGSEGATVQLTPYFPTCGSDCIRHKGTVVLCPQTGVFPPLDNNKSKGWLPLLLLSLLV  
ATWVLVAGIYLMWRHERIKKTSFSTTLLPIKVLVVYPSEICFHHTICYTEFLQNHCRSEVILEKWQKKKIAE  
MGPVQWLATQKKAADKVVFLLSNDVNSVCDGTCGKSEGSPSENSQDSSPCL

100/105  
**FIGURE 99**

CAGCTTCCCACCCTGGGTTTCCGAGGTGCTTCGCCGCTGCCCCACCACTGCAGCCATGATCTCCTAACCGGA  
CACGCAGAAAATGGAATGGGATTAACAGGATTGGAGTGTTCCTGTTGGAAATGATTCTCTTTTGA  
CAAAGCACTACTGGCTATTGGAAATGTTTATTGTAGCCGGCTGGTTTGTAATTGGTTAGAAAGAACATT  
CAGATTCTCTTCAAAAACATAAAATGAAAGCTACAGGTTTTCTGGGGGGTATTGTAGTCCTATTGG  
TTGGCCTTGATAGGCATGATCTCGAAATTATGGATTTCTCTGTTAGGGGCTTCTTCCTGCGTTGT  
TGGCTTATTAGAAGAGTGCCAGTCCTGGATCCCTCTAAATTACCTGGAAATTAGATCATTGTAGATAAAGT  
TGGAGAAAGCAACAATATGGTATAAACAACAAGTGAATTGAAGACTCATTAAAATATTGTGTTATTATAAAGT  
CATTGAAGAATATTACAGCACAAAATTAAATTACATGAAATAGCTTGTAAATGTTCTTACAGGAGTTAAAACGT  
ATAGCCTACAAAGTACCAGCAGCAAATTAGCAAAGAAGCAGTGAACACAGGCTCTACTCAAGTGAACTAAGAAG  
AAGTCAGCAAGCAAACGTAGAGAGGGTAAATCCATGTTAATGATGCTTAAGAAACTCTGAAGGCTATTGTT  
GTTTTCCACAATGTGCGAAACTCAGCCATCCTAGAGAACTGTGGTGCCTGTTCTTCTTAAATTGAAG  
GCTCAGGAGCATCCATAGGCATTGCTTTAGAAGTGTCCACTGCAATGGCAAAAATATTCCAGTTGCACGT  
ATCTCTGGAAGTGTGATGCAATTGATTGTGTCATTAAAGTATTAAACCAAGGAAACCCCAATT  
GATGTATGGATTACTTTTGTAAACATGGTAAAAATAAAACTTTGTGGGTCTCTGAATCTTAATATTCA  
AAGCCAGGTGAAAATCTGAACTAGATATTCTTGTGGAATATGCAAAGGTCACTTACTAACTTTAGTTAC  
TAAATTATAGCTAAGTGTGCACTCCGAAAGTCTCATACTTCTGGGAGTCTGCCCTCTAAGTA  
TCTGTCTATATCATTACGTGTAAGTATTAAACAAAAAGCATTCTGACCATGAATGAAGTAGTTGTT  
ATAGCTTGTCTCATTGAATAGTATTGAAGATACTAAATGATGCAAACCAAATGGATTTTCCATGTCATGA  
TGTAATTTCCTTCTTCTTAAATTAGCAGTGGCTTATTATTGTTCTAATAATTAAATA  
ACTTTGATAATGTTACTTAAGACATGTAACATGTTAAAGGTTAAACTTATGGCTGTTAAAGGGCTATT  
CATTAAATCTGAGTTCCCTTATTTCAGCTTTCTAGCATATAATAGTCATTAGCATTGACATGACATATCCTCA  
TATGATCACTCATCTGAGTTAATTAGAAAATACCTGAGTTCACGTGCTAAAGTCATTCACTGTAATAACTGA  
CTATGGTTCTTAAGAACATGACACTAAAAAGGGTTTTCCACCGTTGCTGATTATTAGACAGTAGG  
AAATAGCTTTCTTAGTTACAAGATGTGACAGCTTAGTGGTAGATGTAGGGAAACATTCAACAGCCAT  
AGTACTATTGTTACCACTGACTATTGTTAAACAGCTTAAAGCTTATTCATGCAAAAGT  
ATAATTGAAATCTGTGGTATTATTACAAACATGTCTACAAAATAGATTACAGCTTATTGTTAGTTAA  
ATCTCTTAATACACAGAGAACTCCCAATTGCTCATCTAAATAAGGAAAGACTTGGGTATAGTGTGATGGTT  
AGTCTTAAGGATTAAGACATTGGTACTGCTTGAATTGACTACGATGTATCTGTGAAAATGGGATGATATTGAC  
AAATGGAGACTCCTACCTCAATAGTTAATGGAATAATAAGAGGCTACTGTTGTCTAATGTTCTCAAAAAGT  
AAATCCTCACTGGAGAGTGTCAAATACATACTTGAGGATTGACTTATATAAGGTGCCCTGAGACTCTGT  
TACACATATTGACCCATATTATTACAATGTTGATAATTCTACCTTTAGAGCAAGAATAGTATCTGCT  
AATGTAAGGGACATCTGTATTAACTCCTTGTAGACATGAATTCTATCAAATGTTCTTGCACGTGAAACAGA  
GATTCTTTCAATAATCTTAAATCAAAAGCATTATTAGACTTGAAGGGTTGATAATCTCCAGTCCTAG  
TAAAGATTGAGAGAGGCTGGAGCAGTTTCAGTTAAATGAGTCTGCAGTTAATATCAAATGTGAGTTGGAC  
TGCCTGGCAACATTATATTCTTATTGAGAACCCCTGATGAGACTATTAAACATACTAGTCTGCTGATAGA  
AAGCACTACATCCTATTGTTCTTCTTCAAAATCAGCCTCTGCTGTAACAAAATGTACTTATAGAG  
ATGGAGGAAAAGGTCTAATACTACATAGCCTTAAGTGTCTGCTGATGTTCAAGTGTATTCTGTAACAGAAA  
CATATTGGAATGTTTCTTCCCCTATAAAATGTAATTCTGAAATACTGCTGTTAAAAGTCCCACTG  
TCAGATTATATTCTAACAAATTGAATATTGTAAAATATACTTGTCTACCTCTAACAAAAGGGTACTTTCTAT  
CAAAAAAA

<sup>101/105</sup>  
**FIGURE 100**

MISLTDTQKIGMGLTGFGVFFLFFG MILFFDKALLAIGNVLVAGLAFVIGLERTRFFFQKHMKATGFFLGGV  
FVVLIGWPLIGMIFEIYGFLLFRGFFPVVVGFIIRRVPVLGSLLNLPGIRSFVDKVGESNNMV

102/105  
**FIGURE 101**

ATTGATTAAGAGATTGCCCTGCAAGGTAATCAGTTAAACCAACCTCTGCCCTGAGTGGATAGGTAGG  
GTTAGGGTTGCCAGATGTCACGAAGTTACAGGATGCTCAGTTAAGGTATATCCCTATACTATAAGGGTTATA  
GTAAAAAAATATTCAATTATGTGAAATTCAAATATAACTGGGTATCAGGTATTCTATGTGGCAACCCTAGGTAGGGG  
AGCACAGGTAGGCAAGCGATTAGAAGATTGCAAGCTCCAAAGTTCTGCACCTCGATGGGACACTAGAACAGG  
AAGGCTCCTGGGCCTTCTGGCTCTGGGAATGAAGCGTGGAAACCCCTCCTAGGCAGGCGAGTCTCAAGTA  
GCCAAGCTCTGACTTCCGAGGGAAAGAAAGGAGGCCATGGCCTCTGCCAGAGCCATGCTCTGCACACTGGGTCA  
GCAGAGTTCAAACGACCTGCAACGTCAGCTGGCGTTAGCTCTAAAGAGGTCTCCAGTCCAGCGCCACGCCAGC  
GGCTAGAGGCCGTCCGCCAGCTCCAAGATGGCGCCACAGCTGCCAGGTGTTAAGATGGCGCGCGGGGCC  
GGCAGCCGCGCTCCCAGGCTCTCCCTCCCCAGCCTCCCTCCGGCTGGCAGCACGACTCGCTAGCCGTGCGCCGAT  
TGCCTCTCGGCCTGGCAATGGTCCCAGGCTGCCGGTGCAGACCGCCCCCGTCATGGCTCCTCGGCTGGTGG  
CAAGTATTGCTGTGGGTCTGGACTTCCCGTCCGCGCGTGGAGGTTGAGAGGAAAGTGGTCGTTATGGTCG  
GAGGAGCAGCCTGCTCACCCCTCTCCAGGTGGGGCTGTGTACCTGGGTGAGGAGGAGCTCTGCATGACCCGATG  
GGCCAGGACAGGGCAGCAGAAGAGGCCAATGCGGTGCTGGGACCCAAGGCATCACATGGTATGCTG  
TCTGTGATTCTGGGAAGCTGAGGACAAAGTGAGTTCAAGGCCATGCCGTACCTGTGGTGTGGAGGAGCG  
GAGGACTCAAGGTGCAACGTCGAGAGAGGCCCTTCTCTGGATGGCGCTGGAGCACACTCCCTGACAGAGAA  
GAGGAGTATTACACAGAGCCAGAAGTGGCGAATCTGACGCAAGCCGACAGAGGACTCCAATAACACTGAAAGT  
CTGAAATCCCCAAAGGTGAACGTGAGGAGAGAAACATTACAGGATTAGAAAATTCACTCTGAAAATTAAAT  
ATGTCACAGGACCTTATGGATTTCTGAACCAAACGGTAGTGAATCTAGTCCTGTTTACACCCGTTG  
TGCCGCTTCTGCCAGTTGGCCCTCACTTAACCTCTGCCCGGGCATTCAGCTTCACTTGGCA  
CTGGATGCATCTCAGCACAGCAGCTTCTACCAGGTTGGCACCGTAGCTTCATAATTATTAAAT  
GGAGCTAAACCAATGGCCAGATTAAATCATACAGATCGAACACTGGAAACACTGAAAATCTTCATTTAAATCAG  
ACAGGTATAGAAGCCAAGAAGATGTGGTGTAACTCAAGCCGACCAAATAGGCCCTCTCCAGCACTTGATA  
AAAAGTGTGGACTGGTTGCTGTATTTCCTTATTCTTTAAATTAGTTATTATGTATGCTACCAATTCA  
GAGAGTATTGGTGGCTAATTCCAGGACAAGAGCAGGAACATGTGGAGTAGTGATGGTCTGAAAGAAGTGGAAA  
GAGGAACCTCAATCCTCGTTCAGAAATTAGTGCTACAGTTCATACATTCTCCAGTGCAGTGTGACTTGA  
AACTTCAGGCAGATTAAAGAACATTTGTGAACAACGTGAATGTATAAAAAAAATTATAAAACTGGTGTAAAC  
TAGTATTGCAATAAGCAAATGCAAAATATTCAATAGATGCACTATTCTGTTTACTGCATGACGTAATCCA  
GTATTGGAAAGTAATCCAGTTGAAATGTGAAGATGTATTCCGGCAGAATAGTGAGTAGAATGACATGCTTACT  
ATACAGAAGGAAAAATTAGGACTCTCAGGTAAATAGTTAAGGAAACCTTGATTCCCTATATATGTTAAGAAGG  
TTAGCTTCTGTTCTTGCAGTTCTAGAGTCATAGCAGGAAAGTGTAAACCAGAATTGGTTAGTGT  
GACCCCTCAAGTAGCAAGTGATGGAAAATAAGAGTCACCATCCTGATGTTGTGATCTCTAAACTCAAAATT  
TGAAGTGTGTTAAGTTGTTCTGGTAAGGGAGATGTAGGAGAAAGGAAATGCTGTAACCAAAGCTCAATTATT  
ATCAGTTCTATGCTAACGTATACTTAAATCATAGTTACCTAACGAGCATGCATTAATTGAAACCTAAATGTT  
CCCAGCAGGCTGGTCTCAAACGCTCAGGTGATCCACCCGCCCTCGGCCTCCAGGTGTTGGATTGCAG  
GTGTGAGCCACTGCGCTGGCTAAACAAACATTGAAAGCTGTTCTAAAAGATTCTAAATTCA  
ACAGCTAATTACCTCATATAAATTACTTTATACTAATTGTTCCAGGGTTAGAGTAGTTGAATGTTATT  
ACAAGGCACCCCTAAATTCTATAGAAATAAACCTCAGATGAGCTC

103/105  
**FIGURE 102**

MVPAAGRRPPRVMRLLGWWQVILLWVLGLPVRGVEVAEESGRLWSEEQPAHPLQVGAVYLGEELLHDPMGQDRAA  
EEANAVLGLGTQGDHMVMLSIPGEAEDKVSSEPSGVTGAGGAEDSRCNVRESLFSLDGAGAHFPDREEYYTE  
PEVAESDAAPTEDSNNTESLKSPKVNCERNITGLENFTLKILNMSQDILMDFLNPNGSDCTLVLFYTPWCRFSAS  
LAPHFNSLPRAFPALHFLALDASQHSSLSTRFGTVAVPNILLFQGAKPMARFNHTDRTLETLKIFIFNQTGIEAK  
KNVVVTQADQIGPLPSTLIKSVDWLLVFSLFFLISFIMYATIRTESIRWLIPGQEQUEHVE

104/105  
**FIGURE 103**

GGCACGAGGCCTGGGTGCGCTGCCGCCACGTCCCCGCCGGCTCAGGCTCCTTCTACTGTCCGAGGGCC  
ACCAGGCCGCCGGGGCTGCTGCCGCCGGATGCGTCTGTTACTAGAGTGGAGAGTCTACCTCGTCTCACATGT  
GCCACAAAGGATGGCATGGCCGGAGTGCCCCACACGTGGCTTCACCCCTGCAAAGCCAGACTCGCCCAG  
CGACACAGTGTCAAGCCCACAGCTCCAAGGAGGAAGATGGTCAGGCTGGGAGCATCCCTAGCAGCAGCCT  
CTGATCCCTGGCCAAGCAGGAGGAACCATTAGCAGCCTGAGGAGCTGGCTGGCTGGGAGCCTGGGACGCC  
CAGCCTTGCTCCCAGCTCACCCACAAGATGTGGACAGCTTGTGCTCATTGGATTTCTCCTGTCTTATCT  
GAAAGCCATGCGGATCCAACGATCCACGCAACTTGTCCCTAACAAATGTGGAAGGGATTAGTCAAGAGGAAT  
GCATCTGTGAAACAGTTGATAATAAAACGTCTGAGGATGTAACCATGGCAGCAGCTCTCTGTACATTGACC  
AAAGGGACTCGGCAGCCCACCTCAACTCTATGGAAGTCACAACAGAGGACACAAGCAGGACAGATGTGAGTGAA  
CCAGCAACTTCAGGAGGTGCAGCTGATGGTGTGACCTCCATTGCTCCCACGGCTGTGGCCTCCAGTACGACTGCG  
GCCTCCATTACGACTCGGCCTCAGTATGACTGTGGCCTCAGTGCTCCACGACTGCAGCCTCAGTACA  
GTGGCCTCATTGCTCCCACGACTGCAGCCTCAGTATGACTGTGGCCTCAGTGCTCCACGACTGCAGCCTCAGTACA  
CCCGCGCCCACGTCCACTTCCACAGGGCGGACCCCGTCCACTACCGCCACTGGCATTCCATCTCAGCACAGCC  
CTCGCACAAGTGCCTAACAGAGCAGCGCTTGCCTAACAGCAACCTGGCACATTGGCACACGTGCTCAGACT  
GTAGCGACCACAGCAAACACAAGCAGCCCCATGAGCAGCTCGTCCAAGTCCTCCAAGCACATGCCAGTGACACC  
GCGGCAAGCCCTGTACCCCTATGCGTCCCCAAGCACAAGGTCCATTAGCCAGGTGTCAGTGGACCAGCCTGTG  
GTTAACACAACAAATAATCCACACCCATGCCCTCAAACACAACCCAGAGGCCGCCCCACCCCCACAGTGGT  
ACCACCAAGGCACAAGCCAGGGAGCCAAGCAGCCACTGCCAGCCCAGTGCCAGTACCTCACACCAGCCAACTCCCTGAG  
ATGGAGGCCATGTCCCCCAGCACACAGCCAAGCCCCATGCCATATACCCAGAGGGCCGTGGCCAGGCACATCC  
CAGGCACCGGAGCAGGTAGAGACTGAAGCCACACCAGGTACTGATTCCACTGGCCAACACCCAGGAGCTCAGGG  
GGCACTAACAGATGCCAGCCACGGACTCGTCCAGCCCAGCACCCAGGCCAGTACATGGTGGTACCAACTGAGCCC  
CTCACCCAGGCCGTGGTAGACAAAATCTCTTCTGGTGGTGTGTTACTCGGGGTGACCCCTTCATCACAGTC  
TTGGTTTGTGAGACCAAAACCAAGTGCTCAAATTCTTTGGTCAATTGAGGAGATATGCCAGATGCTTAAAC  
GGGATGTATGCGGACTCAGAAATGTGAGGGGGGCGGGGGCTGGGGAGGCCGTGGCCCTTCTCGTCTTCC  
TTTGCTTGTGAGACCAAAACCAAGTGCTCAAATTCTTTGGTCAATTGAGGAGATATGCCAGATGCTTAAAC  
ACATTAAATTGCTGTCAAGATTAATTCCATGATCACTAAAGAGTTGCTGCTTTTCAATTATTTGTAAATG  
ATTCTGTGCCCAAGGAGCAGCTGGGGTCCACCTCAGGTGGGGCGAGGACCCGTCTCCCCAGGTGTCGGA  
GCCTGACCTGAATTAAAGTACTGACTGCTCGCC

105/105  
**FIGURE 104**

MWTALVLIWIFSLSLSESHAASNDPRNFVPNKMWKGLVKRNASVETVDNKTSEDVTMAAASPVTLKGTSAAHLN  
SMEVTTEDTSRTDVSEPATSGGAADGVTSIAPTAVASSTTAASITTAASSMTVASSAPTTAASSTTVASIAPTTA  
ASSMTAASSTPMTLALPAPTSTSTGRTPSTTATGHPSLSTALAQVPKSSALPRTATLATLATRAQTVATTANTSS  
PMSTRPSPSKHMPSDTAASPVPPMRPQAQGPISQVSVDQPVVNTTNKSTPMPSNTTPEPAPPTPTVVTTKAQARE  
PTASPVVPVHTSPIPEMEAMSPTTQPSPMPYTQRAAGPGTSQAPEQVETEATPGTDSTGPTPRSSGGTKMPATDS  
CQPSTQGQYMVVTTEPLTQAVVDKTLLLVVLGVTLFITVLVLFALQAYESYKKDYTQVDYLINGMYADSEM