Office de la Propriété Intellectuelle du Canada
Un organisme d'Industrie Canada

Canadian Intellectual Property Office
An agency of Industry Canada

CA 2382302 A1 1999/03/25

(21) 2 382 302

(12) DEMANDE DE BREVET CANADIEN
CANADIAN PATENT APPLICATION

(13) A1

(22) Date de dépôt/Filing Date: 1998/09/16

(41) Mise à la disp. pub./Open to Public Insp.: 1999/03/25

(62) Demande originale/Original Application: 2 303 834

(30) Priorités/Priorities: 1997/09/17 (60/059,184) US;
1997/09/17 (60/059,122) US; 1997/09/17 (60/059,121) US;
1997/09/17 (60/059,113) US; 1997/09/17 (60/059,117) US;
1997/09/17 (60/059,115) US; 1997/09/17 (60/059,119) US;
1997/09/18 (60/059,266) US; 1997/09/18 (60/059,263) US;
1997/10/15 (60/062,125) US; 1997/10/17 (60/062,287) US;
1997/10/17 (60/062,285) US; 1997/10/21 (60/063,486) US;
1997/10/24 (60/063,127) US; 1997/10/24 (60/062,814) US;
1997/10/24 (60/062,816) US; 1997/10/24 (60/063,120) US;
1997/10/24 (60/063,128) US; 1997/10/24 (60/063,121) US;
...

(51) Cl.Int.^{7}/Int.Cl.^{7} C12N 15/12, C12N 15/63, C12N 15/62, C07K 14/47, C12N 5/10, C12P 21/00, C07K 19/00

(71) Demandeur/Applicant:
GENENTECH, INC., US

(72) Inventeurs/Inventors:
WOOD, WILLIAM I., US;
GURNEY, AUSTIN L., US;
GODDARD, AUDREY, US;
PENNICA, DIANE, US;
CHEN, JIAN, US;
YUAN, JEAN, US

(74) Agent: DENNISON ASSOCIATES

(54) Titre : POLYPEPTIDES SECRETES ET TRANSMEMBRANAIRES ET ACIDES NUCLEIQUES LES CODANT
(54) Title: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC ACIDS ENCODING THE SAME

GGCCGGAGCAGCACGGCCGCAGGACCTGGAGCTCCGGCTGCGTCTTCCCGCAGCGCTACCCG
CCATGCGCCTGCCGCGCCGGGCCGCGCTGGGGCTCCTGCCGCTTCTGCTGCTGCTGCCGCCC
GCGCCGGAGGCCGCCAAGAAGCCGACGCCCTGCCACCGGTGCCGGGGGCTGGTGGACAAGTT
TAACCAGGGGATGGTGGACACCGCAAAGAAGAACTTTGGCGGCGGGAACACGGCTTGGGAGG
AAAAGACGCTGTCCAAGTACGAGTCCAGCGAGATTCGCCTGCTGGAGATCCTGGAGGGGCTG
TGCGAGAGCAGCGACTTCGAATGCAATCAGATGCTAGAGGCGCAGGAGGAGCACCTGGAGGC
CTGGTGGCTGCAGCTGAAGAGCGAATATCCTGACTTATTCGAGTGGTTTTGTGTGAAGACAC
TGAAAGTGTGCTGCTCTCCAGGAACCTACGGTCCCGACTGTCTCGCATGCCAGGGCGGATCC
CAGAGGCCCTGCAGCGGGAATGGCCACTGCAGCGGAGATGGGAGCAGACAGGGCGACGGGTC
CTGCCGGTGCCACATGGGGTACCAGGGCCCGCTGTGCACTGACTGCATGGACGGCTACTTCA
GCTCGCTCCGGAACGAGACCCACAGCATCTGCACAGCCTGTGACGAGTCCTGCAAGACGTGC
TCGGGCCTGACCAACAGAGACTGCGGCGAGTGTGAAGTGGGCTGGGTGCTGGACGAGGGCGC
CTGTGTGGATGTGGACGAGTGTGCGGCCGAGCCGCCTCCCTGCAGCGCTGCGCAGT
TCTGTAAGAACGCCAACGGCTCCTACACGTGCGAAGAGTGTGACTCCAGCTGTGTGGGCTGC
ACAGGGGAAGGCCCAGGAAACTGTAAAGAGTGTATCTCTGGCTACGCGAGGGAGCACGGACA
GTGTGCAGATGTGGACGAGTGCTCACTAGCAGAAAAAACCTGTGTGAGGAAAAACGAAAACT
GCTACAATACTCCAGGGAGCTACGTCTGTGTGTGTCCTGACGGCTTCGAAGAAACGGAAGAT
GCCTGTGTGCCGCCGGCAGAGGCTGAAGCCACAGAAGGAGAAAGCCCGACACAGCTGCCCTC
CCGCGAAGACCTGTAATGTGCCGGACTTACCCTTTAAATTATTCAGAAGGATGTCCCGTGGA
AAATGTGGCCCTGAGGATGCCGTCTCCTGCAGTGGACAGCGGCGGGGAGAGGCTGCCTGCTC
TCTAACGGTTGATTCTCATTTGTCCCTTAAACAGCTGCATTTCTTGGTTGTTCTTAAACAGA
CTTGTATATTTTGATACAGTTCTTTGTAATAAAATTGACCATTGTAGGTAATCAGGAGGAAA
AAAAAA

(57) Abrégé/Abstract:
The present invention is directed to novel polypeptides and to nucleic acid molecules encoding those polypeptides. Also provided herein are vectors and host cells comprising those nucleic acid sequences, chimeric polypeptides molecules comprising the polypeptides of the present invention fused to heterologous polypeptide sequences, antibodies which bind to the polypeptides of the present invention and to methods for producing the polypeptides of the present invention.

(21) 2 382 302

(13) A1

(30) Priorités(suite)/Priorities(continued): 1997/10/24 (60/063,045) US; 1997/10/27 (60/063,327) US; 1997/10/27 (60/063,329) US; 1997/10/28 (60/063,549) US; 1997/10/28 (60/063,550) US; 1997/10/28 (60/063,541) US; 1997/10/28 (60/063,544) US; 1997/10/28 (60/063,542) US; 1997/10/28 (60/063,564) US; 1997/10/29 (60/063,435) US; 1997/10/29 (60/063,732) US; 1997/10/29 (60/063,735) US; 1997/10/29 (60/063,738) US; 1997/10/29 (60/064,215) US; 1997/10/29 (60/063,704) US; 1997/10/29 (60/063,734) US; 1997/10/31 (60/064,103) US; 1997/10/31 (60/063,870) US; 1997/11/03 (60/064,248) US; 1997/11/07 (60/064,809) US; 1997/11/12 (60/065,186) US; 1997/11/17 (60/065,846) US; 1997/11/18 (60/065,693) US; 1997/11/21 (60/066,364) US; 1997/11/21 (60/066,120) US; 1997/11/24 (60/066,466) US; 1997/11/24 (60/066,770) US; 1997/11/24 (60/066,511) US; 1997/11/24 (60/066,772) US; 1997/11/24 (60/066,453) US; 1997/11/25 (60/066,840) US

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/12, 15/18, 15/52, C07K 14/47, 14/705, C12N 15/62, C07K 16/18, 16/28

A2

(11) International Publication Number: **WO 99/14328**

(43) International Publication Date: 25 March 1999 (25.03.99)

(21) International Application Number: PCT/US98/19330

(22) International Filing Date: 16 September 1998 (16.09.98)

(30) Priority Data:

60/059,115	17 September 1997 (17.09.97)	US
60/059,184	17 September 1997 (17.09.97)	US
60/059,122	17 September 1997 (17.09.97)	US
60/059,117	17 September 1997 (17.09.97)	US
60/059,113	17 September 1997 (17.09.97)	US
60/059,121	17 September 1997 (17.09.97)	US
60/059,119	17 September 1997 (17.09.97)	US
60/059,263	18 September 1997 (18.09.97)	US
60/059,266	18 September 1997 (18.09.97)	US
60/062,125	15 October 1997 (15.10.97)	US
60/062,287	17 October 1997 (17.10.97)	US
60/062,285	17 October 1997 (17.10.97)	US
60/063,486	21 October 1997 (21.10.97)	US
60/062,816	24 October 1997 (24.10.97)	US
60/062,814	24 October 1997 (24.10.97)	US
60/063,127	24 October 1997 (24.10.97)	US
60/063,120	24 October 1997 (24.10.97)	US
60/063,121	24 October 1997 (24.10.97)	US
60/063,045	24 October 1997 (24.10.97)	US
60/063,128	24 October 1997 (24.10.97)	US
60/063,329	27 October 1997 (27.10.97)	US
60/063,327	27 October 1997 (27.10.97)	US
60/063,549	28 October 1997 (28.10.97)	US
60/063,541	28 October 1997 (28.10.97)	US
60/063,550	28 October 1997 (28.10.97)	US
60/063,542	28 October 1997 (28.10.97)	US
60/063,544	28 October 1997 (28.10.97)	US
60/063,564	28 October 1997 (28.10.97)	US
60/063,734	29 October 1997 (29.10.97)	US
60/063,738	29 October 1997 (29.10.97)	US
60/063,704	29 October 1997 (29.10.97)	US
60/063,435	29 October 1997 (29.10.97)	US
60/064,215	29 October 1997 (29.10.97)	US
60/063,735	29 October 1997 (29.10.97)	US
60/063,732	29 October 1997 (29.10.97)	US
60/064,103	31 October 1997 (31.10.97)	US
60/063,870	31 October 1997 (31.10.97)	US
60/064,248	3 November 1997 (03.11.97)	US
60/064,809	7 November 1997 (07.11.97)	US
60/065,186	12 November 1997 (12.11.97)	US
60/065,846	17 November 1997 (17.11.97)	US
60/065,693	18 November 1997 (18.11.97)	US
60/066,120	21 November 1997 (21.11.97)	US
60/066,364	21 November 1997 (21.11.97)	US
60/066,772	24 November 1997 (24.11.97)	US
60/066,466	24 November 1997 (24.11.97)	US
60/066,770	24 November 1997 (24.11.97)	US
60/066,511	24 November 1997 (24.11.97)	US
60/066,453	24 November 1997 (24.11.97)	US
60/066,840	25 November 1997 (25.11.97)	US

(71) Applicant *(for all designated States except US)*: GENENTECH, INC. [US/US]; One DNA Way, South San Francisco, CA 94080 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: WOOD, William, I. [US/US]; 1400 Tarrytown Street, San Mateo, CA 94402 (US). GURNEY, Austin, L. [US/US]; One Debbie Lane, Belmont, CA 94002 (US). GODDARD, Audrey [CA/US]; 110 Congo Street, San Francisco, CA 94131 (US). PENNICA, Diane [US/US]; 2417 Hale Drive, Burlingame, CA 94010 (US). CHEN, Jian [CN/US]; 1860 Ogden Drive #14, Burlingame, CA 94010 (US). YUAN, Jean [CN/US]; 176 West 37th Avenue, San Mateo, CA 94403 (US).

(74) Agents: DREGER, Walter, H. et al.; Flehr, Hohbach, Test, Albritton & Herbert LLP, Suite 3400, 4 Embarcadero Center, San Francisco, CA 94111–4187 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC ACIDS ENCODING THE SAME

(57) Abstract

The present invention is directed to novel polypeptides and to nucleic acid molecules encoding those polypeptides. Also provided herein are vectors and host cells comprising those nucleic acid sequences, chimeric polypeptides molecules comprising the polypeptides of the present invention fused to heterologous polypeptide sequences, antibodies which bind to the polypeptides of the present invention and to methods for producing the polypeptides of the present invention.

DEMANDES OU BREVETS VOLUMINEUX

LA PRÉSENTE PARTIE DE CETTE DEMANDE OU CE BREVETS COMPREND PLUS D'UN TOME.

CECI EST LE TOME 1 DE 2

NOTE: Pour les tomes additionels, veillez contacter le Bureau Canadien des Brevets.

JUMBO APPLICATIONS / PATENTS

THIS SECTION OF THE APPLICATION / PATENT CONTAINS MORE THAN ONE VOLUME.

THIS IS VOLUME 1 OF 2

NOTE: For additional volumes please contact the Canadian Patent Office.

SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC ACIDS ENCODING THE SAME

FIELD OF THE INVENTION

The present invention relates generally to the identification and isolation of novel DNA and to the recombinant production of novel polypeptides encoded by that DNA.

BACKGROUND OF THE INVENTION

Extracellular and membrane-bound proteins play important roles in the formation, differentiation and maintenance of multicellular organisms. The fate of many individual cells, e.g., proliferation, migration, differentiation, or interaction with other cells, is typically governed by information received from other cells and/or the immediate environment. This information is often transmitted by secreted polypeptides (for instance, mitogenic factors, survival factors, cytotoxic factors, differentiation factors, neuropeptides, and hormones) which are, in turn, received and interpreted by diverse cell receptors or membrane-bound proteins. These secreted polypeptides or signaling molecules normally pass through the cellular secretory pathway to reach their site of action in the extracellular environment, usually at a membrane-bound receptor protein.

Secreted proteins have various industrial applications, including use as pharmaceuticals, diagnostics, biosensors and bioreactors. In fact, most protein drugs available at present, such as thrombolytic agents, interferons, interleukins, erythropoietins, colony stimulating factors, and various other cytokines, are secretory proteins. Their receptors, which are membrane-bound proteins, also have potential as therapeutic or diagnostic agents. Receptor immunoadhesins, for instance, can be employed as therapeutic agents to block receptor-ligand interaction. Membrane-bound proteins can also be employed for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction. Such membrane-bound proteins and cell receptors include, but are not limited to, cytokine receptors, receptor kinases, receptor phosphatases, receptors involved in cell-cell interactions, and cellular adhesin molecules like selectins and integrins. Transduction of signals that regulate cell growth and differentiation is regulated in part by phosphorylation of various cellular proteins. Protein tyrosine kinases, enzymes that catalyze that process, can also act as growth factor receptors. Examples include fibroblast growth factor receptor and nerve growth factor receptor.

Efforts are being undertaken by both industry and academia to identify new, native secreted and membrane-bound receptor proteins. Many efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel secreted and membrane-bound receptor proteins. Examples of screening methods and techniques are described in the literature [see, for example, Klein et al., Proc. Natl. Acad. Sci., 93:7108-7113 (1996); U.S. Patent No. 5,536,637)].

We herein describe the identification and characterization of novel secreted and transmembrane polypeptides and novel nucleic acids encoding those polypeptides.

1. **PRO211 and PRO217**

Epidermal growth factor (EGF) is a conventional mitogenic factor that stimulates the proliferation of various types of cells including epithelial cells and fibroblasts. EGF binds to and activates the EGF receptor (EGFR), which initiates intracellular signaling and subsequent effects. The EGFR is expressed in neurons of the cerebral cortex, cerebellum, and hippocampus in addition to other regions of the central nervous system (CNS). In addition, EGF is also expressed in various regions of the CNS. Therefore, EGF acts not only on mitotic cells, but also on postmitotic neurons. In fact, many studies have indicated that EGF has neurotrophic or neuromodulatory effects on various types of neurons in the CNS. For example, EGF acts directly on cultured cerebral cortical and cerebellar neurons, enhancing neurite outgrowth and survival. On the other hand, EGF also acts on other cell types, including septal cholinergic and mesencephalic dopaminergic neurons, indirectly through glial cells. Evidence of the effects of EGF on neurons in the CNS is accumulating, but the mechanisms of action remain essentially unknown. EGF-induced signaling in mitotic cells is better understood than in postmitotic neurons. Studies of cloned pheochromocytoma PC12 cells and cultured cerebral cortical neurons have suggested that the EGF-induced neurotrophic actions are mediated by sustained activation of the EGFR and mitogen-activated protein kinase (MAPK) in response to EGF. The sustained intracellular signaling correlates with the decreased rate of EGFR down-regulation, which might determine the response of neuronal cells to EGF. It is likely that EGF is a multi-potent growth factor that acts upon various types of cells including mitotic cells and postmitotic neurons.

EGF is produced by the salivary and Brunner's glands of the gastrointestinal system, kidney, pancreas, thyroid gland, pituitary gland, and the nervous system, and is found in body fluids such as saliva, blood, cerebrospinal fluid (CSF), urine, amniotic fluid, prostatic fluid, pancreatic juice, and breast milk, Plata-Salaman, *Peptides* 12: 653-663 (1991).

EGF is mediated by its membrane specific receptor, which contains an intrinsic tyrosine kinase. Stoscheck *et al., J. Cell Biochem.* 31: 135-152 (1986). EGF is believed to function by binding to the extracellular portion of its receptor which induces a transmembrane signal that activates the intrinsic tyrosine kinase.

Purification and sequence analysis of the EGF-like domain has revealed the presence of six conserved cysteine residues which cross-bind to create three peptide loops, Savage *et al., J. Biol. Chem.* 248: 7669-7672 (1979). It is now generally known that several other peptides can react with the EGF receptor which share the same generalized motif $X_nCX_7CX_{4/5}CX_{10}CXCX_5GX_2CX_n$, where X represents any non-cysteine amino acid, and n is a variable repeat number. Non isolated peptides having this motif include TGF-α, amphiregulin, schwannoma-derived growth factor (SDGF), heparin-binding EGF-like growth factors and certain virally encoded peptides (e.g., Vaccinia virus, Reisner, *Nature* 313: 801-803 (1985), Shope fibroma virus, Chang et al., Mol Cell Biol. 7: 535-540 (1987), Molluscum contagiosum, Porter and Archard, *J. Gen. Virol.* 68: 673-682 (1987), and Myxoma virus, Upton *et al., J. Virol.* 61: 1271-1275 (1987), Prigent and Lemoine, *Prog. Growth Factor Res.* 4: 1-24 (1992).

EGF-like domains are not confined to growth factors but have been observed in a variety of cell-surface and extracellular proteins which have interesting properties in cell adhesion, protein-protein interaction and development, Laurence and Gusterson, *Tumor Biol.* 11: 229-261 (1990). These proteins include blood coagulation factors (factors VI, IX, X, XII, protein C, protein S, protein Z, tissue plasminogen activator, urokinase), extracellular matrix components (laminin, cytotactin, entactin), cell surface receptors (LDL receptor, thrombomodulin receptor) and

immunity-related proteins (complement C1r, uromodulin).

Even more interesting, the general structure pattern of EGF-like precursors is preserved through lower organisms as well as in mammalian cells. A number of genes with developmental significance have been identified in invertebrates with EGF-like repeats. For example, the *notch* gene of Drosophila encodes 36 tandemly arranged 40 amino acid repeats which show homology to EGF, Wharton *et al.*, *Cell* 43: 557-581 (1985). Hydropathy plots indicate a putative membrane spanning domain, with the EGF-related sequences being located on the extracellular side of the membrane. Other homeotic genes with EGF-like repeats include Delta, 95F and 5ZD which were identified using probes based on Notch, and the nematode gene *Lin-12* which encodes a putative receptor for a developmental signal transmitted between two specified cells.

Specifically, EGF has been shown to have potential in the preservation and maintenance of gastrointestinal mucosa and the repair of acute and chronic mucosal lesions, Konturek *et al.*, *Eur. J. Gastroenterol Hepatol.* 7 (10), 933-37 (1995), including the treatment of necrotizing enterocolitis, Zollinger-Ellison syndrome, gastrointestinal ulceration gastrointestinal ulcerations and congenital microvillus atrophy, Guglietta and Sullivan, *Eur. J. Gastroenterol Hepatol*, 7(10), 945-50 (1995). Additionally, EGF has been implicated in hair follicle differentiation; du Cros, *J. Invest. Dermatol.* 101 (1 Suppl.), 106S-113S (1993), Hillier, *Clin. Endocrinol.* 33(4), 427-28 (1990); kidney function, Hamm *et al.*, *Semin. Nephrol.* 13 (1): 109-15 (1993), Harris, *Am. J. Kidney Dis.* 17(6): 627-30 (1991); tear fluid, van Setten *et al.*, *Int. Ophthalmol* 15(6); 359-62 (1991); vitamin K mediated blood coagulation, Stenflo *et al.*, *Blood* 78(7): 1637-51 (1991). EGF is also implicated various skin disease characterized by abnormal keratinocyte differentiation, e.g., psoriasis, epithelial cancers such as squamous cell carcinomas of the lung, epidermoid carcinoma of the vulva and gliomas. King *et al.*, *Am. J. Med. Sci.* 296: 154-158 (1988).

Of great interest is mounting evidence that genetic alterations in growth factors signaling pathways are closely linked to developmental abnormalities and to chronic diseases including cancer. Aaronson, *Science* 254: 1146-1153 (1991). For example, c-erb-2 (also known as HER-2), a proto-oncogene with close structural similarity to EGF receptor protein, is overexpressed in human breast cancer. King *et al.*, *Science* 229: 974-976 (1985); Gullick, *Hormones and their actions*, Cooke *et al.*, eds, Amsterdam, Elsevier, pp 349-360 (1986).

We herein describe the identification and characterization of novel polypeptides having homology to EGF, wherein those polypeptides are herein designated PRO211 and PRO217.

2. PRO230

Nephritis is a condition characterized by inflammation of the kidney affecting the structure and normal function of the kidney. This condition can be chronic or acute and is generally caused by infection, degenerative process or vascular disease. In all cases, early detection is desirable so that the patient with nephritis can begin treatment of the condition.

An approach to detecting nephritis is to determine the antigens associated with nephritis and antibodies thereto. In rabbit, a tubulointerstitial nephritis antigen (TIN-ag) has been reported in Nelson, T. R., et al., J. Biol. Chem., 270(27):16265-70 (July 1995) (GENBANK/U24270). This study reports that the rabbit TIN-ag is a basement membrane glycoprotein having a predicted amino acid sequence which has a carboxyl-terminal region exhibiting 30% homology with human preprocathepsin B, a member of the cystein proteinase family of proteins. It is also reported

that the rabbit TIN-ag has a domain in the amino-terminal region containing an epidermal growth factor-like motif that shares homology with laminin A and S chains, alpha 1 chain of type I collagen, von Willebrand's factor and mucin, indicating structural and functional similarities. Studies have also been conducted in mice. However, it is desirable to identify tubulointerstitial nephritis antigens in humans to aid in the development of early detection methods and treatment of nephritis.

Proteins which have homology to tubulointerstitial nephritis antigens are of particular interest to the medical and industrial communities. Often, proteins having homology to each other have similar function. It is also of interest when proteins having homology do not have similar functions, indicating that certain structural motifs identify information other than function, such as locality of function. We herein describe the identification and characterization of a novel polypeptide, designated hgerein as PRO230, which has homology to tubulointerstitial nephritis antigens.

3. PRO232

Stem cells are undifferentiated cells capable of (a) proliferation, (b) self maintenance, (c) the production of a large number of differentiated functional progeny, (d) regeneration of tissue after injury and/or (e) a flexibility in the use of these options. Stem cells often express cell surface antigens which are capable of serving as cell specific markers that can be exploited to identify stem cells, thereby providing a means for identifying and isolating specific stem cell populations.

Having possession of different stem cell populations will allow for a number of important applications. For example, possessing a specific stem cell population will allow for the identification of growth factors and other proteins which are involved in their proliferation and differentiation. In addition, there may be as yet undiscovered proteins which are associated with (1) the early steps of dedication of the stem cell to a particular lineage, (2) prevention of such dedication, and (3) negative control of stem cell proliferation, all of which may be identified if one has possession of the stem cell population. Moreover, stem cells are important and ideal targets for gene therapy where the inserted genes promote the health of the individual into whom the stem cells are transplanted. Finally, stem cells may play important roles in transplantation of organs or tissues, for example liver regeneration and skin grafting.

Given the importance of stem cells in various different applications, efforts are currently being undertaken by both industry and academia to identify new, native stem cell antigen proteins so as to provide specific cell surface markers for identifying stem cell populations as well as for providing insight into the functional roles played by stem cell antigens in cell proliferation and differentiation. We herein describe the identification and characterization of novel polypeptides having homology to a stem cell antigen, wherein those polypeptides are herein designated as PRO232 polypeptides..

4. PRO187

Growth factors are molecular signals or mediators that enhance cell growth or proliferation, alone or in concert, by binding to specific cell surface receptors. However, there are other cellular reactions than only growth upon expression to growth factors. As a result, growth factors are better characterized as multifunctional and potent

cellular regulators. Their biological effects include proliferation, chemotaxis and stimulation of extracellular matrix production. Growth factors can have both stimulatory and inhibitory effects. For example, transforming growth factor (TGF-β) is highly pleiotropic and can stimulate proliferation in some cells, especially connective tissue, while being a potent inhibitor of proliferation in others, such as lymphocytes and epithelial cells.

The physiological effect of growth stimulation or inhibition by growth factors depends upon the state of development and differentiation of the target tissue. The mechanism of local cellular regulation by classical endocrine molecules involves comprehends autocrine (same cell), juxtacrine (neighbor cell), and paracrine (adjacent cells) pathways. Peptide growth factors are elements of a complex biological language, providing the basis for intercellular communication. They permit cells to convey information between each other, mediate interaction between cells and change gene expression. The effect of these multifunctional and pluripotent factors is dependent on the presence or absence of other peptides.

FGF-8 is a member of the fibroblast growth factors (FGFs) which are a family of heparin-binding, potent mitogens for both normal diploid fibroblasts and established cell lines, Gospodarowicz *et al.* (1984), *Proc. Natl. Acad. Sci. USA* 81:6963. The FGF family comprises acidic FGF (FGF-1), basic FGF (FGF-2), INT-2 (FGF-3), K-FGF/HST (FGF-4), FGF-5, FGF-6, KGF (FGF-7), AIGF (FGF-8) among others. All FGFs have two conserved cysteine residues and share 30-50% sequence homology at the amino acid level. These factors are mitogenic for a wide variety of normal diploid mesoderm-derived and neural crest-derived cells, including granulosa cells, adrenal cortical cells, chondrocytes, myoblasts, corneal and vascular endothelial cells (bovine or human), vascular smooth muscle cells, lens, retina and prostatic epithelial cells, oligodendrocytes, astrocytes, chrondocytes, myoblasts and osteoblasts.

Fibroblast growth factors can also stimulate a large number of cell types in a non-mitogenic manner. These activities include promotion of cell migration into wound area (chemotaxis), initiation of new blood vessel formulation (angiogenesis), modulation of nerve regeneration and survival (neurotrophism), modulation of endocrine functions, and stimulation or suppression of specific cellular protein expression, extracellular matrix production and cell survival. Baird & Bohlen, Handbook of Exp. Pharmacol. 95(1): 369-418, Springer, (1990). These properties provide a basis for using fibroblast growth factors in therapeutic approaches to accelerate wound healing, nerve repair, collateral blood vessel formation, and the like. For example, fibroblast growth factors have been suggested to minimize myocardium damage in heart disease and surgery (U.S.P. 4,378,347).

FGF-8, also known as androgen-induced growth factor (AIGF), is a 215 amino acid protein which shares 30-40% sequence homology with the other members of the FGF family. FGF-8 has been proposed to be under androgenic regulation and induction in the mouse mammary carcinoma cell line SC3. Tanaka *et al., Proc. Natl. Acad. Sci. USA* 89: 8928-8932 (1992); Sato *et al., J. Steroid Biochem. Molec. Biol.* 47: 91-98 (1993). As a result, FGF-8 may have a local role in the prostate, which is known to be an androgen-responsive organ. FGF-8 can also be oncogenic, as it displays transforming activity when transfected into NIH-3T3 fibroblasts. Kouhara *et al., Oncogene* 9 455-462 (1994). While FGF-8 has been detected in heart, brain, lung, kidney, testis, prostate and ovary, expression was also detected in the absence of exogenous androgens. Schmitt *et al., J. Steroid Biochem. Mol. Biol.* 57 (3-4): 173-78 (1996).

FGF-8 shares the property with several other FGFs of being expressed at a variety of stages of murine embryogenesis, which supports the theory that the various FGFs have multiple and perhaps coordinated roles in differentiation and embryogenesis. Moreover, FGF-8 has also been identified as a protooncogene that cooperates with Wnt-1 in the process of mammary tumorigenesis (Shackleford *et al., Proc. Natl. Acad. Sci. USA* 90, 740-744 (1993); Heikinheimo *et al., Mech. Dev.* 48: 129-138 (1994)).

In contrast to the other FGFs, FGF-8 exists as three protein isoforms, as a result of alternative splicing of the primary transcript. Tanaka et al., *supra*. Normal adult expression of FGF-8 is weak and confined to gonadal tissue, however northern blot analysis has indicated that FGF-8 mRNA is present from day 10 through day 12 or murine gestation, which suggests that FGF-8 is important to normal development. Heikinheimo *et al., Mech Dev.* 48(2): 129-38 (1994). Further *in situ* hybridization assays between day 8 and 16 of gestation indicated initial expression in the surface ectoderm of the first bronchial arches, the frontonasal process, the forebrain and the midbrain-hindbrain junction. At days 10-12, FGF-8 was expressed in the surface ectoderm of the forelimb and hindlimb buds, the nasal its and nasopharynx, the infundibulum and in the telencephalon, diencephalon and metencephalon. Expression continues in the developing hindlimbs through day 13 of gestation, but is undetectable thereafter. The results suggest that FGF-8 has a unique temporal and spatial pattern in embryogenesis and suggests a role for this growth factor in multiple regions of ectodermal differentiation in the post-gastrulation embryo.

We herein describe the identification of novel poypeptides having homology to FGF-8, wherein those polypeptides are heein designated PRO187 polypeptides.

5. PRO265

Protein-protein interactions include receptor and antigen complexes and signaling mechanisms. As more is known about the structural and functional mechanisms underlying protein-protein interactions, protein-protein interactions can be more easily manipulated to regulate the particular result of the protein-protein interaction. Thus, the underlying mechanisms of protein-protein interactions are of interest to the scientific and medical community.

All proteins containing leucine-rich repeats are thought to be involved in protein-protein interactions. Leucine-rich repeats are short sequence motifs present in a number of proteins with diverse functions and cellular locations. The crystal structure of ribonuclease inhibitor protein has revealed that leucine-rich repeats correspond to beta-alpha structural units. These units are arranged so that they form a parallel beta-sheet with one surface exposed to solvent, so that the protein acquires an unusual, nonglubular shape. These two features have been indicated as responsible for the protein-binding functions of proteins containing leucine-rich repeats. See, Kobe and Deisenhofer, Trends Biochem. Sci., 19(10):415-421 (Oct. 1994).

A study has been reported on leucine-rich proteoglycans which serve as tissue organizers, orienting and ordering collagen fibrils during ontogeny and are involved in pathological processes such as wound healing, tissue repair, and tumor stroma formation. Iozzo, R. V., Crit. Rev. Biochem. Mol. Biol., 32(2):141-174 (1997). Others studies implicating leucine rich proteins in wound healing and tissue repair are De La Salle, C., et al., Vouv. Rev. Fr. Hematol. (Germany), 37(4):215-222 (1995), reporting mutations in the leucine rich motif in a complex associated with the bleeding disorder Bernard-Soulier syndrome and Chlemetson, K. J., Thromb. Haemost. (Germany), 74(1):111-116 (July 1995), reporting that platelets have leucine rich repeats. Another protein of particular interest

which has been reported to have leucine-rich repeats is the SLIT protein which has been reported to be useful in treating neuro-degenerative diseases such as Alzheimer's disease, nerve damage such as in Parkinson's disease, and for diagnosis of cancer, see, Artavanistsakonas, S. and Rothberg, J. M., WO9210518-A1 by Yale University. Other studies reporting on the biological functions of proteins having leucine-rich repeats include: Tayar, N., et al., Mol. Cell Endocrinol., (Ireland), 125(1-2):65-70 (Dec. 1996) (gonadotropin receptor involvement); Miura, Y., et al., Nippon Rinsho (Japan), 54(7):1784-1789 (July 1996) (apoptosis involvement); Harris, P. C., et al., J. Am. Soc. Nephrol., 6(4):1125-1133 (Oct. 1995) (kidney disease involvement); and Ruoslahti, E. I., et al., WO9110727-A by La Jolla Cancer Research Foundation (decorin binding to transforming growth factor-β involvement for treatment for cancer, wound healing and scarring). Also of particular interest is fibromodulin and its use to prevent or reduce dermal scarring. A study of fibromodulin is found in U.S. Patent No. 5,654,270 to Ruoslahti, et al.

Efforts are therefore being undertaken by both industry and academia to identify new proteins having leucine rich repeats to better understand protein-protein interactions. Of particular interest are those proteins having leucine rich repeats and homology to known proteins having leucine rich repeats such as fibromodulin, the SLIT protein and platelet glycoprotein V. Many efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel secreted and membrane-bound proteins having leucine rich repeats. We herein describe the identification and characterization of novel polypeptides having homology to fibromodulin, herein designated as PRO265 polypeptides.

6. PRO219

Human matrilin-2 polypeptide is a member of the von Willebrand factor type A-like module superfamily. von Willebrand factor is a protein which plays an important role in the maintenence of hemostasis. More specifically, von Willebrand factor is a protein which is known to participate in platelet-vessel wall interactions at the site of vascular injury via its ability to interact and form a complex with Factor VIII. The absence of von Willebrand factor in the blood causes an abnormality with the blood platelets that prevents platelet adhesion to the vascular wall at the site of the vascular injury. The result is the propensity for brusing, nose bleeds, intestinal bleeding, and the like comprising von Willebrand's disease.

Given the physiological importance of the blood clotting factors, efforts are currently being undertaken by both industry and academia to identify new, native proteins which may be involved in the coagulation process. We herein describe the identification of a novel full-length polypeptide which possesses homology to the human matrilin-2 precursor polypeptide.

7. PRO246

The cell surface protein HCAR is a membrane-bound protein that acts as a receptor for subgroup C of the adenoviruses and subgroup B of the coxsackieviruses. Thus, HCAR may provide a means for mediating viral infection of cells in that the presence of the HCAR receptor on the cellular surface provides a binding site for viral particles, thereby facilitating viral infection.

In light of the physiological importance of membrane-bound proteins and specficially those which serve a cell surface receptor for viruses, efforts are currently being undertaken by both industry and academia to identify

new, native membrane-bound receptor proteins. Many of these efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel receptor proteins. We herein describe a novel membrane-bound polypeptide (designated herein as PRO246) having homology to the cell surface protein HCAR and to various tumor antigens including A33 and carcinoembryonic antigen, wherein this polypeptide may be a novel cell surface virus receptor or tumor antigen.

8. PRO228

There are a number of known seven transmembrane proteins and within this family is a group which includes CD97 and EMR1. CD97 is a seven-span transmembrane receptor which has a cellular ligand, CD55, DAF. Hamann, et al., J. Exp. Med. (U.S.), 184(3):1189 (1996). Additionally, CD97 has been reported as being a dedifferentiation marker in human thyroid carcinomas and as associated with inflammation. Aust, et al., Cancer Res. (U.S.), 57(9):1798 (1997); Gray, et al., J. Immunol. (U.S.), 157(12):5438 (1996). CD97 has also been reported as being related to the secretin receptor superfamily, but unlike known members of that family, CD97 and EMR1 have extended extracellular regions that possess several EGF domains at the N-terminus. Hamann, et al., Genomics, 32(1):144 (1996); Harmann, et al., J. Immunol., 155(4):1942 (1995). EMR1 is further described in Lin, et al., Genomics, 41(3):301 (1997) and Baud, et al., Genomics, 26(2):334 (1995). While CD97 and EMR1 appear to be related to the secretin receptors, a known member of the secretin family of G protein-coupled receptors includes the alpha-latroxin receptor, latrophilin, which has been described as calcium independent and abundant among neuronal tissues. Lelianova, et al., J. Biol. Chem., 272(34), 21504 (1997); Davletov, et al., J. Biol. Chem. (U.S.), 271(38):23239 (1996). Both members of the secretin receptor superfamily and non-members which are related to the secretin receptor superfamily, or CRF and calcitonin receptors are of interest. In particular, new members of these families, identified by their homology to known proteins, are of interest.

Efforts are being undertaken by both industry and academia to identify new membrane-bound receptor proteins, particularly transmembrane proteins with EGF repeats and large N-terminuses which may belong to the family of seven-transmembrane proteins of which CD97 and EMR1 are members. We herein describe the identification and charactization of novel polypeptides having homology to CD97 and EMR1, designated herein as PRO228 polypeptides.

9. PRO533

Growth factors are molecular signals or mediators that enhance cell growth or proliferation, alone or in concert, by binding to specific cell surface receptors. however, there are other cellular reactions than only growth upon expression to growth factors. As a result, growth factors are better characterized as multifunctional and potent cellular regulators. Their biological effects include proliferation, chemotaxis and stimulation of extracellular matrix production. Growth factors can have both stimulatory and inhibitory effects. For example, transforming growth factors (TGF-β) is highly pleiotropic and can stimulate proliferation in some cells, especially connective tissues, while being a potent inhibitor of proliferation in others, such as lymphocytes and epithelial cells.

The physiological effect of growth stimulation or inhibition by growth factors depends upon the state of development and differentiation of the target tissue. The mechanism of local cellular regulation by classical endocrine

molecules comprehends autocrine (same cell), juxtacrine (neighbor cell), and paracrine (adjacent cell) pathways. Peptide growth factors are elements of a complex biological language, providing the basis for intercellular communication. They permit cells to convey information between each other, mediate interaction between cells and change gene expression. the effect of these multifunctional and pluripotent factors is dependent on the presence or absence of other peptides.

Fibroblast growth factors (FGFs) are a family of heparin-binding, potent mitogens for both normal diploid fibroblasts and established cell lines, Godpodarowicz, D. et al. (1984), Proc. Natl. Acad. Sci. USA 81: 6983. the FGF family comprises acidic FGF (FGF-1), basic FGF (FGF-2), INT-2 (FGF-3), K-FGF/HST (FGF-4), FGF-5, FGF-6, KGF (FGF-7), AIGF (FGF-8) among others. All FGFs have two conserved cysteine residues and share 30-50% sequence homology at the amino acid level. These factors are mitogenic for a wide variety of normal diploid mesoderm-derived and neural crest-derived cells, inducing granulosa cells, adrenal cortical cells, chrondocytes, myoblasts, corneal and vascular endothelial cells (bovine or human), vascular smooth muscle cells, lens, retina and prostatic epithelial cells, oligodendrocytes, astrocytes, chrondocytes, myoblasts and osteoblasts.

Fibroblast growth factors can also stimulate a large number of cell types in a non-mitogenic manner. These activities include promotion of cell migration into a wound area (chemotaxis), initiation of new blood vessel formulation (angiogenesis), modulation of nerve regeneration and survival (neurotrophism), modulation of endocrine functions, and stimulation or suppression of specific cellular protein expression, extracellular matrix production and cell survival. Baird, A. & Bohlen, P., *Handbook of Exp. Phrmacol.* 95(1): 369-418 (1990). These properties provide a basis for using fibroblast growth factors in therapeutic approaches to accelerate wound healing, nerve repair, collateral blood vessel formation, and the like. For example, fibroblast growth factors, have been suggested to minimize myocardium damage in heart disease and surgery (U.S.P. 4,378,437).

We herein describe the identification and characterization of novel polypeptides having homology to FGF, herein designated PRO533 polypeptides.

10. PRO245

Some of the most important proteins involved in the above described regulation and modulation of cellular processes are the enzymes which regulate levels of protein phosphorylation in the cell. For example, it is known that the transduction of signals that regulate cell growth and differentiation is regulated at least in part by phosphorylation and dephosphorylation of various cellular proteins. The enzymes that catalyze these processes include the protein kinases, which function to phosphorylate various cellular proteins, and the protein phosphatases, which function to remove phosphate residues from various cellular proteins. The balance of the level of protein phosphorylation in the cell is thus mediated by the relative activities of these two types of enzymes.

Although many protein kinase enzymes have been identified, the physiological role played by many of these catalytic proteins has yet to be elucidated. It is well known, however, that a number of the known protein kinases function to phosphorylate tyrosine residues in proteins, thereby leading to a variety of different effects. Perhaps most importantly, there has been a great deal of interest in the protein tyrosine kinases since the discovery that many oncogene products and growth factors possess intrinsic protein tyrosine kinase activity. There is, therefore, a desire to identify new members of the protein tyrosine kinase family.

Given the physiological importance of the protein kinases, efforts are being undertaken by both industry and academia to identify new, native kinase proteins. Many of these efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel kinase proteins. We herein describe the identification and characterization of novel polypeptides having homology to tyrosine kinase proteins, designated herein as PRO245 polypeptides.

11. PRO220, PRO221 and PRO227

Protein-protein interactions include receptor and antigen complexes and signaling mechanisms. As more is known about the structural and functional mechanisms underlying protein-protein interactions, protein-protein interactions can be more easily manipulated to regulate the particular result of the protein-protein interaction. Thus, the underlying mechanisms of protein-protein interactions are of interest to the scientific and medical community.

All proteins containing leucine-rich repeats are thought to be involved in protein-protein interactions. Leucine-rich repeats are short sequence motifs present in a number of proteins with diverse functions and cellular locations. The crystal structure of ribonuclease inhibitor protein has revealed that leucine-rich repeats correspond to beta-alpha structural units. These units are arranged so that they form a parallel beta-sheet with one surface exposed to solvent, so that the protein acquires an unusual, nonglubular shape. These two features have been indicated as responsible for the protein-binding functions of proteins containing leucine-rich repeats. See, Kobe and Deisenhofer, Trends Biochem. Sci., 19(10):415-421 (Oct. 1994).

A study has been reported on leucine-rich proteoglycans which serve as tissue organizers, orienting and ordering collagen fibrils during ontogeny and are involved in pathological processes such as wound healing, tissue repair, and tumor stroma formation. Iozzo, R. V., Crit. Rev. Biochem. Mol. Biol., 32(2):141-174 (1997). Others studies implicating leucine rich proteins in wound healing and tissue repair are De La Salle, C., et al., Vouv. Rev. Fr. Hematol. (Germany), 37(4):215-222 (1995), reporting mutations in the leucine rich motif in a complex associated with the bleeding disorder Bernard-Soulier syndrome and Chlemetson, K. J., Thromb. Haemost. (Germany), 74(1):111-116 (July 1995), reporting that platelets have leucine rich repeats. Another protein of particular interest which has been reported to have leucine-rich repeats is the SLIT protein which has been reported to be useful in treating neuro-degenerative diseases such as Alzheimer's disease, nerve damage such as in Parkinson's disease, and for diagnosis of cancer, see, Artavanistsakonas, S. and Rothberg, J. M., WO9210518-A1 by Yale University. Other studies reporting on the biological functions of proteins having leucine-rich repeats include: Tayar, N., et al., Mol. Cell Endocrinol., (Ireland), 125(1-2):65-70 (Dec. 1996) (gonadotropin receptor involvement); Miura, Y., et al., Nippon Rinsho (Japan), 54(7):1784-1789 (July 1996) (apoptosis involvement); Harris, P. C., et al., J. Am. Soc. Nephrol., 6(4):1125-1133 (Oct. 1995) (kidney disease involvement); and Ruoslahti, E. I., et al., WO9110727-A by La Jolla Cancer Research Foundation (decorin binding to transforming growth factorβ involvement for treatment for cancer, wound healing and scarring).

Efforts are therefore being undertaken by both industry and academia to identify new proteins having leucine rich repeats to better understand protein-protein interactions. Of particular interest are those proteins having leucine rich repeats and homology to known proteins having leucine rich repeats such as the SLIT protein and platelet glycoprotein V.

12. PRO258

Immunoglobulins are antibody molecules, the proteins that function both as receptors for antigen on the B-cell membrane and as the secreted products of the plasma cell. Like all antibody molecules, immunoglobulins perform two major functions: they bind specifically to an antigen and they participate in a limited number of biological effector functions. Therefore, new members of the Ig superfamily are always of interest. Molecules which act as receptors by various viruses and those which act to regulate immune function are of particular interest. Also of particular interest are those molecules which have homology to known Ig family members which act as virus receptors or regulate immune function. Thus, molecules having homology to poliovirus receptors, CRTAM and CD166 (a ligand for lymphocyte antigen CD6) are of particular interest.

Extracellular and membrane-bound proteins play important roles in the formation, differentiation and maintenance of multicellular organisms. The fate of many individual cells, e.g., proliferation, migration, differentiation, or interaction with other cells, is typically governed by information received from other cells and/or the immediate environment. This information is often transmitted by secreted polypeptides (for instance, mitogenic factors, survival factors, cytotoxic factors, differentiation factors, neuropeptides, and hormones) which are, in turn, received and interpreted by diverse cell receptors or membrane-bound proteins. These secreted polypeptides or signaling molecules normally pass through the cellular secretory pathway to reach their site of action in the extracellular environment, usually at a membrane-bound receptor protein.

We herein describe the identification and characterization of novel polypeptides having homology to CRTAM, designated herein as PRO258 polypeptides.

13. PRO266

Protein-protein interactions include receptor and antigen complexes and signaling mechanisms. As more is known about the structural and functional mechanisms underlying protein-protein interactions, protein-protein interactions can be more easily manipulated to regulate the particular result of the protein-protein interaction. Thus, the underlying mechanisms of protein-protein interactions are of interest to the scientific and medical community.

All proteins containing leucine-rich repeats are thought to be involved in protein-protein interactions. Leucine-rich repeats are short sequence motifs present in a number of proteins with diverse functions and cellular locations. The crystal structure of ribonuclease inhibitor protein has revealed that leucine-rich repeats correspond to beta-alpha structural units. These units are arranged so that they form a parallel beta-sheet with one surface exposed to solvent, so that the protein acquires an unusual, nonglobular shape. These two features have been indicated as responsible for the protein-binding functions of proteins containing leucine-rich repeats. See, Kobe and Deisenhofer, Trends Biochem. Sci., 19(10):415-421 (Oct. 1994).

A study has been reported on leucine-rich proteoglycans which serve as tissue organizers, orienting and ordering collagen fibrils during ontogeny and are involved in pathological processes such as wound healing, tissue repair, and tumor stroma formation. Iozzo, R. V., Crit. Rev. Biochem. Mol. Biol., 32(2):141-174 (1997). Others studies implicating leucine rich proteins in wound healing and tissue repair are De La Salle, C., et al., Vouv. Rev. Fr. Hematol. (Germany), 37(4):215-222 (1995), reporting mutations in the leucine rich motif in a complex associated with the bleeding disorder Bernard-Soulier syndrome and Chlemetson, K. J., Thromb. Haemost. (Germany),

74(1):111-116 (July 1995), reporting that platelets have leucine rich repeats. Another protein of particular interest which has been reported to have leucine-rich repeats is the SLIT protein which has been reported to be useful in treating neuro-degenerative diseases such as Alzheimer's disease, nerve damage such as in Parkinson's disease, and for diagnosis of cancer, see, Artavanistsakonas, S. and Rothberg, J. M., WO9210518-A1 by Yale University. Other studies reporting on the biological functions of proteins having leucine-rich repeats include: Tayar, N., et al., Mol. Cell Endocrinol., (Ireland), 125(1-2):65-70 (Dec. 1996) (gonadotropin receptor involvement); Miura, Y., et al., Nippon Rinsho (Japan), 54(7):1784-1789 (July 1996) (apoptosis involvement); Harris, P. C., et al., J. Am. Soc. Nephrol., 6(4):1125-1133 (Oct. 1995) (kidney disease involvement); and Ruoslahti, E. I., et al., WO9110727-A by La Jolla Cancer Research Foundation (decorin binding to transforming growth factorβ involvement for treatment for cancer, wound healing and scarring).

Efforts are therefore being undertaken by both industry and academia to identify new proteins having leucine rich repeats to better understand protein-protein interactions, neuronal development and adhesin molecules. Of particular interest are those proteins having leucine rich repeats and homology to known proteins having leucine rich repeats such as the SLIT protein. We herein describe novel polypeptides having homology to SLIT, designated herein as PRO266 polypeptides.

14. PRO269

Thrombomodulin binds to and regulates the activity of thrombin. It is important in the control of blood coagulation. Thrombomodulin functions as a natural anticoagulant by accelerating the activation of protein C by thrombin. Soluble thrombomodulin may have therapeutic use as an antithrombotic agent with reduced risk for hemorrhage as compared with heparin. Thrombomodulin is a cell surface trans-membrane glycoprotein, present on endothelial cells and platelets. A smaller, functionally active form of thrombomodulin circulates in the plasma and is also found in urine. (In Haeberli, A., Human Protein Data, VCH Oub., N.Y., 1992). Peptides having homology to thrombomodulin are particularly desirable.

We herein describe the identification and characterization of novel polypeptides having homology to thrombomodulin, designated herein as PRO269 polypeptides.

15. PRO287

Procollagen C-proteinase enhancer protein binds to and enhances the activity of bone morphogenic protein "BMP1"/procollagen C-proteinase (PCP). It plays a role in extracellular matrix deposition. BMP1 proteins may be used to induce bone and/or cartilage formation and in wound healing and tissue repair. Therefore, procollagen C-proteinase enhancer protein, BMP1 and proteins having homology thereto, are of interest to the scientific and medical communities.

We herein describe the identification and characterization of novel polypeptides having homology to procollagen C-proteinase enhancer protein precursor and procollagen C-proteinase enhancer protein, designated herein as PRO287 polypeptides.

16. PRO214

Growth factors are molecular signals or mediators that enhances cell growth or proliferation, alone or in concert, by binding to specific cell surface receptors. However, there are other cellular reactions than only growth upon expression to growth factors. As a result, growth factors are better characterized as multifunctional and potent cellular regulators. Their biological effects include proliferation, chemotaxis and stimulation of extracellular matrix production. Growth factors can have both stimulatory and inhibitory effects. For example, transforming growth factor β (TGF-β) is highly pleiotropic and can stimulate proliferation in some cells, especially connective tissue, while being a potent inhibitor of proliferation in others, such as lymphocytes and epithelial cells.

The physiological effect of growth stimulation or inhibition by growth factors depends upon the state of development and differentiation of the target tissue. The mechanism of local cellular regulation by classical endocrine molecules involves comprehends autocrine (same cell), juxtacrine (neighbor cell), and paracrine (adjacent cells) pathways. Peptide growth factors are elements of a complex biological language, providing the basis for intercellular communication. They permit cells to convey information between each other, mediate interaction between cells and change gene expression. The effect of these multifunctional and pluripotent factors is dependent on the presence or absence of other peptides.

Epidermal growth factor (EGF) is a conventional mitogenic factor that stimulates the proliferation of various types of cells including epithelial cells and fibroblasts. EGF binds to and activates the EGF receptor (EGFR), which initiates intracellular signaling and subsequent effects. The EGFR is expressed in neurons of the cerebral cortex, cerebellum, and hippocampus in addition to other regions of the central nervous system (CNS). In addition, EGF is also expressed in various regions of the CNS. Therefore, EGF acts not only on mitotic cells, but also on postmitotic neurons. In fact, many studies have indicated that EGF has neurotrophic or neuromodulatory effects on various types of neurons in the CNS. For example, EGF acts directly on cultured cerebral cortical and cerebellar neurons, enhancing neurite outgrowth and survival. On the other hand, EGF also acts on other cell types, including septal cholinergic and mesencephalic dopaminergic neurons, indirectly through glial cells. Evidence of the effects of EGF on neurons in the CNS is accumulating, but the mechanisms of action remain essentially unknown. EGF-induced signaling in mitotic cells is better understood than in postmitotic neurons. Studies of cloned pheochromocytoma PC12 cells and cultured cerebral cortical neurons have suggested that the EGF-induced neurotrophic actions are mediated by sustained activation of the EGFR and mitogen-activated protein kinase (MAPK) in response to EGF. The sustained intracellular signaling correlates with the decreased rate of EGFR down-regulation, which might determine the response of neuronal cells to EGF. It is likely that EGF is a multi-potent growth factor that acts upon various types of cells including mitotic cells and postmitotic neurons.

EGF is produced by the salivary and Brunner's glands of the gastrointestinal system, kidney, pancreas, thyroid gland, pituitary gland, and the nervous system, and is found in body fluids such as saliva, blood, cerebrospinal fluid (CSF), urine, amniotic fluid, prostatic fluid, pancreatic juice, and breast milk, Plata-Salaman, CR *Peptides* 12: 653-663 (1991).

EGF is mediated by its membrane specific receptor, which contains an intrinsic tyrosine kinase. Stoscheck CM *et al., J. Cell Biochem.* 31: 135-152 (1986). EGF is believed to function by binding to the extracellular portion of its receptor which induces a transmembrane signal that activates the intrinsic tyrosine kinase.

Purification and sequence analysis of the EGF-like domain has revealed the presence of six conserved cysteine residues which cross-bind to create three peptide loops, Savage *CR et al., J. Biol. Chem.* 248: 7669-7672 (1979). It is now generally known that several other peptides can react with the EGF receptor which share the same generalized motif $X_nCX_7CX_{4/5}CX_{10}CXCX_5GX_2CX_n$, where X represents any non-cysteine amino acid, and n is a variable repeat number. Non isolated peptides having this motif include TGF-a, amphiregulin, schwannoma-derived growth factor (SDGF), heparin-binding EGF-like growth factors and certain virally encoded peptides (e.g., Vaccinia virus, Reisner AH, *Nature* 313: 801-803 (1985), Shope fibroma virus, Chang W., et al., Mol Cell Biol. 7: 535-540 (1987), Molluscum contagiosum, Porter CD & Archard LC, *J. Gen. Virol.* 68: 673-682 (1987), and Myxoma virus, Upton C *et al., J. Virol.* 61: 1271-1275 (1987). Prigent SA & Lemoine N.R., *Prog. Growth Factor Res.* 4: 1-24 (1992).

EGF-like domains are not confined to growth factors but have been observed in a variety of cell-surface and extracellular proteins which have interesting properties in cell adhesion, protein-protein interaction and development, Laurence DJR & Gusterson BA, *Tumor Biol.* 11: 229-261 (1990). These proteins include blood coagulation factors (factors VI, IX, X, XII, protein C, protein S, protein Z, tissue plasminogen activator, urokinase), extracellular matrix components (laminin, cytotactin, entactin), cell surface receptors (LDL receptor, thrombomodulin receptor) and immunity-related proteins (complement C1r, uromodulin).

Even more interesting, the general structure pattern of EGF-like precursors is preserved through lower organisms as well as in mammalian cells. A number of genes with developmental significance have been identified in invertebrates with EGF-like repeats. For example, the *notch* gene of Drosophila encodes 36 tandemly arranged 40 amino acid repeats which show homology to EGF, Wharton W *et al., Cell* 43: 557-581 (1985). Hydropathy plots indicate a putative membrane spanning domain, with the EGF-related sequences being located on the extracellular side of the membrane. Other homeotic genes with EGF-like repeats include Delta, 95F and 5ZD which were identified using probes based on Notch, and the nematode gene *Lin-12* which encodes a putative receptor for a developmental signal transmitted between two specified cells.

Specifically, EGF has been shown to have potential in the preservation and maintenance of gastrointestinal mucosa and the repair of acute and chronic mucosal lesions, Konturek, *PC et al., Eur. J. Gastroenterol Hepatol.* 7 (10), 933-37 (1995), including the treatment of necrotizing enterocolitis, Zollinger-Ellison syndrome, gastrointestinal ulceration gastrointestinal ulcerations and congenital microvillus atrophy, A. Guglietta & PB Sullivan, *Eur. J. Gastroenterol Hepatol*, 7(10), 945-50 (1995). Additionally, EGF has been implicated in hair follicle differentiation; C.L. du Cros, *J. Invest. Dermatol.* 101 (1 Suppl.), 106S-113S (1993), SG Hillier, *Clin. Endocrinol.* 33(4), 427-28 (1990); kidney function, L.L. Hamm *et al., Semin. Nephrol.* 13 (1): 109-15 (1993), RC Harris, *Am. J. Kidney Dis.* 17(6): 627-30 (1991); tear fluid, GB van Setten *et al., Int. Ophthalmol* 15(6); 359-62 (1991); vitamin K mediated blood coagulation, J. Stenflo *et al., Blood* 78(7): 1637-51 (1991). EGF is also implicated various skin disease characterized by abnormal keratinocyte differentiation, e.g., psoriasis, epithelial cancers such as squamous cell carcinomas of the lung, epidermoid carcinoma of the vulva and gliomas. King, LE *et al., Am. J. Med. Sci.* 296: 154-158 (1988).

Of great interest is mounting evidence that genetic alterations in growth factors signaling pathways are closely linked to developmental abnormalities and to chronic diseases including cancer. Aaronson SA, *Science* 254:

1146-1153 (1991). For example, c-erb-2 (also known as HER-2), a proto-oncogene with close structural similarity to EGF receptor protein, is overexpressed in human breast cancer. King *et al.*, *Science* 229: 974-976 (1985); Gullick, WJ, *Hormones and their actions*, Cooke BA *et al.*, eds, Amsterdam, Elsevier, pp 349-360 (1986).

17. **PRO317**

The TGF-β supergene family, or simply TGF-β superfamily, a group of secreted proteins, includes a large number of related growth and differentiation factors expressed in virtually all phyla. Superfamily members bind to specific cell surface receptors that activate signal transduction mechanisms to elicit their multifunctional cytokine effects. Kolodziejczyk and Hall, Biochem. Cell. Biol., 74: 299-314 (1996); Attisano and Wrana, Cytokine Growth Factor Rev., 7: 327-339 (1996); and Hill, Cellular Signaling, 8: 533-544 (1996).

Members of this family include five distinct forms of TGF-β (Sporn and Roberts, in Peptide Growth Factors and Their Receptors, Sporn and Roberts, eds. (Springer-Verlag: Berlin, 1990) pp. 419-472), as well as the differentiation factors vg1 (Weeks and Melton, Cell, 51: 861-867 (1987)) and DPP-C polypeptide (Padgett *et al.*, Nature, 325: 81-84 (1987)), the hormones activin and inhibin (Mason *et al.*, Nature, 318: 659-663 (1985); Mason *et al.*, Growth Factors, 1: 77-88 (1987)), the Mullerian-inhibiting substance (MIS) (Cate *et al.*, Cell, 45: 685-698 (1986)), the bone morphogenetic proteins (BMPs) (Wozney *et al.*, Science, 242: 1528-1534 (1988); PCT WO 88/00205 published January 14, 1988; U.S. 4,877,864 issued October 31, 1989), the developmentally regulated proteins Vgr-1 (Lyons *et al.*, Proc. Natl. Acad. Sci. USA. 86: 4554-4558 (1989)) and Vgr-2 (Jones *et al.*, Molec. Endocrinol., 6: 1961-1968 (1992)), the mouse growth differentiation factor (GDF), such as GDF-3 and GDF-9 (Kingsley, Genes Dev., 8: 133-146 (1994); McPherron and Lee, J. Biol. Chem., 268: 3444-3449 (1993)), the mouse lefty/Stra1 (Meno *et al.*, Nature, 381: 151-155 (1996); Bouillet *et al.*, Dev. Biol., 170: 420-433 (1995)), glial cell line-derived neurotrophic factor (GDNF) (Lin *et al.*, Science, 260: 1130-1132 (1993), neurturin (Kotzbauer *et al.*, Nature, 384: 467-470 (1996)), and endometrial bleeding-associated factor (EBAF) (Kothapalli *et al.*, J. Clin. Invest., 99: 2342-2350 (1997)). The subset BMP-2A and BMP-2B is approximately 75% homologous in sequence to DPP-C and may represent the mammalian equivalent of that protein.

The proteins of the TGF-β superfamily are disulfide-linked homo- or heterodimers encoded by larger precursor polypeptide chains containing a hydrophobic signal sequence, a long and relatively poorly conserved N-terminal pro region of several hundred amino acids, a cleavage site (usually polybasic), and a shorter and more highly conserved C-terminal region. This C-terminal region corresponds to the processed mature protein and contains approximately 100 amino acids with a characteristic cysteine motif, *i.e.*, the conservation of seven of the nine cysteine residues of TGF-β among all known family members. Although the position of the cleavage site between the mature and pro regions varies among the family members, the C-terminus of all of the proteins is in the identical position, ending in the sequence Cys-X-Cys-X, but differing in every case from the TGF-β consensus C-terminus of Cys-Lys-Cys-Ser. Sporn and Roberts, 1990, *supra*.

There are at least five forms of TGF-β currently identified, TGF-β1, TGF-β2, TGF-β3, TGF-β4, and TGF-β5. The activated form of TGF-β1 is a homodimer formed by dimerization of the carboxy-terminal 112 amino acids of a 390 amino acid precursor. Recombinant TGF-β1 has been cloned (Derynck *et al.*, Nature, 316:701-705 (1985)) and expressed in Chinese hamster ovary cells (Gentry *et al.*, Mol. Cell. Biol., 7: 3418-3427 (1987)). Additionally,

recombinant human TGF-β2 (deMartin *et al.*, EMBO J., 6: 3673 (1987)), as well as human and porcine TGF-β3 (Derynck *et al.*, EMBO J., 7: 3737-3743 (1988); ten Dijke *et al.*, Proc. Natl. Acad. Sci. USA, 85: 4715 (1988)) have been cloned. TGF-β2 has a precursor form of 414 amino acids and is also processed to a homodimer from the carboxy-terminal 112 amino acids that shares approximately 70% homology with the active form of TGF-β1 (Marquardt *et al.*, J. Biol. Chem., 262: 12127 (1987)). See also EP 200,341; 169,016; 268,561; and 267,463; U.S. Pat. No. 4,774,322; Cheifetz *et al.*, Cell, 48: 409-415 (1987); Jakowlew *et al.*, Molecular Endocrin., 2: 747-755 (1988); Derynck *et al.*, J. Biol. Chem., 261: 4377-4379 (1986); Sharples *et al.*, DNA, 6: 239-244 (1987); Derynck *et al.*, Nucl. Acids. Res., 15: 3188-3189 (1987); Derynck *et al.*, Nucl. Acids. Res., 15: 3187 (1987); Seyedin *et al.*, J. Biol. Chem., 261: 5693-5695 (1986); Madisen *et al.*, DNA, 7: 1-8 (1988); and Hanks *et al.*, Proc. Natl. Acad. Sci. (U.S.A.), 85: 79-82 (1988).

TGF-β4 and TGF-β5 were cloned from a chicken chondrocyte cDNA library (Jakowlew *et al.*, Molec. Endocrinol., 2: 1186-1195 (1988)) and from a frog oocyte cDNA library, respectively.

The pro region of TGF-β associates non-covalently with the mature TGF-β dimer (Wakefield *et al.*, J. Biol. Chem., 263: 7646-7654 (1988); Wakefield *et al.*, Growth Factors, 1: 203-218 (1989)), and the pro regions are found to be necessary for proper folding and secretion of the active mature dimers of both TGF-β and activin (Gray and Mason, Science, 247: 1328-1330 (1990)). The association between the mature and pro regions of TGF-β masks the biological activity of the mature dimer, resulting in formation of an inactive latent form. Latency is not a constant of the TGF-β superfamily, since the presence of the pro region has no effect on activin or inhibin biological activity.

A unifying feature of the biology of the proteins from the TGF-β superfamily is their ability to regulate developmental processes. TGF-β has been shown to have numerous regulatory actions on a wide variety of both normal and neoplastic cells. TGF-β is multifunctional, as it can either stimulate or inhibit cell proliferation, differentiation, and other critical processes in cell function (Sporn and Roberts, *supra).*

One member of the TGF-β superfamily, EBAF, is expressed in endometrium only in the late secretory phase and during abnormal endometrial bleeding. Kothapalli *et al.*, J. Clin. Invest., 99: 2342-2350 (1997). Human endometrium is unique in that it is the only tissue in the body that bleeds at regular intervals. In addition, abnormal endometrial bleeding is one of the most common manifestations of gynecological diseases, and is a prime indication for hysterectomy. *In situ* hybridization showed that the mRNA of EBAF was expressed in the stroma without any significant mRNA expression in the endometrial glands or endothelial cells.

The predicted protein sequence of EBAF showed a strong homology to the protein encoded by mouse *lefty/stra3* of the TGF-β superfamily. A motif search revealed that the predicted EBAF protein contains most of the cysteine residues which are conserved among the TGF-β-related proteins and which are necessary for the formation of the cysteine knot structure. The EBAF sequence contains an additional cysteine residue, 12 amino acids upstream from the first conserved cysteine residue. The only other family members known to contain an additional cysteine residue are TGF-βs, inhibins, and GDF-3. EBAF, similar to LEFTY, GDF-3/Vgr2, and GDF-9, lacks the cysteine residue that is known to form the intermolecular disulfide bond. Therefore, EBAF appears to be an additional member of the TGF-β superfamily with an unpaired cysteine residue that may not exist as a dimer. However, hydrophobic contacts between the two monomer subunits may promote dimer formation. Fluorescence *in situ* hybridization showed that the *ebaf* gene is located on human chromosome 1 at band q42.1.

Additional members of the TGF-β superfamily, such as those related to EBAF, are being searched for by industry and academics. We herein describe the identification and characterization of novel polypeptides having homology to EBAF, designated herein as PRO317 polypeptides.

18. PRO301

The widespread occurrence of cancer has prompted the devotion of considerable resources and discovering new treatments of treatment. One particular method involves the creation of tumor or cancer specific monoclonal antibodies (mAbs) which are specific to tumor antigens. Such mAbs, which can distinguish between normal and cancerous cells are useful in the diagnosis, prognosis and treatment of the disease. Particular antigens are known to be associated with neoplastic diseases, such as colorectal cancer.

One particular antigen, the A33 antigen is expressed in more than 90% of primary or metastatic colon cancers as well as normal colon epithelium. Since colon cancer is a widespread disease, early diagnosis and treatment is an important medical goal. Diagnosis and treatment of colon cancer can be implemented using monoclonal antibodies (mAbs) specific therefore having fluorescent, nuclear magnetic or radioactive tags. Radioactive gene, toxins and/or drug tagged mAbs can be used for treatment *in situ* with minimal patient description. mAbs can also be used to diagnose during the diagnosis and treatment of colon cancers. For example, when the serum levels of the A33 antigen are elevated in a patient, a drop of the levels after surgery would indicate the tumor resection was successful. On the other hand, a subsequent rise in serum A33 antigen levels after surgery would indicate that metastases of the original tumor may have formed or that new primary tumors may have appeared. Such monoclonal antibodies can be used in lieu of, or in conjunction with surgery and/or other chemotherapies. For example, U.S.P. 4,579,827 and U.S.S.N. 424,991 (E.P. 199,141) are directed to therapeutic administration of monoclonal antibodies, the latter of which relates to the application of anti-A33 mAb.

Many cancers of epithelial origin have adenovirus receptors. In fact, adenovirus-derived vectors have been proposed as a means of inserting antisense nucleic acids into tumors (U.S.P. 5,518,885). Thus, the association of viral receptors with neoplastic tumors is not unexpected.

We herein describe the identification and characterization of novel polypeptides having homology to certain cancer-associated antigens, designated herein as PRO301 polypeptides.

19. PRO224

Cholesterol uptake can have serious implications on one's health. Cholesterol uptake provides cells with most of the cholesterol they require for membrane synthesis. If this uptake is blocked, cholesterol accumulates in the blood and can contribute to the formation of atherosclerotic plaques in blood vessel walls. Most cholesterol is transported in the blood bound to protein in the form of complexes known as low-density lipoproteins (LDLs). LDLs are endocytosed into cells via LDL receptor proteins. Therefore, LDL receptor proteins, and proteins having homology thereto, are of interest to the scientific and medical communities. Membrane-bound proteins and receptors can play an important role in the formation, differentiation and maintenance of multicellular organisms. The LDL receptors are an example of membrane-bound proteins which are involved in the synthesis and formation of cell membranes, wherein the health of an individual is affected directly and indirectly by its function. Many

membrane-bound proteins act as receptors such as the LDL receptor. These receptors can function to endocytose substrates or they can function as a receptor for a channel. Other membrane-bound proteins function as signals or antigens.

Membrane-bound proteins and receptor molecules have various industrial applications, including as pharmaceutical and diagnostic agents. The membrane-bound proteins can also be employed for screening of potential peptide or small molecule regulators of the relevant receptor/ligand interaction. In the case of the LDL receptor, it is desirable to find molecules which enhance endocytosis so as to lower blood cholesterol levels and plaque formation. It is also desirable to identify molecules which inhibit endocytosis so that these molecules can be avoided or regulated by individuals having high blood cholesterol. Polypeptides which are homologous to lipoprotein receptors but which do not function as lipoprotein receptors are also of interest in the determination of the function of the fragments which show homology.

The following studies report on previously known low density lipoprotein receptors and related proteins including apolipoproteins: Sawamura, et al., Nippon Chemiphar Co, Japan patent application J09098787; Novak, S., et al., J. Biol. Chem., 271:(20)11732-6 (1996); Blaas, D., J. Virol., 69(11)7244-7 (Nov. 1995); Scott, J., J. Inherit. Metab. Dis. (UK), 9/Supp. 1 (3-16) (1986); Yamamoto, et al., Cell, 39:27-38 (1984); Rebece, et al., Neurobiol. Aging, 15:5117 (1994); Novak, S., et al., J. Biol. Chemistry, 271:11732-11736 (1996); and Sestavel and Fruchart, Cell Mol. Biol., 40(4):461-81 (June 1994). These publications and others published prior to the filing of this application provide further background to peptides already known in the art.

Efforts are being undertaken by both industry and academia to identify new, native membrane-bound receptor proteins, particularly those having homology to lipoprotein receptors. We herein describe the identification and characterization of novel polypeptides having homology to lipoprotein receptors, designated herein as PRO224 polypeptides.

20. PRO222

Complement is a group of proteins found in the blood that are important in humoral immunity and inflammation. Complement proteins are sequentially activated by antigen-antibody complexes or by proteolytic enzymes. When activated, complement proteins kill bacteria and other microorganisms, affect vascular permeability, release histamine and attract white blood cells. Complement also enhances phagocytosis when bound to target cells. In order to prevent harm to autologous cells, the complement activation pathway is tightly regulated.

Deficiencies in the regulation of complement activation or in the complement proteins themselves may lead to immune-complex diseases, such as systemic lupus erythematosus, and may result in increased susceptibility to bacterial infection. In all cases, early detection of complement deficiency is desirable so that the patient can begin treatment. Thus, research efforts are currently directed toward identification of soluble and membrane proteins that regulate complement activation.

Proteins known to be important in regulating complement activation in humans include Factor H and Complement receptor type 1 (CR1). Factor H is a 150 kD soluble serum protein that interacts with complement protein C3b to accelerate the decay of C3 convertase and acts as a cofactor for Factor I-mediated cleavage of complement protein C4b. Complement receptor type 1 is a 190-280 kD membrane bound protein found in mast cells

and most blood cells. CR1 interacts with complement proteins C3b, C4b, and iC3b to accelerate dissociation of C3 convertases, acts as a cofactor for Factor I-mediated cleavage of C3b and C4b, and binds immune complexes and promotes their dissolution and phagocytosis.

Proteins which have homology to complement proteins are of particular interest to the medical and industrial communities. Often, proteins having homology to each other have similar function. It is also of interest when proteins having homology do not have similar functions, indicating that certain structural motifs identify information other than function, such as locality of function.

Efforts are being undertaken by both industry and academia to identify new, native secreted and membrane-bound proteins, particularly those having homology to known proteins involved in the complement pathway. Proteins involved in the complement pathway were reviewed in Birmingham DJ (1995), Critical Reviews in Immunology, 15(2):133-154 and in Abbas AK, et al. (1994) Cellular and Molecular Immunology, 2nd Ed. W.B. Saunders Company, Philadelphia, pp 295-315.

We herein describe the identification and characterization of novel polypeptides having homology to complement receptors, designated herein as PRO222 polypeptides.

21. PRO234

The successful function of many systems within multicellular organisms is dependent on cell-cell interactions. Such interactions are affected by the alignment of particular ligands with particular receptors in a manner which allows for ligand-receptor binding and thus a cell-cell adhesion. While protein-protein interactions in cell recognition have been recognized for some time, only recently has the role of carbohydrates in physiologically relevant recognition been widely considered (see B.K. Brandley *et al.*, *J. Leuk. Biol.* 40: 97 (1986) and N. Sharon *et al.*, *Science* 246: 227 (1989). Oligosaccharides are well positioned to act as recognition novel lectins due to their cell surface location and structural diversity. Many oligosaccharide structures can be created through the differential activities of a smaller number of glycosyltransferases. The diverse structures of oligosaccharides can be generated by transcription of relatively few gene products, which suggests that the oligosaccharides are a plausible mechanism by which is directed a wide range of cell-cell interactions. Examples of differential expression of cell surface carbohydrates and putative carbohydrate binding proteins (lectins) on interacting cells have been described (J. Dodd & T.M. Jessel, *J. Neurosci.* 5: 3278 (1985); L.J. Regan *et al.*, *Proc. Natl. Acad. Sci. USA* 83: 2248 (1986); M. Constantine-Paton *et al.*, *Nature* 324: 459 (1986); and M. Tiemeyer *et al.*, *J. Biol. Chem.* 263: 1671 (1989). One interesting member of the lectin family are selectins.

The migration of leukocytes to sites of acute or chronic inflammation involves adhesive interactions between these cells and the endothelium. This specific adhesion is the initial event in the cascade that is initiated by inflammatory insults, and it is, therefore, of paramount importance to the regulated defense of the organism.

The types of cell adhesion molecules that are involved in the interaction between leukocytes and the endothelium during an inflammatory response currently stands at four: (1) selectins; (2) (carbohydrate and glycoprotein) ligands for selectins; (3) integrins; and (4) integrin ligands, which are members of the immunoglobulin gene superfamily.

The selectins are cell adhesion molecules that are unified both structurally and functionally. Structurally, selectins are characterized by the inclusion of a domain with homology to a calcium-dependent lectin (C-lectins), an epidermal growth factor (egf)-like domain and several complement binding-like domains, Bevilacqua, M.P. *et al., Science* 243: 1160-1165 (1989); Johnston *et al., Cell* 56: 1033-1044 (1989); Lasky *et al, Cell* 56: 1045-1055 (1989); Siegalman, M. *et al., Science* 243: 1165-1172 (1989); Stoolman, L.M., *Cell* 56: 907-910 (1989). Functionally, selectins share the common property of their ability to mediate cell binding through interactions between their lectin domains and cell surface carbohydrate ligands (Brandley, B, *et al., Cell* 63, 861-863 (1990); Springer, T. and Lasky, L.A., *Nature* 349, 19-197 (1991); Bevilacqua, M.P. and Nelson, R.M., *J. Clin. Invest.* 91 379-387 (1993) and Tedder *et al., J. Exp. Med.* 170: 123-133 (1989).

There are three members identified so far in the selectin family of cell adhesion molecules: L-selectin (also called peripheral lymph node homing receptor (pnHR), LEC-CAM-1, LAM-1, $gp90^{MEL}$, $gp100^{MEL}$, $gp110^{MEL}$, MEL-14 antigen, Leu-8 antigen, TQ-1 antigen, DREG antigen), E-selectin (LEC-CAM-2, LECAM-2, ELAM-1) and P-selectin (LEC-CAM-3, LECAM-3, GMP-140, PADGEM).

The identification of the C-lectin domain has led to an intense effort to define carbohydrate binding ligands for proteins containing such domains. E-selectin is believed to recognize the carbohydrate sequence NeuNAcα2-3Galβ1-4(Fucα1-3)GlcNAc (sialyl-Lewis x, or sLe^x) and related oligosaccharides, Berg *et al., J. Biol. Chem.* 265: 14869-14872 (1991); Lowe *et al., Cell* 63: 475-484 (1990); Phillips *et al., Science* 250: 1130-1132 (1990); *Tiemeyer et al., Proc. Natl. Acad. Sci. USA* 88: 1138-1142 (1991).

L-selectin, which comprises a lectin domain, performs its adhesive function by recognizing carbohydrate-containing ligands on endothelial cells. L-selectin is expressed on the surface of leukocytes, such as lymphocytes, neutrophils, monocytes and eosinophils, and is involved with the trafficking of lymphocytes to peripheral lymphoid tissues (*Gallatin et al., Nature* 303: 30-34 (1983)) and with acute neutrophil-medicated inflammatory responses (Watson, S.R., *Nature* 349: 164-167 (1991)). The amino acid sequence of L-selectin and the encoding nucleic acid sequence are, for example, disclosed in U.S. patent No. 5,098,833 issued 24 March 1992.

L-selectin (LECAM-1) is particularly interesting because of its ability to block neutrophil influx (Watson *et al., Nature* 349: 164-167 (1991). It is expressed in chronic lymphocytic leukemia cells which bind to HEV (Spertini *et al., Nature* 349: 691-694 (1991). It is also believed that HEV structures at sites of chronic inflammation are associated with the symptoms of diseases such as rheumatoid arthritis, psoriasis and multiple sclerosis.

E-selectin (ELAM-1), is particularly interesting because of its transient expression on endothelial cells in response to IL-1 or TNF. Bevilacqua et al., *Science* 243: 1160 (1989). The time course of this induced expression (2-8 h) suggests a role for this receptor in initial neutrophil induced extravasation in response to infection and injury. It has further been reported that anti-ELAM-1 antibody blocks the influx of neutrophils in a primate asthma model and thus is beneficial for preventing airway obstruction resulting from the inflammatory response. Gundel *et al., J. Clin. Invest.* 88: 1407 (1991).

The adhesion of circulating neutrophils to stimulated vascular endothelium is a primary event of the inflammatory response. P-selectin has been reported to recognize the Lewis x structure (Galβ1-4(Fucα1-3) GlcNAc), Larsen *et al., Cell* 63: 467-474(1990). Others report that an additional terminal linked sialic acid is required for high affinity binding, Moore *et al., J. Cell. Biol.* 112: 491-499 (1991). P-selectin has been shown to be significant in acute

lung injury. Anti-P-selectin antibody has been shown to have strong protective effects in a rodent lung injury model. M.S. Mulligan *et al., J. Clin. Invest.* 90: 1600 (1991).

We herein describe the identification and characterization of novel polypeptides having homology to lectin proteins, herein designated as PRO234 polypeptides.

22. **PRO231**

Some of the most important proteins involved in the above described regulation and modulation of cellular processes are the enzymes which regulate levels of protein phosphorylation in the cell. For example, it is known that the transduction of signals that regulate cell growth and differentiation is regulated at least in part by phosphorylation and dephosphorylation of various cellular proteins. The enzymes that catalyze these processes include the protein kinases, which function to phosphorylate various cellular proteins, and the protein phosphatases, which function to remove phosphate residues from various cellular proteins. The balance of the level of protein phosphorylation in the cell is thus mediated by the relative activities of these two types of enzymes.

Protein phosphatases represent a growing family of enzymes that are found in many diverse forms, including both membrane-bound and soluble forms. While many protein phosphatases have been described, the functions of only a very few are beginning to be understood (Tonks, *Semin. Cell Biol.* 4:373-453 (1993) and Dixon, *Recent Prog. Horm. Res.* 51:405-414 (1996)). However, in general, it appears that many of the protein phosphatases function to modulate the positive or negative signals induced by various protein kinases. Therefore, it is likely that protein phosphatases play critical roles in numerous and diverse cellular processes.

Given the physiological importance of the protein phosphatases, efforts are being undertaken by both industry and academia to identify new, native phosphatase proteins. Many of these efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel phosphatase proteins. Examples of screening methods and techniques are described in the literature [see, for example, Klein et al., Proc. Natl. Acad. Sci., 93:7108-7113 (1996); U.S. Patent No. 5,536,637)].

We herein describe the identification and characterization of novel polypeptides having homology to acid phosphatases, designated herein as PRO231 polypeptides.

23. **PRO229**

Scavenger receptors are known to protect IgG molecules from catabolic degradation. Riechmann and Hollinger, Nature Biotechnology, 15:617 (1997). In particular, studies of the CH2 and CH3 domains have shown that specific sequences of these domains are important in determining the half-lives of antibodies. Ellerson, et al., J. Immunol., 116: 510 (1976); Yasmeen, et al., J. Immunol. 116: 518 (1976; Pollock, et al., Eur. J. Immunol., 20: 2021 (1990). Scavenger receptor proteins and antibodies thereto are further reported in U.S. Patent No. 5,510,466 to Krieger, et al. Due to the ability of scavenger receptors to increase the half-life of polypeptides and their involvement in immune function, molecules having homology to scavenger receptors are of importance to the scientific and medical community.

Efforts are being undertaken by both industry and academia to identify new, native secreted and membrane-bound receptor proteins, particularly those having homology to scavenger receptors. Many efforts are focused on

the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel secreted and membrane-bound receptor proteins. Examples of screening methods and techniques are described in the literature [see, for example, Klein et al., Proc. Natl. Acad. Sci., 93:7108-7113 (1996); U.S. Patent No. 5,536,637)].

We herein describe the identification and characterization of novel polypeptides having homology to scavenger receptors, designated herein as PRO229 polypeptides.

24. **PRO238**

Oxygen free radicals and antioxidants appear to play an important role in the central nervous system after cerebral ischemia and reperfusion. Moreover, cardiac injury, related to ischaemia and reperfusion has been reported to be caused by the action of free radicals. Additionally, studies have reported that the redox state of the cell is a pivotal determinant of the fate of the cells. Furthermore, reactive oxygen species have been reported to be cytotoxic, causing inflammatory disease, including tissue necrosis, organ failure, atherosclerosis, infertility, birth defects, premature aging, mutations and malignancy. Thus, the control of oxidation and reduction is important for a number of reasons including for control and prevention of strokes, heart attacks, oxidative stress and hypertension. In this regard, reductases, and particularly, oxidoreductases, are of interest. Publications further describing this subject matter include Kelsey, et al., Br. J. Cancer, 76(7):852-4 (1997); Friedrich and Weiss, J. Theor. Biol., 187(4):529-40 (1997) and Pieulle, et al., J. Bacteriol., 179(18):5684-92 (1997).

Efforts are being undertaken by both industry and academia to identify new, native secreted and membrane-bound receptor proteins, particularly secreted proteins which have homology to reductase. Many efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel secreted and membrane-bound receptor proteins. Examples of screening methods and techniques are described in the literature [see, for example, Klein et al., Proc. Natl. Acad. Sci., 93:7108-7113 (1996); U.S. Patent No. 5,536,637)].

We herein describe the identification and characterization of novel polypeptides having homology to reductase, designated herein as PRO238 polypeptides.

25. **PRO233**

Studies have reported that the redox state of the cell is an important determinant of the fate of the cell. Furthermore, reactive oxygen species have been reported to be cytotoxic, causing inflammatory disease, including tissue necrosis, organ failure, atherosclerosis, infertility, birth defects, premature aging, mutations and malignancy. Thus, the control of oxidation and reduction is important for a number of reasons, including the control and prevention of strokes, heart attacks, oxidative stress and hypertension. Oxygen free radicals and antioxidants appear to play an important role in the central nervous system after cerebral ischemia and reperfusion. Moreover, cardiac injury, related to ischaemia and reperfusion has been reported to be caused by the action of free radicals. In this regard, reductases, and particularly, oxidoreductases, are of interest. In addition, the transcription factors, NF-kappa B and AP-1, are known to be regulated by redox state and to affect the expression of a large variety of genes thought to be involved in the pathogenesis of AIDS, cancer, atherosclerosis and diabetic complications. Publications further describing this subject matter include Kelsey, et al., Br. J. Cancer, 76(7):852-4 (1997); Friedrich and Weiss, J. Theor. Biol., 187(4):529-40 (1997) and Pieulle, et al., J. Bacteriol., 179(18):5684-92 (1997). Given the physiological

importance of redox reactions *in vivo*, efforts are currently being under taken to identify new, native proteins which are involved in redox reactions. We describe herein the identification of novel polypeptides which have homology to reductase, designated herein as PRO233 polypeptides.

26. PRO223

The carboxypeptidase family of exopeptidases constitutes a diverse group of enzymes that hydrolyze carboxyl-terminal amide bonds in polypeptides, wherein a large number of mammalian tissues produce these enzymes. Many of the carboxypeptidase enzymes that have been identified to date exhibit rather strong cleavage specificities for certain amino acids in polypeptides. For example, carboxypeptidase enzymes have been identified which prefer lysine, arginine, serine or amino acids with either aromatic or branched aliphatic side chains as substrates at the carboxyl terminus of the polypeptide.

With regard to the serine carboxypeptidases, such amino acid specific enzymes have been identified from a variety of different mammalian and non-mammalian organisms. The mammalian serine carboxypeptidase enzymes play important roles in many different biological processes including, for example, protein digestion, activation, inactivation, or modulation of peptide hormone activity, and alteration of the physical properties of proteins and enzymes.

In light of the physiological importance of the serine carboxypeptidases, efforts are being undertaken by both industry and academia to identify new, native secreted and membrane-bound receptor proteins and specifically novel carboxypeptidases. Many of these efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel secreted and membrane-bound receptor proteins. We describe herein novel polypeptides having homology to one or more serine carboxypeptidase polypeptides, designated herein as PRO223 polypeptides.

27. PRO235

Plexin was first identified in Xenopus tadpole nervous system as a membrane glycoprotein which was shown to mediate cell adhesion via a homophilic binding mechanism in the presence of calcium ions. Strong evolutionary conservation between Xenopus, mouse and human homologs of plexin has been observed. [Kaneyama et al., Biochem. And Biophys. Res. Comm. 226: 524-529 (1996)]. Given the physiological importance of cell adhesion mechanisms *in vivo*, efforts are currently being under taken to identify new, native proteins which are involved in cell adhesion. We describe herein the identification of a novel polypeptide which has homology to plexin, designated herein as PRO235.

28. PRO236 and PRO262

β-galactosidase is a well known enzymatic protein which functions to hydrolyze β-galactoside molecules. β-galactosidase has been employed for a variety of different applications, both *in vitro* and *in vivo* and has proven to be an extremely useful research tool. As such, there is an interest in obtaining novel polypeptides which exhibit homology to the β-galactosidase polypeptide.

Given the strong interest in obtaining novel polypeptides having homology to β-galactosidase, efforts are currently being undertaken by both industry and academia to identify new, native β-galactosidase homolog proteins. Many of these efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel β-galactosidase-like proteins. Examples of screening methods and techniques are described in the literature [see, for example, Klein et al., Proc. Natl. Acad. Sci., 93:7108-7113 (1996); U.S. Patent No. 5,536,637)]. We herein describe novel poylpeptides having siginificant homology to the β-galactosidase enzyme, designated herein as PRO236 and PRO262 polypeptides.

29. PRO239

Densin is a glycoprotein which has been isolated from the brain which has all the hallmarks of an adhesion molecule. It is highly concentrated at synaptic sites in the brain and is expressed prominently in dendritic processes in developing neurons. Densin has been characterized as a member of the O-linked sialoglycoproteins. Densin has relevance to medically important processes such as regeneration. Given the physiological importance of synaptic processes and cell adhesion mechanisms *in vivo*, efforts are currently being under taken to identify new, native proteins which are involved in synaptic machinery and cell adhesion. We describe herein the identification of novel polypeptides which have homology to densin, designated herein as PRO239 polypeptides.

30. PRO257

Ebnerin is a cell surface protein associated with von Ebner glands in mammals. Efforts are being undertaken by both industry and academia to identify new, native cell surface receptor proteins and specifically those which possess sequence homology to cell surface proteins such as ebnerin. Many of these efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel receptor proteins. We herein describe the identification of novel polypeptides having significant homology to the von Ebner's gland-associated protein ebnerin, designated herein as PRO257 polypeptides.

31. PRO260

Fucosidases are enzymes that remove fucose residues from fucose containing proteoglycans. In some pathological conditions, such as cancer, rheumatoid arthritis, and diabetes, there is an abnormal fucosylation of serum proteins. Therefore, fucosidases, and proteins having homology to fucosidase, are of importance to the study and abrogation of these conditions. In particular, proteins having homology to the alpha-l-fucosidase precursor are of interest. Fucosidases and fucosidase inhibitors are further described in U.S. Patent Nos. 5,637,490, 5,382,709, 5,240,707, 5,153,325, 5,100,797, 5,096,909 and 5,017,704. Studies are also reported in Valk, et al., J. Virol., 71(9):6796 (1997), Aktogu, et al., Monaldi. Arch. Chest Dis. (Italy), 52(2):118 (1997) and Focarelli, et al., Biochem. Biophys. Res. Commun. (U.S.), 234(1):54 (1997).

Efforts are being undertaken by both industry and academia to identify new, native secreted and membrane-bound receptor proteins. Of particular interest are proteins having homology to the alpha-l-fucosidase precursor. Many efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel secreted and membrane-bound receptor proteins. Examples of screening methods and techniques are

described in the literature [see, for example, Klein et al., Proc. Natl. Acad. Sci., 93:7108-7113 (1996); U.S. Patent No. 5,536,637)].

We herein describe the identification and characterization of novel polypeptides having homology to fucosidases, designated herein as PRO260 polypeptides.

32. **PRO263**

CD44 is a cell surface adhesion molecule involved in cell-cell and cell-matrix interactions. Hyaluronic acid, a component of the extracellular matrix is a major ligand. Other ligands include collagen, fibronectin, laminin, chrondroitin sulfate, mucosal addressin, serglycin and osteoponin. CD44 is also important in regulating cell traffic, lymph node homing, transmission of growth signals, and presentation of chemokines and growth factors to traveling cells. CD44 surface proteins are associated with metastatic tumors and CD44 has been used as a marker for HIV infection. Certain splice variants are associated with metastasis and poor prognosis of cancer patients. Therefore, molecules having homology with CD44 are of particular interest, as their homology indicates that they may have functions related to those functions of CD44. CD44 is further described in U.S. Patent Nos. 5,506,119, 5,504,194 and 5,108,904; Gerberick, et al., Toxicol. Appl. Pharmacol., 146(1):1 (1997); Wittig, et al., Immunol. Letters (Netherlands), 57(1-3):217 (1997); and Oliveira and Odell, Oral Oncol. (England), 33(4):260 (1997).

Efforts are being undertaken by both industry and academia to identify new, native secreted and membrane-bound receptor proteins, particularly transmembrane proteins with homology to CD44 antigen. Many efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel secreted and membrane-bound receptor proteins. Examples of screening methods and techniques are described in the literature [see, for example, Klein et al., Proc. Natl. Acad. Sci., 93:7108-7113 (1996); U.S. Patent No. 5,536,637)].

We herein describe the identification and characterization of novel polypeptides having homology to CD44 antigen, designated herein as PRO263 polypeptides.

33. **PRO270**

Thioredoxins effect reduction-oxidation (redox) state. Many diseases are potentially related to redox state and reactive oxygen species may play a role in many important biological processes. The transcription factors, NF-kappa B and AP-1, are regulated by redox state and are known to affect the expression of a large variety of genes thought to be involved in the pathogenesis of AIDS, cancer, atherosclerosis and diabetic complications. Such proteins may also play a role in cellular antioxidant defense, and in pathological conditions involving oxidative stress such as stroke and inflammation in addition to having a role in apoptosis. Therefore, thioredoxins, and proteins having homology thereto, are of interest to the scientific and medical communities.

We herein describe the identification and characterization of novel polypeptides having homology to thioredoxin, designated herein as PRO270 polypeptides.

34. PRO271

The proteoglycan link protein is a protein which is intimately associated with various extracellular matrix proteins and more specifically with proteins such as collagen. For example, one primary component of collagen is a large proteoglycan called aggrecan. This molecule is retained by binding to the glycosaminoglycan hyaluronan through the amino terminal G1 globular domain of the core protein. This binding is stabilized by the proteoglycan link protein which is a protein that is also associated with other tissues containing hyaluronan binding proteoglycans such as versican.

Link protein has been identified as a potential target for autoimmune antibodies in individuals who suffer from juvenile rheumatoid arthritis (see Guerassimov et al., *J. Rheumatology* 24(5):959-964 (1997)). As such, there is strong interest in identifying novel proteins having homology to link protein. We herein describe the identification and characterization of novel polypeptides having such homology, designated herein as PRO271 polypeptides.

35. PRO272

Reticulocalbin is an endoplasmic reticular protein which may be involved in protein transport and luminal protein processing. Reticulocalbin resides in the lumen of the endopladsmic rerticulum, is known to bind calcium, and may be involved in a luminal retention mechanism of the endoplasmic reticulum. It contains six domains of the EF-hand motif associated with high affinity calcium binding. We describe herein the identification and characterization of a novel polypeptide which has homology to the reticulocalbin protein, designated herein as PRO272.

36. PRO294

Collagen, a naturally occurring protein, finds wide application in industry. Chemically hydrolyzed natural collagen can be denatured and renatured by heating and cooling to produce gelatin, which is used in photographic and medical, among other applications. Collagen has important properties such as the ability to form interchain aggregates having a conformation designated as a triple helix. We herein describe the identification and characterization of a novel polypeptide which has homology to portions of the collagen molecule, designated herein as PRO294.

37. PRO295

The integrins comprise a supergene family of cell-surface glycoprotein receptors that promote cellular adhesion. Each cell has numerous receptors that define its cell adhesive capabilities. Integrins are involved in a wide variety of interaction between cells and other cells or matrix components. The integrins are of particular importance in regulating movement and function of immune system cells The platelet IIb/IIIA integrin complex is of particular importance in regulating platelet aggregation. A member of the integrin family, integrin β-6, is expressed on epithelial cells and modulates epithelial inflammation. Another integrin, leucocyte-associated antigen-1 (LFA-1) is important in the adhesion of lymphocytes during an immune response. The integrins are expressed as heterodimers of non-covalently associated alpha and beta subunits. Given the physiological importance of cell adhesion mechanisms *in vivo*, efforts are currently being under taken to identify new, native proteins which are involved in cell adhesion. We

describe herein the identification and characterization of a novel polypeptide which has homology to integrin, designated herein as PRO295.

38. PRO293

Protein-protein interactions include receptor and antigen complexes and signaling mechanisms. As more is known about the structural and functional mechanisms underlying protein-protein interactions, protein-protein interactions can be more easily manipulated to regulate the particular result of the protein-protein interaction. Thus, the underlying mechanisms of protein-protein interactions are of interest to the scientific and medical community.

All proteins containing leucine-rich repeats are thought to be involved in protein-protein interactions. Leucine-rich repeats are short sequence motifs present in a number of proteins with diverse functions and cellular locations. The crystal structure of ribonuclease inhibitor protein has revealed that leucine-rich repeats correspond to beta-alpha structural units. These units are arranged so that they form a parallel beta-sheet with one surface exposed to solvent, so that the protein acquires an unusual, nonglubular shape. These two features have been indicated as responsible for the protein-binding functions of proteins containing leucine-rich repeats. See, Kobe and Deisenhofer, Trends Biochem. Sci., 19(10):415-421 (Oct. 1994).

A study has been reported on leucine-rich proteoglycans which serve as tissue organizers, orienting and ordering collagen fibrils during ontogeny and are involved in pathological processes such as wound healing, tissue repair, and tumor stroma formation. Iozzo, R. V., Crit. Rev. Biochem. Mol. Biol., 32(2):141-174 (1997). Others studies implicating leucine rich proteins in wound healing and tissue repair are De La Salle, C., et al., Vouv. Rev. Fr. Hematol. (Germany), 37(4):215-222 (1995), reporting mutations in the leucine rich motif in a complex associated with the bleeding disorder Bernard-Soulier syndrome and Chlemetson, K. J., Thromb. Haemost. (Germany), 74(1):111-116 (July 1995), reporting that platelets have leucine rich repeats. Another protein of particular interest which has been reported to have leucine-rich repeats is the SLIT protein which has been reported to be useful in treating neuro-degenerative diseases such as Alzheimer's disease, nerve damage such as in Parkinson's disease, and for diagnosis of cancer, see, Artavanistsakonas, S. and Rothberg, J. M., WO9210518-A1 by Yale University. Other studies reporting on the biological functions of proteins having leucine-rich repeats include: Tayar, N., et al., Mol. Cell Endocrinol., (Ireland), 125(1-2):65-70 (Dec. 1996) (gonadotropin receptor involvement); Miura, Y., et al., Nippon Rinsho (Japan), 54(7):1784-1789 (July 1996) (apoptosis involvement); Harris, P. C., et al., J. Am. Soc. Nephrol., 6(4):1125-1133 (Oct. 1995) (kidney disease involvement); and Ruoslahti, E. I., et al., WO9110727-A by La Jolla Cancer Research Foundation (decorin binding to transforming growth factorβ involvement for treatment for cancer, wound healing and scarring).

Efforts are therefore being undertaken by both industry and academia to identify new proteins having leucine rich repeats to better understand protein-protein interactions. Of particular interest are those proteins having leucine rich repeats and homology to known neuronal leucine rich repeat proteins. Many efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel secreted and membrane-bound proteins having leucine rich repeats. Examples of screening methods and techniques are described in the literature [see, for example, Klein et al., Proc. Natl. Acad. Sci., 93:7108-7113 (1996); U.S. Patent No. 5,536,637)].

We describe herein the identification and characterization of a novel polypeptide which has homology to leucine rich repeat proteins, designated herein as PRO293.

39. **PRO247**

Protein-protein interactions include receptor and antigen complexes and signaling mechanisms. As more is known about the structural and functional mechanisms underlying protein-protein interactions, protein-protein interactions can be more easily manipulated to regulate the particular result of the protein-protein interaction. Thus, the underlying mechanisms of protein-protein interactions are of interest to the scientific and medical community.

All proteins containing leucine-rich repeats are thought to be involved in protein-protein interactions. Leucine-rich repeats are short sequence motifs present in a number of proteins with diverse functions and cellular locations. The crystal structure of ribonuclease inhibitor protein has revealed that leucine-rich repeats correspond to beta-alpha structural units. These units are arranged so that they form a parallel beta-sheet with one surface exposed to solvent, so that the protein acquires an unusual, nonglubular shape. These two features have been indicated as responsible for the protein-binding functions of proteins containing leucine-rich repeats. See, Kobe and Deisenhofer, Trends Biochem. Sci., 19(10):415-421 (Oct. 1994).

A study has been reported on leucine-rich proteoglycans which serve as tissue organizers, orienting and ordering collagen fibrils during ontogeny and are involved in pathological processes such as wound healing, tissue repair, and tumor stroma formation. Iozzo, R. V., Crit. Rev. Biochem. Mol. Biol., 32(2):141-174 (1997). Others studies implicating leucine rich proteins in wound healing and tissue repair are De La Salle, C., et al., Vouv. Rev. Fr. Hematol. (Germany), 37(4):215-222 (1995), reporting mutations in the leucine rich motif in a complex associated with the bleeding disorder Bernard-Soulier syndrome and Chlemetson, K. J., Thromb. Haemost. (Germany), 74(1):111-116 (July 1995), reporting that platelets have leucine rich repeats. Another protein of particular interest which has been reported to have leucine-rich repeats is the SLIT protein which has been reported to be useful in treating neuro-degenerative diseases such as Alzheimer's disease, nerve damage such as in Parkinson's disease, and for diagnosis of cancer, see, Artavanistsakonas, S. and Rothberg, J. M., WO9210518-A1 by Yale University. Other studies reporting on the biological functions of proteins having leucine-rich repeats include: Tayar, N., et al., Mol. Cell Endocrinol., (Ireland), 125(1-2):65-70 (Dec. 1996) (gonadotropin receptor involvement); Miura, Y., et al., Nippon Rinsho (Japan), 54(7):1784-1789 (July 1996) (apoptosis involvement); Harris, P. C., et al., J. Am. Soc. Nephrol., 6(4):1125-1133 (Oct. 1995) (kidney disease involvement); and Ruoslahti, E. I., et al., WO9110727-A by La Jolla Cancer Research Foundation (decorin binding to transforming growth factorβ involvement for treatment for cancer, wound healing and scarring).

Densin is a glycoprotein which has been isolated from the brain which has all the hallmarks of an adhesion molecule. It is highly concentrated at synaptic sites in the brain and is expressed prominently in dendritic processes in developing neurons. Densin has been characterized as a member of the O-linked sialoglycoproteins. Densin has relevance to medically important processes such as regeneration. Given the physiological importance of synaptic processes and cell adhesion mechanisms *in vivo*, efforts are currently being under taken to identify new, native proteins which are involved in synaptic machinery and cell adhesion. Densin is further described in Kennedy, M.B, Trends Neurosci. (England), 20(6):264 (1997) and Apperson, et al., J. Neurosci., 16(21):6839 (1996).

Efforts are therefore being undertaken by both industry and academia to identify new proteins having leucine rich repeats to better understand protein-protein interactions. Of particular interest are those proteins having leucine rich repeats and homology to known proteins having leucine rich repeats such as KIAA0231 and densin. Many efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel secreted and membrane-bound proteins having leucine rich repeats. Examples of screening methods and techniques are described in the literature [see, for example, Klein et al., Proc. Natl. Acad. Sci., 93:7108-7113 (1996); U.S. Patent No. 5,536,637)].

We describe herein the identification and characterization of a novel polypeptide which has homology to leucine rich repeat proteins, designated herein as PRO247.

40. PRO302, PRO303, PRO304, PRO307 and PRO343

Proteases are enzymatic proteins which are involved in a large number of very important biological processes in mammalian and non-mammalian organisms. Numerous different protease enzymes from a variety of different mammalian and non-mammalian organisms have been both identified and characterized. The mammalian protease enzymes play important roles in many different biological processes including, for example, protein digestion, activation, inactivation, or modulation of peptide hormone activity, and alteration of the physical properties of proteins and enzymes.

In light of the important physiological roles played by protease enzymes, efforts are currently being undertaken by both industry and academia to identify new, native protease homologs. Many of these efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel secreted and membrane-bound receptor proteins. Examples of screening methods and techniques are described in the literature [see, for example, Klein et al., Proc. Natl. Acad. Sci., 93:7108-7113 (1996); U.S. Patent No. 5,536,637)]. We herein describe the identification of novel polypeptides having homology to various protease enzymes, designated herein as PRO302, PRO303, PRO304, PRO307 and PRO343 polypeptides.

41. PRO328

The GLIP protein family has been characterized as comprising zinc-finger proteins which play important roles in embryogenesis. These proteins may function as transcriptional regulatory proteins and are known to be amplified in a subset of human tumors. Glioma pathogenesis protein is structurally related to a group of plant pathogenesis-related proteins. It is highly expressed in glioblastoma. See US Pat. Nos. 5,582,981 (issued Dec. 10, 1996) and 5,322,801 (issued June 21, 1996), Ellington, A.D. et al., Nature, 346:818 (1990), Grindley, J.C. et al., Dev. Biol., 188(2):337 (1997), Marine, J.C. et al., Mech. Dev., 63(2):211 (1997), The CRISP or cysteine rich secretory protein family are a group of proteins which are also structurally related to a group of plant pathogenesis proteins. [Schwidetzky, U., Biochem. J., 321:325 (1997), Pfisterer, P., Mol. Cell Biol., 16(11):6160 (1996), Kratzschmar, J., Eur. J. Biochem., 236(3):827 (1996)]. We describe herein the identification of a novel polypeptide which has homology to GLIP and CRISP, designated herein as PRO328 polypeptides.

42. PRO335, PRO331 and PRO326

Protein-protein interactions include receptor and antigen complexes and signaling mechanisms. As more is known about the structural and functional mechanisms underlying protein-protein interactions, protein-protein interactions can be more easily manipulated to regulate the particular result of the protein-protein interaction. Thus, the underlying mechanisms of protein-protein interactions are of interest to the scientific and medical community.

All proteins containing leucine-rich repeats are thought to be involved in protein-protein interactions. Leucine-rich repeats are short sequence motifs present in a number of proteins with diverse functions and cellular locations. The crystal structure of ribonuclease inhibitor protein has revealed that leucine-rich repeats correspond to beta-alpha structural units. These units are arranged so that they form a parallel beta-sheet with one surface exposed to solvent, so that the protein acquires an unusual, nonglubular shape. These two features have been indicated as responsible for the protein-binding functions of proteins containing leucine-rich repeats. See, Kobe and Deisenhofer, Trends Biochem. Sci., 19(10):415-421 (Oct. 1994).

A study has been reported on leucine-rich proteoglycans which serve as tissue organizers, orienting and ordering collagen fibrils during ontogeny and are involved in pathological processes such as wound healing, tissue repair, and tumor stroma formation. Iozzo, R. V., Crit. Rev. Biochem. Mol. Biol., 32(2):141-174 (1997). Others studies implicating leucine rich proteins in wound healing and tissue repair are De La Salle, C., et al., Vouv. Rev. Fr. Hematol. (Germany), 37(4):215-222 (1995), reporting mutations in the leucine rich motif in a complex associated with the bleeding disorder Bernard-Soulier syndrome, Chlemetson, K. J., Thromb. Haemost. (Germany), 74(1):111-116 (July 1995), reporting that platelets have leucine rich repeats and Ruoslahti, E. I., et al., WO9110727-A by La Jolla Cancer Research Foundation reporting that decorin binding to transforming growth factorβ has involvement in a treatment for cancer, wound healing and scarring. Related by function to this group of proteins is the insulin like growth factor (IGF), in that it is useful in wound-healing and associated therapies concerned with re-growth of tissue, such as connective tissue, skin and bone; in promoting body growth in humans and animals; and in stimulating other growth-related processes. The acid labile subunit of IGF (ALS) is also of interest in that it increases the half-life of IGF and is part of the IGF complex in vivo.

Another protein which has been reported to have leucine-rich repeats is the SLIT protein which has been reported to be useful in treating neuro-degenerative diseases such as Alzheimer's disease, nerve damage such as in Parkinson's disease, and for diagnosis of cancer, see, Artavanistsakonas, S. and Rothberg, J. M., WO9210518-A1 by Yale University. Of particular interest is LIG-1, a membrane glycoprotein that is expressed specifically in glial cells in the mouse brain, and has leucine rich repeats and immunoglobulin-like domains. Suzuki, et al., J. Biol. Chem. (U.S.), 271(37):22522 (1996). Other studies reporting on the biological functions of proteins having leucine rich repeats include: Tayar, N., et al., Mol. Cell Endocrinol., (Ireland), 125(1-2):65-70 (Dec. 1996) (gonadotropin receptor involvement); Miura, Y., et al., Nippon Rinsho (Japan), 54(7):1784-1789 (July 1996) (apoptosis involvement); Harris, P. C., et al., J. Am. Soc. Nephrol., 6(4):1125-1133 (Oct. 1995) (kidney disease involvement).

Efforts are therefore being undertaken by both industry and academia to identify new proteins having leucine rich repeats to better understand protein-protein interactions. Of particular interest are those proteins having leucine rich repeats and homology to known proteins having leucine rich repeats such as LIG-1, ALS and decorin. Many efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for

novel secreted and membrane-bound proteins having leucine rich repeats. Examples of screening methods and techniques are described in the literature [see, for example, Klein et al., Proc. Natl. Acad. Sci., 93:7108-7113 (1996); U.S. Patent No. 5,536,637)].

We describe herein the identification and characterization of novel polypeptides which have homology to proteins of the leucine rich repeat superfamily, designated herein as PRO335, PRO331 and PRO326 polypeptides.

43. **PRO332**

Secreted proteins comprising a repeat characterized by an arrangement of conserved leucine residues (leucine-rich repeat motif) have diverse biological roles. Certain proteoglycans, such as biglycan, fibromodulin and decorin, are, for example, characterized by the presence of a leucine-rich repeat of about 24 amino acids [Ruoslahti, Ann. Rev. Cell. Biol. 4 229-255 (1988); Oldberg *et al.*, EMBO J. 8, 2601-2604 (1989)]. In general, proteoglycans are believed to play a role in regulating extracellular matrix, cartilage or bone function. The proteoglycan decorin binds to collagen type I and II and affects the rate of fibril formation. Fibromodulin also binds collagen and delays fibril formation. Both fibromodulin and decorin inhibit the activity of transforming growth factor beta (TGF-β) (U.S. Patent No. 5,583,103 issued December 10, 1996). TGF-β is known to play a key role in the induction of extracellular matrix and has been implicated in the development of fibrotic diseases, such as cancer and glomerulonephritis. Accordingly, proteoglycans have been proposed for the treatment of fibrotic cancer, based upon their ability to inhibit TGF-β's growth stimulating activity on the cancer cell. Proteoglycans have also been described as potentially useful in the treatment of other proliferative pathologies, including rheumatoid arthritis, arteriosclerosis, adult respiratory distress syndrome, cirrhosis of the liver, fibrosis of the lungs, post-myocardial infarction, cardiac fibrosis, post-angioplasty restenosis, renal interstitial fibrosis and certain dermal fibrotic conditions, such as keloids and scarring, which might result from burn injuries, other invasive skin injuries, or cosmetic or reconstructive surgery (U.S. Patent No. 5,654,270, issued August 5, 1997).

We describe herein the identification and characterization of novel polypeptides which have homology to proteins of the leucine rich repeat superfamily, designated herein as PRO332 polypeptides.

44. **PRO334**

Microfibril bundles and proteins found in association with these bundles, particularly attachment molecules, are of interest in the field of dermatology, particularly in the study of skin which has been damaged from aging, injuries or the sun. Fibrillin microfibrils define the continuous elastic network of skin, and are present in dermis as microfibril bundles devoid of measurable elastin extending from the dermal-epithelial junction and as components of the thick elastic fibres present in the deep reticular dermis. Moreover, Marfan syndrome has been linked to mutations which interfere with multimerization of fibrillin monomers or other connective tissue elements.

Fibulin-1 is a modular glycoprotein with amino-terminal anaphlatoxin-like modules followed by nine epidermal growth factor (EGF)-like modules and, depending on alternative splicing, four possible carboxyl termini. Fibulin-2 is a novel extracellular matrix protein frequently found in close association with microfibrils containing either fibronectin or fibrillin. Thus, fibrillin, fibulin, and molecules related thereto are of interest, particularly for the use of preventing skin from being damaged from aging, injuries or the sun, or for restoring skin damaged from

same. Moreover, these molecules are generally of interest in the study of connective tissue and attachment molecules and related mechanisms. Fibrillin, fibulin and related molecules are further described in Adams, et al., J. Mol. Biol., 272(2):226-36 (1997); Kielty and Shuttleworth, Microsc. Res. Tech., 38(4):413-27 (1997); and Child, J. Card. Surg., 12(2Supp.):131-5 (1997).

Currently, efforts are being undertaken by both industry and academia to identify new, native secreted and membrane-bound receptor proteins, particularly secreted proteins which have homology to fibulin and fibrillin. Many efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel secreted and membrane-bound receptor proteins. Examples of screening methods and techniques are described in the literature [see, for example, Klein et al., Proc. Natl. Acad. Sci., 93:7108-7113 (1996); U.S. Patent No. 5,536,637)].

We herein describe the identification and characterization of novel polypeptides having homology to fibulin and fibrillin, designated herein as PRO334 polypeptides.

45. PRO346

The widespread occurrence of cancer has prompted the devotion of considerable resources and discovering new treatments of treatment. One particular method involves the creation of tumor or cancer specific monoclonal antibodies (mAbs) which are specific to tumor antigens. Such mAbs, which can distinguish between normal and cancerous cells are useful in the diagnosis, prognosis and treatment of the disease. Particular antigens are known to be associated with neoplastic diseases, such as colorectal and breast cancer. Since colon cancer is a widespread disease, early diagnosis and treatment is an important medical goal. Diagnosis and treatment of cancer can be implemented using monoclonal antibodies (mAbs) specific therefore having fluorescent, nuclear magnetic or radioactive tags. Radioactive genes, toxins and/or drug tagged mAbs can be used for treatment *in situ* with minimal patient description.

Carcinoembryonic antigen (CEA) is a glycoprotein found in human colon cancer and the digestive organs of a 2-6 month human embryos. CEA is a known human tumor marker and is widely used in the diagnosis of neoplastic diseases, such as colon cancer. For example, when the serum levels of CEA are elevated in a patient, a drop of CEA levels after surgery would indicate the tumor resection was successful. On the other hand, a subsequent rise in serum CEA levels after surgery would indicate that metastases of the original tumor may have formed or that new primary tumors may have appeared. CEA may also be a target for mAb, antisense nucleotides

46. PRO268

Protein disulfide isomerase is an enzymatic protein which is involved in the promotion of correct refolding of proteins through the establishment of correct disulfide bond formation. Protein disulfide isomerase was initially identified based upon its ability to catalyze the renaturation of reduced denatured RNAse (Goldberger et al., *J. Biol. Chem.* 239:1406-1410 (1964) and Epstein et al., *Cold Spring Harbor Symp. Quant. Biol.* 28:439-449 (1963)). Protein disulfide isomerase has been shown to be a resident enzyme of the endoplasmic reticulum which is retained in the endoplasmic reticulum via a -KDEL or -HDEL amino acid sequence at its C-terminus.

Given the importance of disulfide bond-forming enzymes and their potential uses in a number of different applications, for example in increasing the yield of correct refolding of recombinantly produced proteins, efforts are currently being undertaken by both industry and academia to identify new, native proteins having homology to protein disulfide isomerase. Many of these efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel protein disulfide isomerase homologs. We herein describe a novel polypeptide having homology to protein disulfide isomerase, designated herein as PRO268.

47. PRO330

Prolyl 4-hydroxylase is an enzyme which functions to post-translationally hydroxylate proline residues at the Y position of the amino acid sequence Gly-X-Y, which is a repeating three amino acid sequence found in both collagen and procollagen. Hydroxylation of proline residues at the Y position of the Gly-X-Y amino acid triplet to form 4-hydroxyproline residues at those positions is required before newly synthesized collagen polypeptide chains may fold into their proper three-dimensional triple-helical conformation. If hydroxylation does not occur, synthesized collagen polypeptides remain non-helical, are poorly secreted by cells and cannot assemble into stable functional collagen fibrils. Vuorio et al., *Proc. Natl. Acad. Sci. USA* 89:7467-7470 (1992). Prolyl 4-hydroxylase is comprised of at least two different polypeptide subunits, alpha and beta.

Efforts are being undertaken by both industry and academia to identify new, native secreted and membrane-bound receptor proteins. Many efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel secreted and membrane-bound receptor proteins. Examples of screening methods and techniques are described in the literature [see, for example, Klein et al., Proc. Natl. Acad. Sci., 93:7108-7113 (1996); U.S. Patent No. 5,536,637)]. Based upon these efforts, Applicants have herein identified and describe a novel polypeptide having homology to the alpha subunit of prolyl 4-hydroxylase, designated herein as PRO330.

48. PRO339 and PRO310

Fringe is a protein which specifically blocks serrate-mediated activation of notch in the dorsal compartment of the Drosophila wing imaginal disc. Fleming, et al., Development, 124(15):2973-81 (1997). Therefore, fringe is of interest for both its role in development as well as its ability to regulate serrate, particularly serrate's signaling abilities. Also of interest are novel polypeptides which may have a role in development and/or the regulation of serrate-like molecules. Of particular interest are novel polypeptides having homology to fringe as identified and described herein, designated herein as PRO339 and PRO310 polypeptides.

49. PRO244

Lectins are a class of proteins comprising a region that binds carbohydrates specifically and non-covalently. Numerous lectins have been identified in higher animals, both membrane-bound and soluble, and have been implicated in a variety of cell-recognition phenomena and tumor metastasis.

Most lectins can be classified as either C-type (calcium-dependent) or S-type (thiol-dependent).

Lectins are thought to play a role in regulating cellular events that are initiated at the level of the plasma membrane. For example, plasma membrane associated molecules are involved in the activation of various subsets of lymphoid cells, e.g. T-lymphocytes, and it is known that cell surface molecules are responsible for activation of these cells and consequently their response during an immune reaction.

A particular group of cell adhesion molecules, selectins, belong in the superfamily of C-type lectins. This group includes L-selectin (peripheral lymph node homing receptor (pnHR), LEC-CAM-1, LAM-1, gp90^{MEL}, gp100^{MEL}, gp110^{MEL}, MEL-14 antigen, Leu-8 antigen, TQ-1 antigen, DREG antigen), E-selectin (LEC-CAM-2, LECAM-2, ELAM-1), and P-selectin (LEC-CAM-3, LECAM-3, GMP-140, PADGEM). The structure of selectins consists of a C-type lectin (carbohydrate binding) domain, an epidermal growth factor-like (EGF-like) motif, and variable numbers of complement regulatory (CR) motifs. Selectins are associated with leukocyte adhesion, e.g. the attachment of neutrophils to venular endothelial cells adjacent to inflammation (E-selectin), or with the trafficking of lymphocytes from blood to secondary lymphoid organs, e.g. lymph nodes and Peyer's patches (L-selectin).

Another exemplary lectin is the cell-associated macrophage antigen, Mac-2 that is believed to be involved in cell adhesion and immune responses. Macrophages also express a lectin that recognizes Tn Ag, a human carcinoma-associated epitope.

Another C-type lectin is CD95 (Fas antigen/APO-1) that is an important mediator of immunologically relevant regulated or programmed cell death (apoptosis). "Apoptosis" is a non-necrotic cell death that takes place in metazoan animal cells following activation of an intrinsic cell suicide program. The cloning of Fas antigen is described in PCT publication WO 91/10448, and European patent application EP510691. The mature Fas molecule consists of 319 amino acids of which 157 are extracellular, 17 constitute the transmembrane domain, and 145 are intracellular. Increased levels of Fas expression at T cell surface have been associated with tumor cells and HIV-infected cells. Ligation of CD95 triggers apoptosis in the presence of interleukin-1 (IL-2).

C-type lectins also include receptors for oxidized low-density lipoprotein (LDL). This suggests a possible role in the pathogenesis of atherosclerosis.

We herein describe the identification and characterization of novel polypeptides having homology to C-type lectins, designated herein as PRO244 polypeptides.

SUMMARY OF THE INVENTION

1. PRO211 and PRO217

Applicants have identified cDNA clones that encode novel polypeptides having homology to EGF, designated in the present application as "PRO211" and "PRO217" polypeptides.

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO211 or PRO217 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding EGF-like homologue PRO211 and PRO217 polypeptides of Fig. 2 (SEQ ID NO:2) and/or 4 (SEQ ID NO:4) indicated in Fig. 1 (SEQ ID NO: 1) and/or Fig. 3 (SEQ ID NO:3), respectively, or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO211 and PRO217 EGF-like homologue PRO211 and PRO217 polypeptides. In particular, the invention provides isolated native sequence PRO211 and PRO217 EGF-

like homologue polypeptides, which in one embodiment, includes an amino acid sequence comprising residues: 1 to 353 of Fig. 2 (SEQ ID NO:2) or (2) 1 to 379 of Fig. 4 (SEQ ID NO: 4).

2. PRO230

Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is designated in the present application as "PRO230".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO230 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO230 polypeptide having amino acid residues 1 through 467 of Figure 6 (SEQ ID NO:12), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO230 polypeptide. In particular, the invention provides isolated native sequence PRO230 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 through 467 of Figure 6 (SEQ ID NO:12).

In another embodiment, the invention provides an expressed sequence tag (EST) comprising the nucleotide sequence of SEQ ID NO:13 (Figure 7) which is herein designated as DNA20088.

3. PRO232

Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is designated in the present application as "PRO232".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO232 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO232 polypeptide having amino acid residues 1 to 114 of Figure 9 (SEQ ID NO:18), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO232 polypeptide. In particular, the invention provides isolated native sequence PRO232 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 114 of Figure 9 (SEQ ID NO:18).

4. PRO187

Applicants have identified a cDNA clone that encodes a novel polypeptide, designated in the present application as "PRO187".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO187 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO187 polypeptide of Figure 11 (SEQ ID NO:23), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions. In another aspect, the invention provides a nucleic acid comprising the coding sequence of Figure 10 (SEQ ID NO:22) or its complement. In another aspect, the invention provides a nucleic acid of the full length protein of clone DNA27864-1155, deposited with the ATCC under accession number ATCC 209375, alternatively the coding sequence of clone DNA27864-1155,

deposited under accession number ATCC 209375.

In yet another embodiment, the invention provides isolated PRO187 polypeptide. In particular, the invention provides isolated native sequence PRO187 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 205 of Figure 11 (SEQ ID NO:23). Alternatively, the invention provides a polypeptide encoded by the nucleic acid deposited under accession number ATCC 209375.

5. PRO265

Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is designated in the present application as "PRO265".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO265 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO265 polypeptide having amino acid residues 1 to 660 of Figure 13 (SEQ ID NO:28), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO265 polypeptide. In particular, the invention provides isolated native sequence PRO265 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 660 of Figure 13 (SEQ ID NO:28). An additional embodiment of the present invention is directed to an isolated extracellular domain of a PRO265 polypeptide.

6. PRO219

Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is designated in the present application as "PRO219".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO219 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO219 polypeptide having amino acid residues 1 to 915 of Figure 15 (SEQ ID NO:34), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO219 polypeptide. In particular, the invention provides isolated native sequence PRO219 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 915 of Figure 15 (SEQ ID NO:34).

7. PRO246

Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is designated in the present application as "PRO246".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO246 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO246 polypeptide having amino acid residues 1 to 390 of Figure 17 (SEQ ID NO:39), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency

conditions.

In another embodiment, the invention provides isolated PRO246 polypeptide. In particular, the invention provides isolated native sequence PRO246 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 390 of Figure 17 (SEQ ID NO:39). An additional embodiment of the present invention is directed to an isolated extracellular domain of a PRO246 polypeptide.

8. **PRO228**

Applicants have identified a cDNA clone that encodes a novel polypeptide having homology to CD97, EMR1 and latrophilin, wherein the polypeptide is designated in the present application as "PRO228".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO228 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO228 polypeptide having amino acid residues 1 to 690 of Figure 19 (SEQ ID NO:49), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO228 polypeptide. In particular, the invention provides isolated native sequence PRO228 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 690 of Figure 19 (SEQ ID NO:49). An additional embodiment of the present invention is directed to an isolated extracellular domain of a PRO228 polypeptide.

In another embodiment, the invention provides an expressed sequence tag (EST) comprising the nucleotide sequence of SEQ ID NO:50, designated herein as DNA21951.

9. **PRO533**

Applicants have identified a cDNA clone (DNA49435-1219) that encodes a novel polypeptide, designated in the present application as PRO533.

In one embodiment, the invention provides an isolated nucleic acid molecule having at least about 80% sequence identity to (a) a DNA molecule encoding a PRO533 polypeptide comprising the sequence of amino acids 23 to 216 of Figure 22 (SEQ ID NO:59), or (b) the complement of the DNA molecule of (a). The sequence identity preferably is about 85%, more preferably about 90%, most preferably about 95%. In one aspect, the isolated nucleic acid has at least about 80%, preferably at least about 85%, more preferably at least about 90%, and most preferably at least about 95% sequence identity with a polypeptide having amino acid residues 23 to 216 of Figure 22 (SEQ ID NO:59). Preferably, the highest degree of sequence identity occurs within the secreted portion (amino acids 23 to 216 of Figure 22, SEQ ID NO:59). In a further embodiment, the isolated nucleic acid molecule comprises DNA encoding a PRO533 polypeptide having amino acid residues 1 to 216 of Figure 22 (SEQ ID NO:59), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions. In another aspect, the invention provides a nucleic acid of the full length protein of clone DNA49435-1219, deposited with the ATCC under accession number ATCC 209480.

In yet another embodiment, the invention provides isolated PRO533 polypeptide. In particular, the invention provides isolated native sequence PRO533 polypeptide, which in one embodiment, includes an amino acid sequence

comprising residues 23 to 216 of Figure 22 (SEQ ID NO:59). Native PRO533 polypeptides with or without the native signal sequence (amino acids 1 to 22 in Figure 22 (SEQ ID NO:59)), and with or without the initiating methionine are specifically included. Alternatively, the invention provides a PRO533 polypeptide encoded by the nucleic acid deposited under accession number ATCC 209480.

10. PRO245

Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is designated in the present application as "PRO245".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO245 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO245 polypeptide having amino acid residues 1 to 312 of Fig. 24 (SEQ ID NO:64), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO245 polypeptide. In particular, the invention provides isolated native sequence PRO245 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 312 of Figure 24 (SEQ ID NO:64).

11. PRO220, PRO221 and PRO227

Applicants have identified cDNA clones that each encode novel polypeptides, all having leucine rich repeats. These polypeptides are designated in the present application as PRO220, PRO221 and PRO227.

In one embodiment, the invention provides isolated nucleic acid molecules comprising DNA respectively encoding PRO220, PRO221 and PRO227, respectively. In one aspect, provided herein is an isolated nucleic acid comprises DNA encoding the PRO220 polypeptide having amino acid residues 1 through 708 of Figure 26 (SEQ ID NO:69), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions. Also provided herein is an isolated nucleic acid comprises DNA encoding the PRO221 polypeptide having amino acid residues 1 through 259 of Figure 28 (SEQ ID NO:71), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions. Moreover, also provided herein is an isolated nucleic acid comprises DNA encoding the PRO227 polypeptide having amino acid residues 1 through 620 of Figure 30 (SEQ ID NO:73), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO220, PRO221 and PRO227 polypeptides. In particular, provided herein is the isolated native sequence for the PRO220 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 708 of Figure 26 (SEQ ID NO:69). Additionally provided herein is the isolated native sequence for the PRO221 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 259 of Figure 28 (SEQ ID NO:71). Moreover, provided herein is the isolated native sequence for the PRO227 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 620 of Figure 30 (SEQ ID NO:73).

12. PRO258

Applicants have identified a cDNA clone that encodes a novel polypeptide having homology to CRTAM and poliovirus receptor precursors, wherein the polypeptide is designated in the present application as "PRO258".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO258 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO258 polypeptide having amino acid residues 1 to 398 of Figure 32 (SEQ ID NO:84), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO258 polypeptide. In particular, the invention provides isolated native sequence PRO258 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 398 of Figure 32 (SEQ ID NO:84). An additional embodiment of the present invention is directed to an isolated extracellular domain of a PRO258 polypeptide.

13. PRO266

Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is designated in the present application as "PRO266".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO266 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO266 polypeptide having amino acid residues 1 to 696 of Figure 34 (SEQ ID NO:91), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO266 polypeptide. In particular, the invention provides isolated native sequence PRO266 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 696 of Figure 34 (SEQ ID NO:91).

14. PRO269

Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is designated in the present application as PRO269.

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO269 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO269 polypeptide having amino acid residues 1 to 490 of Fig. 36 (SEQ ID NO:96), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO269 polypeptide. In particular, the invention provides isolated native sequence PRO269 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 490 of Figure 36 (SEQ ID NO:96). An additional embodiment of the present invention is directed to an isolated extracellular domain of a PRO269 polypeptide.

15. PRO287

Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is designated in the present application as "PRO287".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO287 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO287 polypeptide having amino acid residues 1 to 415 of Fig. 38 (SEQ ID NO:104), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO287 polypeptide. In particular, the invention provides isolated native sequence PRO287 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 415 of Figure 38 (SEQ ID NO:104).

16. PRO214

Applicants have identified a cDNA clone that encodes a novel polypeptide, designated in the present application as "PRO214".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO214 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO214 polypeptide of Fig. 40 (SEQ ID NO:109), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions. In another aspect, the invention provides a nucleic acid comprising the coding sequence of Fig. 39 (SEQ ID NO:108) or its complement. In another aspect, the invention provides a nucleic acid of the full length protein of clone DNA32286-1191, deposited with ATCC under accession number ATCC 209385.

In yet another embodiment, the invention provides isolated PRO214 polypeptide. In particular, the invention provides isolated native sequence PRO214 polypeptide, which in one embodiment, includes an amino acid sequence comprising the residues of Figure 40 (SEQ ID NO:109). Alternatively, the invention provides a polypeptide encoded by the nucleic acid deposited under accession number ATCC 209385.

17. PRO317

Applicants have identified a cDNA clone that encodes a novel polypeptide, designated in the present application as "PRO317".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding PRO317 polypeptide. In one aspect, the isolated nucleic acid comprises DNA (SEQ ID NO:113) encoding PRO317 polypeptide having amino acid residues 1 to 366 of Fig. 42, or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO317 polypeptide. In particular, the invention provides isolated native-sequence PRO317 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 366 of Figure 42 (SEQ ID NO:114).

In yet another embodiment, the invention supplies a method of detecting the presence of PRO317 in a sample, the method comprising:

a) contacting a detectable anti-PRO317 antibody with a sample suspected of containing PRO317; and

b) detecting binding of the antibody to the sample; wherein the sample is selected from the group consisting of a body fluid, a tissue sample, a cell extract, and a cell culture medium.

In a still further embodiment a method is provided for determining the presence of PRO317 mRNA in a sample, the method comprising:

a) contacting a sample suspected of containing PRO317 mRNA with a detectable nucleic acid probe that hybridizes under moderate to stringent conditions to PRO317 mRNA; and

b) detecting hybridization of the probe to the sample.

Preferably, in this method the sample is a tissue sample and the detecting step is by *in situ* hybridization, or the sample is a cell extract and detection is by Northern analysis.

Further, the invention provides a method for treating a PRO317-associated disorder comprising administering to a mammal an effective amount of the PRO317 polypeptide or a composition thereof containing a carrier, or with an effective amount of a PRO317 agonist or PRO317 antagonist, such as an antibody which binds specifically to PRO317.

18. **PRO301**

Applicants have identified a cDNA clone (DNA40628-1216) that encodes a novel polypeptide, designated in the present application as "PRO301".

In one embodiment, the invention provides an isolated nucleic acid molecule having at least about 80% sequence identity to (a) a DNA molecule encoding a PRO301 polypeptide comprising the sequence of amino acids 28 to 258 of Fig. 44 (SEQ ID NO:119), or (b) the complement of the DNA molecule of (a). The sequence identity preferably is about 85%, more preferably about 90%, most preferably about 95%. In one aspect, the isolated nucleic acid has at least about 80%, preferably at least about 85%, more preferably at least about 90%, and most preferably at least about 95% sequence identity with a polypeptide having amino acid residues 28 to 258 of Fig. 44 (SEQ ID NO:119). Preferably, the highest degree of sequence identity occurs within the extracellular domains (amino acids 28 to 258 of Fig. 44, SEQ ID NO:119). In a further embodiment, the isolated nucleic acid molecule comprises DNA encoding a PRO301 polypeptide having amino acid residues 28 to 299 of Fig. 44 (SEQ ID NO:119), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions. In another aspect, the invention provides a nucleic acid of the full length protein of clone DNA40628-1216, deposited with the ATCC under accession number ATCC 209432, alternatively the coding sequence of clone DNA40628-1216, deposited under accession number ATCC 209432.

In yet another embodiment, the invention provides isolated PRO301 polypeptide. In particular, the invention provides isolated native sequence PRO301 polypeptide, which in one embodiment, includes an amino acid sequence comprising the extracellular domain residues 28 to 258 of Figure 44 (SEQ ID NO:119). Native PRO301 polypeptides with or without the native signal sequence (amino acids 1 to 27 in Figure 44 (SEQ ID NO:119), and with or without the initiating methionine are specifically included. Additionally, the sequences of the invention may also comprise the transmembrane domain (residues 236 to about 258 in Figure 44; SEQ ID NO:119) and/or the intracellular domain (about residue 259 to 299 in Figure 44; SEQ ID NO:119). Alternatively, the invention provides a PRO301

polypeptide encoded by the nucleic acid deposited under accession number ATCC 209432.

19. PRO224

Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is designated in the present application as "PRO224".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO224 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO224 polypeptide having amino acid residues 1 to 282 of Figure 46 (SEQ ID NO:127), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO224 polypeptide. In particular, the invention provides isolated native sequence PRO224 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 282 of Figure 46 (SEQ ID NO:127).

20. PRO222

Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is designated in the present application as "PRO222".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO222 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO222 polypeptide having amino acid residues 1 to 490 of Fig. 48 (SEQ ID NO:132), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO222 polypeptide. In particular, the invention provides isolated native sequence PRO222 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 490 of Figure 48 (SEQ ID NO:132).

21. PRO234

Applicants have identified a cDNA clone that encodes a novel lectin polypeptide molecule, designated in the present application as "PRO234".

In one embodiment, the invention provides an isolated nucleic acid encoding a novel lectin comprising DNA encoding a PRO234 polypeptide. In one aspect, the isolated nucleic acid comprises the DNA encoding PRO234 polypeptides having amino acid residues 1 to 382 of Fig. 50 (SEQ ID NO:137), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions. In another aspect, the invention provides an isolated nucleic acid molecule comprising the nucleotide sequence of Fig. 49 (SEQ ID NO:136).

In another embodiment, the invention provides isolated novel PRO234 polypeptides. In particular, the invention provides isolated native sequence PRO234 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 382 of Figure 50 (SEQ ID NO:137).

In yet another embodiment, the invention provides oligonucleotide probes useful for isolating genomic and cDNA nucleotide sequences.

22. PRO231

Applicants have identified a cDNA clone that encodes a novel polypeptide having homology to a putative acid phosphatase, wherein the polypeptide is designated in the present application as "PRO231".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO231 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO231 polypeptide having amino acid residues 1 to 428 of Fig. 52 (SEQ ID NO:142), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO231 polypeptide. In particular, the invention provides isolated native sequence PRO231 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 428 of Figure 52 (SEQ ID NO:142).

23. PRO229

Applicants have identified a cDNA clone that encodes a novel polypeptide having homology to scavenger receptors wherein the polypeptide is designated in the present application as "PRO229".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO229 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO229 polypeptide having amino acid residues 1 to 347 of Figure 54 (SEQ ID NO:148), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO229 polypeptide. In particular, the invention provides isolated native sequence PRO229 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 347 of Figure 54 (SEQ ID NO:148).

24. PRO238

Applicants have identified a cDNA clone that encodes a novel polypeptide having homology to reductase, wherein the polypeptide is designated in the present application as "PRO238".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO238 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO238 polypeptide having amino acid residues 1 to 310 of Figure 56 (SEQ ID NO:153), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO238 polypeptide. In particular, the invention provides isolated native sequence PRO238 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 310 of Figure 56 (SEQ ID NO:153).

25. **PRO233**

Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is designated in the present application as "PRO233".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO233 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO233 polypeptide having amino acid residues 1 to 300 of Figure 58 (SEQ ID NO:159), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO233 polypeptide. In particular, the invention provides isolated native sequence PRO233 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 300 of Figure 58 (SEQ ID NO:159).

26. **PRO223**

Applicants have identified a cDNA clone that encodes a novel polypeptide having homology to serine carboxypeptidase polypeptides, wherein the polypeptide is designated in the present application as "PRO223".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO223 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO223 polypeptide having amino acid residues 1 to 476 of Figure 60 (SEQ ID NO:164), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO223 polypeptide. In particular, the invention provides isolated native sequence PRO223 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 476 of Figure 60 (SEQ ID NO:164).

27. **PRO235**

Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is designated in the present application as "PRO235".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO235 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO235 polypeptide having amino acid residues 1 to 552 of Figure 62 (SEQ ID NO:170), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO235 polypeptide. In particular, the invention provides isolated native sequence PRO235 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 552 of Figure 62 (SEQ ID NO:170).

28. PRO236 and PRO262

Applicants have identified cDNA clones that encode novel polypeptides having homology to β-galactosidase, wherein those polypeptides are designated in the present application as "PRO236" and "PRO262".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO236 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO236 polypeptide having amino acid residues 1 to 636 of Figure 64 (SEQ ID NO:175), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO262 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO262 polypeptide having amino acid residues 1 to 654 of Figure 66 (SEQ ID NO:177), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO236 polypeptide. In particular, the invention provides isolated native sequence PRO236 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 636 of Figure 64 (SEQ ID NO:175).

In another embodiment, the invention provides isolated PRO262 polypeptide. In particular, the invention provides isolated native sequence PRO262 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 654 of Figure 66 (SEQ ID NO:177).

29. PRO239

Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is designated in the present application as "PRO239".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO239 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO239 polypeptide having amino acid residues 1 to 501 of Figure 68 (SEQ ID NO:185), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO239 polypeptide. In particular, the invention provides isolated native sequence PRO239 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 501 of Figure 68 (SEQ ID NO:185).

30. PRO257

Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is designated in the present application as "PRO257".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO257 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO257 polypeptide having amino acid residues 1 to 607 of Figure 70 (SEQ ID NO:190), or is complementary to such encoding nucleic

acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO257 polypeptide. In particular, the invention provides isolated native sequence PRO257 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 607 of Figure 70 (SEQ ID NO:190). An additional embodiment of the present invention is directed to an isolated extracellular domain of a PRO257 polypeptide.

31. **PRO260**

Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is designated in the present application as "PRO260".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO260 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO260 polypeptide having amino acid residues 1 to 467 of Figure 72 (SEQ ID NO:195), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO260 polypeptide. In particular, the invention provides isolated native sequence PRO260 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 467 of Figure 72 (SEQ ID NO:195).

32. **PRO263**

Applicants have identified a cDNA clone that encodes a novel polypeptide having homology to CD44 antigen, wherein the polypeptide is designated in the present application as "PRO263".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO263 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO263 polypeptide having amino acid residues 1 to 322 of Figure 74 (SEQ ID NO:201), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO263 polypeptide. In particular, the invention provides isolated native sequence PRO263 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 322 of Figure 74 (SEQ ID NO:201). An additional embodiment of the present invention is directed to an isolated extracellular domain of a PRO263 polypeptide.

33. **PRO270**

Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is designated in the present application as "PRO270".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO270 polypeptide. In one aspect, the isolated nucleic acid comprises DNA whivch includes the sequence encoding the PRO270 polypeptide having amino acid residues 1 to 296 of Fig. 76 (SEQ ID NO:207), or is complementary to

such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO270 polypeptide. In particular, the invention provides isolated native sequence PRO270 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 296 of Figure 76 (SEQ ID NO:207).

34. PRO271

Applicants have identified a cDNA clone that encodes a novel polypeptide having homology to the proteoglycan link protein, wherein the polypeptide is designated in the present application as "PRO271".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO271 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO271 polypeptide having amino acid residues 1 to 360 of Figure 78 (SEQ ID NO:213), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO271 polypeptide. In particular, the invention provides isolated native sequence PRO271 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 360 of Figure 78 (SEQ ID NO:213).

35. PRO272

Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is designated in the present application as "PRO272".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO272 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO272 polypeptide having amino acid residues 1 to 328 of Figure 80 (SEQ ID NO:221), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO272 polypeptide. In particular, the invention provides isolated native sequence PRO272 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 328 of Figure 80 (SEQ ID NO:211).

36. PRO294

Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is designated in the present application as "PRO294".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO294 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO294 polypeptide having amino acid residues 1 to 550 of Figure 82 (SEQ ID NO:227), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO294 polypeptide. In particular, the invention provides isolated native sequence PRO294 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 550 of Figure 82 (SEQ ID NO:227).

37. **PRO295**

Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is designated in the present application as "PRO295".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO295 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO295 polypeptide having amino acid residues 1 to 350 of Figure 84 (SEQ ID NO:236), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO295 polypeptide. In particular, the invention provides isolated native sequence PRO295 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 350 of Figure 84 (SEQ ID NO:236).

38. **PRO293**

Applicants have identified a cDNA clone that encodes a novel human neuronal leucine rich repeat polypeptide, wherein the polypeptide is designated in the present application as "PRO293".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO293 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO293 polypeptide having amino acid residues 1 to 713 of Figure 86 (SEQ ID NO:245), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO293 polypeptide. In particular, the invention provides isolated native sequence PRO293 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 713 of Figure 86 (SEQ ID NO:245). An additional embodiment of the present invention is directed to an isolated extracellular domain of a PRO293 polypeptide.

39. **PRO247**

Applicants have identified a cDNA clone that encodes a novel polypeptide having leucine rich repeats wherein the polypeptide is designated in the present application as "PRO247".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO247 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO247 polypeptide having amino acid residues 1 to 546 of Figure 88 (SEQ ID NO:250), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO247 polypeptide. In particular, the invention provides isolated native sequence PRO247 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 546 of Figure 88 (SEQ ID NO:250). An additional embodiment of the present invention is directed to an isolated extracellular domain of a PRO247 polypeptide.

40. **PRO302, PRO303, PRO304, PRO307 and PRO343**

Applicants have identified cDNA clones that encode novel polypeptides having homology to various proteases, wherein those polypeptide are designated in the present application as "PRO302", "PRO303", "PRO304", "PRO307" and "PRO343" polypeptides.

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO302 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO302 polypeptide having amino acid residues 1 to 452 of Figure 90 (SEQ ID NO:255), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO303 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO303 polypeptide having amino acid residues 1 to 314 of Figure 92 (SEQ ID NO:257), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In yet another embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO304 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO304 polypeptide having amino acid residues 1 to 556 of Figure 94 (SEQ ID NO:259), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO307 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO307 polypeptide having amino acid residues 1 to 383 of Figure 96 (SEQ ID NO:261), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO343 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO343 polypeptide having amino acid residues 1 to 317 of Figure 98 (SEQ ID NO:263), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO302 polypeptide. In particular, the invention provides isolated native sequence PRO302 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 452 of Figure 90 (SEQ ID NO:255).

In another embodiment, the invention provides isolated PRO303 polypeptide. In particular, the invention provides isolated native sequence PRO303 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 314 of Figure 92 (SEQ ID NO:257).

In another embodiment, the invention provides isolated PRO304 polypeptide. In particular, the invention provides isolated native sequence PRO304 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 556 of Figure 94 (SEQ ID NO:259).

In another embodiment, the invention provides isolated PRO307 polypeptide. In particular, the invention provides isolated native sequence PRO307 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 383 of Figure 96 (SEQ ID NO:261).

In another embodiment, the invention provides isolated PRO343 polypeptide. In particular, the invention provides isolated native sequence PRO343 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 317 of Figure 98 (SEQ ID NO:263).

41. PRO328

Applicants have identified a cDNA clone that encodes a novel polypeptide, wherein the polypeptide is designated in the present application as "PRO328".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO328 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO328 polypeptide having amino acid residues 1 to 463 of Figure 100 (SEQ ID NO:285), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO328 polypeptide. In particular, the invention provides isolated native sequence PRO328 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 463 of Figure 100 (SEQ ID NO:285). An additional embodiment of the present invention is directed to an isolated extracellular domain of a PRO306 polypeptide.

42. PRO335, PRO331 and PRO326

Applicants have identified three cDNA clones that respectively encode three novel polypeptides, each having leucine rich repeats and homology to LIG-1 and ALS. These polypeptides are designated in the present application as PRO335, PRO331 and PRO326, respectively.

In one embodiment, the invention provides three isolated nucleic acid molecules comprising DNA respectively encoding PRO335, PRO331 and PRO326, respectively. In one aspect, herein is provided an isolated nucleic acid comprising DNA encoding the PRO335 polypeptide having amino acid residues 1 through 1059 of Figure 102 (SEQ ID NO:290), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions. Also provided herein is an isolated nucleic acid comprises DNA encoding the PRO331 polypeptide having amino acid residues 1 through 640 of Figure 104 (SEQ ID NO:292), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions. Additionally provided herein is an

isolated nucleic acid comprises DNA encoding the PRO326 polypeptide having amino acid residues 1 through 1119 of Figure 106 (SEQ ID NO:294), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO335, PRO331 and PRO326 polypeptides or extracellular domains thereof. In particular, the invention provides isolated native sequence for the PRO335 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 through 1059 of Figure 102 (SEQ ID NO:290). Also provided herein is the isolated native sequence for the PRO331 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 through 640 of Figure 104 (SEQ ID NO:292). Also provided herein is the isolated native sequence for the PRO326 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 through 1119 of Figure 106 (SEQ ID NO:294).

43. PRO332

Applicants have identified a cDNA clone (DNA40982-1235) that encodes a novel polypeptide, designated in the present application as "PRO332."

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA having at least about 80% sequence identity to (a) a DNA molecule encoding a PRO358 polypeptide comprising the sequence of amino acids 49 to 642 of Fig. 108 (SEQ ID NO:310), or (b) the complement of the DNA molecule of (a). The sequence identity preferably is about 85%, more preferably about 90%, most preferably about 95%. In one aspect, the isolated nucleic acid has at least about 80%, preferably at least about 85%, more preferably at least about 90%, and most preferably at least about 95% sequence identity with a polypeptide having amino acid residues 1 to 642 of Fig. 108 (SEQ ID NO:310). Preferably, the highest degree of sequence identity occurs within the leucine-rich repeat domains (amino acids 116 to 624 of Fig. 108, SEQ ID NO:310). In a further embodiment, the isolated nucleic acid molecule comprises DNA encoding a PRO332 polypeptide having amino acid residues 49 to 642 of Fig. 108 (SEQ ID NO:310), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO332 polypeptides. In particular, the invention provides isolated native sequence PRO332 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 49 to 624 of Figure 108 (SEQ ID NO:310). Native PRO332 polypeptides with or without the native signal sequence (amino acids 1 to 48 in Figure 108, SEQ ID NO:310), and with or without the initiating methionine are specifically included.

44. PRO334

Applicants have identified a cDNA clone that encodes a novel polypeptide having homology to fibulin and fibrillin, wherein the polypeptide is designated in the present application as "PRO334".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO334 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO334 polypeptide having amino acid residues 1 to 509 of Figure 110 (SEQ ID NO:315), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency

conditions.

In another embodiment, the invention provides isolated PRO334 polypeptide. In particular, the invention provides isolated native sequence PRO334 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 509 of Figure 110 (SEQ ID NO:315).

45. **PRO346**

Applicants have identified a cDNA clone (DNA44167-1243) that encodes a novel polypeptide, designated in the present application as "PRO346."

In one embodiment, the invention provides an isolated nucleic acid molecule having at least about 80% sequence identity to (a) a DNA molecule encoding a PRO346 polypeptide comprising the sequence of amino acids 19 to 339 of Fig. 112 (SEQ ID NO: 320), or (b) the complement of the DNA molecule of (a). The sequence identity preferably is about 85%, more preferably about 90%, most preferably about 95%. In one aspect, the isolated nucleic acid has at least about 80%, preferably at least about 85%, more preferably at least about 90%, and most preferably at least about 95% sequence identity with a polypeptide having amino acid residues 19 to 339 of Fig. 112 (SEQ ID NO:320). Preferably, the highest degree of sequence identity occurs within the extracellular domains (amino acids 19 to 339 of Fig. 112, SEQ ID NO:320). In alternative embodiments, the polypeptide by which the homology is measured comprises the residues 1-339, 19-360 or 19-450 of Fig. 112, SEQ ID NO:320). In a further embodiment, the isolated nucleic acid molecule comprises DNA encoding a PRO346 polypeptide having amino acid residues 19 to 339 of Fig. 112 (SEQ ID NO:320), alternatively residues 1-339, 19-360 or 19-450 of Fig. 112 (SEQ ID NO:320) or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions. In another aspect, the invention provides a nucleic acid of the full length protein of clone DNA44167-1243, deposited with the ATCC under accession number ATCC 209434, alternatively the coding sequence of clone DNA44167-1243, deposited under accession number ATCC 209434.

In yet another embodiment, the invention provides isolated PRO346 polypeptide. In particular, the invention provides isolated native sequence PRO346 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 19 to 339 of Figure 112 (SEQ ID NO:320). Native PRO346 polypeptides with or without the native signal sequence (residues 1 to 18 in Figure 112 (SEQ ID NO:320), with or without the initiating methionine, with or without the transmembrane domain (residues 340 to 360) and with or without the intracellular domain (residues 361 to 450) are specifically included. Alternatively, the invention provides a PRO346 polypeptide encoded by the nucleic acid deposited under accession number ATCC 209434.

46. **PRO268**

Applicants have identified a cDNA clone that encodes a novel polypeptide having homology to protein disulfide isomerase, wherein the polypeptide is designated in the present application as "PRO268".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO268 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO268 polypeptide having amino acid residues 1 to 280 of Figure 114 (SEQ ID NO:325), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency

conditions.

In another embodiment, the invention provides isolated PRO268 polypeptide. In particular, the invention provides isolated native sequence PRO268 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 280 of Figure 114 (SEQ ID NO:325). An additional embodiment of the present invention is directed to an isolated extracellular domain of a PRO268 polypeptide.

47. PRO330

Applicants have identified a cDNA clone that encodes a novel polypeptide having homology to the alpha subunit of prolyl 4-hydroxylase, wherein the polypeptide is designated in the present application as "PRO330".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding a PRO330 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO330 polypeptide having amino acid residues 1 to 533 of Figure 116 (SEQ ID NO:332), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO330 polypeptide. In particular, the invention provides isolated native sequence PRO330 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 533 of Figure 116 (SEQ ID NO:332).

48. PRO339 and PRO310

Applicants have identified two cDNA clones wherein each clone encodes a novel polypeptide having homology to fringe, wherein the polypeptides are designated in the present application as "PRO339" and "PRO310".

In one embodiment, the invention provides isolated nucleic acid molecules comprising DNA encoding a PRO339 and/or a PRO310 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding the PRO339 polypeptide having amino acid residues 1 to 772 of Figure 118 (SEQ ID NO:339), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions. In another aspect, the isolated nucleic acid comprises DNA encoding the PRO310 polypeptide having amino acid residues 1 to 318 of Figure 120 (SEQ ID NO:341), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO339 as well as isolated PRO310 polypeptides. In particular, the invention provides isolated native sequence PRO339 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 772 of Figure 118 (SEQ ID NO:339). The invention further provides isolated native sequence PRO310 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 318 of Figure 120 (SEQ ID NO:341).

49. PRO244

Applicants have identified a cDNA clone that encodes a novel polypeptide, designated in the present application as "PRO244".

In one embodiment, the invention provides an isolated nucleic acid molecule comprising DNA encoding PRO244 polypeptide. In one aspect, the isolated nucleic acid comprises DNA encoding PRO244 polypeptide having amino acid residues 1 to 219 of Fig. 122 (SEQ ID NO:377), or is complementary to such encoding nucleic acid sequence, and remains stably bound to it under at least moderate, and optionally, under high stringency conditions.

In another embodiment, the invention provides isolated PRO244 polypeptide. In particular, the invention provides isolated native sequence PRO244 polypeptide, which in one embodiment, includes an amino acid sequence comprising residues 1 to 219 of Figure 122 (SEQ ID NO:377).

50. Additional Embodiments

In other embodiments of the present invention, the invention provides vectors comprising DNA encoding any of the above or below described polypeptides. A host cell comprising any such vector is also provided. By way of example, the host cells may be CHO cells, *E. coli*, or yeast. A process for producing any of the above or below 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.

In other embodiments, the invention provides chimeric molecules comprising any of the above or below described polypeptides fused to a heterologous polypeptide or amino acid sequence. An example of such a chimeric molecule comprises any of the above or below described polypeptides fused to an epitope tag sequence or a Fc region of an immunoglobulin.

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.

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

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows a nucleotide sequence (SEQ ID NO:1) of a native sequence PRO211 cDNA, wherein SEQ ID NO:1 is a clone designated herein as "UNQ185" and/or "DNA32292-1131".

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 PRO217 cDNA, wherein SEQ ID NO:3 is a clone designated herein as "UNQ191" and/or "DNA33094-1131".

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:11) of a native sequence PRO230 cDNA, wherein SEQ ID NO:11 is a clone designated herein as "UNQ204" and/or "DNA33223-1136".

Figure 6 shows the amino acid sequence (SEQ ID NO:12) derived from the coding sequence of SEQ ID NO:11 shown in Figure 5.

Figure 7 shows a nucleotide sequence designated herein as DNA20088 (SEQ ID NO:13).

Figure 8 shows a nucleotide sequence (SEQ ID NO:17) of a native sequence PRO232 cDNA, wherein SEQ ID NO:17 is a clone designated herein as "UNQ206" and/or "DNA34435-1140".

Figure 9 shows the amino acid sequence (SEQ ID NO:18) derived from the coding sequence of SEQ ID NO:17 shown in Figure 8.

Figure 10 shows a nucleotide sequence (SEQ ID NO:22) of a native sequence PRO187 cDNA, wherein SEQ ID NO:22 is a clone designated herein as "UNQ161" and/or "DNA27864-1155".

Figure 11 shows the amino acid sequence (SEQ ID NO:23) derived from the coding sequence of SEQ ID NO:22 shown in Figure 10.

Figure 12 shows a nucleotide sequence (SEQ ID NO:27) of a native sequence PRO265 cDNA, wherein SEQ ID NO:27 is a clone designated herein as "UNQ232" and/or "DNA36350-1158".

Figure 13 shows the amino acid sequence (SEQ ID NO:28) derived from the coding sequence of SEQ ID NO:27 shown in Figure 12.

Figures 14A-B show a nucleotide sequence (SEQ ID NO:33) of a native sequence PRO219 cDNA, wherein SEQ ID NO:33 is a clone designated herein as "UNQ193" and/or "DNA32290-1164".

Figure 15 shows the amino acid sequence (SEQ ID NO:34) derived from the coding sequence of SEQ ID NO:33 shown in Figures 14A-B.

Figure 16 shows a nucleotide sequence (SEQ ID NO:38) of a native sequence PRO246 cDNA, wherein SEQ ID NO:38 is a clone designated herein as "UNQ220" and/or "DNA35639-1172".

Figure 17 shows the amino acid sequence (SEQ ID NO:39) derived from the coding sequence of SEQ ID NO:38 shown in Figure 16.

Figure 18 shows a nucleotide sequence (SEQ ID NO:48) of a native sequence PRO228 cDNA, wherein SEQ ID NO:48 is a clone designated herein as "UNQ202" and/or "DNA33092-1202".

Figure 19 shows the amino acid sequence (SEQ ID NO:49) derived from the coding sequence of SEQ ID NO:48 shown in Figure 18.

Figure 20 shows a nucleotide sequence designated herein as DNA21951 (SEQ ID NO:50).

Figure 21 shows a nucleotide sequence (SEQ ID NO:58) of a native sequence PRO533 cDNA, wherein SEQ ID NO:58 is a clone designated herein as "UNQ344" and/or "DNA49435-1219".

Figure 22 shows the amino acid sequence (SEQ ID NO:59) derived from the coding sequence of SEQ ID NO:58 shown in Figure 21.

Figure 23 shows a nucleotide sequence (SEQ ID NO:63) of a native sequence PRO245 cDNA, wherein SEQ ID NO:63 is a clone designated herein as "UNQ219" and/or "DNA35638-1141".

Figure 24 shows the amino acid sequence (SEQ ID NO:64) derived from the coding sequence of SEQ ID NO:63 shown in Figure 23.

Figure 25 shows a nucleotide sequence (SEQ ID NO:68) of a native sequence PRO220 cDNA, wherein SEQ ID NO:68 is a clone designated herein as "UNQ194" and/or "DNA32298-1132".

Figure 26 shows the amino acid sequence (SEQ ID NO:69) derived from the coding sequence of SEQ ID NO:68 shown in Figure 25.

Figure 27 shows a nucleotide sequence (SEQ ID NO:70) of a native sequence PRO221 cDNA. wherein SEQ ID NO:70 is a clone designated herein as "UNQ195" and/or "DNA33089-1132".

Figure 28 shows the amino acid sequence (SEQ ID NO:71) derived from the coding sequence of SEQ ID NO:70 shown in Figure 27.

Figure 29 shows a nucleotide sequence (SEQ ID NO:72) of a native sequence PRO227 cDNA, wherein SEQ ID NO:72 is a clone designated herein as "UNQ201" and/or "DNA33786-1132".

Figure 30 shows the amino acid sequence (SEQ ID NO:73) derived from the coding sequence of SEQ ID NO:72 shown in Figure 29.

Figure 31 shows a nucleotide sequence (SEQ ID NO:83) of a native sequence PRO258 cDNA, wherein SEQ ID NO:83 is a clone designated herein as "UNQ225" and/or "DNA35918-1174".

Figure 32 shows the amino acid sequence (SEQ ID NO:84) derived from the coding sequence of SEQ ID NO:83 shown in Figure 31.

Figure 33 shows a nucleotide sequence (SEQ ID NO:90) of a native sequence PRO266 cDNA, wherein SEQ ID NO:90 is a clone designated herein as "UNQ233" and/or "DNA37150-1178".

Figure 34 shows the amino acid sequence (SEQ ID NO:91) derived from the coding sequence of SEQ ID NO:90 shown in Figure 33.

Figure 35 shows a nucleotide sequence (SEQ ID NO:95) of a native sequence PRO269 cDNA, wherein SEQ ID NO:95 is a clone designated herein as "UNQ236" and/or "DNA38260-1180".

Figure 36 shows the amino acid sequence (SEQ ID NO:96) derived from the coding sequence of SEQ ID NO:95 shown in Figure 35.

Figure 37 shows a nucleotide sequence (SEQ ID NO:103) of a native sequence PRO287 cDNA, wherein SEQ ID NO:103 is a clone designated herein as "UNQ250" and/or "DNA39969-1185".

Figure 38 shows the amino acid sequence (SEQ ID NO:104) derived from the coding sequence of SEQ ID NO:103 shown in Figure 37.

Figure 39 shows a nucleotide sequence (SEQ ID NO:108) of a native sequence PRO214 cDNA, wherein SEQ ID NO:108 is a clone designated herein as "UNQ188" and/or "DNA32286-1191".

Figure 40 shows the amino acid sequence (SEQ ID NO:109) derived from the coding sequence of SEQ ID NO:108 shown in Figure 39.

Figure 41 shows a nucleotide sequence (SEQ ID NO:113) of a native sequence PRO317 cDNA, wherein SEQ ID NO:113 is a clone designated herein as "UNQ278" and/or "DNA33461-1199".

Figure 42 shows the amino acid sequence (SEQ ID NO:114) derived from the coding sequence of SEQ ID NO:113 shown in Figure 41.

Figure 43 shows a nucleotide sequence (SEQ ID NO:118) of a native sequence PRO301 cDNA, wherein SEQ ID NO:118 is a clone designated herein as "UNQ264" and/or "DNA40628-1216".

Figure 44 shows the amino acid sequence (SEQ ID NO:119) derived from the coding sequence of SEQ ID NO:118 shown in Figure 43.

Figure 45 shows a nucleotide sequence (SEQ ID NO:126) of a native sequence PRO224 cDNA, wherein SEQ ID NO:126 is a clone designated herein as "UNQ198" and/or "DNA33221-1133".

Figure 46 shows the amino acid sequence (SEQ ID NO:127) derived from the coding sequence of SEQ ID NO:126 shown in Figure 45.

Figure 47 shows a nucleotide sequence (SEQ ID NO:131) of a native sequence PRO222 cDNA, wherein SEQ ID NO:131 is a clone designated herein as "UNQ196" and/or "DNA33107-1135".

Figure 48 shows the amino acid sequence (SEQ ID NO:132) derived from the coding sequence of SEQ ID NO:131 shown in Figure 47.

Figure 49 shows a nucleotide sequence (SEQ ID NO:136) of a native sequence PRO234 cDNA, wherein SEQ ID NO:136 is a clone designated herein as "UNQ208" and/or "DNA35557-1137".

Figure 50 shows the amino acid sequence (SEQ ID NO:137) derived from the coding sequence of SEQ ID NO:136 shown in Figure 49.

Figure 51 shows a nucleotide sequence (SEQ ID NO:141) of a native sequence PRO231 cDNA, wherein SEQ ID NO:141 is a clone designated herein as "UNQ205" and/or "DNA34434-1139".

Figure 52 shows the amino acid sequence (SEQ ID NO:142) derived from the coding sequence of SEQ ID NO:141 shown in Figure 51.

Figure 53 shows a nucleotide sequence (SEQ ID NO:147) of a native sequence PRO229 cDNA, wherein SEQ ID NO:147 is a clone designated herein as "UNQ203" and/or "DNA33100-1159".

Figure 54 shows the amino acid sequence (SEQ ID NO:148) derived from the coding sequence of SEQ ID NO:147 shown in Figure 53.

Figure 55 shows a nucleotide sequence (SEQ ID NO:152) of a native sequence PRO238 cDNA, wherein SEQ ID NO:152 is a clone designated herein as "UNQ212" and/or "DNA35600-1162".

Figure 56 shows the amino acid sequence (SEQ ID NO:153) derived from the coding sequence of SEQ ID NO:152 shown in Figure 55.

Figure 57 shows a nucleotide sequence (SEQ ID NO:158) of a native sequence PRO233 cDNA, wherein SEQ ID NO:158 is a clone designated herein as "UNQ207" and/or "DNA34436-1238".

Figure 58 shows the amino acid sequence (SEQ ID NO:159) derived from the coding sequence of SEQ ID NO:158 shown in Figure 57.

Figure 59 shows a nucleotide sequence (SEQ ID NO:163) of a native sequence PRO223 cDNA, wherein SEQ ID NO:163 is a clone designated herein as "UNQ197" and/or "DNA33206-1165".

Figure 60 shows the amino acid sequence (SEQ ID NO:164) derived from the coding sequence of SEQ ID NO:163 shown in Figure 59.

Figure 61 shows a nucleotide sequence (SEQ ID NO:169) of a native sequence PRO235 cDNA, wherein SEQ ID NO:169 is a clone designated herein as "UNQ209" and/or "DNA35558-1167".

Figure 62 shows the amino acid sequence (SEQ ID NO:170) derived from the coding sequence of SEQ ID NO:169 shown in Figure 61.

Figure 63 shows a nucleotide sequence (SEQ ID NO:174) of a native sequence PRO236 cDNA, wherein SEQ ID NO:174 is a clone designated herein as "UNQ210" and/or "DNA35599-1168".

Figure 64 shows the amino acid sequence (SEQ ID NO:175) derived from the coding sequence of SEQ ID NO:174 shown in Figure 63.

Figure 65 shows a nucleotide sequence (SEQ ID NO:176) of a native sequence PRO262 cDNA, wherein SEQ ID NO:176 is a clone designated herein as "UNQ229" and/or "DNA36992-1168".

Figure 66 shows the amino acid sequence (SEQ ID NO:177) derived from the coding sequence of SEQ ID NO:176 shown in Figure 65.

Figure 67 shows a nucleotide sequence (SEQ ID NO:184) of a native sequence PRO239 cDNA, wherein SEQ ID NO:184 is a clone designated herein as "UNQ213" and/or "DNA34407-1169".

Figure 68 shows the amino acid sequence (SEQ ID NO:185) derived from the coding sequence of SEQ ID NO:184 shown in Figure 67.

Figure 69 shows a nucleotide sequence (SEQ ID NO:189) of a native sequence PRO257 cDNA, wherein SEQ ID NO:189 is a clone designated herein as "UNQ224" and/or "DNA35841-1173".

Figure 70 shows the amino acid sequence (SEQ ID NO:190) derived from the coding sequence of SEQ ID NO:189 shown in Figure 69.

Figure 71 shows a nucleotide sequence (SEQ ID NO:194) of a native sequence PRO260 cDNA, wherein SEQ ID NO:194 is a clone designated herein as "UNQ227" and/or "DNA33470-1175".

Figure 72 shows the amino acid sequence (SEQ ID NO:195) derived from the coding sequence of SEQ ID NO:194 shown in Figure 71.

Figure 73 shows a nucleotide sequence (SEQ ID NO:200) of a native sequence PRO263 cDNA, wherein SEQ ID NO:200 is a clone designated herein as "UNQ230" and/or "DNA34431-1177".

Figure 74 shows the amino acid sequence (SEQ ID NO:201) derived from the coding sequence of SEQ ID NO:200 shown in Figure 73.

Figure 75 shows a nucleotide sequence (SEQ ID NO:206) of a native sequence PRO270 cDNA, wherein SEQ ID NO:206 is a clone designated herein as "UNQ237" and/or "DNA39510-1181".

Figure 76 shows the amino acid sequence (SEQ ID NO:207) derived from the coding sequence of SEQ ID NO:206 shown in Figure 75.

Figure 77 shows a nucleotide sequence (SEQ ID NO:212) of a native sequence PRO271 cDNA, wherein SEQ ID NO:212 is a clone designated herein as "UNQ238" and/or "DNA39423-1182".

Figure 78 shows the amino acid sequence (SEQ ID NO:213) derived from the coding sequence of SEQ ID NO:212 shown in Figure 77.

Figure 79 shows a nucleotide sequence (SEQ ID NO:220) of a native sequence PRO272 cDNA, wherein SEQ ID NO:220 is a clone designated herein as "UNQ239" and/or "DNA40620-1183".

Figure 80 shows the amino acid sequence (SEQ ID NO:221) derived from the coding sequence of SEQ ID NO:220 shown in Figure 79.

Figure 81 shows a nucleotide sequence (SEQ ID NO:226) of a native sequence PRO294 cDNA, wherein SEQ ID NO:226 is a clone designated herein as "UNQ257" and/or "DNA40604-1187".

Figure 82 shows the amino acid sequence (SEQ ID NO:227) derived from the coding sequence of SEQ ID NO:226 shown in Figure 81.

Figure 83 shows a nucleotide sequence (SEQ ID NO:235) of a native sequence PRO295 cDNA, wherein SEQ ID NO:235 is a clone designated herein as "UNQ258" and/or "DNA38268-1188".

Figure 84 shows the amino acid sequence (SEQ ID NO:236) derived from the coding sequence of SEQ ID NO:235 shown in Figure 83.

Figures 85A-B show a nucleotide sequence (SEQ ID NO:244) of a native sequence PRO293 cDNA, wherein SEQ ID NO:244 is a clone designated herein as "UNQ256" and/or "DNA37151-1193".

Figure 86 shows the amino acid sequence (SEQ ID NO:245) derived from the coding sequence of SEQ ID NO:244 shown in Figures 85A-B.

Figures 89A-B show a nucleotide sequence (SEQ ID NO:249) of a native sequence PRO247 cDNA, wherein SEQ ID NO:249 is a clone designated herein as "UNQ221" and/or "DNA35673-1201".

Figure 88 shows the amino acid sequence (SEQ ID NO:250) derived from the coding sequence of SEQ ID NO:249 shown in Figure 87.

Figure 89 shows a nucleotide sequence (SEQ ID NO:254) of a native sequence PRO302 cDNA, wherein SEQ ID NO:254 is a clone designated herein as "UNQ265" and/or "DNA40370-1217".

Figure 90 shows the amino acid sequence (SEQ ID NO:255) derived from the coding sequence of SEQ ID NO:254 shown in Figure 89.

Figure 91 shows a nucleotide sequence (SEQ ID NO:256) of a native sequence PRO303 cDNA, wherein SEQ ID NO:256 is a clone designated herein as "UNQ266" and/or "DNA42551-1217".

Figure 92 shows the amino acid sequence (SEQ ID NO:257) derived from the coding sequence of SEQ ID NO:256 shown in Figure 91.

Figure 93 shows a nucleotide sequence (SEQ ID NO:258) of a native sequence PRO304 cDNA, wherein SEQ ID NO:258 is a clone designated herein as "UNQ267" and/or "DNA39520-1217".

Figure 94 shows the amino acid sequence (SEQ ID NO:259) derived from the coding sequence of SEQ ID NO:258 shown in Figure 93.

Figure 95 shows a nucleotide sequence (SEQ ID NO:260) of a native sequence PRO307 cDNA, wherein SEQ ID NO:260 is a clone designated herein as "UNQ270" and/or "DNA41225-1217".

Figure 96 shows the amino acid sequence (SEQ ID NO:261) derived from the coding sequence of SEQ ID NO:260 shown in Figure 95.

Figure 97 shows a nucleotide sequence (SEQ ID NO:262) of a native sequence PRO343 cDNA, wherein SEQ ID NO:262 is a clone designated herein as "UNQ302" and/or "DNA43318-1217".

Figure 98 shows the amino acid sequence (SEQ ID NO:263) derived from the coding sequence of SEQ ID NO:262 shown in Figure 97.

Figure 99 shows a nucleotide sequence (SEQ ID NO:284) of a native sequence PRO328 cDNA, wherein SEQ ID NO:284 is a clone designated herein as "UNQ289" and/or "DNA40587-1231".

Figure 100 shows the amino acid sequence (SEQ ID NO:285) derived from the coding sequence of SEQ ID NO:284 shown in Figure 99.

Figures 101A-B show a nucleotide sequence (SEQ ID NO:289) of a native sequence PRO335 cDNA, wherein SEQ ID NO:289 is a clone designated herein as "UNQ287" and/or "DNA41388-1234".

Figure 102 shows the amino acid sequence (SEQ ID NO:290) derived from the coding sequence of SEQ ID NO:289 shown in Figures 103A-B.

Figure 103 shows a nucleotide sequence (SEQ ID NO:291) of a native sequence PRO331 cDNA, wherein SEQ ID NO:291 is a clone designated herein as "UNQ292" and/or "DNA40981-1234".

Figure 104 shows the amino acid sequence (SEQ ID NO:292) derived from the coding sequence of SEQ ID NO:291 shown in Figure 103.

Figures 105A-B show a nucleotide sequence (SEQ ID NO:293) of a native sequence PRO326 cDNA, wherein SEQ ID NO:293 is a clone designated herein as "UNQ287" and/or "DNA37140-1234".

Figure 106 shows the amino acid sequence (SEQ ID NO:294) derived from the coding sequence of SEQ ID NO:293 shown in Figures 105A-B.

Figures 107A-B show a nucleotide sequence (SEQ ID NO:309) of a native sequence PRO332 cDNA, wherein SEQ ID NO:309 is a clone designated herein as "UNQ293" or "DNA40982-1235".

Figure 108 shows the amino acid sequence (SEQ ID NO:310) derived from the coding sequence of SEQ ID NO:309 shown in Figure 107.

Figure 109 shows a nucleotide sequence (SEQ ID NO:314) of a native sequence PRO334 cDNA, wherein SEQ ID NO:314 is a clone designated herein as "UNQ295" or "DNA41379-1236".

Figure 110 shows the amino acid sequence (SEQ ID NO:315) derived from the coding sequence of SEQ ID NO:314 shown in Figure 109.

Figure 111 shows a nucleotide sequence (SEQ ID NO:319) of a native sequence PRO346 cDNA, wherein SEQ ID NO:319 is a clone designated herein as "UNQ305" or "DNA44167-1243".

Figure 112 shows the amino acid sequence (SEQ ID NO:320) derived from the coding sequence of SEQ ID NO:319 shown in Figure 111.

Figure 113 shows a nucleotide sequence (SEQ ID NO:324) of a native sequence PRO268 cDNA, wherein SEQ ID NO:324 is a clone designated herein as "UNQ235" or "DNA39427-1179".

Figure 114 shows the amino acid sequence (SEQ ID NO:325) derived from the coding sequence of SEQ ID NO:324 shown in Figure 113.

Figure 115 shows a nucleotide sequence (SEQ ID NO:331) of a native sequence PRO330 cDNA, wherein SEQ ID NO:331 is a clone designated herein as "UNQ290" or "DNA40603-1232".

Figure 116 shows the amino acid sequence (SEQ ID NO:332) derived from the coding sequence of SEQ ID NO:331 shown in Figure 115.

Figure 117 shows a nucleotide sequence (SEQ ID NO:338) of a native sequence PRO339 cDNA, wherein SEQ ID NO:338 is a clone designated herein as "UNQ229" or "DNA43466-1225".

Figure 118 shows the amino acid sequence (SEQ ID NO:339) derived from the coding sequence of SEQ ID NO:338 shown in Figure 117.

Figure 119 shows a nucleotide sequence (SEQ ID NO:340) of a native sequence PRO310 cDNA, wherein SEQ ID NO:340 is a clone designated herein as "UNQ273" or "DNA43046-1225".

Figure 120 shows the amino acid sequence (SEQ ID NO:341) derived from the coding sequence of SEQ ID NO:340 shown in Figure 119.

Figure 121 shows a nucleotide sequence (SEQ ID NO:376) of a native sequence PRO244 cDNA, wherein SEQ ID NO:376 is a clone designated herein as "UNQ218" or "DNA35668-1171".

Figure 122 shows the amino acid sequence (SEQ ID NO:377) derived from the coding sequence of SEQ ID NO:376 shown in Figure 121.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

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" 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.

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 PRO211 is a mature or full-length native sequence PRO211 polypeptide comprising amino acids 1 to 353 of Figure 2 (SEQ ID NO:2), the native sequence PRO217 is a mature or full-length native sequence PRO217 polypeptide comprising amino acids 1 to 379 of Figure 4 (SEQ ID NO:4), the native sequence PRO230 is a mature or full-length native sequence PRO230 polypeptide comprising amino acids 1 to 467 of Figure 6 (SEQ ID NO:12), the native sequence PRO232 polypeptide is a mature or full-length native sequence PRO232 polypeptide comprising amino acids 1 to 114 of Figure 9 (SEQ ID NO:18), the native sequence PRO187 is a mature or full-length native sequence PRO187 comprising amino acids 1 to 205 of Figure 11 (SEQ ID NO:23), the native sequence PRO265 polypeptide is a mature or full-length native sequence PRO265 polypeptide comprising amino acids 1 to 660 of Figure 13 (SEQ ID NO:28) or the native sequence PRO265 polypeptide is an extracellular domain of the full-length PRO265 protein, wherein the putative transmembrane domain of the full-length PRO265 protein is encoded by nucleotides beginning at nucleotide 1969 of SEQ ID NO:31, the native sequence PRO219 polypeptide is a mature or full-length native sequence PRO219 polypeptide comprising amino acids 1 to 915 of Figure 15 (SEQ ID NO:34), the native sequence PRO246 polypeptide is a mature or full-length native sequence PRO246 polypeptide comprising amino acids 1 to 390 of Figure 17 (SEQ ID NO:39) or the native sequence PRO246 polypeptide is an extracellular domain of the full-length PRO246 protein, wherein the putative transmembrane domain of the full-length PRO246 protein is encoded by nucleotides beginning at nucleotide 855 as shown in Figure 16, the native sequence PRO228 polypeptide is a mature or full-length native sequence PRO228 polypeptide comprising amino acids 1 to 690 of Figure 19 (SEQ ID NO:49) or the native sequence PRO228 polypeptide is an extracellular domain of the full-length PRO228 protein, the native sequence PRO533 is a mature or full-length native sequence PRO533 comprising amino acids 1 to 216 of Figure 22 (SEQ ID NO:59), with or without the N-terminal signal sequence, and with or without the initiating methionine at position 1, the native sequence PRO245 polypeptide is a mature or full-length native sequence PRO245 polypeptide

comprising amino acids 1 to 312 of Figure 24 (SEQ ID NO:64), the native sequence of each PRO220, PRO221 and PRO227 polypeptides is a mature or full-length native sequence PRO220, PRO221 and PRO227 polypeptide comprising amino acids 1 through 708 of Figure 26 (SEQ ID NO:69), 1 through 259 of Figure 28 (SEQ ID NO:71), and 1 through 620 of Figure 30 (SEQ ID NO:73), the native sequence PRO258 polypeptide is a mature or full-length native sequence PRO258 polypeptide comprising amino acids 1 to 398 of Figure 32 (SEQ ID NO:84) or the native sequence PRO258 polypeptide is an extracellular domain of the full-length PRO258 protein, wherein the putative transmembrane domain of the full-length PRO258 protein is encoded by nucleotides beginning at nucleotide 1134 of SEQ ID NO:83, the native sequence PRO266 polypeptide is a mature or full-length native sequence PRO266 polypeptide comprising amino acids 1 to 696 of Figure 34 (SEQ ID NO:91) or the native sequence PRO266 polypeptide is an extracellular domain of the full-length PRO266 protein, wherein the putative transmembrane domain of the full-length PRO266 protein is encoded by nucleotides beginning at about nucleotide 2009 of SEQ ID NO:104, the native sequence PRO269 polypeptide is a mature or full-length native sequence PRO269 polypeptide comprising amino acids 1 to 490 of Figure 36 (SEQ ID NO:96) or the native sequence PRO269 polypeptide is an extracellular domain of the full-length PRO269 protein, wherein the putative transmembrane domain of the full-length PRO269 protein is encoded by nucleotides beginning at nucleotide 1502 as shown in Figure 35, the native sequence PRO287 polypeptide is a mature or full-length native sequence PRO287 polypeptide comprising amino acids 1 to 415 of Figure 38 (SEQ ID NO:104), the native sequence PRO214 is a mature or full-length native sequence PRO214 comprising amino acids 1 to 420 of Fig. 40 (SEQ ID NO:109), the native-sequence PRO317 is a full-length native-pre-sequence PRO317 comprising amino acids 1 to 366 of Fig. 42 (SEQ ID NO:114) or a mature native-sequence PRO317 comprising amino acids 19 to 366 of Fig. 42 (SEQ ID NO:114), the native sequence PRO301 is a mature or full-length native sequence PRO301 comprising amino acids 1 to 299 of Fig. 44 (SEQ ID NO:119), with or without the N-terminal signal sequence, with or without the initiating methionine at position 1, with or without the potential transmembrane domain at position 236 to about 258, and with or without the intracellular domain at about position 259 to 299, the native sequence PRO224 polypeptide is a mature or full-length native sequence PRO224 polypeptide comprising amino acids 1 to 282 of Figure 46 (SEQ ID NO:127), the native sequence PRO222 polypeptide is a mature or full-length native sequence PRO222 polypeptide comprising amino acids 1 to 490 of Figure 48 (SEQ ID NO:132), the native sequence PRO234 is a mature or full-length native sequence novel lectin comprising amino acids 1 to 382 of Fig. 50 (SEQ ID NO:137), the native sequence PRO231 polypeptide is a mature or full-length native sequence PRO231 polypeptide comprising amino acids 1 to 428 of Figure 52 (SEQ ID NO:142), the native sequence PRO229 polypeptide is a mature or full-length native sequence PRO229 polypeptide comprising amino acids 1 to 347 of Figure 54 (SEQ ID NO:148), the native sequence PRO238 polypeptide is a mature or full-length native sequence PRO238 polypeptide comprising amino acids 1 to 310 of Figure 56 (SEQ ID NO:153), the native sequence PRO233 polypeptide is a mature or full-length native sequence PRO233 polypeptide comprising amino acids 1 to 300 of Figure 58 (SEQ ID NO:159), the native sequence PRO223 polypeptide is a mature or full-length native sequence PRO223 polypeptide comprising amino acids 1 to 476 of Figure 60 (SEQ ID NO:164), the native sequence PRO235 polypeptide is a mature or full-length native sequence PRO235 polypeptide comprising amino acids 1 to 552 of Figure 62 (SEQ ID NO:170), the native sequence PRO236 polypeptide is a mature or full-length native sequence PRO236 polypeptide comprising amino acids 1 to 636 of Figure 64 (SEQ ID NO:175), the native sequence PRO262

polypeptide is a mature or full-length native sequence PRO262 polypeptide comprising amino acids 1 to 654 of Figure 66 (SEQ ID NO:177), the native sequence PRO239 polypeptide is a mature or full-length native sequence PRO239 polypeptide comprising amino acids 1 to 501 of Figure 68 (SEQ ID NO:185), the native sequence PRO257 polypeptide is a mature or full-length native sequence PRO257 polypeptide comprising amino acids 1 to 607 of Figure 70 (SEQ ID NO:190) or the native sequence PRO257 polypeptide is an extracellular domain of the full-length PRO257 protein, wherein the putative transmembrane domain of the full-length PRO257 protein is encoded by nucleotides beginning at nucleotide 2668 as shown in Figure 69, the native sequence PRO260 polypeptide is a mature or full-length native sequence PRO260 polypeptide comprising amino acids 1 to 467 of Figure 72 (SEQ ID NO:195), the native sequence PRO263 polypeptide is a mature or full-length native sequence PRO263 polypeptide comprising amino acids 1 to 322 of Figure 74 (SEQ ID NO:201) or the native sequence PRO263 polypeptide is an extracellular domain of the full-length PRO263 protein, wherein the putative transmembrane domain of the full-length PRO263 protein is encoded by nucleotides beginning at nucleotide 868 of SEQ ID NO:200, the native sequence PRO270 polypeptide is a mature or full-length native sequence PRO270 polypeptide comprising amino acids 1 to 296 of Figure 76 (SEQ ID NO:207), the native sequence PRO271 polypeptide is a mature or full-length native sequence PRO271 polypeptide comprising amino acids 1 to 360 of Figure 78 (SEQ ID NO:213), the native sequence PRO272 polypeptide is a mature or full-length native sequence PRO272 polypeptide comprising amino acids 1 to 328 of Figure 80 (SEQ ID NO:221), the native sequence PRO294 polypeptide is a mature or full-length native sequence PRO294 polypeptide comprising amino acids 1 to 550 of Figure 82 (SEQ ID NO:227), the native sequence PRO295 polypeptide is a mature or full-length native sequence PRO295 polypeptide comprising amino acids 1 to 350 of Figure 84 (SEQ ID NO:236), the native sequence PRO293 polypeptide is a mature or full-length native sequence PRO293 polypeptide comprising amino acids 1 to 713 of Figure 86 (SEQ ID NO:245) or the native sequence PRO293 polypeptide is an extracellular domain of the full-length PRO293 protein, wherein the putative transmembrane domain of the full-length PRO293 protein is encoded by nucleotides beginning at nucleotide 2771 of SEQ ID NO:244, the native sequence PRO247 polypeptide is a mature or full-length native sequence PRO247 polypeptide comprising amino acids 1 to 546 of Figure 88 (SEQ ID NO:250), the native sequence PRO302 polypeptide is a mature or full-length native sequence PRO302 polypeptide comprising amino acids 1 to 452 of Figure 90 (SEQ ID NO:255), the native sequence PRO303 polypeptide is a mature or full-length native sequence PRO303 polypeptide comprising amino acids 1 to 314 of Figure 92 (SEQ ID NO:257), the native sequence PRO304 polypeptide is a mature or full-length native sequence PRO304 polypeptide comprising amino acids 1 to 556 of Figure 94 (SEQ ID NO:259), the native sequence PRO307 polypeptide is a mature or full-length native sequence PRO307 polypeptide comprising amino acids 1 to 383 of Figure 96 (SEQ ID NO:261), the native sequence PRO343 polypeptide is a mature or full-length native sequence PRO343 polypeptide comprising amino acids 1 to 317 of Figure 98 (SEQ ID NO:263), the native sequence PRO328 polypeptide is a mature or full-length native sequence PRO328 polypeptide comprising amino acids 1 to 463 of Figure 100 (SEQ ID NO:285) or the native sequence PRO306 polypeptide is an extracellular domain of the full-length PRO306 protein, wherein the putative extracellular domain of the full-length PRO306 protein, the native sequence PRO335 polypeptide is a mature or full-length native sequence PRO335 polypeptide comprising amino acids 1 through 1059 of Figure 102 (SEQ ID NO:290), the native sequence PRO331 polypeptide is a mature or full-length native sequence PRO331 polypeptide comprising amino acids 1 through 640 of Figure 104

(SEQ ID NO:292), the native sequence PRO326 polypeptide is a mature or full-length native sequence PRO326 polypeptide comprising amino acids 1 through 1119 of Figure 106 (SEQ ID NO:294), wherein additional embodiments include wherein the transmembrane regions are deleted or the peptides are truncated, so as to not include the transmembrane regions for each of PRO335, PRO331, and PRO326, the native sequence PRO332 is a mature or full-length native sequence PRO332 comprising amino acids 49 to 642 of Fig. 180 (SEQ ID NO:310), without or without the N-terminal signal sequence, and with or without the initiating methionine at position 1, the native sequence PRO334 polypeptide is a mature or full-length native sequence PRO334 polypeptide comprising amino acids 1 to 509 of Figure 110 (SEQ ID NO:315), the native sequence PRO346 is a mature or full-length native sequence PRO346 comprising amino acids 19 to 339 of Fig. 112 (SEQ ID NO:320), with or without the N-terminal signal sequence, with or without the initiating methionine, with or without the transmembrane domain at positions 340 to 360 and with or without the intracellular domain at positions 361 to 450, the native sequence PRO268 polypeptide is a mature or full-length native sequence PRO268 polypeptide comprising amino acids 1 to 280 of Figure 114 (SEQ ID NO:325) or the native sequence PRO268 polypeptide is an extracellular domain of the full-length PRO268 protein, wherein the putative transmembrane domain of the full-length PRO268 protein is encoded by nucleotides beginning at nucleotide 559 as shown in Figure 113, the native sequence PRO330 polypeptide is a mature or full-length native sequence PRO330 polypeptide comprising amino acids 1 to 533 of Figure 116 (SEQ ID NO:332), the native sequence PRO339 polypeptide is a mature or full-length native sequence PRO339 polypeptide comprising amino acids 1 to 772 of Figure 118 (SEQ ID NO:339), the native sequence PRO310 polypeptide is a mature or full-length native sequence PRO310 polypeptide comprising amino acids 1 to 318 of Figure 120 (SEQ ID NO:341) and the native sequence PRO244 is a mature or full-length native sequence PRO244 comprising amino acids 1 to 219 of Fig. 122 (SEQ ID NO:377), wherein the mature, full-length native-sequence PRO244 protein comprises a cytoplasmic domain (about amino acid positions 1 to 20), a transmembrane domain (about amino acid positions 21 to 46), and an extracellular domain (about amino acid positions 47 to 219). Within the extracellular domain, the C-lectin domain is between about amino acid position 55 and about amino acid position 206. Native sequence PRO244 as shown in Figure 122 maps to chromosome 12, bands p12-p13.

"PRO polypeptide variant" means an active PRO polypeptide as defined above or below having at least about 80% amino acid sequence identity with the full-length native sequence 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 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, more preferably at least about 90% amino acid sequence identity, and even more preferably at least about 95% amino acid sequence identity with the amino acid sequence of the full-length native amino acid sequence as disclosed herein.

"PRO317 variants" or "PRO317 sequence variants" as defined herein mean biologically active PRO317s as defined below having less than 100% sequence identity with the PRO317 isolated from recombinant cell culture or from mammalian fetal kidney tissue having the deduced sequence described in Figure 42. Ordinarily, a biologically active PRO317 variant will have an amino acid sequence having at least about 70% amino acid sequence identity with the PRO317 of Figure 42, preferably at least about 75%, more preferably at least about 80%, still more preferably at least about 85%, even more preferably at least about 90%, and most preferably at least about 95% (*i.e.*,

70-100%, 75-100%, 80-100%, 85-100%, 90-100%, and 95-100% sequence identity, respectively). These variants include covalently modified polypeptides, as well as PRO317 fragments and glycosylation variants thereof. PRO317 fragments have a consecutive sequence of at least 10, 15, 20, 25, 30, or 40 amino acid residues, preferably about 10-150 residues, that is identical to the sequence of the PRO317 shown in Figure 42. Other preferred PRO317 fragments include those produced as a result of chemical or enzymatic hydrolysis or digestion of the purified PRO317.

A "chimeric PRO317" is a polypeptide comprising full-length PRO317 or one or more fragments thereof fused or bonded to a second protein or one or more fragments thereof. The chimera will typically share at least one biological property in common with PRO317. The second protein will typically be a cytokine, growth factor, or hormone such as a neurotrophic or angiogenic factor such as GDNF or VEGF, or another member of the TGF-superfamily such as EBAF-1. Another exemplary preferred PRO317 chimera is a "domain chimera" that consists of the N-terminal residues substituted with one or more, but not all, of the residues of the human EBAF-1. In this embodiment, the PRO317 chimera would have individual or blocks of residues from the human EBAF-1 sequence added or substituted into the PRO317 sequence. For example, one or more of those segments of EBAF-1 that are not homologous could be substituted into the corresponding segments of PRO317. It is contemplated that this "PRO317-EBAF-1 domain chimera" will have an agonist biological activity.

"Percent (%) amino acid sequence identity" with respect to the PRO polypeptide sequences 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 the art, for instance, using publicly available computer software such as BLAST, ALIGN or Megalign (DNASTAR) software. The preferred software alignment program is BLAST. 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.

"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 in the art, for instance, using publicly available computer software such as BLAST, 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.

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

"Southern analysis" or "Southern blotting" is a method by which the presence of DNA sequences in a restriction endonuclease digest of DNA or a DNA-containing composition is confirmed by hybridization to a known, labeled oligonucleotide or DNA fragment. Southern analysis typically involves electrophoretic separation of DNA digests on agarose gels, denaturation of the DNA after electrophoretic separation, and transfer of the DNA to nitrocellulose, nylon, or another suitable membrane support for analysis with a radiolabeled, biotinylated, or enzyme-labeled probe as described in sections 9.37-9.52 of Sambrook *et al.*, Molecular Cloning: A Laboratory Manual (New York: Cold Spring Harbor Laboratory Press, 1989).

"Northern analysis" or "Northern blotting" is a method used to identify RNA sequences that hybridize to a known probe such as an oligonucleotide, DNA fragment, cDNA or fragment thereof, or RNA fragment. The probe is labeled with a radioisotope such as ^{32}P, or by biotinylation, or with an enzyme. The RNA to be analyzed is usually electrophoretically separated on an agarose or polyacrylamide gel, transferred to nitrocellulose, nylon, or other suitable membrane, and hybridized with the probe, using standard techniques well known in the art such as those described in sections 7.39-7.52 of Sambrook *et al.*, *supra*.

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. 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; 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 single anti-PRO polypeptide monoclonal antibodies (including agonist, antagonist, and neutralizing antibodies) and anti-PRO polypeptide antibody compositions with polyepitopic specificity. The term "monoclonal antibody" as used herein refers to an antibody obtained from 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.

"Active" or "activity" for the purposes herein refers to form(s) of PRO polypeptide which retain the biologic and/or immunologic activities of the specific native or naturally-occurring PRO polypeptide. The activity of a PRO332 polypeptide preferably involves the regulation of extracellular matrix, cartilage, or bone function.

"PRO317-associated disorder" refers to a pathological condition or disease wherein PRO317 is over- or underexpressed. Such disorders include diseases of the female genital tract or of the endometrium of a mammal, including hyperplasia, endometritis, endometriosis, wherein the patient is at risk for infertility due to endometrial factor, endometrioma, and endometrial cancer, especially those diseases involving abnormal bleeding such as a gynecological disease. They also include diseases involving angiogenesis, wherein the angiogenesis results in a pathological condition, such as cancer involving solid tumors (the therapy for the disorder would result in decreased vascularization and a decline in growth and metastasis of a variety of tumors). Alternatively, the angiogenesis may be beneficial, such as for ischemia, especially coronary ischemia. Hence, these disorders include those found in patients whose hearts are functioning but who have a blocked blood supply due to atherosclerotic coronary artery disease, and those with a functioning but underperfused heart, including patients with coronary arterial disease who are not optimal candidates for angioplasty and coronary artery by-pass surgery. The disorders also include diseases involving the kidney or originating from the kidney tissue, such as polycystic kidney disease and chronic and acute renal failure.

"Treatment" or "treating" refers to both therapeutic treatment and prophylactic or preventative measures. Those in need of treatment include those already with the disorder as well as those prone to have the disorder of those in which the disorder is to be prevented.

"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 sheep, dogs, horses, cats, cows, and the like. Preferably, the mammal herein is a human.

"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 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; 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™.

The term "agonist" is used to refer to peptide and non-peptide analogs of the native PRO polypeptides (where native PRO polypeptide refers to pro-PRO polypeptide, pre-PRO polypeptide, prepro-PRO polypeptide, or

mature PRO polypeptide) of the present invention and to antibodies specifically binding such native PRO polypeptides, provided that they retain at least one biological activity of a native PRO polypeptide. Preferably, the agonists of the present invention retain the qualitative binding recognition properties and receptor activation properties of the native PRO polypeptide.

The term "antagonist" is used to refer to a molecule inhibiting a biological activity of a native PRO polypeptide of the present invention wherein native PRO polypeptide refers to pro-PRO polypeptide, pre-PRO polypeptide, prepro-PRO polypeptide, or mature PRO polypeptide. Preferably, the antagonists herein inhibit the binding of a native PRO polypeptide of the present invention. Preferred antagonists essentially completely block the binding of a native PRO317 polypeptide to a PRO317 polypeptide receptor to which it otherwise binds. Such receptors may include the Type I and Type II, and possibly Type III receptors identified for the TGF- superfamily. Kolodziejczyk and Hall, *supra*. A PRO polypeptide "antagonist" is a molecule which prevents, or interferes with, a PRO antagonist effector function (*e.g.* a molecule which prevents or interferes with binding and/or activation of a PRO polypeptide receptor by PRO polypeptide). Such molecules can be screened for their ability to competitively inhibit PRO polypeptide receptor activation by monitoring binding of native PRO polypeptide in the presence and absence of the test antagonist molecule, for example. Examples of PRO317 polypeptide antagonists include neutralizing antibodies against F-2. An antagonist of the invention also encompasses an antisense polynucleotide against the PRO polypeptide gene, which antisense polynucleotide blocks transcription or translation of the PRO polypeptide gene, thereby inhibiting its expression and biological activity.

"Stringent conditions" means (1) employing low ionic strength and high temperature for washing, for example, 0.015 sodium chloride/0.0015 M sodium citrate/0.1% sodium dodecyl sulfate at 50°C, or (2) employing during hybridization a denaturing agent, such as formamide, for example, 50% (vol/vol) formamide with 0.1% bovine serum albumin/0.1% Ficoll/0.1% polyvinylpyrrolidone/50 nM sodium phosphate buffer at pH 6.5 with 750 mM sodium chloride, 75 mM sodium citrate at 42°C. Another example is use of 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 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 and 0.1% SDS. Yet another example is hybridization using a buffer of 10% dextran sulfate, 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" are described in Sambrook *et al.*, *supra*, and include the use of a washing solution and hybridization conditions (*e.g.*, temperature, ionic strength, and %SDS) less stringent than described above. An example of moderately stringent conditions is a condition such as 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 how to adjust the temperature, ionic strength, *etc.*, as necessary to accommodate factors such as probe length and the like.

II. Compositions and Methods of the Invention

1. **Full-length PRO211 and PRO217 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO211 and PRO217. In particular, Applicants have identified and isolated cDNA encoding PRO211 and PRO217 polypeptides, as disclosed in further detail in the Examples below. Using BLAST (FastA format) sequence alignment computer programs, Applicants found that cDNA sequences encoding full-length native sequence PRO211 and PRO217 have homologies to known proteins having EGF-like domains. Specifically, the cDNA sequence DNA32292-1131 (Figure 1, SEQ ID NO:1) has 36% identify and a Blast score of 209 with PAC6_RAT and 31% identify and a Blast score of 206 with Fibulin-1, isoform c precursor. The cDNA sequence DNA33094-1131 (Figure 3, SEQ ID NO:3) has 36% identity and a Blast score of 336 with eastern newt tenascin, and 37% identity and a Blast score of 331 with human tenascin-X precursor. Accordingly, it is presently believed that PRO211 and PRO217 polypeptides disclosed in the present application are newly identified members of the EGF-like family and possesses properties typical of the EGF-like protein family.

2. **Full-length PRO230 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO230. In particular, Applicants have identified and isolated cDNA encoding a PRO230 polypeptide, as disclosed in further detail in the Examples below. Using known programs such as BLAST and FastA sequence alignment computer programs, Applicants found that a cDNA sequence encoding full-length native sequence PRO230 has 48% amino acid identity with the rabbit tubulointerstitial nephritis antigen precursor. Accordingly, it is presently believed that PRO230 polypeptide disclosed in the present application is a newly identified member of the tubulointerstitial nephritis antigen family and possesses the ability to be recognized by human autoantibodies in certain forms of tubulointerstitial nephritis.

3. **Full-length PRO232 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO232. In particular, Applicants have identified and isolated cDNA encoding a PRO232 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that a portion of the full-length native sequence PRO232 (shown in Figure 9 and SEQ ID NO:18) has 35% sequence identity with a stem cell surface antigen from Gallus gallus. Accordingly, it is presently believed that the PRO232 polypeptide disclosed in the present application may be a newly identified stem cell antigen.

4. **Full-length PRO187 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO187. In particular, Applicants have identified and isolated cDNA encoding a PRO187 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that a full-length native sequence PRO187 (shown in Figure

15) has 74% amino acid sequence identity and BLAST score of 310 with various androgen-induced growth factors and FGF-8. Accordingly, it is presently believed that PRO187 polypeptide disclosed in the present application is a newly identified member of the FGF-8 protein family and may possess identify activity or property typical of the FGF-8-like protein family.

5. **Full-length PRO265 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO265. In particular, Applicants have identified and isolated cDNA encoding a PRO265 polypeptide, as disclosed in further detail in the Examples below. Using programs such as BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO265 polypeptide have significant homology with the fibromodulin protein and fibromodulin precursor protein. Applicants have also found that the DNA encoding the PRO265 polypeptide has significant homology with platelet glycoprotein V, a member of the leucine rich related protein family involved in skin and wound repair. Accordingly, it is presently believed that PRO265 polypeptide disclosed in the present application is a newly identified member of the leucine rich repeat family and possesses protein protein binding capabilities, as well as be involved in skin and wound repair as typical of this family.

6. **Full-length PRO219 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO219. In particular, Applicants have identified and isolated cDNA encoding a PRO219 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO219 polypeptide have significant homology with the mouse and human matrilin-2 precursor polypeptides. Accordingly, it is presently believed that PRO219 polypeptide disclosed in the present application is related to the matrilin-2 precursor polypeptide.

7. **Full-length PRO246 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO246. In particular, Applicants have identified and isolated cDNA encoding a PRO246 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that a portion of the PRO246 polypeptide has significant homology with the human cell surface protein HCAR. Accordingly, it is presently believed that PRO246 polypeptide disclosed in the present application may be a newly identified membrane-bound virus receptor or tumor cell-specific antigen.

8. **Full-length PRO228 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO228. In particular, Applicants have identified and isolated cDNA

encoding a PRO228 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO228 polypeptide have significant homology with the EMR1 protein. Applicants have also found that the DNA encoding the PRO228 polypeptide has significant homology with latrophilin, macrophage-restricted cell surface glycoprotein, B0457.1 and leucocyte antigen CD97 precursor. Accordingly, it is presently believed that PRO228 polypeptide disclosed in the present application is a newly identified member of the seven transmembrane superfamily and possesses characteristics and functional properties typical of this family. In particular, it is believed that PRO228 is a new member of the subgroup within this family to which CD97 and EMR1 belong.

9. Full-length PRO533 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO533. In particular, Applicants have identified and isolated cDNA encoding a PRO533 polypeptide, as disclosed in further detail in the Examples below. Using BLAST-2 and FastA sequence alignment computer programs, Applicants found that a full-length native sequence PRO533 (shown in Figure 22 and SEQ ID NO:59) has a Blast score of 509 and 53% amino acid sequence identity with fibroblast growth factor (FGF). Accordingly, it is presently believed that PRO533 disclosed in the present application is a newly identified member of the fibroblast growth factor family and may possess activity typical of such polypeptides.

10. Full-length PRO245 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO245. In particular, Applicants have identified and isolated cDNA encoding a PRO245 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that a portion of the amino acid sequence of the PRO245 polypeptide has 60% amino acid identity with the human c-myb protein. Accordingly, it is presently believed that the PRO245 polypeptide disclosed in the present application may be a newly identified member of the transmembrane protein tyrosine kinase family.

11. Full-length PRO220, PRO221 and PRO227 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO220, PRO221 and PRO227. In particular, Applicants have identified and isolated cDNAs encoding a PRO220, PRO221 and PRO227 polypeptide, respectively, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, PRO220 has amino acid identity with the amino acid sequence of a leucine rich protein wherein the identity is 87%. PRO220 additionally has amino acid identity with the neuronal leucine rich protein wherein the identity is 55%. The neuronal leucine rich protein is further described in Taguchi, *et al.*, Mol. Brain Res., 35:31-40 (1996).

PRO221 has amino acid identity with the SLIT protein precursor, wherein different portions of these two proteins have the respective percent identities of 39%, 38%, 34%, 31%, and 30%.

PRO227 has amino acid identity with the amino acid sequence of platelet glycoprotein V precursor. The same results were obtained for human glycoprotein V. Different portions of these two proteins show the following percent identities of 30%, 28%, 28%, 31%, 35%, 39% and 27%.

Accordingly, it is presently believed that PRO220, PRO221 and PRO227 polypeptides disclosed in the present application are newly identified members of the leucine rich repeat protein superfamily and that each possesses protein-protein binding capabilities typical of the leucine rich repeat protein superfamily. It is also believed that they have capabilities similar to those of SLIT, the leucine rich repeat protein and human glycoprotein V.

12. Full-length PRO258 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO258. In particular, Applicants have identified and isolated cDNA encoding a PRO258 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO258 polypeptide have significant homology with the CRTAM and poliovirus receptors. Accordingly, it is presently believed that PRO258 polypeptide disclosed in the present application is a newly identified member of the Ig superfamily and possesses virus receptor capabilities or regulates immune function as typical of this family.

13. Full-length PRO266 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO266. In particular, Applicants have identified and isolated cDNA encoding a PRO266 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO266 polypeptide have significant homology with the SLIT protein from Drosophilia. Accordingly, it is presently believed that PRO266 polypeptide disclosed in the present application is a newly identified member of the leucine rich repeat family and possesses ligand-ligand binding activity and neuronal development typical of this family. SLIT has been shown to be useful in the study and treatment of Alzheimer's disease, *supra*, and thus, PRO266 may have involvement in the study and cure of this disease.

14. Full-length PRO269 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO269. In particular, Applicants have identified and isolated cDNA encoding a PRO269 polypeptide, as disclosed in further detail in the Examples below. Using BLAST, FastA and sequence alignment computer programs, Applicants found that the amino acid sequence encoded by nucleotides 314 to 1783 of the full-length native sequence PRO269 (shown in Figure 35 and SEQ ID NO:95) has significant homology to human urinary thrombomodulin and various thrombomodulin analogues respectively, to which it was aligned. Accordingly, it is presently believed that PRO269 polypeptide disclosed in the present application is a newly identified member of the thrombomodulin family.

15. **Full-length PRO287 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO287. In particular, Applicants have identified and isolated cDNA encoding a PRO287 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO287 polypeptide have significant homology with the type 1 procollagen C-proteinase enhancer protein precursor and type 1 procollagen C-proteinase enhancer protein. Accordingly, it is presently believed that PRO287 polypeptide disclosed in the present application is a newly identified member of the C-proteinase enhancer protein family.

16. **Full-length PRO214 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO214. In particular, Applicants have identified and isolated cDNA encoding a PRO214 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that a full-length native sequence PRO214 polypeptide (shown in Figure 40 and SEQ ID NO:109) has 49% amino acid sequence identity with HT protein, a known member of the EGF-family. The comparison resulted in a BLAST score of 920, with 150 matching nucleotides. Accordingly, it is presently believed that the PRO214 polypeptide disclosed in the present application is a newly identified member of the family comprising EGF domains and may possess activities or properties typical of the EGF-domain containing family.

17. **Full-length PRO317 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO317. In particular, cDNA encoding a PRO317 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below. Using BLAST™ and FastA™ sequence alignment computer programs, it was found that a full-length native-sequence PRO317 (shown in Figure 42 and SEQ ID NO:114) has 92% amino acid sequence identity with EBAF-1. Further, it is closely aligned with many other members of the TGF- superfamily.

Accordingly, it is presently believed that PRO317 disclosed in the present application is a newly identified member of the TGF- superfamily and may possess properties that are therapeutically useful in conditions of uterine bleeding, etc. Hence, PRO317 may be useful in diagnosing or treating abnormal bleeding involved in gynecological diseases, for example, to avoid or lessen the need for a hysterectomy. PRO317 may also be useful as an agent that affects angiogenesis in general, so PRO317 may be useful in anti-tumor indications, or conversely, in treating coronary ischemic conditions.

Library sources reveal that ESTs used to obtain the consensus DNA for generating PRO317 primers and probes were found in normal tissues (uterus, prostate, colon, and pancreas), in several tumors (colon, brain (twice), pancreas, and mullerian cell), and in a heart with ischemia. PRO317 has shown up in several tissues as well, but it does look to have a greater concentration in uterus. Hence, PRO317 may have a broader use by the body than EBAF-1. It is contemplated that, at least for some indications, PRO317 may have opposite effects from EBAF-1.

18. Full-length PRO301 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO301. In particular, Applicants have identified and isolated cDNA encoding a PRO301 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that a full-length native sequence PRO301 (shown in Figure 44 and SEQ ID NO:119) has a Blast score of 246 corresponding to 30% amino acid sequence identity with human A33 antigen precursor. Accordingly, it is presently believed that PRO301 disclosed in the present application is a newly identified member of the A33 antigen protein family and may be expressed in human neoplastic diseases such as colorectal cancer.

19. Full-length PRO224 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO224. In particular, Applicants have identified and isolated cDNA encoding a PRO224 polypeptide, as disclosed in further detail in the Examples below. Using known programs such as BLAST and FastA sequence alignment computer programs, Applicants found that full-length native PRO224 (Figure 46, SEQ ID NO:127) has amino acid identity with apolipoprotein E receptor 2906 from homo sapiens. The alignments of different portions of these two polypeptides show amino acid identities of 37%, 36%, 30%, 44%, 44% and 28% respectively. Full-length native PRO224 (Figure 46, SEQ ID NO:127) also has amino acid identity with very low-density lipoprotein receptor precursor from gall. The alignments of different portions of these two polypeptides show amino acid identities of 38%, 37%, 42%, 33%, and 37% respectively. Additionally, full-length native PRO224 (Figure 46, SEQ ID NO:127) has amino acid identity with the chicken oocyte receptor P95 from Gallus gallus. The alignments of different portions of these two polypeptides show amino acid identities of 38%, 37%, 42%, 33%, and 37% respectively. Moreover, full-length native PRO224 (Figure 46, SEQ ID NO:127) has amino acid identity with very low density lipoprotein receptor short form precursor from humans. The alignments of different portions of these two polypeptides show amino acid identities of 32%, 38%, 34%, 45%, and 31%, respectively. Accordingly, it is presently believed that PRO224 polypeptide disclosed in the present application is a newly identified member of the low density lipoprotein receptor family and possesses the structural characteristics required to have the functional ability to recognize and endocytose low density lipoproteins typical of the low density lipoprotein receptor family. (The alignments described above used the following scoring parameters: T=7, S+65, S2=36, Matrix: BLOSUM62.)

20. Full-length PRO222 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO222. In particular, Applicants have identified and isolated cDNA encoding a PRO222 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that a sequence encoding full-length native sequence PRO222 (shown in Figure 48 and SEQ ID NO:132) has 25-26% amino acid identity with mouse complement factor h precursor, has 27-29% amino acid identity with complement receptor, has 25-47% amino acid identity with mouse

complement C3b receptor type 2 long form precursor, has 40% amino acid identity with human hypothetical protein kiaa0247. Accordingly, it is presently believed that PRO222 polypeptide disclosed in the present application is a newly identified member of the complement receptor family and possesses activity typical of the complement receptor family.

21. Full-length PRO234 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO234. In particular, Applicants have identified and isolated cDNA encoding a PRO234 polypeptide, as disclosed in further detail in the Examples below. Using BLAST (FastA-format) sequence alignment computer programs, Applicants found that a cDNA sequence encoding full-length native sequence PRO234 has 31% identity and Blast score of 134 with E-selectin precursor. Accordingly, it is presently believed that the PRO234 polypeptides disclosed in the present application are newly identified members of the lectin/selectin family and possess activity typical of the lectin/selectin family.

22. Full-length PRO231 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO231. In particular, Applicants have identified and isolated cDNA encoding a PRO231 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that the full-length native sequence PRO231 polypeptide (shown in Figure 52 and SEQ ID NO:142) has 30 % and 31 % amino acid identity with human and rat prostatic acid phosphatase precursor proteins, respectively. Accordingly, it is presently believed that the PRO231 polypeptide disclosed in the present application may be a newly identified member of the acid phosphatase protein family.

23. Full-length PRO229 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO229. In particular, Applicants have identified and isolated cDNA encoding a PRO229 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO229 polypeptide have significant homology with antigen wc1.1, M130 antigen, T cell surface glycoprotein CD6 and CD6. It also is related to Sp-alpha. Accordingly, it is presently believed that PRO229 polypeptide disclosed in the present application is a newly identified member of the family containing scavenger receptor homology, a sequence motif found in a number of proteins involved in immune function and thus possesses immune function and /or segments which resist degradation, typical of this family.

24. Full-length PRO238 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO238. In particular, Applicants have identified and isolated cDNA encoding a PRO238 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA

sequence alignment computer programs, Applicants found that various portions of the PRO238 polypeptide have significant homology with reductases, including oxidoreductase and fatty acyl-CoA reductase. Accordingly, it is presently believed that PRO238 polypeptide disclosed in the present application is a newly identified member of the reductase family and possesses reducing activity typical of the reductase family.

25. Full-length PRO233 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO233. In particular, Applicants have identified and isolated cDNA encoding a PRO233 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO233 polypeptide have significant homology with the reductase protein. Applicants have also found that the DNA encoding the PRO233 polypeptide has significant homology with proteins from *Caenorhabditis elegans*. Accordingly, it is presently believed that PRO233 polypeptide disclosed in the present application is a newly identified member of the reductase family and possesses the ability to effect the redox state of the cell typical of the reductase family.

26. Full-length PRO223 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO223. In particular, Applicants have identified and isolated cDNA encoding a PRO223 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that the PRO223 polypeptide has significant homology with various serine carboxypeptidase polypeptides. Accordingly, it is presently believed that PRO223 polypeptide disclosed in the present application is a newly identified serine carboxypeptidase.

27. Full-length PRO235 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO235. In particular, Applicants have identified and isolated cDNA encoding a PRO235 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO235 polypeptide have significant homology with the various plexin proteins. Accordingly, it is presently believed that PRO235 polypeptide disclosed in the present application is a newly identified member of the plexin family and possesses cell adhesion properties typical of the plexin family.

28. Full-length PRO236 and PRO262 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO236 and PRO262. In particular, Applicants have identified and isolated cDNA encoding PRO236 and PRO262 polypeptides, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO236 and PRO262 polypeptides have significant homology with various β-galactosidase and β-galactosidase precursor

polypeptides. Accordingly, it is presently believed that the PRO236 and PRO262 polypeptides disclosed in the present application are newly identified β-galactosidase homologs.

29. Full-length PRO239 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO239. In particular, Applicants have identified and isolated cDNA encoding a PRO239 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO239 polypeptide have significant homology with densin proteins. Accordingly, it is presently believed that PRO239 polypeptide disclosed in the present application is a newly identified member of the densin family and possesses cell adhesion and the ability to effect synaptic processes as is typical of the densin family.

30. Full-length PRO257 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO257. In particular, Applicants have identified and isolated cDNA encoding a PRO257 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO257 polypeptide have significant homology with the ebnerin precursor and ebnerin protein. Accordingly, it is presently believed that PRO257 polypeptide disclosed in the present application is a newly identified protein member which is related to the ebnerin protein.

31. Full-length PRO260 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO260. In particular, Applicants have identified and isolated cDNA encoding a PRO260 polypeptide, as disclosed in further detail in the Examples below. Using programs such as BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO260 polypeptide have significant homology with the alpha-l-fucosidase precursor. Accordingly, it is presently believed that PRO260 polypeptide disclosed in the present application is a newly identified member of the fucosidase family and possesses enzymatic activity related to fucose residues typical of the fucosidase family.

32. Full-length PRO263 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO263. In particular, Applicants have identified and isolated cDNA encoding a PRO263 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO263 polypeptide have significant homology with the CD44 antigen and related proteins. Accordingly, it is presently believed that PRO263 polypeptide disclosed in the present application is a newly identified member of the CD44 antigen family and possesses at least one of the properties associated with these antigens, i.e., cancer and HIV marker, cell-cell or cell-

matrix interactions, regulating cell traffic, lymph node homing, transmission of growth signals, and presentation of chemokines and growth facors to traveling cells.

33. Full-length PRO270 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO270. In particular, Applicants have identified and isolated cDNA encoding a PRO270 polypeptide, as disclosed in further detail in the Examples below. Using BLAST, FastA and sequence alignment computer programs, Applicants found that that various portions of the PRO270 polypeptide have significant homology with various thioredoxin proteins. Accordingly, it is presently believed that PRO270 polypeptide disclosed in the present application is a newly identified member of the thioredoxin family and possesses the ability to effect reduction-oxidation (redox) state typical of the thioredoxin family.

34. Full-length PRO271 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO271. In particular, Applicants have identified and isolated cDNA encoding a PRO271 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that the PRO271 polypeptide has significant homology with various link proteins and precursors thereof. Accordingly, it is presently believed that PRO271 polypeptide disclosed in the present application is a newly identified link protein homolog.

35. Full-length PRO272 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO272. In particular, Applicants have identified and isolated cDNA encoding a PRO272 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO272 polypeptide have significant homology with the human reticulocalbin protein and its precursors. Applicants have also found that the DNA encoding the PRO272 polypeptide has significant homology with the mouse reticulocalbin precursor protein. Accordingly, it is presently believed that PRO272 polypeptide disclosed in the present application is a newly identified member of the reticulocalbin family and possesses the ability to bind calcium typical of the reticulocalbin family.

36. Full-length PRO294 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO294. In particular, Applicants have identified and isolated cDNA encoding a PRO294 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO294 polypeptide have significant homology with the various portions of a number of collagen proteins. Accordingly, it is presently believed that PRO294 polypeptide disclosed in the present application is a newly identified member of the collagen family.

37. **Full-length PRO295 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO295. In particular, Applicants have identified and isolated cDNA encoding a PRO295 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO295 polypeptide have significant homology with integrin proteins. Accordingly, it is presently believed that PRO295 polypeptide disclosed in the present application is a newly identified member of the integrin family and possesses cell adhesion typical of the integrin family.

38. **Full-length PRO293 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO293. In particular, Applicants have identified and isolated cDNA encoding a PRO293 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that portions of the PRO293 polypeptide have significant homology with the neuronal leucine rich repeat proteins 1 and 2, (NLRR-1 and NLRR-2), particularly NLRR-2. Accordingly, it is presently believed that PRO293 polypeptide disclosed in the present application is a newly identified member of the neuronal leucine rich repeat protein family and possesses ligand-ligand binding activity typical of the NRLL protein family.

39. **Full-length PRO247 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO247. In particular, Applicants have identified and isolated cDNA encoding a PRO247 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO247 polypeptide have significant homology with densin. Applicants have also found that the DNA encoding the PRO247 polypeptide has significant homology with a number of other proteins, including KIAA0231. Accordingly, it is presently believed that PRO247 polypeptide disclosed in the present application is a newly identified member of the leucine rich repeat family and possesses ligand binding abilities typical of this family.

40. **Full-length PRO302, PRO303, PRO304, PRO307 and PRO343 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO302, PRO303, PRO304, PRO307 and PRO343. In particular, Applicants have identified and isolated cDNA encoding PRO302, PRO303, PRO304, PRO307 and PRO343 polypeptides, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO302, PRO303, PRO304, PRO307 and PRO343 polypeptides have significant homology with various protease proteins. Accordingly, it is presently believed that the PRO302, PRO303, PRO304, PRO307 and PRO343 polypeptides disclosed in the present application are newly identified protease proteins.

41. **Full-length PRO328 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO328. In particular, Applicants have identified and isolated cDNA encoding a PRO328 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO328 polypeptide have significant homology with the human glioblastoma protein ("GLIP"). Further, Applicants found that various portions of the PRO328 polypeptide have significant homology with the cysteine rich secretory protein ("CRISP") as identified by BLAST homology [ECCRISP3_1, S68683, and CRS3_HUMAN]. Accordingly, it is presently believed that PRO328 polypeptide disclosed in the present application is a newly identified member of the GLIP or CRISP families and possesses transcriptional regulatory activity typical of the GLIP or CRISP families.

42. **Full-length PRO335, PRO331 and PRO326 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO335, PRO331 or PRO326. In particular, Applicants have identified and isolated cDNA encoding a PRO335, PRO331 or PRO326 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO335, PRO331 or PRO326 polypeptide have significant homology with LIG-1, ALS and in the case of PRO331, additionally, decorin. Accordingly, it is presently believed that the PRO335, PRO331 and PRO326 polypeptides disclosed in the present application are newly identified members of the leucine rich repeat superfamily, and particularly, are related to LIG-1 and possess the biological functions of this family as discussed and referenced herein.

43. **Full-length PRO332 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO332. In particular, Applicants have identified and isolated cDNA encoding PRO332 polypeptides, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that a full-length native sequence PRO332 (shown in Figure 108 and SEQ ID NO:310) has about 30-40% amino acid sequence identity with a series of known proteoglycan sequences, including, for example, fibromodulin and fibromodulin precursor sequences of various species (FMOD BOVIN, FMOD_CHICK, FMOD_RAT, FMOD_MOUSE, FMOD_HUMAN, P_R36773), osteomodulin sequences (AB000114_1, AB007848_1), decorin sequences (CFU83141_1, OCU03394_1, P_R42266, P_R42267, P_R42260, P R89439), keratan sulfate proteoglycans (BTU48360_1, AF022890_1), corneal proteoglycan (AF022256_1), and bone/cartilage proteoglycans and proteoglycane precursors (PGS1_BOVIN, PGS2_ MOUSE, PGS2_HUMAN). Accordingly, it is presently believed that PRO332 disclosed in the present application is a new proteoglycan-type molecule, and may play a role in regulating extracellular matrix, cartilage, and/or bone function.

44. Full-length PRO334 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO334. In particular, Applicants have identified and isolated cDNA encoding a PRO334 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO334 polypeptide have significant homology with fibulin and fibrillin. Accordingly, it is presently believed that PRO334 polypeptide disclosed in the present application is a newly identified member of the epidermal growth factor family and possesses properties and activities typical of this family.

45. Full-length PRO346 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO346. In particular, Applicants have identified and isolated cDNA encoding a PRO346 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that a full-length native sequence PRO346 (shown in Figure 112 and SEQ ID NO:320) has 28 % amino acid sequence identity with carcinoembryonic antigen. Accordingly, it is presently believed that PRO346 disclosed in the present application is a newly identified member of the carcinoembryonic protein family and may be expressed in association with neoplastic tissue disorders.

46. Full-length PRO268 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO268. In particular, Applicants have identified and isolated cDNA encoding a PRO268 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that portions of the PRO268 polypeptide have significant homology with the various protein disulfide isomerase proteins. Accordingly, it is presently believed that PRO268 polypeptide disclosed in the present application is a homolog of the protein disulfide isomerase p5 protein.

47. Full-length PRO330 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO330. In particular, Applicants have identified and isolated cDNA encoding a PRO330 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO330 polypeptide have significant homology with the murine prolyl 4-hydroxylase alpha-II subunit protein. Accordingly, it is presently believed that PRO330 polypeptide disclosed in the present application is a novel prolyl 4-hydroxylase subunit polypeptide.

48. Full-length PRO339 and PRO310 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO339 and PRO310. In particular, Applicants have identified and isolated

cDNA encoding a PRO339 polypeptide, as disclosed in further detail in the Examples below. Applicants have also identified and isolated cDNA encoding a PRO310 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO339 and PRO310 polypeptides have significant homology with small secreted proteins from *C. elegans* and are distantly related to fringe. PRO339 also shows homology to collagen-like polymers. Sequences which were used to identify PRO310, designated herein as DNA40533 and DNA42267, also show homology to proteins from *C. elegans*. Accordingly, it is presently believed that the PRO339 and PRO310 polypeptides disclosed in the present application are newly identified member of the family of proteins involved in development, and which may have regulatory abilities similar to the capability of fringe to regulate serrate.

49. Full Length PRO244 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding C-type lectins referred to in the present application as PRO244. In particular, applicants have identified and isolated cDNA encoding PRO244 polypeptides, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that a full-length native sequence PRO244 (shown in Figure 122 and SEQ ID NO:377) has 43% amino acid sequence identity with the hepatic lectin gallus gallus (LECH-CHICK), and 42% amino acid sequence identity with an HIV gp120 binding C-type lectin (A46274). Accordingly, it is presently believed that PRO244 disclosed in the present application is a newly identified member of the C-lectin superfamily and may play a role in immune function, apoptosis, or in the pathogenesis of atherosclerosis. In addition, PRO244 may be useful in identifying tumor-associated epitopes.

50. PRO Polypeptide Variants

In addition to the full-length native sequence PRO polypeptides described herein, it is contemplated that PRO polypeptide variants can be prepared. PRO polypeptide variants can be prepared by introducing appropriate nucleotide changes into the PRO polypeptide DNA, 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 polypeptides, such as changing the number or position of glycosylation sites or altering the membrane anchoring characteristics.

Variations in the native full-length sequence PRO polypeptides or in various domains of the PRO polypeptides 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 polypeptide that results in a change in the amino acid sequence of the PRO polypeptide as compared with the native sequence PRO polypeptide. 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 polypeptide. 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 polypeptide 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 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 in the *in vitro* assay described in the Examples below.

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 desired PRO polypeptide variant DNA.

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. Alanine is also typically preferred because it is the most common amino acid. Further, it is frequently found 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.

51. Modifications of PRO Polypeptides

Covalent modifications of PRO polypeptides are included within the scope of this invention. One type of covalent modification includes reacting targeted amino acid residues of the 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 polypeptide. Derivatization with bifunctional agents is useful, for instance, for crosslinking a PRO polypeptide to a water-insoluble support matrix or surface for use in the method for purifying anti-PRO polypeptide antibodies, and vice-versa. Commonly used crosslinking agents include, *e.g.*, 1,1-bis(diazoacetyl)-2-phenylethane, glutaraldehyde, N-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.

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 α-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 polypeptides included within the scope of this 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 a native sequence PRO polypeptide, and/or adding one or more glycosylation sites that are not present in the native sequence PRO polypeptide.

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 serine or threonine residues to the native sequence PRO polypeptide (for O-linked glycosylation sites). The PRO polypeptide 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 polypeptide is by chemical or enzymatic coupling of glycosides to the polypeptide. Such methods are described in the art, 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, 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 polypeptides of the invention comprises linking the PRO polypeptide to one of a variety of nonproteinaceous polymers, *e.g.*, polyethylene glycol, 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 polypeptides of the present invention may also be modified in a way to form a chimeric molecule comprising a PRO polypeptide fused to another, heterologous polypeptide or amino acid sequence. In one embodiment, such a chimeric molecule comprises a fusion of the PRO polypeptide 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 polypeptide. The presence of such epitope-tagged forms of the PRO polypeptide can be detected using an antibody against the tag polypeptide. Also, provision of the epitope tag enables the PRO polypeptide to be readily purified by affinity purification using an anti-tag antibody or another type of affinity matrix that binds to the epitope tag. In an alternative embodiment, the chimeric molecule may comprise a fusion of the PRO polypeptide with an immunoglobulin or a particular region of an immunoglobulin. For a bivalent form of the chimeric molecule, such a fusion could be to the Fc region of an IgG molecule.

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 α-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)].

52. Modification of PRO317

Amino acid sequence variants of PRO317 are prepared by introducing appropriate nucleotide changes into the PRO317 DNA, or by *in vitro* synthesis of the desired PRO317 polypeptide. Such variants include, for example, deletions from, or insertions or substitutions of, residues within the amino acid sequence shown for human PRO317 in Figure 42. Any combination of deletion, insertion, and substitution is made to arrive at the final construct, provided that the final construct possesses the desired characteristics. The amino acid changes also may alter post-translational processes of the PRO317, such as changing the number or position of glycosylation sites. Moreover, like most mammalian genes, PRO317 is presumably encoded by multi-exon genes. Alternative mRNA constructs which may be attributed to different mRNA splicing events following transcription, and which share large regions of identity with the cDNAs claimed herein, are considered to be within the scope of the present invention.

For the design of amino acid sequence variants of PRO317, the location of the mutation site and the nature of the mutation will depend on the PRO317 characteristic(s) to be modified. For example, candidate PRO317 antagonists or agonists will be initially selected by locating sites that are identical or highly conserved among PRO317, EBAF-1, LEFTY, and other members of the TGF- superfamily. The sites for mutation can be modified individually or in series, *e.g.*, by (1) substituting first with conservative amino acid choices and then with more radical selections depending upon the results achieved, (2) deleting the target residue, or (3) inserting residues of the same or a different class adjacent to the located site, or combinations of options 1-3.

A useful method for identification of certain residues or regions of the PRO317 polypeptide that are preferred locations for mutagenesis is called "alanine scanning mutagenesis," as described by Cunningham and Wells, Science, 244: 1081-1085 (1989). Here, a residue or group of target residues are identified (*e.g.*, charged residues such as arg, asp, his, lys, and glu) and replaced by a neutral or negatively charged amino acid (most preferably alanine or polyalanine) to affect the interaction of the amino acids with the surrounding aqueous environment in or outside the cell. Those domains demonstrating functional sensitivity to the substitutions then are refined by introducing further or other variants at or for the sites of substitution. Thus, while the site for introducing an amino acid sequence variation is predetermined, the nature of the mutation *per se* need not be predetermined. For example, to optimize the performance of a mutation at a given site, alanine scanning or random mutagenesis is conducted at the target codon or region and the PRO317 variants produced are screened for the optimal combination of desired activity.

There are two principal variables in the construction of amino acid sequence variants: the location of the mutation site and the nature of the mutation. These are variants from the Figure 42 sequence, and may represent naturally occurring alleles (which will not require manipulation of the PRO317 DNA) or predetermined mutant forms made by mutating the DNA, either to arrive at an allele or a variant not found in nature. In general, the location and nature of the mutation chosen will depend upon the PRO317 characteristic to be modified.

Amino acid sequence deletions generally range from about 1 to 30 residues, more preferably about 1 to 10 residues, and typically are contiguous. Contiguous deletions ordinarily are made in even numbers of residues, but single or odd numbers of deletions are within the scope hereof. Deletions may be introduced into regions of low

homology among PRO317, EBAF-1, and other members of the TGF- superfamily which share the most sequence identity to the human PRO317 amino acid sequence to modify the activity of PRO317. Deletions from PRO317 in areas of substantial homology with one of the receptor binding sites of other members of the TGF- superfamily will be more likely to modify the biological activity of PRO317 more significantly. The number of consecutive deletions will be selected so as to preserve the tertiary structure of PRO317 in the affected domain, *e.g.*, beta-pleated sheet or alpha helix.

Amino acid sequence insertions include amino- and/or carboxyl-terminal fusions ranging in length from one residue to polypeptides containing a hundred or more residues, as well as intrasequence insertions of single or multiple amino acid residues. Intrasequence insertions (*i.e.*, insertions within the mature PRO317 sequence) may range generally from about 1 to 10 residues, more preferably 1 to 5, most preferably 1 to 3. Insertions are preferably made in even numbers of residues, but this is not required. Examples of terminal insertions include mature PRO317 with an N-terminal methionyl residue, an artifact of the direct production of mature PRO317 in recombinant cell culture, and fusion of a heterologous N-terminal signal sequence to the N-terminus of the mature PRO317 molecule to facilitate the secretion of mature PRO317 from recombinant hosts. Such signal sequences may be obtained from, and thus homologous to, the intended host cell species, but also may be from other members of the TGF- superfamily. Suitable sequences include STII or lpp for *E. coli*, alpha factor for yeast, and viral signals such as herpes gD or the native EBAF-1 sequence for mammalian cells.

Other insertional variants of the PRO317 molecule include the fusion to the N- or C-terminus of PRO317 of immunogenic polypeptides, *e.g.*, bacterial polypeptides such as beta-lactamase or an enzyme encoded by the *E. coli trp* locus, or yeast protein, and C-terminal fusions with proteins having a long half-life such as immunoglobulin constant regions (or other immunoglobulin regions), albumin, or ferritin, as described in WO 89/02922 published 6 April 1989.

A third group of variants are amino acid substitution variants. These variants have at least one amino acid residue in the PRO317 molecule removed and a different residue inserted in its place. The sites of greatest interest for substitutional mutagenesis include sites identified as the active site(s) of PRO317 and sites where the amino acids found in the known analo s are substantially different in terms of side-chain bulk, charge, or hydrophobicity, but where there is also a high degree of sequence identity at the selected site within various animal PRO317 species, or where the amino acids found in known members of the TGF- superfamily and novel PRO317 are substantially different in terms of side-chain bulk, charge, or hydrophobicity, but where there also is a high degree of sequence identity at the selected site within various animal analogues of such members (*e.g.*, among all the animal EBAF-1 molecules). This analysis will highlight residues that may be involved in the modulation of endometrial tissue or angiogenesis, and therefore, variations at these sites may affect such activities.

Other sites of interest are those in which particular residues of the PRO317 obtained from various species are identical among all animal species of PRO317 and other members of the TGF- superfamily, this degree of conservation suggesting importance in achieving biological activity common to these cytokines. These sites, especially those falling within a sequence of at least three other identically conserved sites, are substituted in a relatively conservative manner. Such conservative substitutions are shown in Table 1 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 1, or as further described below in reference to amino acid classes, are introduced and the products screened.

Table 1

Original Residue	Exemplary Substitutions	Preferred Substitutions
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
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
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
Pro (P)	ala	ala
Ser (S)	thr	thr
Thr (T)	ser	ser
Trp (W)	tyr; phe	tyr
Tyr (Y)	trp; phe; thr; ser	phe
Val (V)	ile; leu; met; phe; ala; norleucine	leu

Substantial modifications in function or immunological identity of the PRO317 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;

(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.

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.

In one embodiment of the invention, it is desirable to inactivate one or more protease cleavage sites that are present in the molecule. These sites are identified by inspection of the encoded amino acid sequence, in the case of

trypsin, *e.g.*, for an arginyl or lysinyl residue. When protease cleavage sites are identified, they are rendered inactive to proteolytic cleavage by substituting the targeted residue with another residue, preferably a basic residue such as glutamine or a hydrophilic residue such as serine; by deleting the residue; or by inserting a prolyl residue immediately after the residue.

In another embodiment, any methionyl residues other than the starting methionyl residue of the signal sequence, or any residue located within about three residues N- or C-terminal to each such methionyl residue, is substituted by another residue (preferably in accord with Table 1) or deleted. Alternatively, about 1-3 residues are inserted adjacent to such sites.

Any cysteine residues not involved in maintaining the proper conformation of PRO317 also may be substituted, generally with serine, to improve the oxidative stability of the molecule and prevent aberrant crosslinking.

Nucleic acid molecules encoding amino acid sequence variants of PRO317 are prepared by a variety of methods known in the art. These methods include, but are not limited to, isolation from a natural source (in the case of naturally occurring amino acid sequence variants) or preparation by oligonucleotide-mediated (or site-directed) mutagenesis, PCR mutagenesis, and cassette mutagenesis of an earlier prepared variant or a non-variant version of PRO317.

Oligonucleotide-mediated mutagenesis is a preferred method for preparing substitution, deletion, and insertion variants of PRO317 DNA. This technique is well known in the art as described by Adelman *et al.*, DNA, 2: 183 (1983). Briefly, PRO317 DNA is altered by hybridizing an oligonucleotide encoding the desired mutation to a DNA template, where the template is the single-stranded form of a plasmid or bacteriophage containing the unaltered or native DNA sequence of PRO317. After hybridization, a DNA polymerase is used to synthesize an entire second complementary strand of the template that will thus incorporate the oligonucleotide primer, and will code for the selected alteration in the PRO317 DNA.

Generally, oligonucleotides of at least 25 nucleotides in length are used. An optimal oligonucleotide will have 12 to 15 nucleotides that are completely complementary to the template on either side of the nucleotide(s) coding for the mutation. This ensures that the oligonucleotide will hybridize properly to the single-stranded DNA template molecule. The oligonucleotides are readily synthesized using techniques known in the art such as that described by Crea *et al.*, Proc. Natl. Acad. Sci. USA, 75: 5765 (1978).

The DNA template can be generated by those vectors that are either derived from bacteriophage M13 vectors (the commercially available M13mp18 and M13mp19 vectors are suitable), or those vectors that contain a single-stranded phage origin of replication as described by Viera *et al.* Meth. Enzymol., 153: 3 (1987). Thus, the DNA that is to be mutated may be inserted into one of these vectors to generate single-stranded template. Production of the single-stranded template is described in Sections 4.21-4.41 of Sambrook *et al.*, *supra*.

Alternatively, single-stranded DNA template may be generated by denaturing double-stranded plasmid (or other) DNA using standard techniques.

For alteration of the native DNA sequence (to generate amino acid sequence variants, for example), the oligonucleotide is hybridized to the single-stranded template under suitable hybridization conditions. A DNA polymerizing enzyme, usually the Klenow fragment of DNA polymerase I, is then added to synthesize the complementary strand of the template using the oligonucleotide as a primer for synthesis. A heteroduplex molecule

is thus formed such that one strand of DNA encodes the mutated form of PRO317, and the other strand (the original template) encodes the native, unaltered sequence of PRO317. This heteroduplex molecule is then transformed into a suitable host cell, usually a prokaryote such as *E. coli* JM101. After the cells are grown, they are plated onto agarose plates and screened using the oligonucleotide primer radiolabeled with ^{32}P to identify the bacterial colonies that contain the mutated DNA. The mutated region is then removed and placed in an appropriate vector for protein production, generally an expression vector of the type typically employed for transformation of an appropriate host.

The method described immediately above may be modified such that a homoduplex molecule is created wherein both strands of the plasmid contain the mutation(s). The modifications are as follows: The single-stranded oligonucleotide is annealed to the single-stranded template as described above. A mixture of three deoxyribonucleotides, deoxyriboadenosine (dATP), deoxyriboguanosine (dGTP), and deoxyribothymidine (dTTP), is combined with a modified thio-deoxyribocytosine called dCTP-(aS) (which can be obtained from the Amersham Corporation). This mixture is added to the template-oligonucleotide complex. Upon addition of DNA polymerase to this mixture, a strand of DNA identical to the template except for the mutated bases is generated. In addition, this new strand of DNA will contain dCTP-(aS) instead of dCTP, which serves to protect it from restriction endonuclease digestion.

After the template strand of the double-stranded heteroduplex is nicked with an appropriate restriction enzyme, the template strand can be digested with *Exo*III nuclease or another appropriate nuclease past the region that contains the site(s) to be mutagenized. The reaction is then stopped to leave a molecule that is only partially single-stranded. A complete double-stranded DNA homoduplex is then formed using DNA polymerase in the presence of all four deoxyribonucleotide triphosphates, ATP, and DNA ligase. This homoduplex molecule can then be transformed into a suitable host cell such as *E. coli* JM101, as described above.

DNA encoding PRO317 mutants with more than one amino acid to be substituted may be generated in one of several ways. If the amino acids are located close together in the polypeptide chain, they may be mutated simultaneously using one oligonucleotide that codes for all of the desired amino acid substitutions. If, however, the amino acids are located some distance from each other (separated by more than about ten amino acids), it is more difficult to generate a single oligonucleotide that encodes all of the desired changes. Instead, one of two alternative methods may be employed.

In the first method, a separate oligonucleotide is generated for each amino acid to be substituted. The oligonucleotides are then annealed to the single-stranded template DNA simultaneously, and the second strand of DNA that is synthesized from the template will encode all of the desired amino acid substitutions.

The alternative method involves two or more rounds of mutagenesis to produce the desired mutant. The first round is as described for the single mutants: wild-type DNA is used for the template, an oligonucleotide encoding the first desired amino acid substitution(s) is annealed to this template, and the heteroduplex DNA molecule is then generated. The second round of mutagenesis utilizes the mutated DNA produced in the first round of mutagenesis as the template. Thus, this template already contains one or more mutations. The oligonucleotide encoding the additional desired amino acid substitution(s) is then annealed to this template, and the resulting strand of DNA now encodes mutations from both the first and second rounds of mutagenesis. This resultant DNA can be used as a template in a third round of mutagenesis, and so on.

PCR mutagenesis is also suitable for making amino acid variants of PRO317. While the following discussion refers to DNA, it is understood that the technique also finds application with RNA. The PCR technique generally refers to the following procedure (see Erlich, PCR Technology, (Stockton Press, NY, 1989), the chapter by R. Higuchi, p. 61-70): When small amounts of template DNA are used as starting material in a PCR, primers that differ slightly in sequence from the corresponding region in a template DNA can be used to generate relatively large quantities of a specific DNA fragment that differs from the template sequence only at the positions where the primers differ from the template. For introduction of a mutation into a plasmid DNA, one of the primers is designed to overlap the position of the mutation and to contain the mutation; the sequence of the other primer must be identical to a stretch of sequence of the opposite strand of the plasmid, but this sequence can be located anywhere along the plasmid DNA. It is preferred, however, that the sequence of the second primer is located within 200 nucleotides from that of the first, such that in the end the entire amplified region of DNA bounded by the primers can be easily sequenced. PCR amplification using a primer pair like the one just described results in a population of DNA fragments that differ at the position of the mutation specified by the primer, and possibly at other positions, as template copying is somewhat error-prone.

Another method for preparing variants, cassette mutagenesis, is based on the technique described by Wells *et al.*, Gene, 34: 315 (1985). The starting material is the plasmid (or other vector) comprising the PRO317 DNA to be mutated. The codon(s) in the PRO317 DNA to be mutated are identified. There must be a unique restriction endonuclease site on each side of the identified mutation site(s). If no such restriction sites exist, they may be generated using the above-described oligonucleotide-mediated mutagenesis method to introduce them at appropriate locations in the PRO317 DNA. After the restriction sites have been introduced into the plasmid, the plasmid is cut at these sites to linearize it. A double-stranded oligonucleotide encoding the sequence of the DNA between the restriction sites but containing the desired mutation(s) is synthesized using standard procedures. The two strands are synthesized separately and then hybridized together using standard techniques. This double-stranded oligonucleotide is referred to as the cassette. This cassette is designed to have 3' and 5' ends that are compatible with the ends of the linearized plasmid, such that it can be directly ligated to the plasmid. This plasmid now contains the mutated PRO317 DNA sequence.

Covalent modifications of PRO317 are also included within the scope of this invention. One type of covalent modification includes reacting targeted amino acid residues of the PRO317 with an organic derivatizing agent that is capable of reacting with selected side chains or the N- or C- terminal residues of the PRO317. Derivatization with bifunctional agents is useful, for instance, for crosslinking PRO317 to a water-insoluble support matrix or surface for use in the method for purifying anti-PRO317 antibodies, and vice-versa. Commonly used crosslinking agents include, *e.g.*, 1,1-bis(diazoacetyl)-2-phenylethane, glutaraldehyde, N-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.

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 "-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 PRO317 polypeptide included within the scope of this 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 polypeptide, and/or adding one or more glycosylation sites that are not present in the native-sequence PRO polypeptide. The deduced amino acid sequence of PRO317 shown in Figure 42 (SEQ ID NO:114) has one predicted N-linked glycosylation site at residue 160.

Addition of glycosylation sites to the PRO317 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 serine or threonine residues to the native-sequence PRO317 (for O-linked glycosylation sites). The PRO317 amino acid sequence may optionally be altered through changes at the DNA level, particularly by mutating the DNA encoding the PRO317 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 PRO317 polypeptide is by chemical or enzymatic coupling of glycosides to the polypeptide. Such methods are described in the art, *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 PRO317 polypeptide may be accomplished chemically or enzymatically or by mutational substitution of codons encoding amino acid residues that serve as targets for glycosylation. Chemical deglycosylation techniques are known in the art and described, 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 PRO317 comprises linking the PRO317 polypeptide to one of a variety of nonproteinaceous polymers, *e.g.*, polyethylene glycol, 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 PRO317 of the present invention may also be modified in a way to form a chimeric molecule comprising PRO317 fused to another, heterologous polypeptide or amino acid sequence. In one embodiment, such a chimeric molecule comprises a fusion of the PRO317 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 PRO317. The presence of such epitope-tagged forms of the PRO317 can be detected using an antibody against the tag polypeptide. Also, provision of the epitope tag enables the PRO317 to be readily purified by affinity purification using an anti-tag antibody or another type of affinity matrix that binds to the epitope tag. In an alternative embodiment, the chimeric molecule may comprise a fusion of the PRO317 with an immunoglobulin or a particular region of an immunoglobulin. For a bivalent form of the chimeric molecule, such a fusion could be to the Fc region of an IgG molecule.

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.*, Bio/Technology, 6:1204-1210 (1988)); the KT3 epitope peptide (Martin *et al.*, Science, 255:192-194 (1992)); an "-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)).

53. Preparation of PRO Polypeptides

The description below relates primarily to production of PRO polypeptides by culturing cells transformed or transfected with a vector containing the desired PRO polypeptide nucleic acid. It is, of course, contemplated that alternative methods, which are well known in the art, may be employed to prepare the PRO polypeptide. For instance, the PRO polypeptide 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 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 desired PRO polypeptide may be chemically synthesized separately and combined using chemical or enzymatic methods to produce the full-length PRO polypeptide.

A. Isolation of DNA Encoding PRO Polypeptides

DNA encoding PRO polypeptides may be obtained from a cDNA library prepared from tissue believed to possess the desired PRO polypeptide mRNA and to express it at a detectable level. Accordingly, human PRO polypeptide DNA can be conveniently obtained from a cDNA library prepared from human tissue, such as described in the Examples. The PRO polypeptide-encoding gene may also be obtained from a genomic library or by oligonucleotide synthesis.

Libraries can be screened with probes (such as antibodies to the desired PRO polypeptide 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 the desired PRO polypeptide 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}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 through sequence alignment using computer software programs such as BLAST, ALIGN, DNAstar, and INHERIT which employ various algorithms to measure homology.

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.

B. Selection and Transformation of Host Cells

Host cells are transfected or transformed with expression or cloning vectors described herein for PRO polypeptide 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 transfection are known to the ordinarily skilled artisan, for example, $CaPO_4$ 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 or other cells that contain substantial cell-wall barriers. 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 transformations 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 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 cells include Enterobacteriaceae such as *Escherichia*, *e.g.*, *E. coli*, *Enterobacter*, *Erwinia*, *Klebsiella*, *Proteus*, *Salmonella*, *e.g.*, *Salmonella typhimurium*,

Serratia, *e.g.*, *Serratia marcescans*, and *Shigella*, as well 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*. 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). These examples are 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 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*^{r}; *E. coli* W3110 strain 37D6, which has the complete genotype *tonA ptr3 phoA E15 (argF-lac)169 degP ompT rbs7 ilvG kan*^{r}; *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 cloning or expression hosts for PRO polypeptide-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., 737 [1983]), *K. fragilis* (ATCC 12,424), *K. bulgaricus* (ATCC 16,045), *K. wickeramii* (ATCC 24,178), *K. waltii* (ATCC 56,500), *K. drosophilarum* (ATCC 36,906; Van 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 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 polypeptides 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.

C. Selection and Use of a Replicable Vector

The nucleic acid (*e.g.*, cDNA or genomic DNA) encoding a desired PRO polypeptide 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 polypeptide of interest 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 polypeptide 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* α-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 2μ 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*.

An example of suitable selectable markers for mammalian cells are those that enable the identification of cells competent to take up the PRO polypeptide 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). A suitable selection gene

for use in yeast is the *trp*1 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 *trp*1 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)].

Expression and cloning vectors usually contain a promoter operably linked to the PRO polypeptide 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 β-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)]. Promoters for use in bacterial systems also will contain a Shine-Dalgarno (S.D.) sequence operably linked to the DNA encoding the desired PRO polypeptide.

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.

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.

PRO polypeptide 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.

Transcription of a DNA encoding the desired PRO polypeptide 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, α-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 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 polypeptide 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, 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 polypeptides.

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

D. 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 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.

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 synthetic peptide based on the DNA sequences provided herein or against exogenous sequence fused to a PRO polypeptide DNA and encoding a specific antibody epitope.

E. Purification of Polypeptide

Forms of PRO polypeptides 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 enzymatic cleavage. Cells employed in expression of PRO polypeptides 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 polypeptides 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 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 polypeptide. 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 (1982). The purification step(s) selected will depend, for example, on the nature of the production process used and the particular PRO polypeptide produced.

54. Uses for PRO Polypeptides

Nucleotide sequences (or their complement) encoding the PRO polypeptides of the present invention have various applications in the art of molecular biology, including uses as hybridization probes, in chromosome and gene mapping and in the generation of anti-sense RNA and DNA. PRO polypeptide-encoding nucleic acid will also be useful for the preparation of PRO polypeptides by the recombinant techniques described herein.

The full-length native sequence PRO polypeptide-encoding nucleic acid or portions thereof, may be used as hybridization probes for a cDNA library to isolate the full-length PRO polypeptide gene or to isolate still other genes (for instance, those encoding naturally-occurring variants of the PRO polypeptide or PRO polypeptides from other species) which have a desired sequence identity to the PRO polypeptide nucleic acid sequences. Optionally, the length of the probes will be about 20 to about 50 bases. The hybridization probes may be derived from the nucleotide sequence of any of the DNA molecules disclosed herein or from genomic sequences including promoters, enhancer elements and introns of native sequence PRO polypeptide encoding DNA. By way of example, a screening method will comprise isolating the coding region of the PRO polypeptide gene using the known DNA sequence to synthesize a selected probe of about 40 bases. Hybridization probes may be labeled by a variety of labels, including radionucleotides such as ^{32}P or ^{35}S, or enzymatic labels such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems. Labeled probes having a sequence complementary to that of the specific PRO polypeptide gene of the present invention can be used to screen libraries of human cDNA, genomic DNA or mRNA to determine which members of such libraries the probe hybridizes to. Hybridization techniques are described in further detail in the Examples below.

The ESTs disclosed in the present application may similarly be employed as probes, using the methods disclosed herein.

The probes may also be employed in PCR techniques to generate a pool of sequences for identification of closely related PRO polypeptide sequences.

Nucleotide sequences encoding a PRO polypeptide can also be used to construct hybridization probes for mapping the gene which encodes that PRO polypeptide and for the genetic analysis of individuals with genetic disorders. The nucleotide sequences provided herein may be mapped to a chromosome and specific regions of a chromosome using known techniques, such as *in situ* hybridization, linkage analysis against known chromosomal markers, and hybridization screening with libraries.

The PRO polypeptide can be used in assays to identify its ligands. Similarly, inhibitors of the receptor/ligand binding interaction can be identified. Proteins involved in such binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction. Screening assays can be designed to find lead compounds that mimic the biological activity of a native PRO polypeptide or a ligand for the PRO polypeptide. Such screening assays will include assays amenable to high-throughput screening of chemical libraries, making them particularly suitable for identifying small molecule drug candidates. Small molecules contemplated include synthetic organic or inorganic compounds. 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.

Nucleic acids which encode a PRO polypeptide or its modified forms can also be used to generate either transgenic animals or "knock out" animals which, in turn, are useful in the development and screening of therapeutically useful reagents. A transgenic animal (e.g., a mouse or rat) is an animal having cells that contain a transgene, which transgene was introduced into the animal or an ancestor of the animal at a prenatal, e.g., an embryonic stage. A transgene is a DNA which is integrated into the genome of a cell from which a transgenic animal develops. In one embodiment, cDNA encoding a PRO polypeptide of interest can be used to clone genomic DNA encoding the PRO polypeptide in accordance with established techniques and the genomic sequences used to generate transgenic animals that contain cells which express DNA encoding the PRO polypeptide. Methods for generating transgenic animals, particularly animals such as mice or rats, have become conventional in the art and are described, for example, in U.S. Patent Nos. 4,736,866 and 4,870,009. Typically, particular cells would be targeted for PRO polypeptide transgene incorporation with tissue-specific enhancers. Transgenic animals that include a copy of a transgene encoding a PRO polypeptide introduced into the germ line of the animal at an embryonic stage can be used to examine the effect of increased expression of DNA encoding the PRO polypeptide. Such animals can be used as tester animals for reagents thought to confer protection from, for example, pathological conditions associated with its overexpression. In accordance with this facet of the invention, an animal is treated with the reagent and a reduced incidence of the pathological condition, compared to untreated animals bearing the transgene, would indicate a potential therapeutic intervention for the pathological condition.

Alternatively, non-human homologues of PRO polypeptides can be used to construct a PRO polypeptide "knock out" animal which has a defective or altered gene encoding the PRO polypeptide of interest as a result of homologous recombination between the endogenous gene encoding the PRO polypeptide and altered genomic DNA encoding the PRO polypeptide introduced into an embryonic cell of the animal. For example, cDNA encoding a PRO polypeptide can be used to clone genomic DNA encoding the PRO polypeptide in accordance with established techniques. A portion of the genomic DNA encoding a PRO polypeptide can be deleted or replaced with another gene, such as a 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 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 "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 PRO polypeptide.

With regard to the PRO211 and PRO217 polypeptide, therapeutic indications include disorders associated with the preservation and maintenance of gastrointestinal mucosa and the repair of acute and chronic mucosal lesions

(e.g., enterocolitis, Zollinger-Ellison syndrome, gastrointestinal ulceration and congenital microvillus atrophy), skin diseases associated with abnormal keratinocyte differentiation (e.g., psoriasis, epithelial cancers such as lung squamous cell carcinoma, epidermoid carcinoma of the vulva and gliomas.

Since the PRO232 polypeptide and nucleic acid encoding it possess sequence homology to a cell surface stem cell antigen and its encoding nucleic acid, probes based upon the PRO232 nucleotide sequence may be employed to identify other novel stem cell surface antigen proteins. Soluble forms of the PRO232 polypeptide may be employed as antagonists of membrane bound PRO232 activity both *in vitro* and *in vivo*. PRO232 polypeptides may be employed in screening assays designed to identify agonists or antagonists of the native PRO232 polypeptide, wherein such assays may take the form of any conventional cell-type or biochemical binding assay. Moreover, the PRO232 polypeptide may serve as a molecular marker for the tissues in which the polypeptide is specifically expressed.

With regard to the PRO187 polypeptides disclosed herein, FGF-8 has been implicated in cellular differentiation and embryogenesis, including the patterning which appears during limb formation. FGF-8 and the PRO187 molecules of the invention therefore are likely to have potent effects on cell growth and development. Diseases which relate to cellular growth and differentiation are therefore suitable targets for therapeutics based on functionality similar to FGF-8. For example, diseases related to growth or survival of nerve cells including Parkinson's disease, Alzheimer's disease, ALS, neuropathies. Additionally, disease related to uncontrolled cell growth, e.g., cancer, would also be expected therapeutic targets.

With regard to the PRO265 polypeptides disclosed herein, other methods for use with PRO265 are described in U.S. Patent 5,654,270 to Ruoslahti et al. In particular, PRO265 can be used in comparison with the fibromodulin disclosed therein to compare its effects on reducing dermal scarring and other properties of the fibromodulin described therein including where it is located and with what it binds and does not.

The PRO219 polypeptides of the present invention which play a regulatory role in the blood coagulation cascade may be employed *in vivo* for therapeutic purposes as well as for *in vitro* purposes. Those of ordinary skill in the art will well know how to employ PRO219 polypeptides for such uses.

The PRO246 polypeptides of the present invention which serve as cell surface receptors for one or more viruses will find other uses. For example, extracellular domains derived from these PRO246 polypeptides may be employed therapeutically *in vivo* for lessening the effects of viral infection. Those PRO246 polypeptides which serves as tumor specific antigens may be exploited as therapeutic targets for anti-tumor drugs, and the like. Those of ordinary skill in the art will well know how to employ PRO246 polypeptides for such uses.

Assays in which connective growth factor and other growth factors are usually used should be performed with PRO261. An assay to determine whether TGF beta induces PRO261, indicating a role in cancer is performed as known in the art. Wound repair and tissue growth assays are also performed with PRO261. The results are applied accordingly.

PRO228 polypeptides should be used in assays in which EMR1, CD97 and latrophilin would be used in to determine their relative activities. The results can be applied accordingly. For example, a competitive binding assay with PRO228 and CD97 can be performed with the ligand for CD97, CD55.

Native PRO533 is a 216 amino acid polypeptide of which residues 1-22 are the signal sequence. Residues 3 to 216 have a Blast score of 509, corresponding to 53% homology to fibroblast growth factor. At the nucleotide

level, DNA47412, the EST from which PCR oligos were generated to isolate the full length DNA49435-1219, has been observed to map to 11p15. Sequence homology to the 11p15 locus would indicate that PRO533 may have utility in the treatment of Usher Syndrome or Atrophia areata.

As mentioned previously, fibroblast growth factors can act upon cells in both a mitogenic and non-mitogenic manner. These factors are mitogenic for a wide variety of normal diploid mesoderm-derived and neural crest-derived cells, inducing granulosa cells, adrenal cortical cells, chrondrocytes, myoblasts, corneal and vascular endothelial cells (bovine or human), vascular smooth muscle cells, lens, retina and prostatic epithelial cells, oligodendrocytes, astrocytes, chrondocytes, myoblasts and osteoblasts.

Non-mitogenic actions of fibroblast growth factors include promotion of cell migration into a wound area (chemotaxis), initiation of new blood vessel formulation (angiogenesis), modulation of nerve regeneration and survival (neurotrophism), modulation of endocrine functions, and stimulation or suppression of specific cellular protein expression, extracellular matrix production and cell survival. Baird, A. & Bohlen, P., *Handbook of Exp. Phrmacol.* 95(1): 369-418 (1990). These properties provide a basis for using fibroblast growth factors in therapeutic approaches to accelerate wound healing, nerve repair, collateral blood vessel formation, and the like. For example, fibroblast growth factors, have been suggested to minimize myocardium damage in heart disease and surgery (U.S.P. 4,378,437).

Since the PRO245 polypeptide and nucleic acid encoding it possess sequence homology to a transmembrane protein tyrosine kinase protein and its encoding nucleic acid, probes based upon the PRO245 nucleotide sequence may be employed to identify other novel transmembrane tyrosine kinase proteins. Soluble forms of the PRO245 polypeptide may be employed as antagonists of membrane bound PRO245 activity both *in vitro* and *in vivo*. PRO245 polypeptides may be employed in screening assays designed to identify agonists or antagonists of the native PRO245 polypeptide, wherein such assays may take the form of any conventional cell-type or biochemical binding assay. Moreover, the PRO245 polypeptide may serve as a molecular marker for the tissues in which the polypeptide is specifically expressed.

PRO220, PRO221 and PRO227 all have leucine rich repeats. Additionally, PRO220 and PRO221 have homology to SLIT and leucine rich repeat protein. Therefore, these proteins are useful in assays described in the literature, supra, wherein the SLIT and leucine rich repeat protein are used. Regarding the SLIT protein, PRO227 can be used in an assay to determine the affect of PRO227 on neurodegenerative disease. Additionally, PRO227 has homology to human glycoprotein V. In the case of PRO227, this polypeptide is used in an assay to determine its affect on bleeding, clotting, tissue repair and scarring.

The PRO266 polypeptide can be used in assays to determine if it has a role in neurodegenerative diseases or their reversal.

PRO269 polypeptides and portions thereof which effect the activity of thrombin may also be useful for *in vivo* therapeutic purposes, as well as for various *in vitro* applications. In addition, PRO269 polypeptides and portions thereof may have therapeutic use as an antithrombotic agent with reduced risk for hemorrhage as compared with heparin. Peptides having homology to thrombomodulin are particularly desirable.

PRO287 polypeptides and portions thereof which effect the activity of bone morphogenic protein "BMP1"/procollagen C-proteinase (PCP) may also be useful for *in vivo* therapeutic purposes, as well as for various

in vitro applications. In addition, PRO287 polypeptides and portions thereof may have therapeutic applications in wound healing and tissue repair. Peptides having homology to procollagen C-proteinase enhancer protein and its precursor may also be used to induce bone and/or cartilage formation and are therefore of particular interest to the scientific and medical communities.

Therapeutic indications for PRO214 polypeptides include disorders associated with the preservation and maintenance of gastrointestinal mucosa and the repair of acute and chronic mucosal lesions (e.g., enterocolitis, Zollinger-Ellison syndrome, gastrointestinal ulceration and congenital microvillus atrophy), skin diseases associated with abnormal keratinocyte differentiation (e.g., psoriasis, epithelial cancers such as lung squamous cell carcinoma, epidermoid carcinoma of the vulva and gliomas.

Studies on the generation and analysis of mice deficient in members of the TGF- superfamily are reported in Matzuk, Trends in Endocrinol. and Metabol., 6: 120-127 (1995).

The PRO317 polypeptide, as well as PRO317-specific antibodies, inhibitors, agonists, receptors, or their analogs, herein are useful in treating PRO317-associated disorders. Hence, for example, they may be employed in modulating endometrial bleeding angiogenesis, and may also have an effect on kidney tissue. Endometrial bleeding can occur in gynecological diseases such as endometrial cancer as abnormal bleeding. Thus, the compositions herein may find use in diagnosing and treating abnormal bleeding conditions in the endometrium, as by reducing or eliminating the need for a hysterectomy. The molecules herein may also find use in angiogenesis applications such as anti-tumor indications for which the antibody against vascular endothelial growth factor is used, or, conversely, ischemic indications for which vascular endothelial growth factor is employed.

Bioactive compositions comprising PRO317 or agonists or antagonists thereof may be administered in a suitable therapeutic dose determined by any of several methodologies including clinical studies on mammalian species to determine maximal tolerable dose and on normal human subjects to determine safe dose. Additionally, the bioactive agent may be complexed with a variety of well established compounds or compositions which enhance stability or pharmacological properties such as half-life. It is contemplated that the therapeutic, bioactive composition may be delivered by intravenous infusion into the bloodstream or any other effective means which could be used for treating problems of the kidney, uterus, endometrium, blood vessels, or related tissue, *e.g.*, in the heart or genital tract.

Dosages and administration of PRO317, PRO317 agonist, or PRO317 antagonist in a pharmaceutical composition may be determined by one of ordinary skill in the art of clinical pharmacology or pharmacokinetics. See, for example, Mordenti and Rescigno, Pharmaceutical Research, 9:17-25 (1992); Morenti *et al.*, Pharmaceutical Research, 8:1351-1359 (1991); and Mordenti and Chappell, "The use of interspecies scaling in toxicokinetics" in Toxicokinetics and New Drug Development, Yacobi *et al.* (eds) (Pergamon Press: NY, 1989), pp. 42-96. An effective amount of PRO317, PRO317 agonist, or PRO317 antagonist to be employed therapeutically will depend, for example, upon the therapeutic objectives, the route of administration, and the condition of the mammal. Accordingly, it will be necessary for the therapist to titer the dosage and modify the route of administration as required to obtain the optimal therapeutic effect. A typical daily dosage might range from about 10 ng/kg to up to 100 mg/kg of the mammal's body weight or more per day, preferably about 1 μg/kg/day to 10 mg/kg/day. Typically, the clinician will administer PRO317, PRO317 agonist, or PRO317 antagonist, until a dosage is reached that achieves

the desired effect for treatment of the above mentioned disorders.

PRO317 or an PRO317 agonist or PRO317 antagonist may be administered alone or in combination with another to achieve the desired pharmacological effect. PRO317 itself, or agonists or antagonists of PRO317 can provide different effects when administered therapeutically. Such compounds for treatment will be formulated in a nontoxic, inert, pharmaceutically acceptable aqueous carrier medium preferably at a pH of about 5 to 8, more preferably 6 to 8, although the pH may vary according to the characteristics of the PRO317, agonist, or antagonist being formulated and the condition to be treated. Characteristics of the treatment compounds include solubility of the molecule, half-life, and antigenicity/immunogenicity; these and other characteristics may aid in defining an effective carrier.

PRO317 or PRO317 agonists or PRO317 antagonists may be delivered by known routes of administration including but not limited to topical creams and gels; transmucosal spray and aerosol, transdermal patch and bandage; injectable, intravenous, and lavage formulations; and orally administered liquids and pills, particularly formulated to resist stomach acid and enzymes. The particular formulation, exact dosage, and route of administration will be determined by the attending physician and will vary according to each specific situation.

Such determinations of administration are made by considering multiple variables such as the condition to be treated, the type of mammal to be treated, the compound to be administered, and the pharmacokinetic profile of the particular treatment compound. Additional factors which may be taken into account include disease state (*e.g.* severity) of the patient, age, weight, gender, diet, time of administration, drug combination, reaction sensitivities, and tolerance/response to therapy. Long-acting treatment compound formulations (such as liposomally encapsulated PRO317 or PEGylated PRO317 or PRO317 polymeric microspheres, such as polylactic acid-based microspheres) might be administered every 3 to 4 days, every week, or once every two weeks depending on half-life and clearance rate of the particular treatment compound.

Normal dosage amounts may vary from about 10 ng/kg to up to 100 mg/kg of mammal body weight or more per day, preferably about 1 μg/kg/day to 10 mg/kg/day, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature; see, for example, U.S. Pat. Nos. 4,657,760; 5,206,344; or 5,225,212. It is anticipated that different formulations will be effective for different treatment compounds and different disorders, that administration targeting the uterus, for example, may necessitate delivery in a manner different from that to another organ or tissue, such as cardiac tissue.

Where sustained-release administration of PRO317 is desired in a formulation with release characteristics suitable for the treatment of any disease or disorder requiring administration of PRO317, microencapsulation of PRO317 is contemplated. Microencapsulation of recombinant proteins for sustained release has been successfully performed with human growth hormone (rhGH), interferon- (rhIFN-), interleukin-2, and MN rgp120. Johnson *et al.*, Nat. Med., 2: 795-799 (1996); Yasuda, Biomed. Ther., 27: 1221-1223 (1993); Hora *et al.*, Bio/Technology, 8: 755-758 (1990); Cleland, "Design and Production of Single Immunization Vaccines Using Polylactide Polyglycolide Microsphere Systems," in Vaccine Design: The Subunit and Adjuvant Approach , Powell and Newman, eds, (Plenum Press: New York, 1995), pp. 439-462; WO 97/03692, WO 96/40072, WO 96/07399; and U.S Pat. No. 5,654,010.

The sustained-release formulations of these proteins were developed using poly-lactic-coglycolic acid (PLGA) polymer due to its biocompatibility and wide range of biodegradable properties. The degradation products

of PLGA, lactic and glycolic acids, can be cleared quickly within the human body. Moreover, the degradability of this polymer can be adjusted from months to years depending on its molecular weight and composition. Lewis, "Controlled release of bioactive agents from lactide/glycolide polymer," in: M. Chasin and R. Langer (Eds.), Biodegradable Polymers as Drug Delivery Systems (Marcel Dekker: New York, 1990), pp. 1-41.

For example, for a formulation that can provide a dosing of approximately 80 g/kg/day in mammals with a maximum body weight of 85 kg, the largest dosing would be approximately 6.8 mg PRO317 per day. In order to achieve this dosing level, a sustained- release formulation which contains a maximum possible protein loading (15-20% w/w PRO317) with the lowest possible initial burst (<20%) is necessary. A continuous (zero-order) release of PRO317 from microparticles for 1-2 weeks is also desirable. In addition, the encapsulated protein to be released should maintain its integrity and stability over the desired release period.

It is contemplated that conditions or diseases of the uterus, endometrial tissue, or other genital tissues or cardiac tissues may precipitate damage that is treatable with PRO317 or PRO317 agonist where PRO317 expression is reduced in the diseased state; or with antibodies to PRO317 or other PRO317 antagonists where the expression of PRO317 is increased in the diseased state. These conditions or diseases may be specifically diagnosed by the probing tests discussed above for physiologic and pathologic problems which affect the function of the organ.

The PRO317, PRO317 agonist, or PRO317 antagonist may be administered to a mammal with another biologically active agent, either separately or in the same formulation to treat a common indication for which they are appropriate. For example, it is contemplated that PRO317 can be administered together with EBAF-1 for those indications on which they demonstrate the same qualitative biological effects. Alternatively, where they have opposite effects, EBAF-1 may be administered together with an antagonist to PRO317, such as an anti-PRO317 antibody. Further, PRO317 may be administered together with VEGF for coronary ischemia where such indication is warranted, or with an anti-VEGF for cancer as warranted, or, conversely, an antagonist to PRO317 may be administered with VEGF for coronary ischemia or with anti-VEGF to treat cancer as warranted. These administrations would be in effective amounts for treating such disorders.

Native PRO301 (SEQ ID NO:119) has a Blast score of 246 and 30% homology at residues 24 to 282 of Figure 44 with A33_HUMAN, an A33 antigen precursor. A33 antigen precursor, as explained in the Background is a tumor-specific antigen, and as such, is a recognized marker and therapeutic target for the diagnosis and treatment of colon cancer. The expression of tumor-specific antigens is often associated with the progression of neoplastic tissue disorders. Native PRO301 (SEQ ID NO:119) and A33_HUMAN also show a Blast score of 245 and 30% homology at residues 21 to 282 of Fig. 44 with A33_HUMAN, the variation dependent upon how spaces are inserted into the compared sequences. Native PRO301 (SEQ ID NO:119) also has a Blast score of 165 and 29% homology at residues 60 to 255 of Fig. 44 with HS46KDA_1, a human coxsackie and adenovirus receptor protein, also known as cell surface protein HCAR. This region of PRO301 also shows a similar Blast score and homology with HSU90716_1. Expression of such proteins is usually associated with viral infection and therapeutics for the prevention of such infection may be accordingly conceived. As mentioned in the Background, the expression of viral receptors is often associated with neoplastic tumors.

Therapeutic uses for the PRO234 polypeptides of the invention includes treatments associated with leukocyte homing or the interaction between leukocytes and the endothelium during an inflammatory response. Examples

include asthma, rheumatoid arthritis, psoriasis and multiple sclerosis.

Since the PRO231 polypeptide and nucleic acid encoding it possess sequence homology to a putative acid phosphatase and its encoding nucleic acid, probes based upon the PRO231 nucleotide sequence may be employed to identify other novel phosphatase proteins. Soluble forms of the PRO231 polypeptide may be employed as antagonists of membrane bound PRO231 activity both *in vitro* and *in vivo*. PRO231 polypeptides may be employed in screening assays designed to identify agonists or antagonists of the native PRO231 polypeptide, wherein such assays may take the form of any conventional cell-type or biochemical binding assay. Moreover, the PRO231 polypeptide may serve as a molecular marker for the tissues in which the polypeptide is specifically expressed.

PRO229 polypeptides can be fused with peptides of interest to determine whether the fusion peptide has an increased half-life over the peptide of interest. The PRO229 polypeptides can be used accordingly to increase the half-life of polypeptides of interest. Portions of PRO229 which cause the increase in half-life are an embodiment of the invention herein.

PRO238 can be used in assays which measure its ability to reduce substrates, including oxygen and Aceyl-CoA, and particularly, measure PRO238's ability to produce oxygen free radicals. This is done by using assays which have been previously described. PRO238 can further be used to assay for candidates which block, reduce or reverse its reducing abilities. This is done by performing side by side assays where candidates are added in one assay having PRO238 and a substrate to reduce, and not added in another assay, being the same but for the lack of the presence of the candidate.

PRO233 polypeptides and portions thereof which have homology to reductase may also be useful for *in vivo* therapeutic purposes, as well as for various other applications. The identification of novel reductase proteins and related molecules may be relevant to a number of human disorders such as inflammatory disease, organ failure, atherosclerosis, cardiac injury, infertility, birth defects, premature aging, AIDS, cancer, diabetic complications and mutations in general. Given that oxygen free radicals and antioxidants appear to play important roles in a number of disease processes, the identification of new reductase proteins and reductase-like molecules is of special importance in that such proteins may serve as potential therapeutics for a variety of different human disorders. Such polypeptides may also play important roles in biotechnological and medical research, as well as various industrial applications. As a result, there is particular scientific and medical interest in new molecules, such as PRO233.

The PRO223 polypeptides of the present invention which exhibit serine carboxypeptidease activity may be employed *in vivo* for therapeutic purposes as well as for *in vitro* purposes. Those of ordinary skill in the art will well know how to employ PRO223 polypeptides for such uses.

PRO235 polypeptides and portions thereof which may be involved in cell adhesion are also useful for *in vivo* therapeutic purposes, as well as for various *in vitro* applications. In addition, PRO235 polypeptides and portions thereof may have therapeutic applications in disease states which involve cell adhesion. Given the physiological importance of cell adhesion mechanisms *in vivo*, efforts are currently being under taken to identify new, native proteins which are involved in cell adhesion. Therefore, peptides having homology to plexin are of particular interest to the scientific and medical communities.

Because the PRO236 and PRO262 polypeptides disclosed herein are homologous to various known β-galactosidase proteins, the PRO236 and PRO262 polypeptides disclosed herein will find use in conjugates of

monoclonal antibodies and the polypeptide for specific killing of tumor cells by generation of active drug from a galactosylated prodrug (e.g., the generation of 5-fluorouridine from the prodrug β-D-galactosyl-5-fluorouridine). The PRO236 and PRO262 polypeptides disclosed herein may also find various uses both *in vivo* and *in vitro*, wherein those uses will be similar or identical to uses for which β-galactosidase proteins are now employed. Those of ordinary skill in the art will well know how to employ PRO236 and PRO262 polypeptides for such uses.

PRO239 polypeptides and portions thereof which have homology to densin may also be useful for *in vivo* therapeutic purposes, as well as for various *in vitro* applications. In addition, PRO239 polypeptides and portions thereof may have therapeutic applications in disease states which involve synaptic mechanisms, regeneration or cell adhesion. Given the physiological importance of synaptic processes, regeneration and cell adhesion mechanisms *in vivo*, efforts are currently being under taken to identify new, native proteins which are involved in synaptic machinery and cell adhesion. Therefore, peptides having homology to densin are of particular interest to the scientific and medical communities.

The PRO260 polypeptides described herein can be used in assays to determine their relation to fucosidase. In particular, the PRO260 polypeptides can be used in assays in determining their ability to remove fucose or other sugar residues from proteoglycans. The PRO260 polypeptides can be assayed to determine if they have any functional or locational similarities as fucosidase. The PRO260 polypeptides can then be used to regulate the systems in which they are integral.

PRO263 can be used in assays wherein CD44 antigen is generally used to determine PRO263 activity relative to that of CD44. The results can be used accordingly.

PRO270 polypeptides and portions thereof which effect reduction-oxidation (redox) state may also be useful for *in vivo* therapeutic purposes, as well as for various *in vitro* applications. More specifically, PRO270 polypeptides may affect the expression of a large variety of genes thought to be involved in the pathogenesis of AIDS, cancer, atherosclerosis, diabetic complications and in pathological conditions involving oxidative stress such as stroke and inflammation. In addition, PRO270 polypeptides and portions thereof may affect the expression of a genes which have a role in apoptosis. Therefore, peptides having homology to thioredoxin are particularly desirable to the scientific and medical communities.

PRO272 polypeptides and portions thereof which possess the ability to bind calcium may also have numerous *in vivo* therapeutic uses, as well as various *in vitro* applications. Therefore, peptides having homology to reticulocalbin are particularly desirable. Those with ordinary skill in the art will know how to employ PRO272 polypeptides and portions thereof for such purposes.

PRO294 polypeptides and portions thereof which have homology to collagen may also be useful for *in vivo* therapeutic purposes, as well as for various other applications. The identification of novel collagens and collage-like molecules may have relevance to a number of human disorders. Thus, the identification of new collagens and collage-like molecules is of special importance in that such proteins may serve as potential therapeutics for a variety of different human disorders. Such polypeptides may also play important roles in biotechnological and medical research as well as various industrial applications. Given the large number of uses for collagen, there is substantial interest in polypeptides with homology to the collagen molecule.

PRO295 polypeptides and portions thereof which have homology to integrin may also be useful for *in vivo* therapeutic purposes, as well as for various other applications. The identification of novel integrins and integrin-like molecules may have relevance to a number of human disorders such as modulating the binding or activity of cells of the immune system. Thus, the identification of new integrins and integrin-like molecules is of special importance in that such proteins may serve as potential therapeutics for a variety of different human disorders. Such polypeptides may also play important roles in biotechnological and medical research as well as various industrial applications. As a result, there is particular scientific and medical interest in new molecules, such as PRO295.

As the PRO293 polypeptide is clearly a leucine rich repeat polypeptide homologue, the peptide can be used in all applications that the known NLRR-1 and NLRR-2 polypeptides are used. The activity can be compared between these peptides and thus applied accordingly.

The PRO247 polypeptides described herein can be used in assays in which densin is used to determine the activity of PRO247 relative to densin or these other proteins. The results can be used accordingly in diagnostics and/or therapeutic applications with PRO247.

PRO302, PRO303, PRO304, PRO307 and PRO343 polypeptides of the present invention which possess protease activity may be employed both *in vivo* for therapeutic purposes and *in vitro*. Those of ordinary skill in the art will well know how to employ the PRO302, PRO303, PRO304, PRO307 and PRO343 polypeptides of the present invention for such purposes.

PRO328 polypeptides and portions thereof which have homology to GLIP and CRISP may also be useful for *in vivo* therapeutic purposes, as well as for various other applications. The identification of novel GLIP and CRISP-like molecules may have relevance to a number of human disorders which involve transcriptional regulation or are over expressed in human tumors. Thus, the identification of new GLIP and CRISP-like molecules is of special importance in that such proteins may serve as potential therapeutics for a variety of different human disorders. Such polypeptides may also play important roles in biotechnological and medical research as well as in various industrial applications. As a result, there is particular scientific and medical interest in new molecules, such as PRO328.

Uses for PRO335, PRO331 or PRO326 including uses in competitive assays with LIG-1, ALS and decorin to determine their relative activities. The results can be used accordingly. PRO335, PRO331 or PRO326 can also be used in assays where LIG-1 would be used to determine if the same effects are incurred.

PRO332 contains GAG repeat (GKEK) at amino acid positions 625-628 in Fig. 108 (SEQ ID NO:310). Slippage in such repeats can be associated with human disease. Accordingly, PRO332 can use useful for the treatment of such disease conditions by gene therapy, i.e. by introduction of a gene containing the correct GKEK sequence motif.

Other uses of PRO334 include use in assays in which fibrillin or fibulin would be used to determine the relative activity of PRO334 to fibrillin or fibulin. In particular, PRO334 can be used in assays which require the mechanisms imparted by epidermal growth factor repeats.

Native PRO346 (SEQ ID NO:320) has a Blast score of 230, corresponding to 27% homology between amino acid residues 21 to 343 with residues 35 to 1040 CGM6_HUMAN, a carcinoembryonic antigen cgm6 precursor. This homology region includes nearly all but 2 N-terminal extracellular domain residues, including an immunoglobulin superfamily homology at residues 148 to 339 of PRO346 in addition to several transmembrane

residues (340-343). Carcinoembryonic antigen precursor, as explained in the Background is a tumor-specific antigen, and as such, is a recognized marker and therapeutic target for the diagnosis and treatment of colon cancer. The expression of tumor-specific antigens is often associated with the progression of neoplastic tissue disorders. Native PRO346 (SEQ ID NO:320) and P_W06874, a human carcinoembryonic antigen CEA-d have a Blast score of 224 and homology of 28% between residues 2 to 343 and 67 to 342, respectively. This homology includes the entire extracellular domain residues of native PRO346, minus the initiator methionine (residues 2 to 18) as well as several transmembrane residues (340-343).

PRO268 polypeptides which have protein disulfide isomerase activity will be useful for many applications where protein disulfide isomerase activity is desirable including, for example, for use in promoting proper disulfide bond formation in recombinantly produced proteins so as to increase the yield of correctly folded protein. Those of ordinary skill in the art will readily know how to employ such PRO268 polypeptides for such purposes.

PRO330 polypeptides of the present invention which possess biological activity related to that of the prolyl 4-hydroxylase alpha subunit protein may be employed both *in vivo* for therapeutic purposes and *in vitro*. Those of ordinary skill in the art will well know how to employ the PRO330 polypeptides of the present invention for such purposes.

55. Anti-PRO Polypeptide Antibodies

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

A. Polyclonal Antibodies

The anti-PRO polypeptide 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 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 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.

B. Monoclonal Antibodies

The anti-PRO polypeptide antibodies may, alternatively, be monoclonal antibodies. Monoclonal antibodies may 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 of interest or a fusion protein thereof. 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 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 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, Rockville, Maryland. 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].

The culture medium in which the hybridoma cells are cultured can then be assayed for the presence of monoclonal antibodies directed against the PRO polypeptide of interest. 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).

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.

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.

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 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 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.

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.

C. Humanized Antibodies

The anti-PRO polypeptide 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', $F(ab')_2$ 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)].

D. 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 PRO polypeptide, 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*, 305:537-539 (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).

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 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).

E. 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 by forming a thioether bond. Examples of suitable reagents for this purpose include iminothiolate and methyl-4-mercaptobutyrimidate and those disclosed, for example, in U.S. Patent No. 4,676,980.

56. Uses for Anti-Pro Polypeptide Antibodies

The anti-PRO polypeptide antibodies of the invention have various utilities. For example, anti-PRO polypeptide antibodies may be used in diagnostic assays for a PRO polypeptide, *e.g.*, detecting its expression in specific cells, tissues, or serum. Various diagnostic assay techniques known in the art may be used, such as competitive binding assays, direct or indirect sandwich assays and immunoprecipitation assays conducted in either heterogeneous or homogeneous phases [Zola, Monoclonal Antibodies: A Manual of Techniques, CRC Press, Inc. (1987) pp. 147-158]. The antibodies used in the diagnostic assays can be labeled with a detectable moiety. The detectable moiety should be capable of producing, either directly or indirectly, a detectable signal. For example, the detectable moiety may be a radioisotope, such as ^{3}H, ^{14}C, ^{32}P, ^{35}S, or ^{125}I, a fluorescent or chemiluminescent compound, such as fluorescein isothiocyanate, rhodamine, or luciferin, or an enzyme, such as alkaline phosphatase, beta-galactosidase or horseradish peroxidase. Any method known in the art for conjugating the antibody to the detectable moiety may be employed, including those methods described by Hunter *et al., Nature*, 144:945 (1962); David *et al., Biochemistry*, 13:1014 (1974); Pain *et al., J. Immunol. Meth.*, 40:219 (1981); and Nygren, *J. Histochem. and Cytochem.*, 30:407 (1982).

Anti-PRO polypeptide antibodies also are useful for the affinity purification of PRO polypeptide from recombinant cell culture or natural sources. In this process, the antibodies against the PRO polypeptide are immobilized on a suitable support, such a Sephadex resin or filter paper, using methods well known in the art. The immobilized antibody then is contacted with a sample containing the PRO polypeptide to be purified, and thereafter the support is washed with a suitable solvent that will remove substantially all the material in the sample except the PRO polypeptide, which is bound to the immobilized antibody. Finally, the support is washed with another suitable solvent that will release the PRO polypeptide from the antibody.

With regard to PRO211 and PRO217, therapeutic indications include disorders associated with the preservation and maintenance of gastrointestinal mucosa and the repair of acute and chronic mucosal lesions (e.g., enterocolitis, Zollinger-Ellison syndrome, gastrointestinal ulceration and congenital microvillus atrophy), skin diseases associated with abnormal keratinocyte differentiation (e.g., psoriasis, epithelial cancers such as lung squamous cell carcinoma, epidermoid carcinoma of the vulva and gliomas.

With regard to anti-PRO187 antibodies, FGF-8 has been implicated in cellular differentiation and embryogenesis, including the patterning which appears during limb formation. FGF-8 and the PRO187 molecules of the invention therefore are likely to have potent effects on cell growth and development. Diseases which relate to cellular growth and differentiation are therefore suitable targets for therapeutics based on functionality similar to FGF-8. For example, diseases related to growth or survival of nerve cells including Parkinson's disease, Alzheimer's disease, ALS, neuropathies. Additionally, disease related to uncontrolled cell growth, e.g., cancer, would also be expected therapeutic targets.

Native PRO533 is a 216 amino acid polypeptide of which residues 1-22 are the signal sequence. Residues 3 to 216 have a Blast score of 509, corresponding to 53% homology to fibroblast growth factor. At the nucleotide level, DNA47412, the EST from which PCR oligos were generated to isolate the full length DNA49435-1219, has been observed to map to 11p15. Sequence homology to the 11p15 locus would indicate that PRO533 may have utility in the treatment of Usher Syndrome or Atrophia areata.

As mentioned previously, fibroblast growth factors can act upon cells in both a mitogenic and non-mitogenic manner. These factors are mitogenic for a wide variety of normal diploid mesoderm-derived and neural crest-derived cells, inducing granulosa cells, adrenal cortical cells, chrondrocytes, myoblasts, corneal and vascular endothelial cells (bovine or human), vascular smooth muscle cells, lens, retina and prostatic epithelial cells, oligodendrocytes, astrocytes, chrondocytes, myoblasts and osteoblasts. Antibodies to these factors can be generated to modulate such effects.

Non-mitogenic actions of fibroblast growth factors include promotion of cell migration into a wound area (chemotaxis), initiation of new blood vessel formulation (angiogenesis), modulation of nerve regeneration and survival (neurotrophism), modulation of endocrine functions, and stimulation or suppression of specific cellular protein expression, extracellular matrix production and cell survival. Baird, A. & Bohlen, P., *Handbook of Exp. Phrmacol.* 95(1): 369-418 (1990). These properties provide a basis for using fibroblast growth factors in therapeutic approaches to accelerate wound healing, nerve repair, collateral blood vessel formation, and the like. For example, fibroblast growth factors, have been suggested to minimize myocardium damage in heart disease and surgery (U.S.P. 4,378,437). Antibodies to these factors can be generated to modulate such effects.

Therapeutic indications for PRO214 polypeptides include disorders associated with the preservation and maintenance of gastrointestinal mucosa and the repair of acute and chronic mucosal lesions (e.g., enterocolitis, Zollinger-Ellison syndrome, gastrointestinal ulceration and congenital microvillus atrophy), skin diseases associated with abnormal keratinocyte differentiation (e.g., psoriasis, epithelial cancers such as lung squamous cell carcinoma, epidermoid carcinoma of the vulva and gliomas.

Anti-PRO317 antibodies find use in anti-tumor indications if they are angiostatic, or in coronary ischemic indications if they are angiogenic.

Native PRO301 (SEQ ID NO:119) has a Blast score of 246 and 30% homology at residues 24 to 282 of Fig. 44 with A33_HUMAN, an A33 antigen precursor. A33 antigen precursor, as explained in the Background is a tumor-specific antigen, and as such, is a recognized marker and therapeutic target for the diagnosis and treatment of colon cancer. The expression of tumor-specific antigens is often associated with the progression of neoplastic tissue disorders. Native PRO301 (SEQ ID NO:119) and A33_HUMAN also show a Blast score of 245 and 30% homology at residues 21 to 282 of Fig. 44 with A33_HUMAN, the variation dependent upon how spaces are inserted into the compared sequences. Native PRO301 (SEQ ID NO:119) also has a Blast score of 165 and 29% homology at residues 60 to 255 of Fig. 44 with HS46KDA_1, a human coxsackie and adenovirus receptor protein, also known as cell surface protein HCAR. This region of PRO301 also shows a similar Blast score and homology with HSU90716_1. Expression of such proteins is usually associated with viral infection and therapeutics for the prevention of such infection may be accordingly conceived. Accordingly, antibodies to the above identified antigens and receptors have therapeutic potential as diagnostic and treatment techniques.

Therapeutic uses for the PRO234 polypeptides of the invention includes treatments associated with leukocyte homing or the interaction between leukocytes and the endothelium during an inflammatory response. Examples include asthma, rheumatoid arthritis, psoriasis and multiple sclerosis.

Cancer-associated or specific antigens permit the creation of tumor or cancer specific monoclonal antibodies (mAbs) which are specific to such tumor antigens. Such mAbs, which can distinguish between normal and cancerous

cells are useful in the diagnosis, prognosis and treatment of the disease.

Cancer specific monoclonal antibodies (mAbs) which are specific to tumor antigens. Such mAbs, which can distinguish between normal and cancerous cells are useful in the diagnosis, prognosis and treatment of the disease. Particular antigens are known to be associated with neoplastic diseases, such as colorectal and breast cancer. Since colon cancer is a widespread disease, early diagnosis and treatment is an important medical goal. Diagnosis and treatment of cancer can be implemented using monoclonal antibodies (mAbs) specific therefore having fluorescent, nuclear magnetic or radioactive tags. Radioactive genes, toxins and/or drug tagged mAbs can be used for treatment *in situ* with minimal patient description.

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

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

EXAMPLES

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, Rockville, Maryland.

EXAMPLE 1: Extracellular Domain Homology Screening to Identify Novel Polypeptides and cDNA Encoding Therefor

The extracellular domain (ECD) sequences (including the secretion signal sequence, if any) from about 950 known secreted proteins from the Swiss-Prot public database were used to search EST databases. The EST databases included public databases (e.g., Dayhoff, GenBank), and proprietary databases (e.g. LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA). The search was performed using the computer program BLAST or BLAST2 (Altschul, and Gish, Methods in Enzymology 266: 460-80 (1996); http://blast.wustl/edu/blast/README.html) as a comparison of the ECD protein sequences to a 6 frame translation of the EST sequences. Those comparisons with a Blast score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into consensus DNA sequences with the program "phrap" (Phil Green, University of Washington, Seattle, WA; (http://bozeman.mbt.washington.edu/phrap.docs/phrap.html).

Using this extracellular domain homology screen, consensus DNA sequences were assembled relative to the other identified EST sequences. In addition, the consensus DNA sequences obtained were often (but not always) extended using repeated cycles of BLAST and phrap to extend the consensus sequence as far as possible using the sources of EST sequences discussed above.

Based upon the consensus sequences obtained as described above, oligonucleotides were then synthesized and used to identify by PCR a cDNA library that contained the sequence of interest and for use as probes to isolate a clone of the full-length coding sequence for a PRO polypeptide. Forward (.f) and reverse (.r) PCR primers generally range from 20 to 30 nucleotides and are often designed to give a PCR product of about 100-1000 bp in length. The probe (.p) sequences are typically 40-55 bp in length. In some cases, additional oligonucleotides are

synthesized when the consensus sequence is greater than about 1-1.5kbp. In order to screen several libraries for a full-length clone, DNA from the libraries was screened by PCR amplification, as per Ausubel et al., Current Protocols in Molecular Biology, with the PCR primer pair. A positive library was then used to isolate clones encoding the gene of interest using the probe oligonucleotide and one of the primer pairs.

The cDNA libraries used to isolate the cDNA clones were constructed by standard methods using commercially available reagents such as those from Invitrogen, San Diego, CA. The cDNA was primed with oligo dT containing a NotI site, linked with blunt to SalI hemikinased adaptors, cleaved with NotI, sized appropriately by gel electrophoresis, and cloned in a defined orientation into a suitable cloning vector (such as pRKB or pRKD; pRK5B is a precursor of pRK5D that does not contain the SfiI site; see, Holmes et al., Science, 253:1278-1280 (1991)) in the unique XhoI and NotI sites.

EXAMPLE 2: Isolation of cDNA Clones Encoding PRO211 and PRO217

Consensus DNA sequences were assembled as described in Example 1 above and were designated as DNA28730 and DNA28760, respectively. Based on these consensus sequences, oligonucleotides were synthesized and used to identify by PCR a cDNA library that contained the sequences of interest and for use as probes to isolate a clone of the full-length coding sequence for the PRO211 and PRO217 polypeptides. The libraries used to isolate DNA32292-1131 and DNA33094-1131 were fetal lung libraries.

cDNA clones were sequenced in their entirety. The entire nucleotide sequences of PRO211 (DNA32292-1131; UNQ185) and PRO217 (UNQ191; DNA33094-1131) are shown in Figure 1 (SEQ ID NO: 1) and Figure 3 (SEQ ID NO:3), respectively. The predicted polypeptides are 353 and 379 amino acid in length, respectively, with respective molecular weights of approximately 38,190 and 41,520 daltons.

The oligonucleotide sequences used in the above procedures were the following:

28730.p (OLI 516) (SEQ ID NO:5)

5'-AGGGAGCACGGACAGTGTGCAGATGTGGACGAGTGCTCACTAGCA-3'

28730.f (OLI 517) (SEQ ID NO:6)

5'-AGAGTGTATCTCTGGCTACGC-3'

28730.r (OLI 518) (SEQ ID NO:7)

5'-TAAGTCCGGCACATTACAGGTC-3'

28760.p (OLI 617) (SEQ ID NO:8)

5'-CCCACGATGTATGAATGGTGGACTTTGTGTGACTCCTGGTTTCTGCATC-3'

28760.f (OLI 618) (SEQ ID NO:9)

5'-AAAGACGCATCTGCGAGTGTCC-3'

28760.r (OLI 619) (SEQ ID NO:10)

5'-TGCTGATTTCACACTGCTCTCCC-3'

EXAMPLE 3: Isolation of cDNA Clones Encoding Human PRO230

A consensus DNA sequence was assembled relative to the other identified EST sequences as described in Example 1 above, wherein the consensus sequence is designated herein as DNA30857. An EST proprietary to

Genentech was employed in the consensus assembly. The EST is designated as DNA20088 and has the nucleotide sequence shown in Figure 7 (SEQ ID NO:13).

Based on the DNA30857 consensus sequence, oligonucleotides were synthesized to identify by PCR a cDNA library that contained the sequence of interest and for use as probes to isolate a clone of the full-length coding sequence for PRO230.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-TTCGAGGCCTCTGAGAAGTGGCCC-3' (SEQ ID NO:14)

reverse PCR primer 5'-GGCGGTATCTCTCTGGCCTCCC-3' (SEQ ID NO:15)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA30857 sequence which had the following nucleotide sequence

hybridization probe

5'-TTCTCCACAGCAGCTGTGGCATCCGATCGTGTCTCAATCCATTCTCTGGG-3' (SEQ ID NO:16)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO230 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal lung tissue. DNAsequencing of the clones isolated as described above gave the full-length DNA sequence for PRO230 (herein designated as UNQ204 (DNA33223-1136)) and the derived protein sequence for PRO230.

The entire nucleotide sequence of UNQ204 (DNA33223-1136) is shown in Figure 5 (SEQ ID NO:11). Clone UNQ204 (DNA33223-1136) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 100-103 and ending at the stop codon at nucleotide positions 1501-1503 (Figure 5; SEQ ID NO:11). The predicted polypeptide precursor is 467 amino acids long (Figure 6).

EXAMPLE 4: Isolation of cDNA Clones Encoding Human PRO232

A consensus DNA sequence was assembled relative to the other identified EST sequences as described in Example 1 above, wherein the consensus sequence is designated herein as DNA30935. Based on the DNA30935 consensus sequence, oligonucleotides were synthesized to identify by PCR a cDNA library that contained the sequence of interest and for use as probes to isolate a clone of the full-length coding sequence for PRO232.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-TGCTGTGCTACTCCTGCAAAGCCC-3' (SEQ ID NO:19)

reverse PCR primer 5'-TGCACAAGTCGGTGTCACAGCACG-3' (SEQ ID NO:20)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA30935 sequence which had the following nucleotide sequence

hybridization probe

5'-AGCAACGAGGACTGCCTGCAGGTGGAGAACTGCACCCAGCTGGG-3' (SEQ ID NO:21)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO232 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal kidney tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO232 [herein designated as UNQ206 (DNA34435-1140)] and the derived protein sequence for PRO232.

The entire nucleotide sequence of UNQ206 (DNA34435-1140) is shown in Figure 8 (SEQ ID NO:17). Clone UNQ206 (DNA34435-1140) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 17-19 and ending at the stop codon at nucleotide positions 359-361 (Fig. 8; SEQ ID NO:17). The predicted polypeptide precursor is 114 amino acids long (Fig. 9). Clone UNQ206 (DNA34435-1140) has been deposited with ATCC on September 16, 1997 and is assigned ATCC deposit no. ATCC 209250.

Analysis of the amino acid sequence of the full-length PRO232 suggests that it possesses 35% sequence identity with a stem cell surface antigen from Gallus gallus.

EXAMPLE 5: Isolation of cDNA Clones Encoding PRO187

A proprietary expressed sequence tag (EST) DNA database (LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA) was searched and an EST (#843193) was identified which showed homology to fibroblast growth factor (FGF-8) also known as androgen-induced growth factor. mRNA was isolated from human fetal lung tissue using reagents and protocols from Invitrogen, San Diego, CA (Fast Track 2). The cDNA libraries used to isolate the cDNA clones were constructed by standard methods using commercially available reagents (e.g., Invitrogen, San Diego, CA, Life Technologies, Gaithersburg, MD). The cDNA was primed with oligo dT containing a NotI site, linked with blunt to SalI hemikinased adaptors, cleaved with NotI, sized appropriately by gel electrophoresis, and cloned in a defined orientation into the cloning vector pRK5D using reagents and protocols from Life Technologies, Gaithersburg, MD (Super Script Plasmid System). The double-stranded cDNA was sized to greater than 1000 bp and the SalI/NotI linkered cDNA was cloned into XhoI/NotI cleaved vector. pRK5D is a cloning vector that has an sp6 transcription initiation site followed by an SfiI restriction enzyme site preceding the XhoI/NotI cDNA cloning sites.

Several libraries from various tissue sources were screened by PCR amplification with the following oligonucleotide probes:

IN843193.f (OLI315) (SEQ ID NO:24)
5'-CAGTACGTGAGGGACCAGGGCGCCATGA-3'

IN843193.r (OLI 317) (SEQ ID NO:25)
5'-CCGGTGACCTGCACGTGCTTGCCA-3'

A positive library was then used to isolate clones encoding the PRO187 gene using one of the above oligonucleotides and the following oligonucleotide probe:

IN843193.p (OLI 316) (SEQ ID NO:26)
5'-GCGGATCTGCCGCCTGCTCANCTGGTCGGTCATGGCGCCCT-3'

A cDNA clone was sequenced in entirety. The entire nucleotide sequence of PRO187 (DNA27864-1155) is shown in Figure 10 (SEQ ID NO:22). Clone DNA27864-1155 contains a single open reading frame with an apparent translational initiation site at nucleotide position 1 (Figure 10; SEQ ID NO:22). The predicted polypeptide precursor is 205 amino acids long. Clone DNA27864-1155 has been deposited with the ATCC (designation: DNA27864-1155) and is assigned ATCC deposit no. ATCC 209375.

Based on a BLAST and FastA sequence alignment analysis (using the ALIGN computer program) of the full-length sequence, the PRO187 polypeptide shows 74% amino acid sequence identity (Blast score 310) to human fibroblast growth factor-8 (androgen-induced growth factor).

EXAMPLE 6: Isolation of cDNA Clones Encoding PRO265

A consensus DNA sequence was assembled relative to other EST sequences as described in Example 1 above using phrap. This consensus sequence is herein designated DNA33679. Based on the DNA33679 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO265.

PCR primers (two forward and one reverse) were synthesized:

forward PCR primer A: 5'-CGGTCTACCTGTATGGCAACC-3' (SEQ ID NO:29);

forward PCR primer B: 5'-GCAGGACAACCAGATAAACCAC-3' (SEQ ID NO:30);

reverse PCR primer 5'-ACGCAGATTTGAGAAGGCTGTC-3' (SEQ ID NO:31)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA33679 sequence which had the following nucleotide sequence

hybridization probe

5'-TTCACGGGCTGCTCTTGCCCAGCTCTTGAAGCTTGAAGAGCTGCAC-3' (SEQ ID NO:32)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with PCR primer pairs identified above. A positive library was then used to isolate clones encoding the PRO265 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human a fetal brain library.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO265 [herein designated as UNQ232 (DNA36350-1158] (SEQ ID NO:27) and the derived protein sequence for PRO265.

The entire nucleotide sequence of UNQ232 (DNA36350-1158) is shown in Figure 12 (SEQ ID NO:27). Clone UNQ232 (DNA36350-1158) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 352-354 and ending at the stop codon at positions 2332-2334 (Figure 12). The predicted polypeptide precursor is 660 amino acids long (Figure 13). Clone UNQ232 (DNA36350-1158) has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209378.

Analysis of the amino acid sequence of the full-length PRO265 polypeptide suggests that portions of it possess significant homology to the fibromodulin and the fibromodulin precursor, thereby indicating that PRO265 may be a novel member of the leucine rich repeat family, particularly related to fibromodulin.

EXAMPLE 7: Isolation of cDNA Clones Encoding Human PRO219

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA28729. Based on the DNA28729 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO219.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-GTGACCCTGGTTGTGAATACTCC-3' (SEQ ID NO:35)

reverse PCR primer 5'-ACAGCCATGGTCTATAGCTTGG-3' (SEQ ID NO:36)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA28729 sequence which had the following nucleotide sequence

hybridization probe

5'-GCCTGTCAGTGTCCTGAGGGACACGTGCTCCGCAGCGATGGGAAG-3' (SEQ ID NO:37)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO219 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal kidney tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO219 [herein designated as UNQ193 (DNA32290-1164)] (SEQ ID NO:33) and the derived protein sequence for PRO219.

The entire nucleotide sequence of UNQ193 (DNA32290-1164) is shown in Figures 14A-B (SEQ ID NO:33). Clone UNQ193 (DNA32290-1164) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 204-206 and ending at the stop codon at nucleotide positions 2949-2951 (Figures 14A-B). The predicted polypeptide precursor is 915 amino acids long (Figure 15). Clone UNQ193 (DNA32290-1164) has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209384.

Analysis of the amino acid sequence of the full-length PRO219 polypeptide suggests that portions of it possess significant homology to the mouse and human matrilin-2 precursor polypeptides.

EXAMPLE 8: Isolation of cDNA Clones Encoding Human PRO246

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA30955. Based on the DNA30955 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO246.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-AGGGTCTCCAGGAGAAAGACTC-3' (SEQ ID NO:40)

reverse PCR primer 5'-ATTGTGGGCCTTGCAGACATAGAC-3' (SEQ ID NO:41)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA30955 sequence which had the following nucleotide sequence

hybridization probe

5'-GGCCACAGCATCAAAACCTTAGAACTCAATGTACTGGTTCCTCCAGCTCC-3' (SEQ ID NO:42)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO246 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal liver tissue. DNAsequencing of the clones isolated as described above gave the full-length DNA sequence for PRO246 [herein designated as

UNQ220 (DNA35639-1172)] (SEQ ID NO:38) and the derived protein sequence for PRO246.

The entire nucleotide sequence of UNQ220 (DNA35639-1172) is shown in Figure 16 (SEQ ID NO:38). Clone UNQ220 (DNA35639-1172) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 126-128 and ending at the stop codon at nucleotide positions 1296-1298 (Figure 16). The predicted polypeptide precursor is 390 amino acids long (Figure 17). Clone UNQ220 (DNA35639-1172) has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209396.

Analysis of the amino acid sequence of the full-length PRO246 polypeptide suggests that it possess significant homology to the human cell surface protein HCAR, thereby indicating that PRO246 may be a novel cell surface virus receptor.

EXAMPLE 9: Isolation of cDNA Clones Encoding Human PRO228

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA28758. An EST proprietary to Genentech was employed in the consensus assembly. This EST is shown in Figure 20 (SEQ ID NO:50) and is herein designated as DNA21951.

Based on the DNA28758 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO228.

PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-GGTAATGAGCTCCATTACAG-3' (SEQ ID NO:51)

forward PCR primer 5'-GGAGTAGAAAGCGCATGG-3' (SEQ ID NO:52)

forward PCR primer 5'-CACCTGATACCATGAATGGCAG-3' (SEQ ID NO:53)

reverse PCR primer 5'-CGAGCTCGAATTAATTCG-3' (SEQ ID NO:54)

reverse PCR primer 5'-GGATCTCCTGAGCTCAGG-3' (SEQ ID NO:55)

reverse PCR primer 5'-CCTAGTTGAGTGATCCTTGTAAG-3' (SEQ ID NO:56)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA28758 sequence which had the following nucleotide sequence

hybridization probe

5'-ATGAGACCCACACCTCATGCCGCTGTAATCACCTGACACATTTTGCAATT-3' (SEQ ID NO:57)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pairs identified above. A positive library was then used to isolate clones encoding the PRO228 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal kidney tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO228 [herein designated as UNQ202 (DNA33092-1202)] (SEQ ID NO:48) and the derived protein sequence for PRO228.

The entire nucleotide sequence of UNQ202 (DNA33092-1202) is shown in Figure 18 (SEQ ID NO:48). Clone UNQ202 (DNA33092-1202) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 24-26 of SEQ ID NO:48 and ending at the stop codon after nucleotide position 2093 of SEQ

ID NO:48. The predicted polypeptide precursor is 690 amino acids long (Figure 19). Clone UNQ202 (DNA33092-1202) has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209420.

Analysis of the amino acid sequence of the full-length PRO228 polypeptide suggests that portions of it possess significant homology to the secretin-related proteins CD97 and EMR1 as well as the secretin member, latrophilin,. thereby indicating that PRO228 may be a new member of the secretin related proteins.

EXAMPLE 10: Isolation of cDNA Clones Encoding Human PRO533

The EST sequence accession number AF007268, a murine fibroblast growth factor (FGF-15) was used to search various public EST databases (e.g., GenBank, Dayhoff, etc.). The search was performed using the computer program BLAST or BLAST2 [Altschul et al., Methods in Enzymology, 266:460-480 (1996); http://blast.wustl/edu/blast/README.html] as a comparison of the ECD protein sequences to a 6 frame translation of the EST sequences. The search resulted in a hit with GenBank EST AA220994, which has been identified as stratagene NT2 neuronal precursor 937230.

Based on the Genbank EST AA220994 sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence. Forward and reverse PCR primers may range from 20 to 30 nucleotides (typically about 24), and are designed to give a PCR product of 100-1000 bp in length. The probe sequences are typically 40-55 bp (typically about 50) in length. In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification, as per *Ausubel et al., Current Protocols in Molecular Biology*, with the PCR primer pair. A positive library was then used to isolate clones encoding the gene of interest using the probe oligonucleotide and one of the PCR primers.

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified below. A positive library was then used to isolate clones encoding the PRO533 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal retina. The cDNA libraries used to isolated the cDNA clones were constructed by standard methods using commercially available reagents (e.g., Invitrogen, San Diego, CA; Clontech, etc.) The cDNA was primed with oligo dT containing a NotI site, linked with blunt to SalI hemikinased adaptors, cleaved with NotI, sized appropriately by gel electrophoresis, and cloned in a defined orientation into a suitable cloning vector (such as pRKB or pRKD; pRK5B is a precursor of pRK5D that does not contain the SfiI site; see, Holmes et al., *Science*, 253:1278-1280 (1991)) in the unique XhoI and NotI sites.

A cDNA clone was sequenced in its entirety. The full length nucleotide sequence of PRO533 is shown in Figure 21 (SEQ ID NO:58). Clone DNA49435-1219 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 459-461 (Figure 21; SEQ ID NO:58). The predicted polypeptide precursor is 216 amino acids long. Clone DNA47412-1219 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209480.

Based on a BLAST-2 and FastA sequence alignment analysis of the full-length sequence, PRO533 shows amino acid sequence identity to fibroblast growth factor (53%).

The oligonucleotide sequences used in the above procedure were the following:

FGF15.forward: 5'-ATCCGCCCAGATGGCTACAATGTGTA-3' (SEQ ID NO:60);

FGF15.probe: 5'-GCCTCCCGGTCTCCCTGAGCAGTGCCAAACAGCGGCAGTGTA-3' (SEQ ID NO:61);

FGF15.reverse: 5'-CCAGTCCGGTGACAAGCCCAAA-3' (SEQ ID NO:62).

EXAMPLE 11: Isolation of cDNA Clones Encoding Human PRO245

A consensus DNA sequence was assembled relative to the other identified EST sequences as described in Example 1 above, wherein the consensus sequence is designated herein as DNA30954.

Based on the DNA30954 consensus sequence, oligonucleotides were synthesized to identify by PCR a cDNA library that contained the sequence of interest and for use as probes to isolate a clone of the full-length coding sequence for PRO245.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-ATCGTTGTGAAGTTAGTGCCCC-3' (SEQ ID NO:65)

reverse PCR primer 5'-ACCTGCGATATCCAACAGAATTG-3' (SEQ ID NO:66)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA30954 sequence which had the following nucleotide sequence

hybridization probe

5'-GGAAGAGGATACAGTCACTCTGGAAGTATTAGTGGCTCCAGCAGTTCC-3' (SEQ ID NO:67)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO245 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal liver tissue. DNAsequencing of the clones isolated as described above gave the full-length DNA sequence for PRO245 [herein designated as UNQ219 (DNA35638-1141)] and the derived protein sequence for PRO245.

The entire nucleotide sequence of UNQ219 (DNA35638-1141) is shown in Figure 23 (SEQ ID NO:63). Clone UNQ219 (DNA35638-1141) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 89-91 and ending at the stop codon at nucleotide positions 1025-1027 (Fig. 23; SEQ ID NO:63). The predicted polypeptide precursor is 312 amino acids long (Fig. 24). Clone UNQ219 (DNA35638-1141) has been deposited with ATCC on September 16, 1997 and is assigned ATCC deposit no. ATCC 209265.

Analysis of the amino acid sequence of the full-length PRO245 suggests that a portion of it possesses 60% amino acid identity with the human c-myb protein and, therefore, may be a new member of the transmembrane protein receptor tyrosine kinase family.

EXAMPLE 12: Isolation of cDNA Clones Encoding Human PRO220, PRO221 and PRO227

(a) PRO220

A consensus DNA sequence was assembled relative to the other identified EST sequences as described in Example 1 above, wherein the consensus sequence is designated herein as DNA28749. Based on the DNA28749 consensus sequence, oligonucleotides were synthesized to identify by PCR a cDNA library that contained the

sequence of interest and for use as probes to isolate a clone of the full-length coding sequence for PRO220.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-TCACCTGGAGCCTTTATTGGCC-3' (SEQ ID NO:74)

reverse PCR primer 5'-ATACCAGCTATAACCAGGCTGCG-3' (SEQ ID NO:75)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA28749 sequence which had the following nucleotide sequence:

hybridization probe

5'-CAACAGTAAGTGGTTTGATGCTCTTCCAAATCTAGAGATTCTGATGATTGGG-3' (SEQ ID NO:76).

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO220 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal lung tissue. DNAsequencing of the clones isolated as described above gave the full-length DNA sequence for PRO220 [herein designated as UNQ194 (DNA32298-1132) and the derived protein sequence for PRO220.

The entire nucleotide sequence of UNQ194 (DNA32298-1132) is shown in Figure 25 (SEQ ID NO:68). Clone UNQ194 (DNA32298-1132) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 480-482 and ending at the stop codon at nucleotide positions 2604-2606 (Figure 25). The predicted polypeptide precursor is 708 amino acids long (Figure 26). Clone UNQ194 (DNA32298-1132) has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209257.

Analysis of the amino acid sequence of the full-length PRO220 shows it has homology to member of the leucine rich repeat protein superfamily, including the leucine rich repeat protein and the neuronal leucine-rich repeat protein 1.

(b) PRO221

A consensus DNA sequence was assembled relative to the other identified EST sequences as described in Example 1 above, wherein the consensus sequence is designated herein as DNA28756. Based on the DNA28756 consensus sequence, oligonucleotides were synthesized to identify by PCR a cDNA library that contained the sequence of interest and for use as probes to isolate a clone of the full-length coding sequence for PRO221.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-CCATGTGTCTCCTCCTACAAAG-3' (SEQ ID NO:77)

reverse PCR primer 5'-GGGAATAGATGTGATCTGATTGG-3' (SEQ ID NO:78)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA28756 sequence which had the following nucleotide sequence:

hybridization probe

5'-CACCTGTAGCAATGCAAATCTCAAGGAAATACCTAGAGATCTTCCTCCTG-3' (SEQ ID NO:79)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO221 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal lung tissue. DNAsequencing of the clones isolated as described above gave the full-length DNA sequence for PRO221 [herein designated as UNQ195 (DNA33089-1132) and the derived protein sequence for PRO221.

The entire nucleotide sequence of UNQ195 (DNA33089-1132) is shown in Figure 27 (SEQ ID NO:70). Clone UNQ195 (DNA33089-1132) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 179-181 and ending at the stop codon at nucleotide positions 956-958 (Figure 27). The predicted polypeptide precursor is 259 amino acids long (Figure 28). PRO221 is believed to have a transmembrane region at amino acids 206-225. Clone UNQ195 (DNA33089-1132) has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209262.

Analysis of the amino acid sequence of the full-length PRO221 shows it has homology to member of the leucine rich repeat protein superfamily, including the SLIT protein.

(c) PRO227

A consensus DNA sequence was assembled relative to the other identified EST sequences as described in Example 1 above, wherein the consensus sequence is designated herein as DNA28740. Based on the DNA28740 consensus sequence, oligonucleotides were synthesized to identify by PCR a cDNA library that contained the sequence of interest and for use as probes to isolate a clone of the full-length coding sequence for PRO227.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-AGCAACCGCCTGAAGCTCATCC-3' (SEQ ID NO:80)

reverse PCR primer 5'-AAGGCGCGGTGAAAGATGTAGACG-3' (SEQ ID NO:81)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA28740 sequence which had the following nucleotide sequence:

hybridization probe

5'GACTACATGTTTCAGGACCTGTACAACCTCAAGTCACTGGAGGTTGGCGA-3' (SEQ ID NO:82).

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO227 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal lung tissue. DNAsequencing of the clones isolated as described above gave the full-length DNA sequence for PRO227 [herein designated as UNQ201 (DNA33786-1132) and the derived protein sequence for PRO227.

The entire nucleotide sequence of UNQ201 (DNA33786-1132) is shown in Figure 29 (SEQ ID NO:72). Clone UNQ201 (DNA33786-1132) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 117-119 and ending at the stop codon at nucleotide positions 1977-1979 (Figure 29). The predicted polypeptide precursor is 620 amino acids long (Figure 30). PRO227 is believed to have a transmembrane region. Clone UNQ201 (DNA33786-1132) has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209253.

Analysis of the amino acid sequence of the full-length PRO221 shows it has homology to member of the leucine rich repeat protein superfamily, including the platelet glycoprotein V precursor and the human glycoprotein V.

EXAMPLE 13: Isolation of cDNA Clones Encoding Human PRO258

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA28746.

Based on the DNA28746 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO258.

PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-GCTAGGAATTCCACAGAAGCCC-3' (SEQ ID NO:85)

reverse PCR primer 5'-AACCTGGAATGTCACCGAGCTG-3' (SEQ ID NO:86)

reverse PCR primer 5'-CCTAGCACAGTGACGAGGGACTTGGC-3' (SEQ ID NO:87)

Additionally, synthetic oligonucleotide hybridization probes were constructed from the consensus DNA28740 sequence which had the following nucleotide sequence:

hybridization probe

5'-AAGACACAGCCACCCTAAACTGTCAGTCTTCTGGGAGCAAGCCTGCAGCC-3' (SEQ ID NO:88)

5'-GCCCTGGCAGACGAGGGCGAGTACACCTGCTCAATCTTCACTATGCCTGT-3' (SEQ ID NO:89)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO258 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal lung tissue. DNAsequencing of the clones isolated as described above gave the full-length DNA sequence for PRO258 [herein designated as UNQ225 (DNA35918-1174)] (SEQ ID NO:83) and the derived protein sequence for PRO258.

The entire nucleotide sequence of UNQ225 (DNA35918-1174) is shown in Figure 31 (SEQ ID NO:83). Clone UNQ225 (DNA35918-1174) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 147-149 of SEQ ID NO:83 and ending at the stop codon after nucleotide position 1340 of SEQ ID NO:83 (Figure 31). The predicted polypeptide precursor is 398 amino acids long (Figure 32). Clone UNQ225 (DNA35918-1174) has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209402.

Analysis of the amino acid sequence of the full-length PRO258 polypeptide suggests that portions of it possess significant homology to the CRTAM and the poliovirus receptor and have an Ig domain, thereby indicating that PRO258 is a new member of the Ig superfamily.

EXAMPLE 14: Isolation of cDNA Clones Encoding Human PRO266

An expressed sequence tag database was searched for ESTs having homology to SLIT, resulting in the identification of a single EST sequence designated herein as T73996. Based on the T73996 EST sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and

2) for use as probes to isolate a clone of the full-length coding sequence for PRO266.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-GTTGGATCTGGGCAACAATAAC-3' (SEQ ID NO:92)

reverse PCR primer 5'-ATTGTTGTGCAGGCTGAGTTTAAG-3' (SEQ ID NO:93)

Additionally, a synthetic oligonucleotide hybridization probe was constructed which had the following nucleotide sequence

hybridization probe

5'-GGTGGCTATACATGGATAGCAATTACCTGGACACGCTGTCCCGGG-3' (SEQ ID NO:94)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO266 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal brain tissue. DNAsequencing of the clones isolated as described above gave the full-length DNA sequence for PRO266 [herein designated as UNQ233 (DNA37150-1178)] (SEQ ID NO:90) and the derived protein sequence for PRO266.

The entire nucleotide sequence of UNQ233 (DNA37150-1178) is shown in Figure 33 (SEQ ID NO:90). Clone UNQ233 (DNA37150-1178) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 167-169 and ending at the stop codon after nucleotide position 2254 of SEQ ID NO:90. The predicted polypeptide precursor is 696 amino acids long (Figure 34). Clone UNQ233 (DNA37150-1178) has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209401.

Analysis of the amino acid sequence of the full-length PRO266 polypeptide suggests that portions of it possess significant homology to the SLIT protein, thereby indicating that PRO266 may be a novel leucine rich repeat protein.

EXAMPLE 15: Isolation of cDNA Clones Encoding Human PRO269

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA35705. Based on the DNA35705 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO269.

Forward and reverse PCR primers were synthesized:

forward PCR primer (.f1) 5'-TGGAAGGAGATGCGATGCCACCTG -3'
(SEQ ID NO:97)

forward PCR primer (.f2) 5'-TGACCAGTGGGGAAGGACAG-3' (SEQ ID NO:98)

forward PCR primer (.f3) 5'-ACAGAGCAGAGGGTGCCTTG-3' (SEQ ID NO:99)

reverse PCR primer (.r1) 5'-TCAGGGACAAGTGGTGTCTCTCCC-3'
(SEQ ID NO:100)

reverse PCR primer (.r2) 5'-TCAGGGAAGGAGTGTGCAGTTCTG-3'
(SEQ ID NO:101)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA35705 sequence which had the following nucleotide sequence:

hybridization probe

5'-ACAGCTCCCGATCTCAGTTACTTGCATCGCGGACGAAATCGGCGCTCGCT-3' (SEQ ID NO:102)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pairs identified above. A positive library was then used to isolate clones encoding the PRO269 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal kidney tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO269 [herein designated as UNQ236 (DNA38260-1180)] (SEQ ID NO:95) and the derived protein sequence for PRO269.

The entire nucleotide sequence of UNQ236 (DNA38260-1180) is shown in Figure 35 (SEQ ID NO:95). Clone UNQ236 (DNA38260-1180) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 314-316 and ending at the stop codon at nucleotide positions 1784-1786 (Fig. 35; SEQ ID NO:95). The predicted polypeptide precursor is 490 amino acids long (Fig. 36). Clone UNQ236 (DNA38260-1180) has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209397.

Analysis of the amino acid sequence of the full-length PRO269 suggests that portions of it possess significant homology to the human thrombomodulin proteins, thereby indicating that PRO269 may possess one or more thrombomodulin-like domains.

EXAMPLE 16: Isolation of cDNA Clones Encoding Human PRO287

A consensus DNA sequence encoding PRO287 was assembled relative to the other identified EST sequences as described in Example 1 above, wherein the consensus sequence is designated herein as DNA28728. Based on the DNA28728 consensus sequence, oligonucleotides were synthesized to identify by PCR a cDNA library that contained the sequence of interest and for use as probes to isolate a clone of the full-length coding sequence for PRO287.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-CCGATTCATAGACCTCGAGAGT-3' (SEQ ID NO:105)

reverse PCR primer 5'-GTCAAGGAGTCCTCCACAATAC-3' (SEQ ID NO:106)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA28728 sequence which had the following nucleotide sequence

hybridization probe

5'-GTGTACAATGGCCATGCCAATGGCCAGCGCATTGGCCGCTTCTGT-3'

(SEQ ID NO:107)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO287 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal kidney tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO287 [herein designated as UNQ250 (DNA39969-1185), SEQ ID NO:103] and the derived protein sequence for PRO287.

The entire nucleotide sequence of UNQ250 (DNA39969-1185) is shown in Figure 37 (SEQ ID NO:103). Clone UNQ250 (DNA39969-1185) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 307-309 and ending at the stop codon at nucleotide positions 1552-1554 (Fig. 37; SEQ ID NO:103). The predicted polypeptide precursor is 415 amino acids long (Fig. 38). Clone UNQ250 (DNA39969-1185) has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209400.

Analysis of the amino acid sequence of the full-length PRO287 suggests that it may possess one or more procollagen C-proteinase enhancer protein precursor or procollagen C-proteinase enhancer protein-like domains. Based on a BLAST and FastA sequence alignment analysis of the full-length sequence, PRO287 shows nucleic acid sequence identity to procollagen C-proteinase enhancer protein precursor and procollagen C-proteinase enhancer protein (47 and 54%, respectively).

EXAMPLE 17: Isolation of cDNA Clones Encoding Human PRO214

A consensus DNA sequence was assembled using phrap as described in Example 1 above. This consensus DNA sequence is designated herein as DNA28744. Based on this consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence.

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified below. A positive library was then used to isolate clones encoding the PRO214 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal lung tissue. A cDNA clone was sequenced in its entirety. The full length nucleotide sequence of DNA32286-1191 is shown in Figure 39 (SEQ ID NO:108). DNA32286-1191 contains a single open reading frame with an apparent translational initiation site at nucleotide position 103 (Fig. 39; SEQ ID NO:108). The predicted polypeptide precursor is 420 amino acids long (SEQ ID NO:109).

Based on a BLAST and FastA sequence alignment analysis of the full-length sequence, PRO214 polypeptide shows amino acid sequence identity to HT protein and/or Fibulin (49% and 38%, respectively).

The oligonucleotide sequences used in the above procedure were the following:

28744.p (OLI555)

5'-CCTGGCTATCAGCAGGTGGGCTCCAAGTGTCTCGATGTGGATGAGTGTGA-3' (SEQ ID NO:110)

28744.f (OLI556)

5'-ATTCTGCGTGAACACTGAGGGC-3' (SEQ ID NO:111)

28744.r (OLI557)

5'-ATCTGCTTGTAGCCCTCGGCAC-3' (SEQ ID NO:112)

EXAMPLE 18: Isolation of cDNA Clones Encoding Human PRO317

A consensus DNA sequence was assembled using phrap as described in Example 1 above, wherein the consensus sequence is herein designated as DNA28722. Based on this consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes

to isolate a clone of the full-length coding sequence. The forward and reverse PCR primers, respectively, synthesized for this purpose were:

5'-AGGACTGCCATAACTTGCCTG (OLI489) (SEQ ID NO:115) and

5'-ATAGGAGTTGAAGCAGCGCTGC (OLI490) (SEQ ID NO:116).

The probe synthesized for this purpose was:

5'-TGTGTGGACATAGACGAGTGCCGCTACCGCTACTGCCAGCACCGC (OLI488) (SEQ ID NO:117)

mRNA for construction of the cDNA libraries was isolated from human fetal kidney tissue.

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification, as per Ausubel *et al., Current Protocols in Molecular Biology* (1989), with the PCR primer pair identified above. A positive library was then used to isolate clones containing the PRO317 gene using the probe oligonucleotide identified above and one of the PCR primers.

A cDNA clone was sequenced in its entirety. The entire nucleotide sequence of DNA33461-1199 (encoding PRO317) is shown in Figure 41 (SEQ ID NO:113). Clone DNA33461-1199 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 68-70 (Fig. 41; SEQ ID NO:113). The predicted polypeptide precursor is 366 amino acids long. The predicted signal sequence is amino acids 1-18 of Figure 42 (SEQ ID NO:114). There is one predicted N-linked glycosylation site at amino acid residue 160. Clone DNA33461-1199 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209367.

Based on BLAST™ and FastA™ sequence alignment analysis (using the ALIGN™ computer program) of the full-length PRO317sequence, PRO317 shows the most amino acid sequence identity to EBAF-1 (92%). The results also demonstrate a significant homology between human PRO317 and mouse LEFTY protein. The C-terminal end of the PRO317 protein contains many conserved sequences consistent with the pattern expected of a member of the TGF- superfamily.

In situ expression analysis in human tissues performed as described below evidences that there is distinctly strong expression of the PRO317 polypeptide in pancreatic tissue.

EXAMPLE 19: Isolation of cDNA clones Encoding Human PRO301

A consensus DNA sequence designated herein as DNA35936 was assembled using phrap as described in Example 1 above. Based on this consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence.

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified below. A positive library was then used to isolate clones encoding the PRO301 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal kidney.

A cDNA clone was sequenced in its entirety. The full length nucleotide sequence of native sequence PRO301 is shown in Figure 43 (SEQ ID NO:118). Clone DNA40628-1216 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 52-54 (Fig. 43; SEQ ID NO:118). The predicted polypeptide precursor is 299 amino acids long with a predicted molecular weight of 32,583 daltons and pI of 8.29.

Clone DNA40628-1216 has been deposited with ATCC and is assigned ATCC deposit No. ATCC 209432.

Based on a BLAST and FastA sequence alignment analysis of the full-length sequence, PRO301 shows amino acid sequence identity to A33 antigen precursor (30%) and coxsackie and adenovirus receptor protein (29%).

The oligonucleotide sequences used in the above procedure were the following:

OLI2162 (35936.f1) 5'-TCGCGGAGCTGTGTTCTGTTTCCC-3' (SEQ ID NO:120)
OLI2163 (35936.p1)
5'-TGATCGCGATGGGGACAAAGGCGCAAGCTCGAGAGGAAACTGTTGTGCCT-3' (SEQ ID NO:121)
OLI2164 (35936.f2)
5'-ACACCTGGTTCAAAGATGGG-3' (SEQ ID NO:122)
OLI2165 (35936.r1)
5'-TAGGAAGAGTTGCTGAAGGCACGG-3' (SEQ ID NO:123)
OLI2166 (35936.f3)
5'-TTGCCTTACTCAGGTGCTAC-3' (SEQ ID NO:124)
OLI2167 (35936.r2)
5'-ACTCAGCAGTGGTAGGAAAG-3' (SEQ ID NO:125)

EXAMPLE 20: Isolation of cDNA Clones Encoding Human PRO224

A consensus DNA sequence assembled relative to the other identified EST sequences as described in Example 1, wherein the consensus sequence is designated herein as DNA30845. Based on the DNA30845 consensus sequence, oligonucleotides were synthesized to identify by PCR a cDNA library that contained the sequence of interest and for use as probes to isolate a clone of the full-length coding sequence for PRO224.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-AAGTTCCAGTGCCGCACCAGTGGC-3' (SEQ ID NO:128)

reverse PCR primer 5'-TTGGTTCCACAGCCGAGCTCGTCG-3' (SEQ ID NO:129)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA30845 sequence which had the following nucleotide sequence

hybridization probe

5'-GAGGAGGAGTGCAGGATTGAGCCATGTACCCAGAAAGGGCAATGCCCACC-3' (SEQ ID NO:130)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO224 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal liver tissue. DNAsequencing of the clones isolated as described above gave the full-length DNA sequence for PRO224 [herein designated as UNQ198 (DNA33221-1133)] and the derived protein sequence for PRO224.

The entire nucleotide sequence of UNQ198 (DNA33221-1133) is shown in Figure 45 (SEQ ID NO:126). Clone UNQ198 (DNA33221-1133) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 96-98 and ending at the stop codon at nucleotide positions 942-944 (Figure 45; SEQ ID NO:126). The start of a transmembrane region begins at nucleotide position 777. The predicted polypeptide precursor is 282 amino acids long (Figure 46). Clone UNQ198 (DNA33221-1133) has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209263.

Analysis of the amino acid sequence of the full-length PRO224 suggests that it has homology to very low-

density lipoprotein receptors, apolipoprotein E receptor and chicken oocyte receptors P95. Based on a BLAST and FastA sequence alignment analysis of the full-length sequence, PRO224 has amino acid identity to portions of these proteins in the range from 28% to 45%, and overall identity with these proteins in the range from 33% to 39%.

EXAMPLE 21: Isolation of cDNA Clones Encoding Human PRO222

A consensus DNA sequence was assembled relative to the other identified EST sequences as described in Example 1 above, wherein the consensus sequence is designated herein as DNA28771. Based on the DNA28771 consensus sequence, oligonucleotides were synthesized to identify by PCR a cDNA library that contained the sequence of interest and for use as probes to isolate a clone of the full-length coding sequence for PRO222.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-ATCTCCTATCGCTGCTTTCCCGG-3' (SEQ ID NO:133)

reverse PCR primer 5'-AGCCAGGATCGCAGTAAAACTCC-3' (SEQ ID NO:134)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA28771 sequence which had the following nucleotide sequence:

hybridization probe

5'-ATTTAAACTTGATGGGTCTGCGTATCTTGAGTGCTTACAAAACCTTATCT-3' (SEQ ID NO:135)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO222 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal kidney tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO222 [herein designated as UNQ196 (DNA33107-1135)] and the derived protein sequence for PRO222.

The entire nucleotide sequence of UNQ196 (DNA33107-1135) is shown in Figure 47 (SEQ ID NO:131). Clone UNQ196 (DNA33107-1135) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 159-161 and ending at the stop codon at nucleotide positions 1629-1631 (Fig. 47; SEQ ID NO:131). The predicted polypeptide precursor is 490 amino acids long (Fig. 48). Clone UNQ196 (DNA33107-1135) has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209251.

Based on a BLAST and FastA sequence alignment analysis of the full-length sequence, PRO222 shows amino acid sequence identity to mouse complement factor h precursor (25-26%), complement receptor (27-29%), mouse complement C3b receptor type 2 long form precursor (25-47%) and human hypothetical protein kiaa0247 (40%).

EXAMPLE 22: Isolation of cDNA clones Encoding PRO234

A consensus DNA sequence was assembled (DNA30926) using phrap as described in Example 1 above. Based on this consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence.

RNA for the construction of the cDNA libraries was isolated using standard isolation protocols, e.g., Ausubel *et al.*, *Current Protocols in Molecular Biology*, from tissue or cell line sources or it was purchased from

commercial sources (e.g., Clontech). The cDNA libraries used to isolate the cDNA clones were constructed by standard methods (e.g., Ausubel *et al.*) using commercially available reagents (e.g., Invitrogen). This library was derived from 22 week old fetal brain tissue.

A cDNA clone was sequenced in its entirety. The entire nucleotide sequence of PRO234 is shown in Figure 49 (SEQ ID NO:136). The predicted polypeptide precursor is 382 amino acids long and has a calculated molecular weight of approximately 43.1 kDa.

The oligonucleotide sequences used in the above procedure were the following:

30926.p (OLI826) (SEQ ID NO:138): 5'-GTTCATTGAAAACCTCTTGCCATCT
GATGGTGACTTCTGGATTGGGCTCA-3'

30926.f (OLI827) (SEQ ID NO:139): 5'-AAGCCAAAGAAGCCTGCAGGAGGG-3'

30926.r (OLI828) (SEQ ID NO:140): 5'-CAGTCCAAGCATAAAGGTCCTGGC-3'

EXAMPLE 23: Isolation of cDNA Clones Encoding Human PRO231

A consensus DNA sequence was assembled relative to the other identified EST sequences as described in Example 1 above, wherein the consensus sequence was designated herein as DNA30933. Based on the DNA30933 consensus sequence, oligonucleotides were synthesized to identify by PCR a cDNA library that contained the sequence of interest and for use as probes to isolate a clone of the full-length coding sequence for PRO231.

Three PCR primers (two forward and one reverse) were synthesized:

forward PCR primer 1 5'-CCAACTACCAAAGCTGCTGGAGCC-3' (SEQ ID NO:143)

forward PCR primer 2 5'-GCAGCTCTATTACCACGGGAAGGA-3' (SEQ ID NO:144)

reverse PCR primer 5'-TCCTTCCCGTGGTAATAGAGCTGC-3' (SEQ ID NO:145)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA30933 sequence which had the following nucleotide sequence

hybridization probe

5'-GGCAGAGAACCAGAGGCCGGAGGAGACTGCCTCTTTACAGCCAGG-3' (SEQ ID NO:146)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pairs identified above. A positive library was then used to isolate clones encoding the PRO231 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal liver tissue. DNAsequencing of the clones isolated as described above gave the full-length DNA sequence for PRO231 [herein designated as UNQ205 (DNA34434-1139)] and the derived protein sequence for PRO231.

The entire nucleotide sequence of UNQ205 (DNA34434-1139) is shown in Figure 51 (SEQ ID NO:141). Clone UNQ205 (DNA34434-1139) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 173-175 and ending at the stop codon at nucleotide positions 1457-1459 (Fig. 51; SEQ ID NO:141). The predicted polypeptide precursor is 428 amino acids long (Fig. 52). Clone UNQ205 (DNA34434-1139) has been deposited with ATCC on September 16, 1997 and is assigned ATCC deposit no. ATCC 209252.

Analysis of the amino acid sequence of the full-length PRO231 suggests that it possesses 30% and 31% amino acid identity with the human and rat prostatic acid phosphatase precursor proteins, respectively.

EXAMPLE 24: Isolation of cDNA Clones Encoding Human PRO229

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA28762. Based on the DNA28762 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO229.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-TTCAGCTCATCACCTTCACCTGCC-3' (SEQ ID NO:149)

reverse PCR primer 5'-GGCTCATACAAAATACCACTAGGG-3' (SEQ ID NO:150)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA28762 sequence which had the following nucleotide sequence

hybridization probe

5'-GGGCCTCCACCGCTGTGAAGGGCGGGTGGAGGTGGAACAGAAAGGCCAGT-3' (SEQ ID NO:151)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO229 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal liver tissue. DNAsequencing of the clones isolated as described above gave the full-length DNA sequence for PRO229 [herein designated as UNQ203 (DNA33100-1159)] (SEQ ID NO:147) and the derived protein sequence for PRO229.

The entire nucleotide sequence of UNQ203 (DNA33100-1159) is shown in Figure 53 (SEQ ID NO:147). Clone UNQ203 (DNA33100-1159) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 98-100 and ending at the stop codon at nucleotide positions 1139-1141 (Figure 53). The predicted polypeptide precursor is 347 amino acids long (Figure 54). Clone UNQ203 (DNA33100-1159) has been deposited with ATCC and is assigned ATCC deposit no.ATCC 209377

Analysis of the amino acid sequence of the full-length PRO229 polypeptide suggests that portions of it possess significant homology to antigen wc1.1, M130 antigen and CD6.

EXAMPLE 25: Isolation of cDNA Clones Encoding Human PRO238

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described above in Example 1. This consensus sequence is herein designated DNA30908. Based on the DNA30908 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO238.

PCR primers (forward and reverse) were synthesized:

forward PCR primer 1 5'-GGTGCTAAACTGGTGCTCTGTGGC-3' (SEQ ID NO:154)

forward PCR primer 2 5'-CAGGGCAAGATGAGCATTCC-3' (SEQ ID NO:155)

reverse PCR primer 5'-TCATACTGTTCCATCTCGGCACGC-3' (SEQ ID NO:156)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA30908 sequence which had the following nucleotide sequence

hybridization probe

5'-AATGGTGGGGCCCTAGAAGAGCTCATCAGAGAACTCACCGCTTCTCATGC-3' (SEQ ID NO:157)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO238 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal liver tissue. DNAsequencing of the clones isolated as described above gave the full-length DNA sequence for PRO238 and the derived protein sequence for PRO238.

The entire nucleotide sequence of DNA35600-1162 is shown in Figure 55 (SEQ ID NO:152). Clone DNA35600-1162 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 134-136 and ending prior to the stop codon at nucleotide positions 1064-1066 (Figure 55). The predicted polypeptide precursor is 310 amino acids long (Figure 56). Clone DNA35600-1162 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209370.

Analysis of the amino acid sequence of the full-length PRO238 polypeptide suggests that portions of it possess significant homology to reductase, particularly oxidoreductase, thereby indicating that PRO238 may be a novel reductase.

EXAMPLE 26: Isolation of cDNA Clones Encoding Human PRO233

The extracellular domain (ECD) sequences (including the secretion signal, if any) of from about 950 known secreted proteins from the Swiss-Prot public protein database were used to search expressed sequence tag (EST) databases. The EST databases included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA). The search was performed using the computer program BLAST or BLAST2 (Altshul et al., Methods in Enzymology 266:460-480 (1996)) as a comparison of the ECD protein sequences to a 6 frame translation of the EST sequence. Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into consensus DNA sequences with the program "phrap" (Phil Green, University of Washington, Seattle, Washington; http://bozeman.mbt.washington.edu/phrap.docs/phrap.html).

An expressed sequence tag (EST) was identified by the EST database search and a consensus DNA sequence was assembled relative to other EST sequences using phrap. This consensus sequence is herein designated DNA30945. Based on the DNA30945 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO233.

Forward and reverse PCR primers were synthesized:

forward PCR primer	5'-GGTGAAGGCAGAAATTGGAGATG-3'	(SEQ ID NO:160)
reverse PCR primer	5'-ATCCCATGCATCAGCCTGTTTACC-3'	(SEQ ID NO:161)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA30945 sequence which had the following nucleotide sequence

hybridization probe

5'-GCTGGTGTAGTCTATACATCAGATTTGTTTGCTACACAAGATCCTCAG-3'

(SEQ ID NO:162)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO233 gene using the probe oligonucleotide.

RNA for construction of the cDNA libraries was isolated from human fetal brain tissue. DNAsequencing of the clones isolated as described above gave the full-length DNA sequence for PRO233 [herein designated as UNQ207 (DNA34436-1238)] (SEQ ID NO:158) and the derived protein sequence for PRO233.

The entire nucleotide sequence of UNQ207 (DNA34436-1238) is shown in Figure 57 (SEQ ID NO:158). Clone UNQ207 (DNA34436-1238) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 101-103 and ending at the stop codon at nucleotide positions 1001-1003 (Figure 57). The predicted polypeptide precursor is 300 amino acids long (Figure 58). The full-length PRO233 protein shown in Figure 58 has an estimated molecular weight of about 32,964 daltons and a pI of about 9.52. Clone UNQ207 (DNA34436-1238) has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209523.

Analysis of the amino acid sequence of the full-length PRO233 polypeptide suggests that portions of it possess significant homology to reductase proteins, thereby indicating that PRO233 may be a novel reductase.

EXAMPLE 27: Isolation of cDNA Clones Encoding Human PRO223

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA30836. Based on the DNA30836 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO223.

PCR primer pairs (one forward and two reverse) were synthesized:

forward PCR primer 5'-TTCCATGCCACCTAAGGGAGACTC-3' (SEQ ID NO:165)

reverse PCR primer 1 5'-TGGATGAGGTGTGCAATGGCTGGC-3' (SEQ ID NO:166)

reverse PCR primer 2 5'-AGCTCTCAGAGGCTGGTCATAGGG-3' (SEQ ID NO:167)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA30836 sequence which had the following nucleotide sequence

hybridization probe

5'-GTCGGCCCTTTCCCAGGACTGAACATGAAGAGTTATGCCGGCTTCCTCAC-3' (SEQ ID NO:168)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO223 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal liver tissue. DNAsequencing of the clones isolated as described above gave the full-length DNA sequence for PRO223 [herein designated as UNQ197 (DNA33206-1165)] (SEQ ID NO:163) and the derived protein sequence for PRO223.

The entire nucleotide sequence of UNQ197 (DNA33206-1165) is shown in Figure 59 (SEQ ID NO:163). Clone UNQ197 (DNA33206-1165) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 97-99 and ending at the stop codon at nucleotide positions 1525-1527 (Figure 59). The predicted polypeptide precursor is 476 amino acids long (Figure 60). Clone UNQ197 (DNA33206-1165) has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209372.

Analysis of the amino acid sequence of the full-length PRO223 polypeptide suggests that it possesses significant homology to various serine carboxypeptidase proteins, thereby indicating that PRO223 may be a novel serine carboxypeptidase.

EXAMPLE 28: Isolation of cDNA Clones Encoding Human PRO235

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated "DNA30927". Based on the DNA30927 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO235.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-TGGAATACCGCCTCCTGCAG-3' (SEQ ID NO:171)

reverse PCR primer 5'-CTTCTGCCCTTTGGAGAAGATGGC-3' (SEQ ID NO:172)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA30927 sequence which had the following nucleotide sequence

hybridization probe

5'-GGACTCACTGGCCCAGGCCTTCAATATCACCAGCCAGGACGAT-3' (SEQ ID NO:173)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO235 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal liver tissue. DNAsequencing of the clones isolated as described above gave the full-length DNA sequence for PRO235 [herein designated as UNQ209 (DNA35558-1167)] (SEQ ID NO:169) and the derived protein sequence for PRO235.

The entire nucleotide sequence of UNQ209 (DNA35558-1167) is shown in Figure 61 (SEQ ID NO:169). Clone UNQ209 (DNA35558-1167) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 667-669 and ending at the stop codon at nucleotide positions 2323-2325 (Figure 61). The predicted polypeptide precursor is 552 amino acids long (Figure 62). Clone UNQ209 (DNA35558-1167) has been deposited with ATCC and is assigned ATCC deposit no. 209374.

Analysis of the amino acid sequence of the full-length PRO235 polypeptide suggests that portions of it possess significant homology to the human, mouse and *Xenopus* plexin protein, thereby indicating that PRO235 may be a novel plexin protein.

EXAMPLE 29: Isolation of cDNA Clones Encoding Human PRO236 and Human PRO262

Consensus DNA sequences were assembled relative to other EST sequences using phrap as described in Example 1 above. These consensus sequences are herein designated DNA30901 and DNA30847. Based on the DNA30901 and DNA30847 consensus sequences, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO236 and PRO262, respectively.

Based upon the DNA30901 consensus sequence, a pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-TGGCTACTCCAAGACCCTGGCATG-3' (SEQ ID NO:178)

reverse PCR primer 5'-TGGACAAATCCCCTTGCTCAGCCC-3' (SEQ ID NO:179)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA30901 sequence which had the following nucleotide sequence

hybridization probe

5'-GGGCTTCACCGAAGCAGTGGACCTTTATTTTGACCACCTGATGTCCAGGG-3' (SEQ ID NO:180)

Based upon the DNA30847 consensus sequence, a pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-CCAGCTATGACTATGATGCACC-3' (SEQ ID NO:181)

reverse PCR primer 5'-TGGCACCCAGAATGGTGTTGGCTC-3' (SEQ ID NO:182)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA30847 sequence which had the following nucleotide sequence

hybridization probe

5'-CGAGATGTCATCAGCAAGTTCCAGGAAGTTCCTTTGGGACCTTTACCTCC-3' (SEQ ID NO:183)

In order to screen several libraries for a source of full-length clones, DNA from the libraries was screened by PCR amplification with the PCR primer pairs identified above. Positive libraries were then used to isolate clones encoding the PRO236 and PRO262 genes using the probe oligonucleotides and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal lung tissue for PRO236 and human fetal liver tissue for PRO262.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO236 [herein designated as UNQ210 (DNA35599-1168)] (SEQ ID NO:174), the derived protein sequence for PRO236, the full-length DNA sequence for PRO262 [herein designated as UNQ229 (DNA36992-1168)] (SEQ ID NO:176) and the derived protein sequence for PRO262.

The entire nucleotide sequence of UNQ210 (DNA35599-1168) is shown in Figure 63 (SEQ ID NO:174). Clone UNQ210 (DNA35599-1168) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 69-71 and ending at the stop codon at nucleotide positions 1977-1979 (Figure 63). The predicted polypeptide precursor is 636 amino acids long (Figure 64). Clone UNQ210 (DNA35599-1168) has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209373.

The entire nucleotide sequence of UNQ229 (DNA36992-1168) is shown in Figure 65 (SEQ ID NO:176). Clone UNQ229 (DNA36992-1168) contains a single open reading frame with an apparent translational initiation site

at nucleotide positions 240-242 and ending at the stop codon at nucleotide positions 2202-2204 (Figure 65). The predicted polypeptide precursor is 654 amino acids long (Figure 66). Clone UNQ229 (DNA36992-1168) has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209382.

Analysis of the amino acid sequence of the full-length PRO236 and PRO262 polypeptides suggests that portions of those polypeptides possess significant homology to β-galactosidase proteins derived from various sources, thereby indicating that PRO236 and PRO262 may be novel β-galactosidase homologs.

EXAMPLE 30: Isolation of cDNA Clones Encoding Human PRO239

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA30909. Based on the DNA30909 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO239.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer	5'-CCTCCCTCTATTACCCATGTC-3'	(SEQ ID NO:186)
reverse PCR primer	5'-GACCAACTTTCTCTGGGAGTGAGG-3'	(SEQ ID NO:187)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA30909 sequence which had the following nucleotide sequence

hybridization probe

5'-GTCACTTTATTTCTCTAACAACAAGCTCGAATCCTTACCAGTGGCAG-3'

(SEQ ID NO:188)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO239 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal lung tissue. DNAsequencing of the clones isolated as described above gave the full-length DNA sequence for PRO239 [herein designated as UNQ213 (DNA34407-1169)] (SEQ ID NO:184) and the derived protein sequence for PRO239.

The entire nucleotide sequence of UNQ213 (DNA34407-1169) is shown in Figure 67 (SEQ ID NO:184). Clone UNQ213 (DNA34407-1169) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 72-74 and ending at the stop codon at nucleotide positions 1575-1577 (Figure 67). The predicted polypeptide precursor is 501 amino acids long (Figure 68). Clone UNQ213 (DNA34407-1169) has been deposited with ATCC and is assigned ATCC deposit no.ATCC 209383.

Analysis of the amino acid sequence of the full-length PRO239 polypeptide suggests that portions of it possess significant homology to the densin protein, thereby indicating that PRO239 may be a novel molecule in the densin family.

EXAMPLE 31: Isolation of cDNA Clones Encoding Human PRO257

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA28731. Based on the DNA28731 consensus

sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO257.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-TCTCTATTCCAAACTGTGGCG-3' (SEQ ID NO:191)

reverse PCR primer 5'-TTTGATGACGATTCGAAGGTGG-3' (SEQ ID NO:192)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA28731 sequence which had the following nucleotide sequence

hybridization probe

5'-GGAAGGATCCTTCACCAGCCCCAATTACCCAAAGCCGCATCCTGAGC-3' (SEQ ID NO:193)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO257 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal kidney tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO257 [herein designated as UNQ224 (DNA35841-1173) (SEQ ID NO:189) and the derived protein sequence for PRO257.

The entire nucleotide sequence of UNQ224 (DNA35841-1173) is shown in Figure 69 (SEQ ID NO:189). Clone UNQ224 (DNA35841-1173) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 964-966 and ending at the stop codon at nucleotide positions 2785-2787 (Figure 69). The predicted polypeptide precursor is 607 amino acids long (Figure 70). Clone UNQ224 (DNA35841-1173) has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209403.

Analysis of the amino acid sequence of the full-length PRO257 polypeptide suggests that portions of it possess significant homology to the ebnerin protein, thereby indicating that PRO257 may be a novel protein member related to the ebnerin protein.

EXAMPLE 32: Isolation of cDNA Clones Encoding Human PRO260

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA30834. Based on the DNA30834 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO260.

PCR primers (forward and two reverse) were synthesized:

forward PCR primer: 5'-TGGTTTGACCAGGCCAAGTTCGG-3' (SEQ ID NO:196);

reverse PCR primer A: 5'-GGATTCATCCTCAAGGAAGAGCGG-3' (SEQ ID NO:197); and

reverse PCR primer B: 5'AACTTGCAGCATCAGCCACTCTGC-3' (SEQ ID NO:198)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA30834 sequence which had the following nucleotide sequence:

hybridization probe:

5'-TTCCGTGCCCAGCTTCGGTAGCGAGTGGTTCTGGTGGTATTGGCA-3' (SEQ ID NO:199)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO260 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal kidney tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO260 [herein designated as UNQ227 (DNA33470-1175)] (SEQ ID NO:194) and the derived protein sequence for PRO260.

The entire nucleotide sequence of UNQ227 (DNA33470-1175) is shown in Figure 71 (SEQ ID NO:194). Clone UNQ227 (DNA33470-1175) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 67-69 and ending at the stop codon 1468-1470 (see Figure 71). The predicted polypeptide precursor is 467 amino acids long (Figure 72). Clone UNQ227 (DNA33470-1175) has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209398.

Analysis of the amino acid sequence of the full-length PRO260 polypeptide suggests that portions of it possess significant homology to the alpha-l-fucosidase precursor, thereby indicating that PRO260 may be a novel fucosidase.

EXAMPLE 33: Isolation of cDNA Clones Encoding Human PRO263

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA30914. Based on the DNA30914 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO263.

PCR primers (tow forward and one reverse) were synthesized:

forward PCR primer 1: 5'-GAGCTTTCCATCCAGGTGTCATGC-3' (SEQ ID NO:202);

forward PCR primer 2: 5'-GTCAGTGACAGTACCTACTCGG-3' (SEQ ID NO:203); reverse PCR primer: 5'-TGGAGCAGGAGGAGTAGTAGTAGG-3' (SEQ ID NO:204)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA30914 sequence which had the following nucleotide sequence:

hybridization probe:

5'-AGGAGGCCTGTAGGCTGCTGGGACTAAGTTTGGCCGGCAAGGACCAAGTT-3' (SEQ ID NO:205)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO263 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal liver tissue. DNAsequencing of the clones isolated as described above gave the full-length DNA sequence for PRO263 [herein designated as UNQ230 (DNA34431-1177)] (SEQ ID NO:200) and the derived protein sequence for PRO263.

The entire nucleotide sequence of UNQ230 (DNA34431-1177) is shown in Figure 73 (SEQ ID NO:200). Clone UNQ230 (DNA34431-1177) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 160-162 of SEQ ID NO:200 and ending at the stop codon after the nucleotide at position 1126-1128 of SEQ ID NO:200 (Figure 73). The predicted polypeptide precursor is 322 amino acids long (Figure 74).

Clone UNQ230 (DNA34431-1177) has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209399.

Analysis of the amino acid sequence of the full-length PRO263 polypeptide suggests that portions of it possess significant homology to CD44 antigen, thereby indicating that PRO263 may be a novel cell surface adhesion molecule.

EXAMPLE 34: Isolation of cDNA Clones Encoding Human PRO270

A consensus DNA sequence was assembled relative to the other identified EST sequences as described in Example 1 above, wherein the consensus sequence was designated herein as DNA35712. Based on the DNA35712 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO270.

Forward and reverse PCR primers were synthesized:

forward PCR primer (.f1) 5'-GCTTGGATATTCGCATGGGCCTAC-3' (SEQ ID NO:208)

forward PCR primer (.f2) 5'-TGGAGACAATATCCCTGAGG-3' (SEQ ID NO:209)

reverse PCR primer (.r1) 5'-AACAGTTGGCCACAGCATGGCAGG-3' (SEQ ID NO:210)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA35712 sequence which had the following nucleotide sequence

hybridization probe

5'-CCATTGATGAGGAACTAGAACGGGACAAGAGGGTCACTTGGATTGTGGAG-3'

(SEQ ID NO:211)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO270 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal lung tissue. DNAsequencing of the clones isolated as described above gave the full-length DNA sequence for PRO270 [herein designated as UNQ237, DNA39510-1181] (SEQ ID NO:206) and the derived protein sequence for PRO270.

The entire nucleotide sequence of UNQ237, DNA39510-1181 is shown in Figure 75 (SEQ ID NO:206). Clone UNQ237 (DNA39510-1181) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 3-5 and ending at the stop codon at nucleotide positions 891-893 (Fig. 75; SEQ ID NO:206). The predicted polypeptide precursor is 296 amino acids long (Fig. 76). Clone UNQ237 (DNA39510-1181) has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209392.

Analysis of the amino acid sequence of the full-length PRO270 suggests that portions of it possess significant homology to the thioredoxin-protein, thereby indicating that the PRO270 protein may be a novel member of the thioredoxin family.

EXAMPLE 35: Isolation of cDNA Clones Encoding Human PRO271

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA35737. Based on the DNA35737 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of

interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO271.

Forward and reverse PCR primers were synthesized:

forward PCR primer 1 5'-TGCTTCGCTACTGCCCTC-3' (SEQ ID NO:214)

forward PCR primer 2 5'-TTCCCTTGTGGGTTGGAG-3' (SEQ ID NO:215)

forward PCR primer 3 5'-AGGGCTGGAAGCCAGTTC-3' (SEQ ID NO:216)

reverse PCR primer 1 5'-AGCCAGTGAGGAAATGCG-3' (SEQ ID NO:217)

reverse PCR primer 2 5'-TGTCCAAAGTACACACACCTGAGG-3' (SEQ ID NO:218)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA35737 sequence which had the following nucleotide sequence

hybridization probe

5'-GATGCCACGATCGCCAAGGTGGGACAGCTCTTTGCCGCCTGGAAG-3' (SEQ ID NO:219)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO271 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal brain tissue. DNAsequencing of the clones isolated as described above gave the full-length DNA sequence for PRO271 [herein designated as UNQ238 (DNA39423-1182)] (SEQ ID NO:212) and the derived protein sequence for PRO271.

The entire nucleotide sequence of UNQ238 (DNA39423-1182) is shown in Figure 77 (SEQ ID NO:212). Clone UNQ238 (DNA39423-1182) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 101-103 and ending at the stop codon at nucleotide positions 1181-1183 (Figure 77). The predicted polypeptide precursor is 360 amino acids long (Figure 78). Clone UNQ238 (DNA39423-1182) has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209387.

Analysis of the amino acid sequence of the full-length PRO271 polypeptide suggests that it possess significant homology to the proteoglycan link protein, thereby indicating that PRO271 may be a link protein homolog.

EXAMPLE 36: Isolation of cDNA Clones Encoding Human PRO272

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA36460. Based on the DNA36460 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO272.

Forward and reverse PCR primers were synthesized:

forward PCR primer (.f1) 5'-CGCAGGCCCTCATGGCCAGG-3' (SEQ ID NO:222)

forward PCR primer (.f2) 5'-GAAATCCTGGGTAATTGG-3' (SEQ ID NO:223)

reverse PCR primer 5'-GTGCGCGGTGCTCACAGCTCATC-3' (SEQ ID NO:224)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA36460 sequence which had the following nucleotide sequence

hybridization probe

5'-CCCCCCTGAGCGACGCTCCCCCATGATGACGCCCACGGGAACTTC-3' (SEQ ID NO:225)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pairs identified above. A positive library was then used to isolate clones encoding the PRO272 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal lung tissue. DNAsequencing of the clones isolated as described above gave the full-length DNA sequence for PRO272 [herein designated as UNQ239 (DNA40620-1183)] (SEQ ID NO:220) and the derived protein sequence for PRO272.

The entire nucleotide sequence of UNQ239 (DNA40620-1183) is shown in Figure 79 (SEQ ID NO:220). Clone UNQ239 (DNA40620-1183) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 35-37 and ending at the stop codon at nucleotide positions 1019-1021 (Figure 79). The predicted polypeptide precursor is 328 amino acids long (Figure 80). Clone UNQ239 (DNA40620-1183) has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209388.

Analysis of the amino acid sequence of the full-length PRO272 polypeptide suggests that portions of it possess significant homology to the human and mouse reticulocalbin proteins, respectively, thereby indicating that PRO272 may be a novel reticulocalbin protein.

EXAMPLE 37: Isolation of cDNA Clones Encoding Human PRO294

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA35731. Based on the DNA35731 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO294.

Forward and reverse PCR primers were synthesized:

forward PCR primer (.f1) 5'-TGGTCTCGCACACCGATC-3' (SEQ ID NO:228)

forward PCR primer (.f2) 5'-CTGCTGTCCACAGGGGAG-3' (SEQ ID NO:229)

forward PCR primer (.f3) 5'-CCTTGAAGCATACTGCTC-3' (SEQ ID NO:230)

forward PCR primer (.f4) 5'-GAGATAGCAATTTCCGCC-3' (SEQ ID NO:231)

reverse PCR primer (.r1) 5'-TTCCTCAAGAGGGCAGCC-3' (SEQ ID NO:232)

reverse PCR primer (.r2) 5'-CTTGGCACCAATGTCCGAGATTTC-3'

(SEQ ID NO:233)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA35731 sequence which had the following nucleotide sequence

hybridization probe

5'-GCTCTGAGGAAGGTGACGCGCGGGGCCTCCGAACCCTTGGCCTTG-3'

(SEQ ID NO:234)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pairs identified above. A positive library was then used to isolate clones encoding the PRO294 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal brain tissue. DNAsequencing of the clones isolated as described above gave the full-length DNA sequence for PRO294 [herein designated as

UNQ257 (DNA40604-1187)] (SEQ ID NO:226) and the derived protein sequence for PRO294.

The entire nucleotide sequence of UNQ257 (DNA40604-1187) is shown in Figure 81 (SEQ ID NO:226). Clone UNQ257 (DNA40604-1187) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 396-398 and ending at the stop codon at nucleotide positions 2046-2048 (Figure 81). The predicted polypeptide precursor is 550 amino acids long (Figure 82). Clone UNQ257 (DNA40604-1187) has been deposited with ATCC and is assigned ATCC deposit no. 209394.

Analysis of the amino acid sequence of the full-length PRO294 polypeptide suggests that portions of it possess significant homology to portions of various collagen proteins, thereby indicating that PRO294 may be collagen-like molecule.

EXAMPLE 38: Isolation of cDNA Clones Encoding Human PRO295

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA35814. Based on the DNA35814 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO295.

Forward and reverse PCR primers were synthesized:

forward PCR primer (.f1) 5'-GCAGAGCGGAGATGCAGCGGCTTG-3'
(SEQ ID NO:238)

forward PCR primer (.f2) 5'-CCCAGCATGTACTGCCAG-3' (SEQ ID NO:239)

forward PCR primer (.f3) 5'-TTGGCAGCTTCATGGAGG-3' (SEQ ID NO:240)

forward PCR primer (.f4) 5'-CCTGGGCAAAAATGCAAC-3' (SEQ ID NO:241)

reverse PCR primer (.r1) 5'-CTCCAGCTCCTGGCGCACCTCCTC-3' (SEQ ID NO:242)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA35814 sequence which had the following nucleotide sequence

hybridization probe

5'-GGCTCTCAGCTACCGCGCAGGAGCGAGGCCACCCTCAATGAGATG-3'
(SEQ ID NO:243)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pairs identified above. A positive library was then used to isolate clones encoding the PRO295 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal lung tissue. DNAsequencing of the clones isolated as described above gave the full-length DNA sequence for PRO295 [herein designated as UNQ258 (DNA38268-1188)] (SEQ ID NO:235) and the derived protein sequence for PRO295.

The entire nucleotide sequence of UNQ258 (DNA38268-1188) is shown in Figure 83 (SEQ ID NO:235). Clone UNQ258 (DNA38268-1188) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 153-155 and ending at the stop codon at nucleotide positions 1202-1204 (Figure 83). The predicted polypeptide precursor is 350 amino acids long (Figure 84). Clone UNQ258 (DNA38268-1188) has been deposited with ATCC and is assigned ATCC deposit no. 209421.

Analysis of the amino acid sequence of the full-length PRO295 polypeptide suggests that portions of it possess significant homology to the integrin proteins, thereby indicating that PRO295 may be a novel integrin.

EXAMPLE 39: Isolation of cDNA Clones Encoding Human PRO293

The extracellular domain (ECD) sequences (including the secretion signal, if any) of from about 950 known secreted proteins from the Swiss-Prot public protein database were used to search expressed sequence tag (EST) databases. The EST databases included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA). The search was performed using the computer program BLAST or BLAST2 (Altshul et al., Methods in Enzymology 266:460-480 (1996)) as a comparison of the ECD protein sequences to a 6 frame translation of the EST sequence. Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into consensus DNA sequences with the program "phrap" (Phil Green, University of Washington, Seattle, Washington; http://bozeman.mbt.washington.edu/phrap.docs/phrap.html).

Based on an expression tag sequence designated herein as T08294 identified in the above analysis, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO293.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-AACAAGGTAAGATGCCATCCTG-3' (SEQ ID NO:246)

reverse PCR primer 5'-AAACTTGTCGATGGAGACCAGCTC-3' (SEQ ID NO:247)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the expression sequence tag which had the following nucleotide sequence

hybridization probe

5'-AGGGGCTGCAAAGCCTGGAGAGCCTCTCCTTCTATGACAACCAGC-3' (SEQ ID NO:248)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO293 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal brain tissue. DNAsequencing of the clones isolated as described above gave the full-length DNA sequence for PRO293 [herein designated as UNQ256 (DNA37151-1193)] (SEQ ID NO:244) and the derived protein sequence for PRO293.

The entire nucleotide sequence of UNQ256 (DNA37151-1193) is shown in Figures 85A-B (SEQ ID NO:244). Clone UNQ256 (DNA37151-1193) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 881-883 and ending at the stop codon after nucleotide position 3019 of SEQ ID NO:244, Figures 85A-B). The predicted polypeptide precursor is 713 amino acids long (Figure 86). Clone UNQ256 (DNA37151-1193) has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209393.

Analysis of the amino acid sequence of the full-length PRO293 polypeptide suggests that portions of it possess significant homology to the NLRR proteins, thereby indicating that PRO293 may be a novel NLRR protein.

EXAMPLE 40: Isolation of cDNA Clones Encoding Human PRO247

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA33480. Based on the DNA33480 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO247.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-CAACAATGAGGGCACCAAGC-3' (SEQ ID NO:251)

reverse PCR primer 5'-GATGGCTAGGTTCTGGAGGTTCTG-3' (SEQ ID NO:252)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the DNA33480 expression sequence tag which had the following nucleotide sequence

hybridization probe

5'-CAACCTGCAGGAGATTGACCTCAAGGACAACAACCTCAAGACCATCG-3' (SEQ ID NO:253)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO247 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal brain tissue. DNAsequencing of the clones isolated as described above gave the full-length DNA sequence for PRO247 [herein designated as UNQ221 (DNA35673-1201)] (SEQ ID NO:249) and the derived protein sequence for PRO247.

The entire nucleotide sequence of UNQ221 (DNA35673-1201) is shown in Figures 89A-B (SEQ ID NO:249). Clone UNQ221 (DNA35673-1201) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 80-82 of SEQ ID NO:249 and ending at the stop codon after nucleotide position 1717 of SEQ ID NO:249 (Figures 89A-B). The predicted polypeptide precursor is 546 amino acids long (Figure 88). Clone UNQ221 (DNA35673-1201) has been deposited with ATCC and is assigned ATCC deposit no. 209418.

Analysis of the amino acid sequence of the full-length PRO247 polypeptide suggests that portions of it possess significant homology to the densin molecule and KIAA0231, thereby indicating that PRO247 may be a novel leucine rich repeat protein.

EXAMPLE 41: Isolation of cDNA Clones Encoding Human PRO302, PRO303, PRO304, PRO307 and PRO343

Consensus DNA sequences were assembled relative to other EST sequences using phrap as described in Example 1 above. These consensus sequences are herein designated DNA35953, DNA35955, DNA35958, DNA37160 and DNA30895. Based on the DNA35953 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO302.

PCR primers (forward and reverse) were synthesized:

forward PCR primer 1 5'-GTCCGCAAGGATGCCTACATGTTC-3' (SEQ ID NO:264)

forward PCR primer 2 5'-GCAGAGGTGTCTAAGGTTG-3' (SEQ ID NO:265)

reverse PCR primer 5'-AGCTCTAGACCAATGCCAGCTTCC-3' (SEQ ID NO:266)

Also, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA35953 sequence which had the following nucleotide sequence

hybridization probe

5'-GCCACCAACTCCTGCAAGAACTTCTCAGAACTGCCCCTGGTCATG-3' (SEQ ID NO:267)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pairs identified above. A positive library was then used to isolate clones encoding the PRO302 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal kidney tissue (LIB228).

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO302 [herein designated as UNQ265 (DNA40370-1217)] (SEQ ID NO:254) and the derived protein sequence for PRO302.

The entire nucleotide sequence of UNQ265 (DNA40370-1217) is shown in Figure 89 (SEQ ID NO:254). Clone UNQ265 (DNA40370-1217) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 34-36 and ending at the stop codon at nucleotide positions 1390-1392 (Figure 89). The predicted polypeptide precursor is 452 amino acids long (Figure 90). Various unique aspects of the PRO302 protein are shown in Figure 90. Clone UNQ265 (DNA40370-1217) has been deposited with the ATCC on November 21, 1997 and is assigned ATCC deposit no. ATCC 209485.

Based on the DNA35955 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO303.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-GGGGAATTCACCCTATGACATTGCC-3' (SEQ ID NO:268)

reverse PCR primer 5'-GAATGCCCTGCAAGCATCAACTGG-3' (SEQ ID NO:269)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA35955 sequence which had the following nucleotide sequence:

hybridization probe

5'-GCACCTGTCACCTACACTAAACACATCCAGCCCATCTGTCTCCAGGCCTC-3' (SEQ ID NO:270)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pairs identified above. A positive library was then used to isolate clones encoding the PRO303 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal lung tissue (LIB25).

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO303 [herein designated as UNQ266 (DNA42551-1217)] (SEQ ID NO:256) and the derived protein sequence for PRO303.

The entire nucleotide sequence of UNQ266 (DNA42551-1217) is shown in Figure 91 (SEQ ID NO:256). Clone UNQ266 (DNA42551-1217) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 20-22 and ending at the stop codon at nucleotide positions 962-964 (Figure 91). The predicted polypeptide precursor is 314 amino acids long (Figure 92). Various unique aspects of the PRO303 protein are shown in Figure 92. Clone UNQ266 (DNA42551-1217) has been deposited on November 21, 1997 with the ATCC and is assigned ATCC deposit no. ATCC 209483.

Based on the DNA35958 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO304.

Pairs of PCR primers (forward and reverse) were synthesized:

forward PCR primer 1 5'-GCGGAAGGGCAGAATGGGACTCCAAG-3' (SEQ ID NO:271)

forward PCR primer 2 5'-CAGCCCTGCCACATGTGC-3' (SEQ ID NO:272)

forward PCR primer 3 5'-TACTGGGTGGTCAGCAAC-3' (SEQ ID NO:273)

reverse PCR primer 5'-GGCGAAGAGCAGGGTGAGACCCCG-3' (SEQ ID NO:274)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA35958 sequence which had the following nucleotide sequence

hybridization probe

5'-GCCCTCATCCTCTCTGGCAAATGCAGTTACAGCCCGGAGCCCGAC-3' (SEQ ID NO:275)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pairs identified above. A positive library was then used to isolate clones encoding the PRO304 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from 22 week human fetal brain tissue (LIB153).

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO304 [herein designated as UNQ267 (DNA39520-1217)] (SEQ ID NO:258) and the derived protein sequence for PRO304.

The entire nucleotide sequence of UNQ267 (DNA39520-1217) is shown in Figure 93 (SEQ ID NO:258). Clone UNQ267 (DNA39520-1217) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 34-36 and ending at the stop codon at nucleotide positions 1702-1704 (Figure 93). The predicted polypeptide precursor is 556 amino acids long (Figure 94). Various unique aspects of the PRO304 protein are shown in Figure 94. Clone UNQ267 (DNA39520-1217) has been deposited with ATCC on November 21, 1997 and is assigned ATCC deposit no. ATCC 209482.

Based on the DNA37160 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO307.

Pairs of PCR primers (forward and reverse) were synthesized:

forward PCR primer 1 5'-GGGCAGGGATTCCAGGGCTCC-3' (SEQ ID NO:276)

forward PCR primer 2 5'-GGCTATGACAGCAGGTTC-3' (SEQ ID NO:277)

forward PCR primer 3 5'-TGACAATGACCGACCAGG-3' (SEQ ID NO:278)

reverse PCR primer 5'-GCATCGCATTGCTGGTAGAGCAAG-3' (SEQ ID NO:279)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA37160 sequence which had the following nucleotide sequence

hybridization probe

5'-TTACAGTGCCCCCTGGAAACCCACTTGGCCTGCATACCGCCTCCC-3' (SEQ ID NO:280)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pairs identified above. A positive library was then used to isolate clones

encoding the PRO307 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal liver tissue (LIB229).

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO307 [herein designated as UNQ270 (DNA41225-1217)] (SEQ ID NO:260) and the derived protein sequence for PRO307.

The entire nucleotide sequence of UNQ270 (DNA41225-1217) is shown in Figure 95 (SEQ ID NO:260). Clone UNQ270 (DNA41225-1217) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 92-94 and ending at the stop codon at nucleotide positions 1241-1243 (Figure 95). The predicted polypeptide precursor is 383 amino acids long (Figure 96). Various unique aspects of the PRO307 protein are shown in Figure 96. Clone UNQ270 (DNA41225-1217) has been deposited with ATCC on November 21, 1997 and is assigned ATCC deposit no. ATCC 209491.

Based on the DNA30895 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO343.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-CGTCTCGAGCGCTCCATACAGTTCCCTTGCCCCA-3' (SEQ ID NO:281)

reverse PCR primer

5'-TGGAGGGGGAGCGGGATGCTTGTCTGGGCGACTCCGGGGGCC

CCCTCATGTGCCAGGTGGA-3' (SEQ ID NO:282)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA30895 sequence which had the following nucleotide sequence

hybridization probe

5'-CCCTCAGACCCTGCAGAAGCTGAAGGTTCCTATCATCGAC

TCGGAAGTCTGCAGCCATCTGTACTGGCGGGGAGCAGGACAGGGACCCATCACTGAGGACATGCTGT

GTGCCGGCTACT-3' (SEQ ID NO:283)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pairs identified above. A positive library was then used to isolate clones encoding the PRO343 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal lung tissue (LIB26).

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO343 [herein designated as UNQ302 (DNA43318-1217)] (SEQ ID NO:262) and the derived protein sequence for PRO343.

The entire nucleotide sequence of UNQ302 (DNA43318-1217) is shown in Figure 97 (SEQ ID NO:262). Clone UNQ302 (DNA43318-1217) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 53-55 and ending at the stop codon at nucleotide positions 1004-1006 (Figure 97). The predicted polypeptide precursor is 317 amino acids long (Figure 98). Various unique aspects of the PRO343 protein are shown in Figure 98. Clone UNQ302 (DNA43318-1217) has been deposited with ATCC on November 21, 1997 and is assigned ATCC deposit no. ATCC 209481.

EXAMPLE 42: Isolation of cDNA Clones Encoding Human PRO328

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA35615. Based on the DNA35615 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO328.

Forward and reverse PCR primers were synthesized:

forward PCR primer	5'-TCCTGCAGTTTCCTGATGC-3'	(SEQ ID NO:286)
reverse PCR primer	5'-CTCATATTGCACACCAGTAATTCG-3'	(SEQ ID NO:287)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA35615 sequence which had the following nucleotide sequence

hybridization probe

5'-ATGAGGAGAAACGTTTGATGGTGGAGCTGCACAACCTCTACCGGG-3'

(SEQ ID NO:288)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO328 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal kidney tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO328 [herein designated as UNQ289 (DNA40587-1231)] (SEQ ID NO:284) and the derived protein sequence for PRO328.

The entire nucleotide sequence of UNQ289 (DNA40587-1231) is shown in Figure 99 (SEQ ID NO:284). Clone UNQ289 (DNA40587-1231) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 15-17 and ending at the stop codon at nucleotide positions 1404-1406 (Figure 99). The predicted polypeptide precursor is 463 amino acids long (Figure 100). Clone UNQ289 (DNA40587-1231) has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209438.

Analysis of the amino acid sequence of the full-length PRO328 polypeptide suggests that portions of it possess significant homology to the human glioblastoma protein and to the cysteine rich secretory protein thereby indicating that PRO328 may be a novel glioblastoma protein or cysteine rich secretory protein.

EXAMPLE 43: Isolation of cDNA Clones Encoding Human PRO335, PRO331 or PRO326

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA36685. Based on the DNA36685 consensus sequence, and Incyte EST sequence no. 2228990, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO335, PRO331 or PRO326.

Forward and reverse PCR primers were synthesized for the determination of PRO335:

forward PCR primer	5'-GGAACCGAATCTCAGCTA-3'	(SEQ ID NO:295)
forward PCR primer	5'-CCTAAACTGAACTGGACCA-3'	(SEQ ID NO:296)
forward PCR primer	5'-GGCTGGAGACACTGAACCT-3'	(SEQ ID NO:297)

forward PCR primer 5'-ACAGCTGCACAGCTCAGAACAGTG-3' (SEQ ID NO:298)

reverse PCR primer 5'-CATTCCCAGTATAAAAATTTTC-3' (SEQ ID NO:299)

reverse PCR primer 5'-GGGTCTTGGTGAATGAGG-3' (SEQ ID NO:300)

reverse PCR primer 5'-GTGCCTCTCGGTTACCACCAATGG-3' (SEQ ID NO:301)

Additionally, a synthetic oligonucleotide hybridization probe was constructed for the determination of PRO335 which had the following nucleotide sequence

hybridization probe

5'-GCGGCCACTGTTGGACCGAACTGTAACCAAGGGAGAAACAGCCGTCCTAC-3'

(SEQ ID NO:302)

Forward and reverse PCR primers were synthesized for the determination of PRO331:

forward PCR primer 5'-GCCTTTGACAACCTTCAGTCACTAGTGG-3' (SEQ ID NO:303)

reverse PCR primer 5'-CCCCATGTGTCCATGACTGTTCCC-3' (SEQ ID NO:304)

Additionally, a synthetic oligonucleotide hybridization probe was constructed for the determination of PRO331 which had the following nucleotide sequence

hybridization probe

5'-TACTGCCTCATGACCTCTTCACTCCCTTGCATCATCTTAGAGCGG-3'

(SEQ ID NO:305)

Forward and reverse PCR primers were synthesized for the determination of PRO326:

forward PCR primer 5'-ACTCCAAGGAAATCGGATCCGTTC-3' (SEQ ID NO:306)

reverse PCR primer 5'-TTAGCAGCTGAGGATGGGCACAAC-3' (SEQ ID NO:307)

Additionally, a synthetic oligonucleotide hybridization probe was constructed for the determination of PRO331 which had the following nucleotide sequence

hybridization probe

5'-GCCTTCACTGGTTTGGATGCATTGGAGCATCTAGACCTGAGTGACAACGC-3'

(SEQ ID NO:308)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pairs identified above. A positive library was then used to isolate clones encoding the PRO335, PRO331 or PRO326 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal kidney tissue (PRO335 and PRO326) and human fetal brain (PRO331).

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO335, PRO331 or PRO326 [herein designated as SEQ ID NOS:289, 291 and 293, respectively; see Figures 103A-B, 105 and 107, respectively], and the derived protein sequence for PRO335, PRO331 or PRO326 (see Figures 104, 106 and 108, respectively; SEQ ID NOS:290, 292 and 294, respectively).

The entire nucleotide sequences are shown in Figures 103A-B, 105 and 107, deposited with the ATCC on June 2, 1998, November 7, 1997 and November 21, 1997, respectively.

Analysis of the amino acid sequence of the full-length PRO335, PRO331 or PRO326 polypeptide suggests that portions of it possess significant homology to the LIG-1 protein, thereby indicating that PRO335, PRO331 and

PRO326 may be a novel LIG-1-related protein.

EXAMPLE 44: Isolation of cDNA clones Encoding Human PRO332

Based upon an ECD homology search performed as described in Example 1 above, a consensus DNA sequence designated herein as DNA36688 was assembled. Based on the DNA36688 consensus sequence, oligonucleotides were synthesized to identify by PCR a cDNA library that contained the sequence of interest and for use as probes to isolate a clone of the full-length coding sequence for PRO332.

A pair of PCR primers (forward and reverse) were synthesized:

5'-GCATTGGCCGCGAGACTTTGCC-3' (SEQ ID NO:311)

5'-GCGGCCACGGTCCTTGGAAATG-3' (SEQ ID NO:312)

A probe was also synthesized:

5'-TGGAGGAGCTCAACCTCAGCTACAACCGCATCACCAGCCCACAGG-3'

(SEQ ID NO:313)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO332 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from a human fetal liver library (LIB229).

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for DNA40982-1235 and the derived protein sequence for PRO332.

The entire nucleotide sequence of DNA40982-1235 is shown in Figures 109A-B (SEQ ID NO:309). Clone DNA40982-1235 contains a single open reading frame (with an apparent translational initiation site at nucleotide positions 342-344, as indicated in Figures 109A-B). The predicted polypeptide precursor is 642 amino acids long, and has a calculated molecular weight of 72,067 (pI: 6.60). Clone DNA40982-1235 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209433.

Based on a BLAST and FastA sequence alignment analysis of the full-length sequence, PRO332 shows about 30-40% amino acid sequence identity with a series of known proteoglycan sequences, including, for example, fibromodulin and fibromodulin precursor sequences of various species (FMOD_BOVIN, FMOD_CHICK, FMOD RAT, FMOD_MOUSE, FMOD_HUMAN, P_R36773), osteomodulin sequences (AB000114_1, AB007848_1), decorin sequences (CFU83141_1, OCU03394_1, P_R42266, P_R42267, P_R42260, P_R89439), keratan sulfate proteoglycans (BTU48360_1, AF022890_1), corneal proteoglycan (AF022256_1), and bone/cartilage proteoglycans and proteoglycane precursors (PGS1_BOVIN, PGS2_ MOUSE, PGS2_HUMAN).

EXAMPLE 45: Isolation of cDNA clones Encoding Human PRO334

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. Based on the consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO334.

Forward and reverse PCR primers were synthesized for the determination of PRO334:

forward PCR primer 5'-GATGGTTCCTGCTCAAGTGCCCTG-3' (SEQ ID NO:316)

reverse PCR primer 5'-TTGCACTTGTAGGACCCACGTACG-3' (SEQ ID NO:317)

Additionally, a synthetic oligonucleotide hybridization probe was constructed for the determination of PRO334 which had the following nucleotide sequence

hybridization probe

5'-CTGATGGGAGGACCTGTGTAGATGTTGATGAATGTGCTACAGGAAGAGCC-3'

(SEQ ID NO:318)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO334 gene using the probe oligonucleotide and one of the PCR primers.

Human fetal kidney cDNA libraries used to isolate the cDNA clones were constructed by standard methods using commercially available reagents such as those from Invitrogen, San Diego, CA.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO334 [herein designated as DNA41379-1236] (SEQ ID NO:314) and the derived protein sequence for PRO334.

The entire nucleotide sequence of DNA41379-1236 (also referred to as UNQ295) is shown in Figure 109 (SEQ ID NO:314). Clone DNA41379-1236 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 203-205 and ending at the stop codon at nucleotide positions 1730-1732 (Figure 109). The predicted polypeptide precursor is 509 amino acids long (Figure 110). Clone DNA41379-1236 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209488.

Analysis of the amino acid sequence of the full-length PRO334 polypeptide suggests that portions of it possess significant homology to the fibulin and fibrillin proteins, thereby indicating that PRO334 may be a novel member of the EGF protein family.

EXAMPLE 46: Isolation of cDNA Clones Encoding Human PRO346

A consensus DNA sequence was identified using phrap as described in Example 1 above. Specifically, this consensus sequence is herein designated DNA38240. Based on the DNA38240 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length PRO346 coding sequence.

RNA for construction of the cDNA libraries was isolated from human fetal liver. The cDNA libraries used to isolated the cDNA clones were constructed by standard methods using commercially available reagents (e.g., Invitrogen, San Diego, CA; Clontech, etc.) The cDNA was primed with oligo dT containing a NotI site, linked with blunt to SalI hemikinased adaptors, cleaved with NotI, sized appropriately by gel electrophoresis, and cloned in a defined orientation into a suitable cloning vector (such as pRKB or pRKD; pRK5B is a precursor of pRK5D that does not contain the SfiI site; see, Holmes et al., *Science*, 253:1278-1280 (1991)) in the unique XhoI and NotI sites.

A cDNA clone was sequenced in entirety. The entire nucleotide sequence of DNA44167-1243 is shown in Figure 111 (SEQ ID NO:319). Clone DNA44167-1243 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 64-66 (Fig. 113; SEQ ID NO:319). The predicted polypeptide

precursor is 450 amino acids long. Clone DNA44167-1243 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209434 (designation DNA44167-1243).

Based on a BLAST, BLAST-2 and FastA sequence alignment analysis (using the ALIGN computer program) of the full-length sequence, PRO346 shows amino acid sequence identity to carcinoembryonic antigen (28%).

The oligonucleotide sequences used in the above procedure were the following:

OLI2691 (38240.f1)

5'-GATCCTGTCACAAAGCCAGTGGTGC-3' (SEQ ID NO:321)

OLI2693 (38240.r1)

5'-CACTGACAGGGTTCCTCACCCAGG-3' (SEQ ID NO:322)

OLI2692 (38240.p1)

5'-CTCCCTCTGGGCTGTGGAGTATGTGGGGAACATGACCCTGACATG-3' (SEQ ID NO:323)

EXAMPLE 47: Isolation of cDNA Clones Encoding Human PRO268

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA35698. Based on the DNA35698 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO268.

Forward and reverse PCR primers were synthesized:

forward PCR primer 1 5'-TGAGGTGGGCAAGCGGCGAAATG-3' (SEQ ID NO:326)

forward PCR primer 2 5'-TATGTGGATCAGGACGTGCC-3' (SEQ ID NO:327)

forward PCR primer 3 5'-TGCAGGGTTCAGTCTAGATTG-3' (SEQ ID NO:328)

reverse PCR primer 5'-TTGAAGGACAAAGGCAATCTGCCAC-3' (SEQ ID NO:329)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA35698 sequence which had the following nucleotide sequence

hybridization probe

5'-GGAGTCTTGCAGTTCCCCTGGCAGTCCTGGTGCTGTTGCTTTGGG-3' (SEQ ID NO:330)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO268 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal lung tissue. DNAsequencing of the clones isolated as described above gave the full-length DNA sequence for PRO268 [herein designated as UNQ235 (DNA39427-1179)] (SEQ ID NO:324) and the derived protein sequence for PRO268.

The entire nucleotide sequence of UNQ235 (DNA39427-1179) is shown in Figure 113 (SEQ ID NO:324). Clone UNQ235 (DNA39427-1179) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 13-15 and ending at the stop codon at nucleotide positions 853-855 (Figure 113). The predicted polypeptide precursor is 280 amino acids long (Figure 114). Clone UNQ235 (DNA39427-1179) has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209395.

Analysis of the amino acid sequence of the full-length PRO268 polypeptide suggests that it possess significant homology to protein disulfide isomerase, thereby indicating that PRO268 may be a novel protein disulfide isomerase.

EXAMPLE 48: Isolation of cDNA Clones Encoding Human PRO330

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA35730. Based on the DNA35730 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO330.

Forward and reverse PCR primers were synthesized:

forward PCR primer 1 5'-CCAGGCACAATTTCCAGA-3' (SEQ ID NO:333)

forward PCR primer 2 5'-GGACCCTTCTGTGTGCCAG-3' (SEQ ID NO:334)

reverse PCR primer 1 5'-GGTCTCAAGAACTCCTGTC-3' (SEQ ID NO:335)

reverse PCR primer 2 5'-ACACTCAGCATTGCCTGGTACTTG-3' (SEQ ID NO:336)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus sequence which had the following nucleotide sequence

hybridization probe

5'-GGGCACATGACTGACCTGATTTATGCAGAGAAAGAGCTGGTGCAG-3' (SEQ ID NO:337)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO330 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal liver tissue. DNAsequencing of the clones isolated as described above gave the full-length DNA sequence for PRO330 [herein designated as UNQ290 (DNA40603-1232)] (SEQ ID NO:331) and the derived protein sequence for PRO330.

The entire nucleotide sequence of UNQ290 (DNA40603-1232) is shown in Figure 115 (SEQ ID NO:331). Clone UNQ290 (DNA40603-1232) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 167-169 and ending at the stop codon at nucleotide positions 1766-1768 (Figure 115). The predicted polypeptide precursor is 533 amino acids long (Figure 116). Clone UNQ290 (DNA40603-1232) has been deposited with ATCC and is assigned ATCC deposit no.ATCC 209486 on November 21, 1997.

Analysis of the amino acid sequence of the full-length PRO330 polypeptide suggests that portions of it possess significant homology to the mouse prolyl 4-hydroxylase alpha subunit protein, thereby indicating that PRO330 may be a novel prolyl 4-hydroxylase alpha subunit polypeptide.

EXAMPLE 49: Isolation of cDNA Clones Encoding Human PRO310

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA40553. Based on the DNA40553 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO310.

Forward and reverse PCR primers were synthesized:

forward PCR primer 1 5'-TCCCCAAGCCGTTCTAGACGCGG-3' (SEQ ID NO:342)

forward PCR primer 2 5'-CTGGTTCTTCCTTGCACG-3' (SEQ ID NO:343)

reverse PCR primer 5'-GCCCAAATGCCCTAAGGCGGTATACCCC-3' (SEQ ID NO:344)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus sequence which had the following nucleotide sequence

hybridization probe

5'-GGGTGTGATGCTTGGAAGCATTTTCTGTGCTTTGATCACTATGCTAGGAC-3' (SEQ ID NO:345)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO310 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal liver tissue. DNAsequencing of the clones isolated as described above gave the full-length DNA sequence for PRO310 [herein designated as DNA43046-1225 (SEQ ID NO:340) and the derived protein sequence for PRO310 (SEQ ID NO:341).

The entire nucleotide sequence of DNA43046-1225 is shown in Figure 119 (SEQ ID NO:340). Clone DNA43046-1225 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 81-83 and ending at the stop codon at nucleotide positions 1035-1037 (Figure 119). The predicted polypeptide precursor is 318 amino acids long (Figure 120) and has a calculated molecular weight of approximately 36,382 daltons. Clone DNA43046-1225 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209484.

Analysis of the amino acid sequence of the full-length PRO310 polypeptide suggests that portions of it possess homology to *C. elegans* proteins and to fringe, thereby indicating that PRO310 may be involved in development.

EXAMPLE 50: Isolation of cDNA clones Encoding Human PRO339

An expressed sequence tag (EST) DNA database (LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA) was searched and ESTs were identified. An assembly of Incyte clones and a consensus sequence was formed using phrap as described in Example 1 above.

Forward and reverse PCR primers were synthesized based upon the assembly-created consensus sequence:

forward PCR primer 1 5'-GGGATGCAGGTGGTGTCTCATGGGG-3' (SEQ ID NO:346)

forward PCR primer 2 5'-CCCTCATGTACCGGCTCC-3' (SEQ ID NO:347)

forward PCR primer 3 5'-GTGTGACACAGCGTGGGC-3' (SEQ ID NO:43)

forward PCR primer 4 5'-GACCGGCAGGCTTCTGCG-3' (SEQ ID NO:44)

reverse PCR primer 1 5'-CAGCAGCTTCAGCCACCAGGAGTGG-3' (SEQ ID NO:45)

reverse PCR primer 2 5'-CTGAGCCGTGGGCTGCAGTCTCGC-3' (SEQ ID NO:46)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus sequence which had the following nucleotide sequence

hybridization probe

5'-CCGACTACGACTGGTTCTTCATCATGCAGGATGACACATATGTGC-3' (SEQ ID NO:47)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pairs identified above. A positive library was then used to isolate clones encoding the PRO339 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal liver tissue. A cDNA clone was sequenced in entirety. The entire nucleotide sequence of DNA43466-1225 is shown in Figure 117 (SEQ ID NO:338). Clone DNA43466-1225 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 333-335 and ending at the stop codon found at nucleotide positions 2649-2651 (Figure 117; SEQ ID NO:338). The predicted polypeptide precursor is 772 amino acids long and has a calculated molecular weight of approximately 86,226 daltons. Clone DNA43466-1225 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209490.

Based on a BLAST and FastA sequence alignment analysis (using the ALIGN computer program) of the full-length sequence, PRO339 has homology to *C. elegans* proteins and collagen-like polymer sequences as well as to fringe, thereby indicating that PRO339 may be involved in development or tissue growth.

EXAMPLE 51: Isolation of cDNA Clones Encoding Human PRO244

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. Based on this consensus sequence, oligonucleotides were synthesized to identify by PCR a cDNA library that contained the sequence of interest and for use as probes to isolate a clone of the full-length coding sequence for PRO244.

A pair of PCR primers (forward and reverse) were synthesized:

5'-TTCAGCTTCTGGGATGTAGGG-3' (30923.f1) (SEQ ID NO:378)

5'-TATTCCTACCATTTCACAAATCCG-3' (30923.r1) (SEQ ID NO:379)

A probe was also synthesized:

5'-GGAGGACTGTGCCACCATGAGAGACTCTTCAAACCCAAGGCAAAATTGG-3' (30923.p1) (SEQ ID NO:380)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO244 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from a human fetal kidney library. DNA sequencing of the clones isolated as described above gave the full-length DNA sequence and the derived protein sequence for PRO244.

The entire nucleotide sequence of PRO244 is shown in Figure 121 (SEQ ID NO:376). Clone DNA35668-1171 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 106-108 (Fig. 121). The predicted polypeptide precursor is 219 amino acids long. Clone DNA35668-1171 has been deposited with ATCC (designated as DNA35663-1171) and is assigned ATCC deposit no. ATCC209371. The protein has a cytoplasmic domain (aa 1-20), a transmembrane domain (aa 21-46), and an extracellular domain (aa 47-219), with

a C-lectin domain at aa 55-206.

Based on a BLAST and FastA sequence alignment analysis of the full-length sequence, PRO244 shows notable amino acid sequence identity to hepatic lectin gallus gallus (43%), HIC hp120-binding C-type lectin (42%), macrophage lectin 2 (HUMHML2-1, 41%), and sequence PR32188 (44%).

EXAMPLE 52: Use of PRO Polypeptide-Encoding Nucleic Acid as Hybridization Probes

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

DNA comprising the coding sequence of of a PRO polypeptide of interest as disclosed herein may be employed as a probe or used as a basis from which to prepare probes to screen for homologous DNAs (such as those encoding naturally-occurring variants of the PRO polypeptide) 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 polypeptide-encoding nucleic acid-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 polypeptide can then be identified using standard techniques known in the art.

EXAMPLE 53: Expression of PRO Polypeptides in *E. coli*

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

The DNA sequence encoding the desired PRO polypeptide is initially amplified using selected PCR primers. The primers 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 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 specific PRO polypeptide 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 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 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 polypeptide can then be purified using a metal chelating column under conditions that allow tight binding of the protein.

PRO187, PRO317, PRO301, PRO224 and PRO238 were successfully expressed in *E. coli* in a poly-His tagged form, using the following procedure. The DNA encoding PRO187, PRO317, PRO301, PRO224 or PRO238 was initially amplified using selected PCR primers. The primers contained 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 were then ligated into an expression vector, which was used to transform an *E. coli* host based on strain 52 (W3110 fuhA(tonA) lon galE rpoHts(htpRts) clpP(lacIq). Transformants were first grown in LB containing 50 mg/ml carbenicillin at 30°C with shaking until an O.D.600 of 3-5 was reached. Cultures were then diluted 50-100 fold into CRAP media (prepared by mixing 3.57 g $(NH_4)_2SO_4$, 0.71 g sodium citrate·2H2O, 1.07 g KCl, 5.36 g Difco yeast extract, 5.36 g Sheffield hycase SF in 500 mL water, as well as 110 mM MPOS, pH 7.3, 0.55% (w/v) glucose and 7 mM $MgSO_4$) and grown for approximately 20-30 hours at 30°C with shaking. Samples were removed to verify expression by SDS-PAGE analysis, and the bulk culture is centrifuged to pellet the cells. Cell pellets were frozen until purification and refolding.

E. coli paste from 0.5 to 1 L fermentations (6-10 g pellets) was 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 was stirred overnight at 4°C. This step results in a denatured protein with all cysteine residues blocked by sulfitolization. The solution was centrifuged at 40,000 rpm in a Beckman Ultracentifuge for 30 min. The supernatant was 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. Depending the clarified extract was loaded onto a 5 ml Qiagen Ni-NTA metal chelate column equilibrated in the metal chelate column buffer. The column was washed with additional buffer containing 50 mM imidazole (Calbiochem, Utrol grade), pH 7.4. The protein was eluted with buffer containing 250 mM imidazole. Fractions containing the desired protein were pooled and stored at 4°C. Protein concentration was estimated by its absorbance at 280 nm using the calculated extinction coefficient based on its amino acid sequence.

The proteins were refolded by diluting 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 EDTA. Refolding volumes were chosen so that the final protein concentration was between 50 to 100 micrograms/ml. The refolding solution was stirred gently at 4°C for 12-36 hours. The refolding reaction was quenched by the addition of TFA to a final concentration of 0.4% (pH of approximately 3). Before further purification of the protein, the solution was filtered through a 0.22 micron filter and acetonitrile was added to 2-10% final concentration. The refolded protein was 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 were analyzed on SDS polyacrylamide gels and fractions containing homogeneous refolded protein were pooled. Generally, the properly refolded species of most 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 PRO187, PRO317, PRO301, PRO224 and PRO238 proteins, respectively, were pooled and the acetonitrile removed using a gentle stream of nitrogen directed at the solution. Proteins were 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 (Pharmacia) resins equilibrated in the formulation buffer and sterile filtered.

EXAMPLE 54: Expression of PRO Polypeptides in Mammalian Cells

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

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

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 polypeptide 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 $CaCl_2$. To this mixture is added, dropwise, 500 μl of 50 mM HEPES (pH 7.35), 280 mM NaCl, 1.5 mM $NaPO_4$, and a precipitate is allowed to form for 10 minutes 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.

Approximately 24 hours after the transfections, the culture medium is removed and replaced with culture medium (alone) or culture medium containing 200 μCi/ml ^{35}S-cysteine and 200 μCi/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 further incubation (in serum free medium) and the medium is tested in selected bioassays.

In an alternative technique, PRO polypeptide may be introduced into 293 cells transiently using the dextran sulfate method described by Somparyrac et al., Proc. Natl. Acad. Sci., 12:7575 (1981). 293 cells are grown to maximal density in a spinner flask and 700 μg pRK5-PRO polypeptide DNA is added. The cells are first concentrated from the spinner flask by centrifugation and washed with PBS. The DNA-dextran precipitate is incubated 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 μg/ml bovine insulin and 0.1 μg/ml bovine transferrin. After about four days, the conditioned media is centrifuged and filtered to remove cells and debris. The sample containing expressed PRO polypeptide can then be concentrated and purified by any selected

method, such as dialysis and/or column chromatography.

In another embodiment, PRO polypeptides can be expressed in CHO cells. The pRK5-PRO polypeptide can be transfected into CHO cells using known reagents such as $CaPO_4$ 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 the conditioned medium is harvested. The medium containing the expressed PRO polypeptide can then be concentrated and purified by any selected method.

Epitope-tagged PRO polypeptide may also be expressed in host CHO cells. The PRO polypeptide 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 polypeptide insert can then be subcloned into a SV40 driven 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 driven vector. Labeling may be performed, as described above, to verify expression. The culture medium containing the expressed poly-His tagged PRO polypeptide can then be concentrated and purified by any selected method, such as by Ni^{2+}-chelate affinity chromatography.

PRO211, PRO217, PRO230, PRO219, PRO245, PRO221, PRO258, PRO301, PRO224, PRO222, PRO234, PRO229, PRO223, PRO328 and PRO332 were successfully expressed in CHO cells by both a transient and a stable expression procedure. In addition, PRO232, PRO265, PRO246, PRO228, PRO227, PRO220, PRO266, PRO269, PRO287, PRO214, PRO231, PRO233, PRO238, PRO244, PRO235, PRO236, PRO262, PRO239, PRO257, PRO260, PRO263, PRO270, PRO271, PRO272, PRO294, PRO295, PRO293, PRO247, PRO303 and PRO268 were successfully transiently expressed in CHO cells.

Stable expression in CHO cells was performed using the following procedure. The proteins were expressed as an IgG construct (immunoadhesin), in which the coding sequences for the soluble forms (e.g. extracellular domains) of the respective proteins were 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 were subcloned in a CHO expression vector using standard techniques as described in Ausubel et al., *Current Protocols of Molecular Biology*, Unit 3.16, 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.

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

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

For the poly-His tagged constructs, the proteins were purified using a Ni-NTA column (Qiagen). Before purification, imidazole was added to the conditioned media to a concentration of 5 mM. The conditioned media was 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 was washed with additional equilibration buffer and the protein eluted with equilibration buffer containing 0.25 M imidazole. The highly purified protein was subsequently desalted 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.

Immunoadhesin (Fc containing) constructs of were purified from the conditioned media as follows. The conditioned medium was 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 was washed extensively with equilibration buffer before elution with 100 mM citric acid, pH 3.5. The eluted protein was immediately neutralized by collecting 1 ml fractions into tubes containing 275 μL of 1 M Tris buffer, pH 9. The highly purified protein was subsequently desalted into storage buffer as described above for the poly-His tagged proteins. The homogeneity was assessed by SDS polyacrylamide gels and by N-terminal amino acid sequencing by Edman degradation.

PRO211, PRO217, PRO230, PRO232, PRO187, PRO265, PRO219, PRO246, PRO228, PRO533, PRO245, PRO221, PRO227, PRO220, PRO258, PRO266, PRO269, PRO287, PRO214, PRO317, PRO301, PRO224, PRO222, PRO234, PRO231, PRO229, PRO233, PRO238, PRO223, PRO235, PRO236, PRO262, PRO239, PRO257, PRO260, PRO263, PRO270, PRO271, PRO272, PRO294, PRO295, PRO293, PRO247, PRO304, PRO302, PRO307, PRO303, PRO343, PRO328, PRO326, PRO331, PRO332, PRO334, PRO346, PRO268, PRO330, PRO310 and PRO339 were also successfully transiently expressed in COS cells.

EXAMPLE 55: Expression of PRO Polypeptides in Yeast

The following method describes recombinant expression of a desired PRO polypeptide in yeast.

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

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.

Recombinant PRO polypeptide 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 the PRO polypeptide may further be purified using selected column chromatography resins.

EXAMPLE 56: Expression of PRO Polypeptides in Baculovirus-Infected Insect Cells

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

The desired PRO polypeptide is fused upstream of an epitope tag contained with 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 PRO polypeptide or the desired portion of the PRO polypeptide (such as the sequence encoding the extracellular domain of a transmembrane protein) is amplified by PCR with primers complementary to the 5' and 3' regions. The 5' primer may incorporate flanking (selected) restriction enzyme sites. The product is then digested with those selected restriction enzymes and subcloned into the expression vector.

Recombinant baculovirus is generated by co-transfecting the above plasmid and BaculoGold™ 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 is performed as described by O'Reilley et al., Baculovirus expression vectors: A laboratory Manual, Oxford: Oxford University Press (1994).

Expressed poly-his tagged PRO polypeptide can then be purified, for example, by Ni^{2+}-chelate affinity chromatography as follows. Extracts are prepared from recombinant virus-infected Sf9 cells as described by 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 $MgCl_2$; 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 μm filter. A Ni^{2+}-NTA agarose column (commercially available from Qiagen) is prepared 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 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 Ni^{2+}-NTA-conjugated to alkaline phosphatase (Qiagen). Fractions containing the eluted His_{10}-tagged PRO polypeptide are pooled and dialyzed against loading buffer.

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

PRO211, PRO217, PRO230, PRO187, PRO265, PRO246, PRO228, PRO533, PRO245, PRO221, PRO220, PRO258, PRO266, PRO269, PRO287, PRO214, PRO301, PRO224, PRO222, PRO234, PRO231, PRO229, PRO235, PRO239, PRO257, PRO272, PRO294, PRO295, PRO328, PRO326, PRO331, PRO334, PRO346 and PRO310 were successfully expressed in baculovirus infected Sf9 or high5 insect cells. While the expression was actually performed in a 0.5-2 L scale, it can be readily scaled up for larger (e.g. 8 L) preparations. The proteins were expressed as an IgG construct (immunoadhesin), in which the protein extracellular region was fused to an IgG1 constant region sequence containing the hinge, CH2 and CH3 domains and/or in poly-His tagged forms.

Following PCR amplification, the respective coding sequences were subcloned into a baculovirus expression vector (pb.PH.IgG for IgG fusions and pb.PH.His.c for poly-His tagged proteins), and the vector and Baculogold® baculovirus DNA (Pharmingen) were co-transfected into 105 *Spodoptera frugiperda* ("Sf9") cells (ATCC CRL 1711), using Lipofectin (Gibco BRL). pb.PH.IgG and pb.PH.His are modifications of the commercially available baculovirus expression vector pVL1393 (Pharmingen), with modified polylinker regions to include the His or Fc tag sequences. The cells were grown in Hink's TNM-FH medium supplemented with 10% FBS (Hyclone). Cells were incubated for 5 days at 28°C. The supernatant was harvested and subsequently used for the first viral amplification by infecting Sf9 cells in Hink's TNM-FH medium supplemented with 10% FBS at an approximate multiplicity of infection (MOI) of 10. Cells were incubated for 3 days at 28°C. The supernatant was harvested and the expression of the constructs in the baculovirus expression vector was determined by batch binding of 1 ml of supernatant to 25 mL of Ni-NTA beads (QIAGEN) for histidine tagged proteins or Protein-A Sepharose CL-4B beads (Pharmacia) for IgG tagged proteins followed by SDS-PAGE analysis comparing to a known concentration of protein standard by Coomassie blue staining.

The first viral amplification supernatant was used to infect a spinner culture (500 ml) of Sf9 cells grown in ESF-921 medium (Expression Systems LLC) at an approximate MOI of 0.1. Cells were incubated for 3 days at 28°C. The supernatant was harvested and filtered. Batch binding and SDS-PAGE analysis was repeated, as necessary, until expression of the spinner culture was confirmed.

The conditioned medium from the transfected cells (0.5 to 3 L) was harvested by centrifugation to remove the cells and filtered through 0.22 micron filters. For the poly-His tagged constructs, the protein construct were purified using a Ni-NTA column (Qiagen). Before purification, imidazole was added to the conditioned media to a concentration of 5 mM. The conditioned media were 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 was washed with additional equilibration buffer and the protein eluted with equilibration buffer containing 0.25 M imidazole. The highly purified protein was subsequently desalted 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.

Immunoadhesin (Fc containing) constructs of proteins were purified from the conditioned media as follows. The conditioned media were 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 was washed extensively with equilibration buffer before elution with 100 mM citric acid, pH 3.5. The eluted protein was immediately neutralized by collecting 1 ml fractions into tubes containing 275 mL of 1 M Tris buffer, pH 9. The highly purified protein was subsequently desalted into storage buffer as described above for the poly-His tagged proteins. The homogeneity of the proteins was verified by SDS polyacrylamide gel (PEG) electrophoresis and N-terminal amino acid sequencing by Edman degradation.

EXAMPLE 57: Preparation of Antibodies that Bind to PRO Polypeptides

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

Techniques for producing the monoclonal antibodies are known in the art and are described, for instance, in Goding, supra. Immunogens that may be employed include purified PRO polypeptide, fusion proteins containing the PRO polypeptide, and cells expressing recombinant PRO polypeptide 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 polypeptide immunogen emulsified in complete Freund's adjuvant and injected subcutaneously or intraperitoneally in an amount from 1-100 micrograms. 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 polypeptide antibodies.

After a suitable antibody titer has been detected, the animals "positive" for antibodies can be injected with a final intravenous injection of PRO polypeptide. 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 the PRO polypeptide. Determination of "positive" hybridoma cells secreting the desired monoclonal antibodies against the PRO polypeptide is within the skill in the art.

The positive hybridoma cells can be injected intraperitoneally into syngeneic Balb/c mice to produce ascites containing the anti-PRO polypeptide 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 chromatography. Alternatively, affinity chromatography based upon binding of antibody to protein A or protein G can be employed.

EXAMPLE 58: Chimeric PRO Polypeptides

PRO polypeptides may be expressed as chimeric proteins with one or more additional polypeptide domains added to facilitate protein purification. Such purification facilitating domains include, but are not limited to, metal chelating peptides such as histidine-tryptophan modules that allow purification on immobilized metals, protein A domains that allow purification on immobilized immunoglobulin, and the domain utilized in the FLAGS™ extension/affinity purification system (Immunex Corp., Seattle Wash.). The inclusion of a cleavable linker sequence such as Factor XA or enterokinase (Invitrogen, San Diego Calif.) between the purification domain and the PRO polypeptide sequence may be useful to facilitate expression of DNA encoding the PRO polypeptide.

EXAMPLE 59: Purification of PRO Polypeptides Using Specific Antibodies

Native or recombinant PRO polypeptides may be purified by a variety of standard techniques in the 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 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™ (Pharmacia LKB Biotechnology). The antibody is coupled to the resin, the resin is blocked, and the derivative resin is washed 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 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 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 60: Drug Screening

This invention is particularly useful for screening compounds by using PRO polypeptides or 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 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 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 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 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 the peptide and immobilize it on the solid support.

This invention also contemplates the use of competitive drug screening assays in which neutralizing 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 61: Rational Drug Design

The goal of rational drug design is to produce structural analogs of biologically active polypeptide 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 an PRO 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, 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 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 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.

EXAMPLE 62: Diagnostic Test Using PRO317 Polypeptide-Specific Antibodies

Particular anti-PRO317 polypeptide antibodies are useful for the diagnosis of prepathologic conditions, and chronic or acute diseases such as gynecological diseases or ischemic diseases which are characterized by differences in the amount or distribution of PRO317. PRO317 has been found to be expressed in human kidney and is thus likely to be associated with abnormalities or pathologies which affect this organ. Further, since it is so closely related to EBAF-1, it is likely to affect the endometrium and other genital tissues. Further, due to library sources of certain ESTs, it appears that PRO317 may be involved as well in forming blood vessels and hence to be a modulator of angiogenesis.

Diagnostic tests for PRO317 include methods utilizing the antibody and a label to detect PRO317 in human body fluids, tissues, or extracts of such tissues. The polypeptide and antibodies of the present invention may be used with or without modification. Frequently, the polypeptide and antibodies will be labeled by joining them, either covalently or noncovalently, with a substance which provides for a detectable signal. A wide variety of labels and conjugation techniques are known and have been reported extensively in both the scientific and patent literature. Suitable labels include radionuclides, enzymes, substrates, cofactors, inhibitors, fluorescent agents, chemiluminescent agents, magnetic particles, and the like. Patents teaching the use of such labels include U.S. Pat. Nos. 3,817,837; 3,850,752; 3,939,350; 3,996,345; 4,277,437; 4,275,149; and 4,366,241. Also, recombinant immunoglobulins may be produced as shown in U.S. Pat. No. 4,816,567.

A variety of protocols for measuring soluble or membrane-bound PRo317, using either polyclonal or monoclonal antibodies specific for that PRO317, are known in the art. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA), radioreceptor assay (RRA), and fluorescent activated cell sorting (FACS). A two-site monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on PRO317 is preferred, but a competitive binding assay may be employed. These assays

are described, among other places, in Maddox *et al.* J Exp. Med., 158:1211 (1983).

EXAMPLE 63: Identification of PRO317 Receptors

Purified PRO317 is useful for characterization and purification of specific cell surface receptors and other binding molecules. Cells which respond to PRO317 by metabolic changes or other specific responses are likely to express a receptor for PRO317. Such receptors include, but are not limited to, receptors associated with and activated by tyrosine and serine/threonine kinases. See Kolodziejczyk and Hall, *supra*, for a review on known receptors for the TGF- superfamily. Candidate receptors for this superfamily fall into two primary groups, termed type I and type II receptors. Both types are serine/threonine kinases. Upon activation by the appropriate ligand, type I and type II receptors physically interact to form hetero-oligomers and subsequently activate intracellular signaling cascades, ultimately regulating gene transcription and expression. In addition, TGF- binds to a third receptor class, type III, a membrane-anchored proteoglycan lacking the kinase activity typical of signal transducing molecules.

PRO317 receptors or other PRO317-binding molecules may be identified by interaction with radiolabeled PRO317. Radioactive labels may be incorporated into PRO317 by various methods known in the art. A preferred embodiment is the labeling of primary amino groups in PRO317 with ^{125}I Bolton-Hunter reagent (Bolton and Hunter, Biochem. J., 133:529 (1973)), which has been used to label other polypeptides without concomitant loss of biological activity (Hebert *et al.*, J. Biol. Chem., 266:18989 (1991); McColl *et al.*, J. Immunol., 150:4550-4555 (1993)). Receptor-bearing cells are incubated with labeled PRO317. The cells are then washed to removed unbound PRO317, and receptor-bound PRO317 is quantified. The data obtained using different concentrations of PRO317 are used to calculate values for the number and affinity of receptors.

Labeled PRO317 is useful as a reagent for purification of its specific receptor. In one embodiment of affinity purification, PRO317 is covalently coupled to a chromatography column. Receptor-bearing cells are extracted, and the extract is passed over the column. The receptor binds to the column by virtue of its biological affinity for PRO317. The receptor is recovered from the column and subjected to N-terminal protein sequencing. This amino acid sequence is then used to design degenerate oligonucleotide probes for cloning the receptor gene.

In an alternative method, mRNA is obtained from receptor-bearing cells and made into a cDNA library. The library is transfected into a population of cells, and those cells expressing the receptor are selected using fluorescently labeled PRO317. The receptor is identified by recovering and sequencing recombinant DNA from highly labeled cells.

In another alternative method, antibodies are raised against the surface of receptor bearing cells, specifically monoclonal antibodies. The monoclonal antibodies are screened to identify those which inhibit the binding of labeled PRO317. These monoclonal antibodies are then used in affinity purification or expression cloning of the receptor.

Soluble receptors or other soluble binding molecules are identified in a similar manner. Labeled PRO317 is incubated with extracts or other appropriate materials derived from the uterus. After incubation, PRO317 complexes larger than the size of purified PRO317 are identified by a sizing technique such as size-exclusion chromatography or density gradient centrifugation and are purified by methods known in the art. The soluble receptors or binding protein(s) are subjected to N-terminal sequencing to obtain information sufficient for database identification, if the soluble protein is known, or for cloning, if the soluble protein is unknown.

EXAMPLE 64: Determination of PRO317-Induced Cellular Response

The biological activity of PRO317 is measured, for example, by binding of an PRO317 of the invention to an PRO317 receptor. A test compound is screened as an antagonist for its ability to block binding of PRO317 to the receptor. A test compound is screened as an agonist of the PRO317 for its ability to bind an PRO317 receptor and influence the same physiological events as PRO317 using, for example, the KIRA-ELISA assay described by Sadick *et al.*, Analytical Biochemistry, 235:207-214 (1996) in which activation of a receptor tyrosine kinase is monitored by immuno-capture of the activated receptor and quantitation of the level of ligand-induced phosphorylation. The assay may be adapted to monitor PRO317-induced receptor activation through the use of an PRO317 receptor-specific antibody to capture the activated receptor. These techniques are also applicable to other PRO polypeptides described herein.

EXAMPLE 65: Use of PRO224 for Screening Compounds

PRO224 is expressed in a cell stripped of membrane proteins and capable of expressing PRO224. Low density lipoproteins having a detectable label are added to the cells and incubated for a sufficient time for endocytosis. The cells are washed. The cells are then analysed for label bound to the membrane and within the cell after cell lysis. Detection of the low density lipoproteins within the cell determines that PRO224 is within the family of low density lipoprotein receptor proteins. Members found within this family are then used for screening compounds which affect these receptors, and particularly the uptake of cholesterol via these receptors.

EXAMPLE 66: Ability of PRO Polypeptides to Inhibit Vascular Endothelial Growth Factor (VEGF) Stimulated Proliferation of Endothelial Cell Growth

The ability of various PRO polypeptides to inhibit VEGF stimulated proliferation of endothelial cells was tested. Specifically, bovine adrenal cortical capillary endothelial (ACE) cells (from primary culture, maximum 12-14 passages) were plated on 96-well microtiter plates (Amersham Life Science) at a density of 500 cells/well per 100 μL in low glucose DMEM, 10% calf serum, 2 mM glutamine, 1x pen/strept and fungizone, supplemented with 3 ng/mL VEGF. Controls were plated the same way but some did not include VEGF. A test sample of the PRO polypeptide of interest was added in a 100 μl volume for a 200 μl final volume. Cells were incubated for 6-7 days at 37°C. The media was aspirated and the cells washed 1x with PBS. An acid phosphatase reaction mixture (100 μL, 0.1M sodium acetate, pH 5.5, 0.1% Triton-100, 10 mM p-nitrophenyl phosphate) was added. After incubation for 2 hours at 37°C, the reaction was stopped by addition of 10 μl 1N NaOH. OD was measured on microtiter plate reader at 405 nm. Controls were no cells, cells alone, cells + FGF (5 ng/mL), cells + VEGF (3 ng/mL), cells + VEGF (3 ng/ml) + TGF-β (1 ng/ml), and cells + VEGF (3ng/mL) + LIF (5 ng/mL). (TGF-β at a 1 ng/ml concentration is known to block 70-90% of VEGF stimulated cell proliferation.)

The results were assessed by calculating the percentage inhibition of VEGF (3 ng/ml) stimulated cells proliferation, determined by measuring acid phosphatase activity at OD405 nm, (1) relative to cells without stimulation, and (2) relative to the reference TGF-β inhibition of VEGF stimulated activity. The results, as shown in Table 2 below, are indicative of the utility of the PRO polypeptides in cancer therapy and specifically in inhibiting tumor angiogenesis. The numerical values (relative inhibition) shown in Table 2 are determined by calculating the

percent inhibition of VEGF stimulated proliferation by the PRO polypeptide relative to cells without stimulation and then dividing that percentage into the percent inhibition obtained by TGF-β at 1 ng/ml which is known to block 70-90% of VEGF stimulated cell proliferation.

Table 2

PRO Name	PRO Concentration	Relative Inhibition
PRO211	0.01%	99.0
PRO211	0.01%	1.09
PRO211	0.1%	0.95
PRO211	0.1%	67.0
PRO211	1.0%	0.27
PRO211	1.0%	20.0
PRO217	0.01%	1.06
PRO217	0.1%	0.84
PRO217	1.0%	0.39
PRO217	2.5 μM	0.2
PRO217	25 nM	0.88
PRO217	250 nM	0.58
PRO187	0.01%	0.91
PRO187	0.1%	0.82
PRO187	1.0%	0.44
PRO219	5.7 μM	0.61
PRO219	57 nM	1.09
PRO219	570 nM	0.97
PRO246	0.01%	1.04
PRO246	0.1%	1.0
PRO246	1.0%	0.49
PRO228	0.01%	0.99
PRO228	0.1%	0.93
PRO228	1.0%	0.57
PRO228	0.01%	0.95
PRO228	0.01%	0.98
PRO228	0.1%	0.77
PRO228	0.1%	0.88
PRO228	1.0%	0.16
PRO228	1.0%	0.48
PRO245	0.01%	0.76
PRO245	0.1%	0.35
PRO245	1.0%	0.11
PRO245	0.48 nM	1.03
PRO245	4.8 nM	0.95
PRO245	48 nM	0.49
PRO221	0.01%	1.03
PRO221	0.01%	1.06
PRO221	0.1%	0.82
PRO221	0.1%	0.93
PRO221	1.0%	0.31
PRO221	1.0%	0.43
PRO258	0.01%	0.98
PRO258	0.01%	1.06
PRO258	0.1%	0.95
PRO258	0.1%	1.02
PRO258	1.0%	0.6
PRO258	1.0%	0.69

Table 2 cont'

PRO Name	PRO Concentration	Relative Inhibition
PRO301	7.0 μM	1.02
PRO301	70 μM	0.88
PRO301	700 μM	0.44
PRO301	0.01%	0.92
PRO301	0.1%	0.85
PRO301	1.0%	0.68
PRO224	0.01%	101.0
PRO224	0.1%	65.0
PRO224	1.0%	23.0
PRO272	0.01%	0.95
PRO272	0.1%	0.57
PRO272	1.0%	0.18
PRO328	0.01%	0.98
PRO328	0.1%	0.96
PRO328	1.0%	0.6
PRO331	0.01%	0.88
PRO331	0.1%	0.82
PRO331	1.0%	0.56

EXAMPLE 67: Retinal Neuron Survival

This example demonstrates that PRO220 polypeptides have efficacy in enhancing the survival of retinal neuron cells.

Sprague Dawley rat pups at postnatal day 7 (mixed population: glia and retinal neuronal types) are killed by decapitation following CO_2 anesthesia and the eyes are removed under sterile conditions. The neural retina is dissected away from the pigment epithelium and other ocular tissue and then dissociated into a single cell suspension using 0.25% trypsin in Ca^{2+}, Mg^{2+}-free PBS. The retinas are incubated at 37°C for 7-10 minutes after which the trypsin is inactivated by adding 1 ml soybean trypsin inhibitor. The cells are plated at 100,000 cells per well in 96 well plates in DMEM/F12 supplemented with N2 and with or without the specific test PRO polypeptide. Cells for all experiments are grown at 37°C in a water saturated atmosphere of 5% CO_2. After 2-3 days in culture, cells are stained with calcein AM then fixed using 4% paraformaldehyde and stained with DAPI for determination of total cell count. The total cells (fluorescent) are quantified at 20X objective magnification using CCD camera and NIH image software for MacIntosh. Fields in the well are chosen at random.

The effect of various concentration of PRO220 polypeptides are reported in Table 3 below where percent survival is calculated by dividing the total number of calcein AM positive cells at 2-3 days in culture by the total number of DAPI-labeled cells at 2-3 days in culture. Anything above 30% survival is considered positive.

Table 3

PRO Name	PRO Concentration	Percent Survival
PRO220	0.01%	2.4%
PRO220	0.01%	4.1%
PRO220	0.1%	3.0%
PRO220	0.1%	3.1%
PRO220	1.0%	72.4%
PRO220	1.0%	42.1%

EXAMPLE 68: Rod Photoreceptor Survival

This example demonstrates that PRO220 polypeptides have efficacy in enhancing the survival of rod photoreceptor cells.

Sprague Dawley rat pups at 7 day postnatal (mixed population: glia and retinal neuronal cell types) are killed by decapitation following CO_2 anesthesis and the eyes are removed under sterile conditions. The neural retina is dissected away form the pigment epithelium and other ocular tissue and then dissociated into a single cell suspension using 0.25% trypsin in Ca^{2+}, Mg^{2+}-free PBS. The retinas are incubated at 37°C for 7-10 minutes after which the trypsin is inactivated by adding 1 ml soybean trypsin inhibitor. The cells are plated at 100,000 cells per well in 96 well plates in DMEM/F12 supplemented with N2 and with or without the specific test PRO polypeptide. Cells for all experiments are grown at 37°C in a water saturated atmosphere of 5% CO2. After 2-3 days in culture, cells are fixed using 4% paraformaldehyde, and then stained using CellTracker Green CMFDA. Rho 4D2 (ascites or IgG 1:100), a monoclonal antibody directed towards the visual pigment rhodopsin is used to detect rod photoreceptor cells by indirect immunofluorescence. The results are reported as % survival: total number of calcein/CellTracker - rhodopsin positive cells at 2-3 days in culture, divided by the total number of rhodopsin positive cells at time 2-3 days in culture. The total cells (fluorescent) are quantified at 20x objective magnification using a CCD camera and NIH image software for MacIntosh. Fields in the well are chosen at random.

The effect of various concentration of PRO220 polypeptides are reported in Table 4 below. Anything above 10% survival is considered positive..

Table 4

PRO Name	PRO Concentration	Percent Survival
PRO220	0.01%	0.0%
PRO220	0.1%	0.0%
PRO220	2.0%	0.0%
PRO220	10%	0.0%
PRO220	20%	66.9%
PRO220	1.0%	56.9%

EXAMPLE 69: Induction of Endothelial Cell Apoptosis

The ability of PRO228 polypeptides to induce apoptosis in endothelial cells was tested in human venous umbilical vein endothelial cells (HUVEC, Cell Systems), using a 96-well format, in 0% serum media supplemented with 100 ng/ml VEGF. (As HUVEC cells are easily dislodged from the plating surface, all pipetting in the wells must be done as gently as practicable.)

The media was aspired and the cells washed once with PBS. 5 ml of 1 x trypsin was added to the cells in a T-175 flask, and the cells were allowed to stand until they were released from the plate (about 5-10 minutes). Trypsinization was stopped by adding 5 ml of growth media. The cells were spun at 1000 rpm for 5 minutes at 4°C. The media was aspirated and the cells were resuspended in 10 ml of 10% serum complemented medium (Cell Systems), 1 x penn/strep.

The cells were plated on 96-well microtiter plates (Amersham Life Science, cytostar-T scintillating microplate, RPNQ160, sterile, tissue-culture treated, individually wrapped), in 10% serum (CSG-medium, Cell

Systems), at a density of 2 x 10^4 cells per well in a total volume of 100 μl. The PRO228 polypeptide was added in triplicate at dilutions of 1%, 0.33% and 0.11%. Wells without cells were used as a blank and wells with cells only as a negative control. As a positive control 1:3 serial dilutions of 50 μl of a 3x stock of staurosporine were used. The ability of the PRO228 polypeptide to induce apoptosis was determined using Annexin V, a member of the calcium and phospholipid binding proteins, to detect apoptosis.

0.2 ml Annexin V - Biotin stock solution (100 μg/ml) were diluted in 4.6 ml 2 x Ca^{2+} binding buffer and 2.5% BSA (1:25 dilution). 50 μls of the diluted Annexin V - Biotin solution were added to each well (except controls) to a final concentration of 1.0 μg/ml. The samples were incubated for 10-15 minutes with Annexin-Biotin prior to direct addition of ^{35}S-Streptavidin. ^{35}S-Streptavidin was diluted in 2x Ca^{2+} binding buffer, 2.5% BSA and was added to all wells at a final concentration of 3 x 10^4 cpm/well. The plates were then sealed, centrifuged at 1000 rpm for 15 minutes and placed on orbital shaker for 2 hours. The analysis was performed on 1450 Microbeta Trilux (Wallac). The results are shown in Table 5 below where percent above background represents the percentage amount of counts per minute above the negative controls. Percents greater than or equal to 30% above background are considered positive.

Table 5

PRO Name	PRO Concentration	Percent Above Background
PRO228	0.11%	0.7%
PRO228	0.11%	47.6%
PRO228	0.33%	92.2%
PRO228	0.33%	123.7%
PRO228	1.0%	51.4%
PRO228	1.0%	95.3%

EXAMPLE 70: PDB12 Cell Inhibition

This example demonstrates that various PRO polypeptides have efficacy in inhibiting protein production by PDB12 pancreatic ductal cells.

PDB12 pancreatic ductal cells are plated on fibronectin coated 96 well plates at 1.5x10^3 cells per well in 100 μL/180 μL of growth media. 100 μL of growth media with the PRO polypeptide test sample or negative control lacking the PRO polypeptide is then added to well, for a final volume of 200 μL. Controls contain growth medium containing a protein shown to be inactive in this assay. Cells are incubated for 4 days at 37°C. 20 μL of Alamar Blue Dye (AB) is then added to each well and the flourescent reading is measured at 4 hours post addition of AB, on a microtiter plate reader at 530 nm excitation and 590 nm emission. The standard employed is cells without Bovine Pituitary Extract (BPE) and with various concentrations of BPE. Buffer or CM controls from unknowns are run 2 times on each 96 well plate.

The results from these assays are shown in Table 6 below wherein percent decrease in protein production is calculated by comparing the Alamar Blue Dye calculated protein concentration produced by the PRO polypeptide-treated cells with the Alamar Blue Dye calculated protein concentration produced by the negative control cells. A percent decrease in protein production of greater than or equal to 25% as compared to the negative control cells is considered positive.

Table 6

PRO Name	PRO Concentration	Percent Decrease in Protein Production
PRO211	0.1%	0.0%
PRO211	0.01%	0.6%
PRO211	1.0%	59.7%
PRO287	2.0%	22.3%
PRO287	10%	18.2%
PRO287	50%	67.5%
PRO287	2.0%	45.53%
PRO287	10%	57.3%
PRO287	50%	52.24%
PRO301	2.0%	0.0%
PRO301	10%	59.8%
PRO301	50%	65.6%
PRO293	2.0%	0.0%
PRO293	10%	40.4%
PRO293	50%	56.7%

EXAMPLE 71: Stimulation of Adult Heart Hypertrophy

This assay is designed to measure the ability of various PRO polypeptides to stimulate hypertrophy of adult heart.

Ventricular myocytes freshly isolated from adult (250g) Sprague Dawley rats are plated at 2000 cell/well in 180 μl volume. Cells are isolated and plated on day 1, the PRO polypeptide-containing test samples or growth medium only (negative control) (20 μl volume) is added on day 2 and the cells are then fixed and stained on day 5. After staining, cell size is visualized wherein cells showing no growth enhancement as compared to control cells are given a value of 0.0, cells showing small to moderate growth enhancement as compared to control cells are given a value of 1.0 and cells showing large growth enhancement as compared to control cells are given a value of 2.0. Any degree of growth enhancement as compared to the negative control cells is considered positive for the assay. The results are shown in Table 7 below.

Table 7

PRO Name	PRO Concentration	Growth Enhancement Score
PRO287	20%	1.0
PRO287	20%	1.0
PRO301	20%	1.0
PRO301	20%	1.0
PRO293	20%	1.0
PRO293	20%	1.0
PRO303	20%	1.0
PRO303	20%	1.0

EXAMPLE 72: PDB12 Cell Proliferation

This example demonstrates that various PRO polypeptides have efficacy in inducing proliferation of PDB12 pancreatic ductal cells.

PDB12 pancreatic ductal cells are plated on fibronectin coated 96 well plates at $1.5x10^3$ cells per well in 100 μL/180 μL of growth media. 100 μL of growth media with the PRO polypeptide test sample or negative control

lacking the PRO polypeptide is then added to well, for a final volume of 200 μL. Controls contain growth medium containing a protein shown to be inactive in this assay. Cells are incubated for 4 days at 37°C. 20 μL of Alamar Blue Dye (AB) is then added to each well and the flourescent reading is measured at 4 hours post addition of AB, on a microtiter plate reader at 530 nm excitation and 590 nm emission. The standard employed is cells without Bovine Pituitary Extract (BPE) and with various concentrations of BPE. Buffer or growth medium only controls from unknowns are run 2 times on each 96 well plate.

The results from these assays are shown in Table 8 below wherein percent increase in protein production is calculated by comparing the Alamar Blue Dye calculated protein concentration produced by the PRO polypeptide-treated cells with the Alamar Blue Dye calculated protein concentration produced by the negative control cells. A percent increase in protein production of greater than or equal to 25% as compared to the negative control cells is considered positive.

Table 8

PRO Name	PRO Concentration	Percent Increase in Protein Production
PRO301	2.0%	44.0%
PRO301	10%	67.4%
PRO301	50%	185.8%
PRO303	2.0%	27.9%
PRO303	10%	174.9%
PRO303	50%	193.1%

EXAMPLE 73: Enhancement of Heart Neonatal Hypertrophy Induced by PRO224

This assay is designed to measure the ability of PRO224 polypeptides to stimulate hypertrophy of neonatal heart.

Cardiac myocytes from 1-day old Harlan Sprague Dawley rats were obtained. Cells (180 μl at 7.5 x 10^4/ml, serum <0.1%, freshly isolated) are added on day 1 to 96-well plates previously coated with DMEM/F12 + 4% FCS. Test samples containing the test PRO224 polypeptide or growth medium only (hegative control) (20 μl/well) are added directly to the wells on day 1. PGF (20 μl/well) is then added on day 2 at final concentration of 10^{-6} M. The cells are then stained on day 4 and visually scored on day 5, wherein cells showing no increase in size as compared to negative controls are scored 0.0, cells showing a small to moderate increase in size as compared to negative controls are scored 1.0 and cells showing a large increase in size as compared to negative controls are scored 2.0. The results are shown in Table 9 below.

Table 9

PRO Name	PRO Concentration	Growth Enhancement Score
PRO224	0.01%	0.0
PRO224	0.1%	0.0
PRO224	1.0%	1.0

EXAMPLE 74: *In situ* Hybridization

In situ hybridization is a powerful and versatile technique for the detection and localization of nucleic acid sequences within cell or tissue preparations. It may be useful, for example, to identify sites of gene expression,

analyze the tissue distribution of transcription, identify and localize viral infection, follow changes in specific mRNA synthesis and aid in chromosome mapping.

In situ hybridization was performed following an optimized version of the protocol by Lu and Gillett, Cell Vision 1:169-176 (1994), using PCR-generated ^{33}P-labeled riboprobes. Briefly, formalin-fixed, paraffin-embedded human tissues were sectioned, deparaffinized, deproteinated in proteinase K (20 g/ml) for 15 minutes at 37°C, and further processed for *in situ* hybridization as described by Lu and Gillett, *supra*. A [^{33}-P] UTP-labeled antisense riboprobe was generated from a PCR product and hybridized at 55°C overnight. The slides were dipped in Kodak NTB2 nuclear track emulsion and exposed for 4 weeks.

^{33}P-Riboprobe synthesis

6.0 μl (125 mCi) of ^{33}P-UTP (Amersham BF 1002, SA<2000 Ci/mmol) were speed vac dried. To each tube containing dried ^{33}P-UTP, the following ingredients were added:

2.0 μl 5x transcription buffer

1.0 μl DTT (100 mM)

2.0 μl NTP mix (2.5 mM : 10 μ; each of 10 mM GTP, CTP & ATP + 10 μl H_2O)

1.0 μl UTP (50 μM)

1.0 μl Rnasin

1.0 μl DNA template (1μg)

1.0 μl H_2O

1.0 μl RNA polymerase (for PCR products T3 = AS, T7 = S, usually)

The tubes were incubated at 37°C for one hour. 1.0 μl RQ1 DNase were added, followed by incubation at 37°C for 15 minutes. 90 μl TE (10 mM Tris pH 7.6/1mM EDTA pH 8.0) were added, and the mixture was pipetted onto DE81 paper. The remaining solution was loaded in a Microcon-50 ultrafiltration unit, and spun using program 10 (6 minutes). The filtration unit was inverted over a second tube and spun using program 2 (3 minutes). After the final recovery spin, 100 μl TE were added. 1 μl of the final product was pipetted on DE81 paper and counted in 6 ml of Biofluor II.

The probe was run on a TBE/urea gel. 1-3 μl of the probe or 5 μl of RNA Mrk III were added to 3 μl of loading buffer. After heating on a 95°C heat block for three minutes, the gel was immediately placed on ice. The wells of gel were flushed, the sample loaded, and run at 180-250 volts for 45 minutes. The gel was wrapped in saran wrap and exposed to XAR film with an intensifying screen in -70°C freezer one hour to overnight.

^{33}P-Hybridization

A. Pretreatment of frozen sections

The slides were removed from the freezer, placed on aluminium trays and thawed at room temperature for 5 minutes. The trays were placed in 55°C incubator for five minutes to reduce condensation. The slides were fixed for 10 minutes in 4% paraformaldehyde on ice in the fume hood, and washed in 0.5 x SSC for 5 minutes, at room temperature (25 ml 20 x SSC + 975 ml SQ H_2O). After deproteination in 0.5 μg/ml proteinase K for 10 minutes at 37°C (12.5 μl of 10 mg/ml stock in 250 ml prewarmed RNase-free RNAse buffer), the sections were washed in 0.5 x SSC for 10 minutes at room temperature. The sections were dehydrated in 70%, 95%, 100% ethanol, 2 minutes each.

B. Pretreatment of paraffin-embedded sections

The slides were deparaffinized, placed in SQ H_2O, and rinsed twice in 2 x SSC at room temperature, for 5 minutes each time. The sections were deproteinated in 20 μg/ml proteinase K (500 μl of 10 mg/ml in 250 ml RNase-free RNase buffer; 37°C, 15 minutes) - human embryo, or 8 x proteinase K (100 μl in 250 ml Rnase buffer, 37°C, 30 minutes) - formalin tissues. Subsequent rinsing in 0.5 x SSC and dehydration were performed as described above.

C. Prehybridization

The slides were laid out in a plastic box lined with Box buffer (4 x SSC, 50% formamide) - saturated filter paper. The tissue was covered with 50 μl of hybridization buffer (3.75g Dextran Sulfate + 6 ml SQ H_2O), vortexed and heated in the microwave for 2 minutes with the cap loosened. After cooling on ice, 18.75 ml formamide, 3.75 ml 20 x SSC and 9 ml SQ H_2O were added, the tissue was vortexed well, and incubated at 42°C for 1-4 hours.

D. Hybridization

1.0×10^6 cpm probe and 1.0 μl tRNA (50 mg/ml stock) per slide were heated at 95°C for 3 minutes. The slides were cooled on ice, and 48 μl hybridization buffer were added per slide. After vortexing, 50 μl ^{33}P mix were added to 50 μl prehybridization on slide. The slides were incubated overnight at 55°C.

E. Washes

Washing was done 2 x 10 minutes with 2xSSC, EDTA at room temperature (400 ml 20 x SSC + 16 ml 0.25M EDTA, V_f=4L), followed by RNaseA treatment at 37°C for 30 minutes (500 μl of 10 mg/ml in 250 ml Rnase buffer = 20 μg/ml), The slides were washed 2 x 10 minutes with 2 x SSC, EDTA at room temperature. The stringency wash conditions were as follows: 2 hours at 55°C, 0.1 x SSC, EDTA (20 ml 20 x SSC + 16 ml EDTA, V_f=4L).

F. Oligonucleotides

In situ analysis was performed on a variety of DNA sequences disclosed herein. The oligonucleotides employed for these analyses are as follows.

(1) DNA33094-1131 (PRO217)

p1 5'-GGATTCTAATACGACTCACTATAGGGCTCAGAAAAGCGCAACAGAGAA-3' (SEQ ID NO:348)

p2 5'-CTATGAAATTAACCCTCACTAAAGGGATGTCTTCCATGCCAACCTTC-3' (SEQ ID NO:349)

(2) DNA33223-1136 (PRO230)

p1 5'-GGATTCTAATACGACTCACTATAGGGCGGCGATGTCCACTGGGGCTAC-3' (SEQ ID NO:350)

p2 5'-CTATGAAATTAACCCTCACTAAAGGGACGAGGAAGATGGGCGGATGGT-3' (SEQ ID NO:351)

(3) DNA34435-1140 (PRO232)

p1 5'-GGATTCTAATACGACTCACTATAGGGCACCCACGCGTCCGGCTGCTT-3' (SEQ ID NO:352)

p2 5'-CTATGAAATTAACCCTCACTAAAGGGACGGGGGACACCACGGACCAGA-3' (SEQ ID NO:353)

(4) DNA35639-1172 (PRO246)

p1 5'-GGATTCTAATACGACTCACTATAGGGCTTGCTGCGGTTTTTGTTCCTG-3' (SEQ ID NO:354)

p2 5'-CTATGAAATTAACCCTCACTAAAGGGAGCTGCCGATCCCACTGGTATT-3' (SEQ ID NO:355)

(5) DNA49435-1219 (PRO533)

p1 5'-GGATTCTAATACGACTCACTATAGGGCGGATCCTGGCCGGCCTCTG-3' (SEQ ID NO:356)

p2 5'-CTATGAAATTAACCCTCACTAAAGGGAGCCCGGGCATGGTCTCAGTTA-3' (SEQ ID NO:357)

(6) DNA35638-1141 (PRO245)

p1 5'-GGATTCTAATACGACTCACTATAGGGCGGGAAGATGGCGAGGAGGAG-3' (SEQ ID NO:358)

p2 5'-CTATGAAATTAACCCTCACTAAAGGGACCAAGGCCACAAACGGAAATC-3' (SEQ ID NO:359)

(7) DNA33089-1132 (PRO221)

p1 5'-GGATTCTAATACGACTCACTATAGGGCTGTGCTTTCATTCTGCCAGTA-3' (SEQ ID NO:360)

p2 5'-CTATGAAATTAACCCTCACTAAAGGGAGGGTACAATTAAGGGGTGGAT-3' (SEQ ID NO:361)

(8) DNA35918-1174 (PRO258)

p1 5'-GGATTCTAATACGACTCACTATAGGGCCCGCCTCGCTCCTGCTCCTG-3' (SEQ ID NO:362)

p2 5'-CTATGAAATTAACCCTCACTAAAGGGAGGATTGCCGCGACCCTCACAG-3' (SEQ ID NO:363)

(9) DNA32286-1191 (PRO214)

p1 5'-GGATTCTAATACGACTCACTATAGGGCCCCTCCTGCCTTCCCTGTCC-3' (SEQ ID NO:364)

p2 5'-CTATGAAATTAACCCTCACTAAAGGGAGTGGTGGCCGCGATTATCTGC-3' (SEQ ID NO:365)

(10) DNA33221-1133 (PRO224)

p1 5'-GGATTCTAATACGACTCACTATAGGGCGCAGCGATGGCAGCGATGAGG-3' (SEQ ID NO:366)

p2 5'-CTATGAAATTAACCCTCACTAAAGGGACAGACGGGGCAGAGGGAGTG-3' (SEQ ID NO:367)

(11) DNA35557-1137 (PRO234)

p1 5'-GGATTCTAATACGACTCACTATAGGGCCAGGAGGCGTGAGGAGAAAC-3' (SEQ ID NO:368)

p2 5'-CTATGAAATTAACCCTCACTAAAGGGAAAGACATGTCATCGGGAGTGG-3' (SEQ ID NO:369)

(12) DNA33100-1159 (PRO229)

p1 5'-GGATTCTAATACGACTCACTATAGGGCCGGGTGGAGGTGGAACAGAAA-3' (SEQ ID NO:370)

p2 5'-CTATGAAATTAACCCTCACTAAAGGGACACAGACAGAGCCCCATACGC-3' (SEQ ID NO:371)

(13) DNA34431-1177 (PRO263)

p1 5'-GGATTCTAATACGACTCACTATAGGGCCAGGGAAATCCGGATGTCTC-3' (SEQ ID NO:372)

p2 5'-CTATGAAATTAACCCTCACTAAAGGGAGTAAGGGGATGCCACCGAGTA-3' (SEQ ID NO:373)

(14) DNA38268-1188 (PRO295)

p1 5'-GGATTCTAATACGACTCACTATAGGGCCAGCTACCCGCAGGAGGAGG-3' (SEQ ID NO:374)

p2 5'-CTATGAAATTAACCCTCACTAAAGGGATCCCAGGTGATGAGGTCCAGA-3' (SEQ ID NO:375)

G. Results

In situ analysis was performed on a variety of DNA sequences disclosed herein. The results from these analyses are as follows.

(1) DNA33094-1131 (PRO217)

Highly distinctive expression pattern, that does not indicate an obvious biological function. In the human embryo it was expressed in outer smooth muscle layer of the GI tract, respiratiry cartilage, branching respiratory epithelium, osteoblasts, tendons, gonad, in the optic nerve head and developing dermis. In the adult expression was observed in the epidermal pegs of the chimp tongue, the basal epithelial/myoepithelial cells of the prostate and urinary bladder. Also expressed in the alveolar lining cells of the adult lung, mesenchymal cells juxtaposed to erectile tissue in the penis and the cerebral cortex (probably glial cells). In the kidney, expression was only seen in disease, in cells surrounding thyroidized renal tubules.

Human fetal tissues examined (E12-E16 weeks) include: Placenta, umbilical cord, liver, kidney, adrenals, thyroid, lungs, heart, great vessels, oesophagus, stomach, small intestine, spleen, thymus, pancreas, brain, eye, spinal cord, body wall, pelvis and lower limb.

Adult human tissues examined: Kidney (normal and end-stage), adrenal, myocardium, aorta, spleen, lymph node, gall bladder, pancreas, lung, skin, eye (inc. retina), prostate, bladder, liver (normal, cirrhotic, acute failure).

Non-human primate tissues examined:

(a) Chimp Tissues: Salivary gland, stomach, thyroid, parathyroid, skin, thymus, ovary, lymph node.

(b) Rhesus Monkey Tissues: Cerebral cortex, hippocampus, cerebellum, penis.

(2) DNA33223-1136 (PRO230)

Sections show an intense signal associated with arterial and venous vessels in the fetus. In arteries the signal appeared to be confined to smooth-muscle/pericytic cells. The signal is also seen in capillary vessels and in glomeruli. It is not clear whether or not endothelial cells are expressing this mRNA. Expression is also observed in epithelial cells in the fetal lens. Strong expression was also seen in cells within placental trophoblastic villi, these cells lie between the trophoblast and the fibroblast-like cells that express HGF - uncertain histogenesis. In the adult, there was no evidence of expression and the wall of the aorta and most vessels appear to be negative. However, expression was seen over vascular channels in the normal prostate and in the epithelium lining the gallbladder. Insurers expression was seen in the vessels of the soft-tissue sarcoma and a renal cell carcinoma. In summary, this is a molecule that shows relatively specific vascular expression in the fetus as well as in some adult organs. Expression was also

observed in the fetal lens and the adult gallbladder.

In a secondary screen, vascular expression was observed, similar to that observed above, seen in fetal blocks. Expression is on vascular smooth muscle, rather than endothelium. Expression also seen in smooth muscle of the developing oesophagus, so as reported previously, this molecule is not vascular specific. Expression was examined in 4 lung and 4 breast carcinomas. Substantial expression was seen in vascular smooth muscle of at least 3/4 lung cancers and 2/4 breast cancers. In addition, in one breast carcinoma, expression was observed in peritumoral stromal cells of uncertain histogenesis (possibly myofibroblasts). No endothelial cell expression was observed in this study.

(3) DNA34435-1140 (PRO232)

Strong expression in prostatic epithelium and bladder epithelium, lower level of expression in bronchial epithelium. High background / low level expression seen in a number of sites, including among others, bone, blood, chondrosarcoma, adult heart and fetal liver. It is felt that this level of signal represents background, partly because signal at this level was seen over the blood. All other tissues negative.

Human fetal tissues examined (E12-E16 weeks) include: Placenta, umbilical cord, liver, kidney, adrenals, thyroid, lungs, heart, great vessels, oesophagus, stomach, small intestine,
spleen, thymus, pancreas, brain, eye, spinal cord, body wall, pelvis, testis and lower limb.

Adult human tissues examined: Kidney (normal and end-stage), adrenal, spleen, lymph node, pancreas, lung, eye (inc. retina), bladder, liver (normal, cirrhotic, acute failure).

Non-human primate tissues examined:

Chimp Tissues: adrenal

Rhesus Monkey Tissues: Cerebral cortex, hippocampus

In a secondary screen, expression was observed in the epithelium of the prostate, the superficial layers of the urethelium of the urinary bladder, the urethelium lining the renal pelvis and the urethelium of the ureter (1 out of 2 experiments). The urethra of a rhesus monkey was negative; it is unclear whether this represents a true lack of expression by the urethra, or if it is the result of a failure of the probe to cross react with rhesus tissue. The findings in the prostate and bladder are similar to those previously described using an isotopic detection technique. Expression of the mRNA for this antigen is NOT prostate epithelial specific. The antigen may serve as a useful marker for urethelial derived tissues. Expression in the superficial, post-mitotic cells, of the urinary tract epithelium also suggest that it is unlikely to represent a specific stem cell marker, as this would be expected to be expressed specifically in basal epithelium.

(4) DNA35639-1172 (PRO246)

Strongly expressed in fetal vascular endothelium, including tissues of the CNS. Lower level of expression in adult vasculature, including the CNS. Not obviously expressed at higher levels in tumor vascular endothelium. Signal also seen over bone matrix and adult spleen, not obviously cell associated, probably related to non-specific background at these sites.

Human fetal tissues examined (E12-E16 weeks) include: Placenta, umbilical cord, liver, kidney, adrenals, thyroid,

lungs, heart, great vessels, oesophagus, stomach, small intestine, spleen, thymus, pancreas, brain, eye, spinal cord, body wall, pelvis, testis and lower limb.

Adult human tissues examined: Kidney (normal and end-stage), adrenal, spleen, lymph node, pancreas, lung, eye (inc. retina), bladder, liver (normal, cirrhotic, acute failure).

Non-human primate tissues examined:

Chimp Tissues: adrenal

Rhesus Monkey Tissues: Cerebral cortex, hippocampus

(5) DNA49435-1219 (PRO533)

Moderate expression over cortical neurones in the fetal brain. Expression over the inner aspect of the fetal retina, possible expression in the developing lens. Expression over fetal skin, cartilage, small intestine, placental villi and umbilical cord. In adult tissues there is an extremely high level of expression over the gallbladder epithelium. Moderate expression over the adult kidney, gastric and colonic epithelia. Low-level expression was observed over many cell types in many tissues, this may be related to stickiness of the probe, these data should therefore be interpreted with a degree of caution.

Human fetal tissues examined (E12-E16 weeks) include: Placenta, umbilical cord, liver, kidney, adrenals, thyroid, lungs, heart, great vessels, oesophagus, stomach, small intestine, spleen, thymus, pancreas, brain, eye, spinal cord, body wall, pelvis, testis and lower limb.

Adult human tissues examined: Kidney (normal and end-stage), adrenal, spleen, lymph node, pancreas, lung, eye (inc. retina), bladder, liver (normal, cirrhotic, acute failure).

Non-human primate tissues examined:

Chimp Tissues: adrenal

Rhesus Monkey Tissues: Cerebral cortex, hippocampus, cerebellum.

(6) DNA35638-1141 (PRO245)

Expression observed in the endothelium lining a subset of fetal and placental vessels. Endothelial expression was confined to these tissue blocks. Expression also observed over intermediate trophoblast cells of placenta. All other tissues negative.

Fetal tissues examined (E12-E16 weeks) include: Placenta, umbilical cord, liver, kidney, adrenals, thyroid, lungs, heart, great vessels, oesophagus, stomach, small intestine, spleen, thymus, pancreas, brain, eye, spinal cord, body wall, pelvis and lower limb.

Adult tissues examined: Liver, kidney, adrenal, myocardium, aorta, spleen, lymph node,
pancreas, lung, skin, cerebral cortex (rm), hippocampus(rm), cerebellum(rm), penis, eye, bladder, stomach, gastric carcinoma, colon, colonic carcinoma, thyroid (chimp), parathyroid (chimp) ovary (chimp) and chondrosarcoma. Acetominophen induced liver injury and hepatic
cirrhosis

(7) DNA33089-1132 (PRO221)

Specific expression over fetal cerebral white and grey matter, as well as over neurones in the spinal cord. Probe appears to cross react with rat. Low level of expression over cerebellar neurones in adult rhesus brain. All other tissues negative.

Fetal tissues examined (E12-E16 weeks) include: Placenta, umbilical cord, liver, kidney, adrenals, thyroid, lungs, heart, great vessels, oesophagus, stomach, small intestine, spleen, thymus, pancreas, brain, eye, spinal cord, body wall, pelvis and lower limb.

Adult tissues examined: Liver, kidney, adrenal, myocardium, aorta, spleen, lymph node, pancreas, lung, skin, cerebral cortex (rm), hippocampus(rm), cerebellum(rm), penis, eye, bladder, stomach, gastric carcinoma, colon, colonic carcinoma and chondrosarcoma. Acetominophen induced liver injury and hepatic cirrhosis

(8) DNA35918-1174 (PRO258)

Strong expression in the nervous system. In the rhesus monkey brain expression is observed in cortical, hippocampal and cerebellar neurones. Expression over spinal neurones in the fetal spinal cord, the developing brain and the inner aspects of the fetal retina. Expression over developing dorsal root and autonomic ganglia as well as enteric nerves. Expression observed over ganglion cells in the adult prostate. In the rat, there is strong expression over the developing hind brain and spinal cord. Strong expression over interstitial cells in the placental villi. All other tissues were negative.

Fetal tissues examined (E12-E16 weeks) include: Placenta, umbilical cord, liver, kidney, adrenals, thyroid, lungs, heart, great vessels, oesophagus, stomach, small intestine, spleen, thymus, pancreas, brain, eye, spinal cord, body wall, pelvis and lower limb.

Adult tissues examined: Liver, kidney, renal cell carcinoma, adrenal, aorta, spleen, lymph node, pancreas, lung, myocardium, skin, cerebral cortex (rm), hippocampus(rm), cerebellum(rm), bladder, prostate, stomach, gastric carcinoma, colon, colonic carcinoma, thyroid (chimp), parathyroid (chimp) ovary (chimp) and chondrosarcoma. Acetominophen induced liver injury and hepatic cirrhosis.

(9) DNA32286-1191 (PRO214)

Fetal tissue: Low level throughout mesenchyme. Moderate expression in placental stromal cells in membranous tissues and in thyroid. Low level expression in cortical neurones. Adult tissue: all negative.

Fetal tissues examined (E12-E16 weeks) include: Placenta, umbilical cord, liver, kidney, adrenals, thyroid, lungs, heart, great vessels, oesophagus, stomach, small intestine, spleen, thymus, pancreas, brain, eye, spinal cord, body wall, pelvis and lower limb.

Adult tissues examined include: Liver, kidney, adrenal, myocardium, aorta, spleen, lymph node, pancreas, lung and skin.

(10) DNA33221-1133 (PRO224)

Expression limited to vascular endothelium in fetal spleen, adult spleen, fetal liver, adult thyroid and adult lymph node (chimp). Additional site of expression is the developing

spinal ganglia. All other tissues negative.

Human fetal tissues examined (E12-E16 weeks) include: Placenta, umbilical cord, liver, kidney, adrenals, thyroid, lungs, heart, great vessels, oesophagus, stomach, small intestine, spleen, thymus, pancreas, brain, eye, spinal cord, body wall, pelvis and lower limb.

Adult human tissues examined: Kidney (normal and end-stage), adrenal, myocardium, aorta, spleen, lymph node, pancreas, lung, skin, eye (inc. retina), bladder, liver (normal, cirrhotic, acute failure).

Non-human primate tissues examined:

Chimp Tissues: Salivary gland, stomach, thyroid, parathyroid, skin, thymus, ovary, lymph node.

Rhesus Monkey Tissues: Cerebral cortex, hippocampus, cerebellum, penis.

(11) DNA35557-1137 (PRO234)

Specific expression over developing motor neurones in ventral aspect of the fetal spinal cord (will develop into ventral horns of spinal cord). All other tissues negative. Possible role in growth, differentiation and/or development of spinal motor neurons.

Fetal tissues examined (E12-E16 weeks) include: Placenta, umbilical cord, liver, kidney, adrenals, thyroid, lungs, heart, great vessels, oesophagus, stomach, small intestine, spleen, thymus, pancreas, brain, eye, spinal cord, body wall, pelvis and lower limb.

Adult tissues examined: Liver, kidney, adrenal, myocardium, aorta, spleen, lymph node, pancreas, lung, skin, cerebral cortex (rm), hippocampus(rm), cerebellum(rm), penis, eye, bladder, stomach, gastric carcinoma, colon, colonic carcinoma and chondrosarcoma. Acetominophen induced liver injury and hepatic cirrhosis

(12) DNA33100-1159 (PRO229)

Striking expression in mononuclear phagocytes (macrophages) of fetal and adult spleen, liver, lymph node and adult thymus (in tingible body macrophages). The highest expression is in the spleen. All other tissues negative. Localisation and homology are entirely consistent with a role as a scavenger receptor for cells of the reticuloendothelial system. Expression also observed in placental mononuclear cells.

Human fetal tissues examined (E12-E16 weeks) include: Placenta, umbilical cord, liver, kidney, adrenals, thyroid, lungs, heart, great vessels, oesophagus, stomach, small intestine, spleen, thymus, pancreas, brain, eye, spinal cord, body wall, pelvis and lower limb.

Adult human tissues examined: Kidney (normal and end-stage), adrenal, myocardium, aorta, spleen, lymph node, gall bladder, pancreas, lung, skin, eye (inc. retina), prostate, bladder, liver (normal, cirrhotic, acute failure).

Non-human primate tissues examined:

Chimp Tissues: Salivary gland, stomach, thyroid, parathyroid, skin, thymus, ovary, lymph node.

Rhesus Monkey Tissues: Cerebral cortex, hippocampus, cerebellum, penis.

(13) DNA34431-1177 (PRO263)

Widepread expression in human fetal tissues and placenta over mononuclear cells, probably macrophages +/- lymphocytes. The cellular distribution follows a perivascular pattern in many tissues. Strong expression also seen

in epithelial cells of the fetal adrenal cortex. All adult tissues were negative.

Fetal tissues examined (E12-E16 weeks) include: Placenta, umbilical cord, liver, kidney, adrenals, thyroid, lungs, heart, great vessels, oesophagus, stomach, small intestine, spleen, thymus, pancreas, brain, eye, spinal cord, body wall, pelvis and lower limb.

Adult tissues examined: Liver, kidney, adrenal, spleen, lymph node, pancreas, lung, skin, cerebral cortex (rm), hippocampus(rm), cerebellum(rm), bladder, stomach, colon and colonic carcinoma. Acetominophen induced liver injury and hepatic cirrhosis.

A secondary screen evidenced expression over stromal mononuclear cells probably histiocytes.

(14) DNA38268-1188 (PRO295)

High expression over ganglion cells in human fetal spinal ganglia and over large neurones in the anterior horns of the developing spinal cord. In the adult there is expression in the chimp adrenal medulla (neural), neurones of the rhesus monkey brain (hippocampus [+++] and cerebral cortex) and neurones in ganglia in the normal adult human prostate (the only section that contains ganglion cells, ie expression in this cell type is presumed NOT to be confined to the prostate). All other tissues negative.

Human fetal tissues examined (E12-E16 weeks) include: Placenta, umbilical cord, liver, kidney, adrenals, thyroid, lungs, great vessels, stomach, small intestine, spleen, thymus, pancreas, brain, eye, spinal cord, body wall, pelvis, testis and lower limb.

Adult human tissues examined: Kidney (normal and end-stage), adrenal, spleen, lymph node, pancreas, lung, eye (inc. retina), bladder, liver (normal, cirrhotic, acute failure).

Non-human primate tissues examined:

Chimp Tissues: adrenal

Rhesus Monkey Tissues: Cerebral cortex, hippocampus, cerebellum.

Deposit of Material

The following materials have been deposited with the American Type Culture Collection, 12301 Parklawn Drive, Rockville, MD, USA (ATCC):

Material	ATCC Dep. No.	Deposit Date
DNA32292-1131	ATCC 209258	September 16, 1997
DNA33094-1131	ATCC 209256	September 16, 1997
DNA33223-1136	ATCC 209264	September 16, 1997
DNA34435-1140	ATCC 209250	September 16, 1997
DNA27864-1155	ATCC 209375	October 16, 1997
DNA36350-1158	ATCC 209378	October 16, 1997
DNA32290-1164	ATCC 209384	October 16, 1997
DNA35639-1172	ATCC 209396	October 17, 1997
DNA33092-1202	ATCC 209420	October 28, 1997
DNA49435-1219	ATCC 209480	November 21, 1997
DNA35638-1141	ATCC 209265	September 16, 1997
DNA32298-1132	ATCC 209257	September 16, 1997
DNA33089-1132	ATCC 209262	September 16, 1997
DNA33786-1132	ATCC 209253	September 16, 1997
DNA35918-1174	ATCC 209402	October 17, 1997

DNA37150-1178	ATCC 209401	October 17, 1997
DNA38260-1180	ATCC 209397	October 17, 1997
DNA39969-1185	ATCC 209400	October 17, 1997
DNA32286-1191	ATCC 209385	October 16, 1997
DNA33461-1199	ATCC 209367	October 15, 1997
DNA40628-1216	ATCC 209432	November 7, 1997
DNA33221-1133	ATCC 209263	September 16, 1997
DNA33107-1135	ATCC 209251	September 16, 1997
DNA35557-1137	ATCC 209255	September 16, 1997
DNA34434-1139	ATCC 209252	September 16, 1997
DNA33100-1159	ATCC 209373	October 16, 1997
DNA35600-1162	ATCC 209370	October 16, 1997
DNA34436-1238	ATCC 209523	December 10, 1997
DNA33206-1165	ATCC 209372	October 16, 1997
DNA35558-1167	ATCC 209374	October 16, 1997
DNA35599-1168	ATCC 209373	October 16, 1997
DNA36992-1168	ATCC 209382	October 16, 1997
DNA34407-1169	ATCC 209383	October 16, 1997
DNA35841-1173	ATCC 209403	October 17, 1997
DNA33470-1175	ATCC 209398	October 17, 1997
DNA34431-1177	ATCC 209399	October 17, 1997
DNA39510-1181	ATCC 209392	October 17, 1997
DNA39423-1182	ATCC 209387	October 17, 1997
DNA40620-1183	ATCC 209388	October 17, 1997
DNA40604-1187	ATCC 209394	October 17, 1997
DNA38268-1188	ATCC 209421	October 28, 1997
DNA37151-1193	ATCC 209393	October 17, 1997
DNA35673-1201	ATCC 209418	October 28, 1997
DNA40370-1217	ATCC 209485	November 21, 1997
DNA42551-1217	ATCC 209483	November 21, 1997
DNA39520-1217	ATCC 209482	November 21, 1997
DNA41225-1217	ATCC 209491	November 21, 1997
DNA43318-1217	ATCC 209481	November 21, 1997
DNA40587-1231	ATCC 209438	November 7, 1997
DNA41338-1234	ATCC 209927	June 2, 1998
DNA40981-1234	ATCC 209439	November 7, 1997
DNA37140-1234	ATCC 209489	November 21, 1997
DNA40982-1235	ATCC 209433	November 7, 1997
DNA41379-1236	ATCC 209488	November 21, 1997
DNA44167-1243	ATCC 209434	November 7, 1997
DNA39427-1179	ATCC 209395	October 17, 1997
DNA40603-1232	ATCC 209486	November 21, 1997
DNA43466-1225	ATCC 209490	November 21, 1997
DNA43046-1225	ATCC 209484	November 21, 1997
DNA35668-1171	ATCC 209371	October 16, 1997

These deposit were made under the provisions of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purpose of Patent Procedure and the Regulations thereunder (Budapest Treaty). This assures maintenance of a viable culture of the deposit for 30 years from the date of deposit. The deposits will be made available by ATCC under the terms of the Budapest Treaty, and subject to an agreement between Genentech, Inc. and ATCC, which assures permanent and unrestricted availability of the progeny of the culture of the deposit to the public upon issuance of the pertinent U.S. patent or upon laying open to the public of any U.S. or foreign patent application, whichever comes first, and assures availability of the progeny to one determined

by the U.S. Commissioner of Patents and Trademarks to be entitled thereto according to 35 USC § 122 and the Commissioner's rules pursuant thereto (including 37 CFR § 1.14 with particular reference to 886 OG 638).

The assignee of the present application has agreed that if a culture of the materials on deposit should die or be lost or destroyed when cultivated under suitable conditions, the materials will be promptly replaced on notification with another of the same. Availability of the deposited material is not to be construed as a license to practice the invention in contravention of the rights granted under the authority of any government in accordance with its patent laws.

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 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 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.

SEQUENCE LISTING

<110> GENENTECH, INC. et al.

<120> NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING THE SAME

<130> 10792 CA

<140> PCT/US 98/19330
<141> 1998-09-16

<160> 379

<170> PatentIn Ver. 2.0

<210> 1
<211> 1364
<212> DNA
<213> Homo sapiens

<400> 1
ggccggagca gcacggccgc aggacctgga gctccggctg cgtcttcccg cagcgctacc 60
cgccatgcgc ctgccgcgcc gggccgcgct ggggctcctg ccgcttctgc tgctgctgcc 120
gcccgcgccg gaggccgcca agaagccgac gccctgccac cggtgccggg ggctggtgga 180
caagtttaac caggggatgg tggacaccgc aaagaagaac tttggcggcg ggaacacggc 240
ttgggaggaa aagacgctgt ccaagtacga gtccagcgag attcgcctgc tggagatcct 300
ggaggggctg tgcgagagca gcgacttcga atgcaatcag atgctagagg cgcaggagga 360
gcacctggag gcctggtggc tgcagctgaa gagcgaatat cctgacttat tcgagtggtt 420
ttgtgtgaag acactgaaag tgtgctgctc tccaggaacc tacggtcccg actgtctcgc 480
atgccagggc ggatcccaga ggccctgcag cgggaatggc cactgcagcg gagatgggag 540
cagacagggc gacgggtcct gccggtgcca catggggtac cagggcccgc tgtgcactga 600
ctgcatggac ggctacttca gctcgctccg gaacgagacc cacagcatct gcacagcctg 660
tgacgagtcc tgcaagacgt gctcgggcct gaccaacaga gactgcggcg agtgtgaagt 720
gggctgggtg ctggacgagg gcgcctgtgt ggatgtggac gagtgtgcgg ccgagccgcc 780
tccctgcagc gctgcgcagt tctgtaagaa cgccaacggc tcctacacgt gcgaagagtg 840
tgactccagc tgtgtgggct gcacagggga aggcccagga aactgtaaag agtgtatctc 900
tggctacgcg agggagcacg gacagtgtgc agatgtggac gagtgctcac tagcagaaaa 960
aacctgtgtg aggaaaaacg aaaactgcta caatactcca gggagctacg tctgtgtgtg 1020
tcctgacggc ttcgaagaaa cggaagatgc ctgtgtgccg ccggcagagg ctgaagccac 1080
agaaggagaa agcccgacac agctgccctc ccgcgaagac ctgtaatgtg ccggacttac 1140
cctttaaatt attcagaagg atgtcccgtg gaaaatgtgg ccctgaggat gccgtctcct 1200
gcagtggaca gcggcgggga gaggctgcct gctctctaac ggttgattct catttgtccc 1260
ttaaacagct gcatttcttg gttgttctta aacagacttg tatattttga tacagttctt 1320
tgtaataaaa ttgaccattg taggtaatca ggaggaaaaa aaaa 1364

<210> 2
<211> 353
<212> PRT
<213> Homo sapiens

<400> 2
Met Arg Leu Pro Arg Arg Ala Ala Leu Gly Leu Leu Pro Leu Leu Leu
1 5 10 15

Leu Leu Pro Pro Ala Pro Glu Ala Ala Lys Lys Pro Thr Pro Cys His
20 25 30

Arg Cys Arg Gly Leu Val Asp Lys Phe Asn Gln Gly Met Val Asp Thr
35 40 45

Ala Lys Lys Asn Phe Gly Gly Gly Asn Thr Ala Trp Glu Glu Lys Thr
50 55 60

Leu Ser Lys Tyr Glu Ser Ser Glu Ile Arg Leu Leu Glu Ile Leu Glu
65 70 75 80

Gly Leu Cys Glu Ser Ser Asp Phe Glu Cys Asn Gln Met Leu Glu Ala
85 90 95

Gln Glu Glu His Leu Glu Ala Trp Trp Leu Gln Leu Lys Ser Glu Tyr
100 105 110

Pro Asp Leu Phe Glu Trp Phe Cys Val Lys Thr Leu Lys Val Cys Cys
115 120 125

Ser Pro Gly Thr Tyr Gly Pro Asp Cys Leu Ala Cys Gln Gly Gly Ser
130 135 140

Gln Arg Pro Cys Ser Gly Asn Gly His Cys Ser Gly Asp Gly Ser Arg
145 150 155 160

Gln Gly Asp Gly Ser Cys Arg Cys His Met Gly Tyr Gln Gly Pro Leu
165 170 175

Cys Thr Asp Cys Met Asp Gly Tyr Phe Ser Ser Leu Arg Asn Glu Thr
180 185 190

His Ser Ile Cys Thr Ala Cys Asp Glu Ser Cys Lys Thr Cys Ser Gly
195 200 205

Leu Thr Asn Arg Asp Cys Gly Glu Cys Glu Val Gly Trp Val Leu Asp
210 215 220

Glu Gly Ala Cys Val Asp Val Asp Glu Cys Ala Ala Glu Pro Pro Pro
225 230 235 240

Cys Ser Ala Ala Gln Phe Cys Lys Asn Ala Asn Gly Ser Tyr Thr Cys
245 250 255

Glu Glu Cys Asp Ser Ser Cys Val Gly Cys Thr Gly Glu Gly Pro Gly
260 265 270

Asn Cys Lys Glu Cys Ile Ser Gly Tyr Ala Arg Glu His Gly Gln Cys
275 280 285

Ala Asp Val Asp Glu Cys Ser Leu Ala Glu Lys Thr Cys Val Arg Lys
290 295 300

Asn Glu Asn Cys Tyr Asn Thr Pro Gly Ser Tyr Val Cys Val Cys Pro
305 310 315 320

Asp Gly Phe Glu Glu Thr Glu Asp Ala Cys Val Pro Pro Ala Glu Ala
325 330 335

Glu Ala Thr Glu Gly Glu Ser Pro Thr Gln Leu Pro Ser Arg Glu Asp
340 345 350

Leu

<210> 3
<211> 2033
<212> DNA
<213> Homo sapiens

<400> 3
ccaggccggg aggcgacgcg cccagccgtc taaacgggaa cagccctggc tgagggagct 60
gcagcgcagc agagtatctg acggcgccag gttgcgtagg tgcggcacga ggagttttcc 120
cggcagcgag gaggtcctga gcagcatggc ccggaggagc gccttccctg ccgccgcgct 180
ctggctctgg agcatcctcc tgtgcctgct ggcactgcgg gcggaggccg ggccgccgca 240
ggaggagagc ctgtacctat ggatcgatgc tcaccaggca agagtactca taggatttga 300
agaagatatc ctgattgttt cagaggggaa aatggcacct tttacacatg atttcagaaa 360
agcgcaacag agaatgccag ctattcctgt caatatccat tccatgaatt ttacctggca 420
agctgcaggg caggcagaat acttctatga attcctgtcc ttgcgctccc tggataaagg 480
catcatggca gatccaaccg tcaatgtccc tctgctggga acagtgcctc acaaggcatc 540
agttgttcaa gttggtttcc catgtcttgg aaaacaggat ggggtggcag catttgaagt 600
ggatgtgatt gttatgaatt ctgaaggcaa caccattctc caaacacctc aaaatgctat 660
cttctttaaa acatgtcaac aagctgagtg cccaggcggg tgccgaaatg gaggcttttg 720
taatgaaaga cgcatctgcg agtgtcctga tgggttccac ggacctcact gtgagaaagc 780
cctttgtacc ccacgatgta tgaatggtgg actttgtgtg actcctggtt tctgcatctg 840
cccacctgga ttctatggag tgaactgtga caaagcaaac tgctcaacca cctgctttaa 900
tggagggacc tgtttctacc ctggaaaatg tatttgccct ccaggactag agggagagca 960
gtgtgaaatc agcaaatgcc cacaaccctg tcgaaatgga ggtaaatgca ttggtaaaag 1020
caaatgtaag tgttccaaag gttaccaggg agacctctgt tcaaagcctg tctgcgagcc 1080
tggctgtggt gcacatggaa cctgccatga acccaacaaa tgccaatgtc aagaaggttg 1140
gcatggaaga cactgcaata aaaggtacga agccagcctc atacatgccc tgaggccagc 1200
aggcgcccag ctcaggcagc acacgccttc acttaaaaag gccgaggagc ggcgggatcc 1260
acctgaatcc aattacatct ggtgaactcc gacatctgaa acgttttaag ttacaccaag 1320
ttcatagcct ttgttaacct ttcatgtgtt gaatgttcaa ataatgttca ttacacttaa 1380
gaatactggc ctgaatttta ttagcttcat tataaatcac tgagctgata tttactcttc 1440
cttttaagtt ttctaagtac gtctgtagca tgatggtata gattttcttg tttcagtgct 1500
ttgggacaga ttttatatta tgtcaattga tcaggttaaa attttcagtg tgtagttggc 1560
agatattttc aaaattacaa tgcatttatg gtgtctgggg gcaggggaac atcagaaagg 1620
ttaaattggg caaaaatgcg taagtcacaa gaatttggat ggtgcagtta atgttgaagt 1680
tacagcattt cagattttat tgtcagatat ttagatgttt gttacatttt taaaaattgc 1740
tcttaatttt taaactctca atacaatata ttttgacctt accattattc cagagattca 1800
gtattaaaaa aaaaaaaatt acactgtggt agtggcattt aaacaatata atatattcta 1860
aacacaatga aatagggaat ataatgtatg aacttttttgc attggcttga agcaatataa 1920
tatattgtaa acaaaacaca gctcttacct aataaacatt ttatactgtt tgtatgtata 1980
aaataaaggt gctgctttag ttttttggaa aaaaaaaaaa aaaaaaaaaa aaa 2033

<210> 4
<211> 379
<212> PRT
<213> Homo sapiens

<400> 4
Met Ala Arg Arg Ser Ala Phe Pro Ala Ala Ala Leu Trp Leu Trp Ser
1 5 10 15

Ile Leu Leu Cys Leu Leu Ala Leu Arg Ala Glu Ala Gly Pro Pro Gln
20 25 30

Glu Glu Ser Leu Tyr Leu Trp Ile Asp Ala His Gln Ala Arg Val Leu
35 40 45

Ile Gly Phe Glu Glu Asp Ile Leu Ile Val Ser Glu Gly Lys Met Ala
50 55 60

Pro Phe Thr His Asp Phe Arg Lys Ala Gln Gln Arg Met Pro Ala Ile
65 70 75 80

```
Pro Val Asn Ile His Ser Met Asn Phe Thr Trp Gln Ala Ala Gly Gln
                   85                  90                  95

Ala Glu Tyr Phe Tyr Glu Phe Leu Ser Leu Arg Ser Leu Asp Lys Gly
               100                 105                 110

Ile Met Ala Asp Pro Thr Val Asn Val Pro Leu Leu Gly Thr Val Pro
           115                 120                 125

His Lys Ala Ser Val Val Gln Val Gly Phe Pro Cys Leu Gly Lys Gln
       130                 135                 140

Asp Gly Val Ala Ala Phe Glu Val Asp Val Ile Val Met Asn Ser Glu
145                 150                 155                 160

Gly Asn Thr Ile Leu Gln Thr Pro Gln Asn Ala Ile Phe Phe Lys Thr
                   165                 170                 175

Cys Gln Gln Ala Glu Cys Pro Gly Gly Cys Arg Asn Gly Gly Phe Cys
               180                 185                 190

Asn Glu Arg Arg Ile Cys Glu Cys Pro Asp Gly Phe His Gly Pro His
           195                 200                 205

Cys Glu Lys Ala Leu Cys Thr Pro Arg Cys Met Asn Gly Gly Leu Cys
       210                 215                 220

Val Thr Pro Gly Phe Cys Ile Cys Pro Pro Gly Phe Tyr Gly Val Asn
225                 230                 235                 240

Cys Asp Lys Ala Asn Cys Ser Thr Thr Cys Phe Asn Gly Gly Thr Cys
                   245                 250                 255

Phe Tyr Pro Gly Lys Cys Ile Cys Pro Pro Gly Leu Glu Gly Glu Gln
               260                 265                 270

Cys Glu Ile Ser Lys Cys Pro Gln Pro Cys Arg Asn Gly Gly Lys Cys
           275                 280                 285

Ile Gly Lys Ser Lys Cys Lys Cys Ser Lys Gly Tyr Gln Gly Asp Leu
       290                 295                 300

Cys Ser Lys Pro Val Cys Glu Pro Gly Cys Gly Ala His Gly Thr Cys
305                 310                 315                 320

His Glu Pro Asn Lys Cys Gln Cys Gln Glu Gly Trp His Gly Arg His
                   325                 330                 335

Cys Asn Lys Arg Tyr Glu Ala Ser Leu Ile His Ala Leu Arg Pro Ala
               340                 345                 350

Gly Ala Gln Leu Arg Gln His Thr Pro Ser Leu Lys Lys Ala Glu Glu
           355                 360                 365

Arg Arg Asp Pro Pro Glu Ser Asn Tyr Ile Trp
       370                 375
```

<210> 5
<211> 45

<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 5
agggagcacg gacagtgtgc agatgtggac gagtgctcac tagca 45

<210> 6
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 6
agagtgtatc tctggctacg c 21

<210> 7
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 7
taagtccggc acattacagg tc 22

<210> 8
<211> 49
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 8
cccacgatgt atgaatggtg gactttgtgt gactcctggt ttctgcatc 49

<210> 9
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 9
aaagacgcat ctgcgagtgt cc 22

<210> 10
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 10
tgctgatttc acactgctct ccc 23

<210> 11
<211> 2217
<212> DNA
<213> Homo sapiens

<400> 11
cggacgcgtg ggcgtccggc ggtcgcagag ccaggaggcg gaggcgcgcg ggccagcctg 60
ggccccagcc cacaccttca ccagggccca ggagccacca tgtggcgatg tccactgggg 120
ctactgctgt tgctgccgct ggctggccac ttggctctgg gtgcccagca gggtcgtggg 180
cgccgggagc tagcaccggg tctgcacctg cggggcatcc gggacgcggg aggccggtac 240
tgccaggagc aggacctgtg ctgccgcggc cgtgccgacg actgtgccct gccctacctg 300
ggcgccatct gttactgtga cctcttctgc aaccgcacgg tctccgactg ctgccctgac 360
ttctgggact tctgcctcgg cgtgccaccc ccttttcccc cgatccaagg atgtatgcat 420
ggaggtcgta tctatccagt cttgggaacg tactgggaca actgtaaccg ttgcacctgc 480
caggagaaca ggcagtggca gtgtgaccaa gaaccatgcc tggtggatcc agacatgatc 540
aaagccatca accagggcaa ctatggctgg caggctggga accacagcgc cttctggggc 600
atgaccctgg atgagggcat tcgctaccgc ctgggcacca tccgcccatc ttcctcggtc 660
atgaacatgc atgaaattta tacagtgctg aacccagggg aggtgcttcc cacagccttc 720
gaggcctctg agaagtggcc caacctgatt catgagcctc ttgaccaagg caactgtgca 780
ggctcctggg ccttctccac agcagctgtg gcatccgatc gtgtctcaat ccattctctg 840
ggacacatga cgcctgtcct gtcgccccag aacctgctgt cttgtgacac ccaccagcag 900
cagggctgcc gcggtgggcg tctcgatggt gcctggtggt tcctgcgtcg ccgaggggtg 960
gtgtctgacc actgctaccc cttctcgggc cgtgaacgag acgaggctgg ccctgcgccc 1020
ccctgtatga tgcacagccg agccatgggt cggggcaagc gccaggccac tgcccactgc 1080
cccaacagct atgttaataa caatgacatc taccaggtca ctcctgtcta ccgcctcggc 1140
tccaacgaca aggagatcat gaaggagctg atggagaatg gccctgtcca agccctcatg 1200
gaggtgcatg aggacttctt cctatacaag ggaggcatct acagccacac gccagtgagc 1260
cttgggaggc cagagagata ccgccggcat gggacccact cagtcaagat cacaggatgg 1320
ggagaggaga cgctgccaga tggaaggacg ctcaaatact ggactgcggc caactcctgg 1380
ggcccagcct ggggcgagag gggccacttc cgcatcgtgc gcggcgtcaa tgagtgcgac 1440
atcgagagct tcgtgctggg cgtctggggc cgcgtgggca tggaggacat gggtcatcac 1500
tgaggctgcg ggcaccacgc ggggtccggc ctgggatcca ggctaagggc cggcggaaga 1560
ggccccaatg gggcggtgac cccagcctcg cccgacagag cccggggcgc aggcgggcgc 1620
cagggcgcta atcccggcgc gggttccgct gacgcagcgc cccgcctggg agccgcgggc 1680
aggcgagact ggcggagccc ccagacctcc cagtggggac ggggcagggc ctggcctggg 1740
aagagcacag ctgcagatcc caggcctctg gcgcccccac tcaagactac caaagccagg 1800
acacctcaag tctccagccc caatacccca ccccaatccc gtattctttt ttttttttct 1860
ttagacaggg tcttgctccg ttgcccaggt tggagtgcag tggcccatca gggctcactg 1920
taacctccga ctcctgggtt caagtgaccc tcccacctca gcctctcaag tagctgggac 1980
tacaggtgca ccaccacacc tggctaattt ttgtattttt tgtaaagagg ggggtctcac 2040
tgtgttgccc aggctggttt cgaactcctg ggctcaagcg gtccacctgc ctccgcctcc 2100
caaagtgctg ggattgcagg catgagccac tgcacccagc cctgtattct tattcttcag 2160
atatttattt ttctttttcac tgttttaaaa taaaaccaaa gtattgataa aaaaaaa 2217

<210> 12
<211> 467
<212> PRT
<213> Homo sapiens

<400> 12
Met Trp Arg Cys Pro Leu Gly Leu Leu Leu Leu Leu Pro Leu Ala Gly
1 5 10 15

His Leu Ala Leu Gly Ala Gln Gln Gly Arg Gly Arg Arg Glu Leu Ala
20 25 30

Pro Gly Leu His Leu Arg Gly Ile Arg Asp Ala Gly Gly Arg Tyr Cys
35 40 45

Gln Glu Gln Asp Leu Cys Cys Arg Gly Arg Ala Asp Asp Cys Ala Leu
50 55 60

Pro Tyr Leu Gly Ala Ile Cys Tyr Cys Asp Leu Phe Cys Asn Arg Thr
65 70 75 80

Val Ser Asp Cys Cys Pro Asp Phe Trp Asp Phe Cys Leu Gly Val Pro
85 90 95

Pro Pro Phe Pro Pro Ile Gln Gly Cys Met His Gly Gly Arg Ile Tyr
100 105 110

Pro Val Leu Gly Thr Tyr Trp Asp Asn Cys Asn Arg Cys Thr Cys Gln
115 120 125

Glu Asn Arg Gln Trp Gln Cys Asp Gln Glu Pro Cys Leu Val Asp Pro
130 135 140

Asp Met Ile Lys Ala Ile Asn Gln Gly Asn Tyr Gly Trp Gln Ala Gly
145 150 155 160

Asn His Ser Ala Phe Trp Gly Met Thr Leu Asp Glu Gly Ile Arg Tyr
165 170 175

Arg Leu Gly Thr Ile Arg Pro Ser Ser Ser Val Met Asn Met His Glu
180 185 190

Ile Tyr Thr Val Leu Asn Pro Gly Glu Val Leu Pro Thr Ala Phe Glu
195 200 205

Ala Ser Glu Lys Trp Pro Asn Leu Ile His Glu Pro Leu Asp Gln Gly
210 215 220

Asn Cys Ala Gly Ser Trp Ala Phe Ser Thr Ala Ala Val Ala Ser Asp
225 230 235 240

Arg Val Ser Ile His Ser Leu Gly His Met Thr Pro Val Leu Ser Pro
245 250 255

Gln Asn Leu Leu Ser Cys Asp Thr His Gln Gln Gln Gly Cys Arg Gly
260 265 270

Gly Arg Leu Asp Gly Ala Trp Trp Phe Leu Arg Arg Arg Gly Val Val
275 280 285

Ser Asp His Cys Tyr Pro Phe Ser Gly Arg Glu Arg Asp Glu Ala Gly
290 295 300

Pro Ala Pro Pro Cys Met Met His Ser Arg Ala Met Gly Arg Gly Lys
305 310 315 320

Arg Gln Ala Thr Ala His Cys Pro Asn Ser Tyr Val Asn Asn Asn Asp
325 330 335

Ile Tyr Gln Val Thr Pro Val Tyr Arg Leu Gly Ser Asn Asp Lys Glu
340 345 350

Ile Met Lys Glu Leu Met Glu Asn Gly Pro Val Gln Ala Leu Met Glu
355 360 365

Val His Glu Asp Phe Phe Leu Tyr Lys Gly Gly Ile Tyr Ser His Thr
370 375 380

Pro Val Ser Leu Gly Arg Pro Glu Arg Tyr Arg Arg His Gly Thr His
385 390 395 400

Ser Val Lys Ile Thr Gly Trp Gly Glu Glu Thr Leu Pro Asp Gly Arg
405 410 415

Thr Leu Lys Tyr Trp Thr Ala Ala Asn Ser Trp Gly Pro Ala Trp Gly
420 425 430

Glu Arg Gly His Phe Arg Ile Val Arg Gly Val Asn Glu Cys Asp Ile
435 440 445

Glu Ser Phe Val Leu Gly Val Trp Gly Arg Val Gly Met Glu Asp Met
450 455 460

Gly His His
465

<210> 13
<211> 533
<212> DNA
<213> Homo sapiens

<400> 13
aggctccttg gccctttttc cacagcaagc ttntgcnatc ccgattcgtt gtctcaaatc 60
caattctctt gggacacatn acgcctgtcc tttngcccca gaacctgctg tcttgtacac 120
ccaccagcag cagggctgcc gcgntgggcg tctcgatggt gcctggtggt tcctgcgtcg 180
ccgagggntg gtgtctgacc actgctaccc cttctcgggc cgtgaacgag acgaggctgg 240
ccctgcgccc ccctgtatga tgcacagccg agccatgggt cggggcaagc gccaggccac 300
tgcccactgc cccaacagct atgttaataa caatgacatc taccaggtca ctcctgtcta 360
ccgcctcggc tccaacgaca aggagatcat gaaggagctg atggagaatg gccctgtcca 420
agcccctcatg gaggtgcatg aggacttctt cctatacaag ggaggcatct acagccacac 480
gccagtgagc cttgggaggc cagagagata ccgccggcat gggacccact cag 533

<210> 14
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 14
ttcgaggcct ctgagaagtg gccc 24

<210> 15
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 15
ggcggtatct ctctggcctc cc 22

<210> 16

<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 16
ttctccacag cagctgtggc atccgatcgt gtctcaatcc attctctggg 50

<210> 17
<211> 960
<212> DNA
<213> Homo sapiens

<400> 17
gctgcttgcc ctgttgatgg caggcttggc cctgcagcca ggcactgccc tgctgtgcta 60
ctcctgcaaa gcccaggtga gcaacgagga ctgcctgcag gtggagaact gcacccagct 120
gggggagcag tgctggaccg cgcgcatccg cgcagttggc ctcctgaccg tcatcagcaa 180
aggctgcagc ttgaactgcg tggatgactc acaggactac tacgtgggca agaagaacat 240
cacgtgctgt gacaccgact tgtgcaacgc cagcggggcc catgccctgc agccggctgc 300
cgccatcctt gcgctgctcc ctgcactcgg cctgctgctc tggggacccg gccagctata 360
ggctctgggg ggccccgctg cagcccacac tgggtgtggt gccccaggcc tctgtgccac 420
tcctcacaga cctggcccag tgggagcctg tcctggttcc tgaggcacat cctaacgcaa 480
gtctgaccat gtatgtctgc acccctgtcc cccaccctga ccctcccatg gccctctcca 540
ggactcccac ccggcagatc agctctagtg acacagatcc gcctgcagat ggccccctcca 600
accctctctg ctgctgtttc catggcccag cattctccac ccttaaccct gtgctcaggc 660
acctcttccc ccaggaagcc ttccctgccc accccatcta tgacttgagc caggtctggt 720
ccgtggtgtc ccccgcaccc agcaggggac aggcactcag gagggcccag taaaggctga 780
gatgaagtgg actgagtaga actggaggac aagagtcgac gtgagttcct gggagtctcc 840
agagatgggg cctggaggcc tggaggaagg ggccaggcct cacattcgtg gggctccctg 900
aatggcagcc tgagcacagc gtaggcccctt aataaacacc tgttggataa gccaaaaaaa 960

<210> 18
<211> 114
<212> PRT
<213> Homo sapiens

<400> 18
Met Ala Gly Leu Ala Leu Gln Pro Gly Thr Ala Leu Leu Cys Tyr Ser
1 5 10 15

Cys Lys Ala Gln Val Ser Asn Glu Asp Cys Leu Gln Val Glu Asn Cys
20 25 30

Thr Gln Leu Gly Glu Gln Cys Trp Thr Ala Arg Ile Arg Ala Val Gly
35 40 45

Leu Leu Thr Val Ile Ser Lys Gly Cys Ser Leu Asn Cys Val Asp Asp
50 55 60

Ser Gln Asp Tyr Tyr Val Gly Lys Lys Asn Ile Thr Cys Cys Asp Thr
65 70 75 80

Asp Leu Cys Asn Ala Ser Gly Ala His Ala Leu Gln Pro Ala Ala Ala
85 90 95

Ile Leu Ala Leu Leu Pro Ala Leu Gly Leu Leu Leu Trp Gly Pro Gly
100 105 110

Gln Leu

<210> 19
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 19
tgctgtgcta ctcctgcaaa gccc 24

<210> 20
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 20
tgcacaagtc ggtgtcacag cacg 24

<210> 21
<211> 44
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 21
agcaacgagg actgcctgca ggtggagaac tgcacccagc tggg 44

<210> 22
<211> 1175
<212> DNA
<213> Homo sapiens

<400> 22
atgggagccg cccgcctgct gcccaacctc actctgtgct tacagctgct gattctctgc 60
tgtcaaactc agtacgtgag ggaccagggc gccatgaccg accagctgag caggcggcag 120
atccgcgagt accaactcta cagcaggacc agtggcaagc acgtgcaggt caccgggcgt 180
cgcatctccg ccaccgccga ggacggcaac aagtttgcca agctcatagt ggagacggac 240
acgtttggca gccgggttcg catcaaaggg gctgagagtg agaagtacat ctgtatgaac 300
aagaggggca agctcatcgg gaagcccagc gggaagagca aagactgcgt gttcacggag 360
atcgtgctgg agaacaacta tacggccttc cagaacgccc ggcacgaggg ctggttcatg 420
gccttcacgc ggcaggggcg gccccgccag gcttcccgca gccgccagaa ccagcgcgag 480
gcccacttca tcaagcgcct ctaccaaggc cagctgccct tccccaacca cgccgagaag 540
cagaagcagt tcgagtttgt gggctccgcc cccacccgcc ggaccaagcg cacacggcgg 600
ccccagcccc tcacgtagtc tgggaggcag ggggcagcag cccctgggcc gcctccccac 660
ccctttccct tcttaatcca aggactgggc tggggtggcg ggaggggagc cagatccccg 720
agggaggacc ctgagggccg cgaagcatcc gagcccccag ctgggaaggg gcaggccggt 780
gccccagggg cggctggcac agtgcccccc tcccggacgg gtggcaggcc ctggagagga 840
actgagtgtc accctgatct caggccacca gcctttgccg gcctcccagc cgggctcctg 900
aagcccgctg aaaggtcagc gactgaaggc cttgcagaca accgtctgga ggtggctgtc 960
ctcaaaatct gcttctcgga tctccctcag tctgccccca gcccccaaac tcctcctggc 1020
tagactgtag gaagggactt ttgtttgttt gtttgtttca ggaaaaaaga aagggagaga 1080
gaggaaaata gagggttgtc cactcctcac attccacgac ccaggcctgc accccacccc 1140

caactcccag ccccggaata aaaccatttt cctgc 1175

<210> 23
<211> 205
<212> PRT
<213> Homo sapiens

<400> 23

Met	Gly	Ala	Ala	Arg	Leu	Leu	Pro	Asn	Leu	Thr	Leu	Cys	Leu	Gln	Leu
1				5				10					15		
Leu	Ile	Leu	Cys	Cys	Gln	Thr	Gln	Tyr	Val	Arg	Asp	Gln	Gly	Ala	Met
			20				25					30			
Thr	Asp	Gln	Leu	Ser	Arg	Arg	Gln	Ile	Arg	Glu	Tyr	Gln	Leu	Tyr	Ser
		35					40					45			
Arg	Thr	Ser	Gly	Lys	His	Val	Gln	Val	Thr	Gly	Arg	Arg	Ile	Ser	Ala
	50					55					60				
Thr	Ala	Glu	Asp	Gly	Asn	Lys	Phe	Ala	Lys	Leu	Ile	Val	Glu	Thr	Asp
65					70					75					80
Thr	Phe	Gly	Ser	Arg	Val	Arg	Ile	Lys	Gly	Ala	Glu	Ser	Glu	Lys	Tyr
				85					90					95	
Ile	Cys	Met	Asn	Lys	Arg	Gly	Lys	Leu	Ile	Gly	Lys	Pro	Ser	Gly	Lys
			100					105					110		
Ser	Lys	Asp	Cys	Val	Phe	Thr	Glu	Ile	Val	Leu	Glu	Asn	Asn	Tyr	Thr
		115					120					125			
Ala	Phe	Gln	Asn	Ala	Arg	His	Glu	Gly	Trp	Phe	Met	Ala	Phe	Thr	Arg
	130					135					140				
Gln	Gly	Arg	Pro	Arg	Gln	Ala	Ser	Arg	Ser	Arg	Gln	Asn	Gln	Arg	Glu
145					150					155					160
Ala	His	Phe	Ile	Lys	Arg	Leu	Tyr	Gln	Gly	Gln	Leu	Pro	Phe	Pro	Asn
				165					170					175	
His	Ala	Glu	Lys	Gln	Lys	Gln	Phe	Glu	Phe	Val	Gly	Ser	Ala	Pro	Thr
			180					185					190		
Arg	Arg	Thr	Lys	Arg	Thr	Arg	Arg	Pro	Gln	Pro	Leu	Thr			
		195					200					205			

<210> 24
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 24
cagtacgtga gggaccaggg cgccatga 28

<210> 25
<211> 24

DEMANDES OU BREVETS VOLUMINEUX

LA PRÉSENTE PARTIE DE CETTE DEMANDE OU CE BREVETS COMPREND PLUS D'UN TOME.

CECI EST LE TOME 1 DE 2

NOTE: Pour les tomes additionels, veillez contacter le Bureau Canadien des Brevets.

JUMBO APPLICATIONS / PATENTS

THIS SECTION OF THE APPLICATION / PATENT CONTAINS MORE THAN ONE VOLUME.

THIS IS VOLUME 1 OF 2

NOTE: For additional volumes please contact the Canadian Patent Office.

JJ-10 792-3CA

THE EMBODIMENTS OF THE INVENTION IN WHICH AN EXCLUSIVE PROPERTY OR PRIVILEGE IS CLAIMED ARE DEFINED AS FOLLOWS:

1. Isolated nucleic acid having at least 80% sequence identity to a nucleotide sequence that encodes a polypeptide comprising an amino acid sequence shown in Figure 44 (SEQ ID NO.: 119).

2. The nucleic acid of Claim 1, wherein said nucleotide sequence comprises a nucleotide sequence shown in Figure 43 (SEQ ID NO.: 118) or the complement thereof.

3. The nucleic acid of Claim 1, wherein said nucleotide sequence comprises the full-length coding sequence of the sequence shown in Figure 43 (SEQ ID NO: 118) or the complement thereof.

4. Isolated nucleic acid which comprises the full-length coding sequence of the DNA deposited under accession number ATCC 209432.

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.

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

8. The host cell of claim 7 wherein said cell is CHO cell.

9. The host cell of claim 7 wherein said cell is an *E.coli*.

10. The host cell of claim 7 wherein said cell is a yeast cell.

11. A process for producing a PRO 301 polypeptides comprising culturing the host cell of claim 7 under conditions suitable for expression of said PRO 301 polypeptide and recovering said PRO 301 polypeptide from the cell culture.

12. Isolated native sequence PRO 301 polypeptide having at least 80% sequence identity to an amino acid sequence selected from the group consisting of the amino acid sequence shown in Figure 44 (SEQ ID NO.: 119).

13. Isolated PRO 301 polypeptide having at least 80% sequence identity to the amino acid sequence encoded by the nucleotide deposited under accession number ATCC 209432.

14. A chimeric molecule comprising a polypeptide according to claim 12 fused to a heterologous amino acid sequence.

15. The chimeric molecule of claim 14 wherein said heterologous amino acid sequence is an epitope tag sequence.

16. The chimeric molecule of claim 14 wherein said heterologous amino acid sequence is an Fc region of an immunoglobulin.

FIGURE 1

GGCCGGAGCAGCACGGCCGCAGGACCTGGAGCTCCGGCTGCGTCTTCCCGCAGCGCTACCCG
CCATGCGCCTGCCGCGCCGGGCCGCGCTGGGGCTCCTGCCGCTTCTGCTGCTGCTGCCGCCC
GCGCCGGAGGCCGCCAAGAAGCCGACGCCCTGCCACCGGTGCCGGGGGCTGGTGGACAAGTT
TAACCAGGGGATGGTGGACACCGCAAAGAAGAACTTTGGCGGCGGGAACACGGCTTGGGAGG
AAAAGACGCTGTCCAAGTACGAGTCCAGCGAGATTCGCCTGCTGGAGATCCTGGAGGGGCTG
TGCGAGAGCAGCGACTTCGAATGCAATCAGATGCTAGAGGCGCAGGAGGAGCACCTGGAGGC
CTGGTGGCTGCAGCTGAAGAGCGAATATCCTGACTTATTCGAGTGGTTTTGTGTGAAGACAC
TGAAAGTGTGCTGCTCTCCAGGAACCTACGGTCCCGACTGTCTCGCATGCCAGGGCGGATCC
CAGAGGCCCTGCAGCGGGAATGGCCACTGCAGCGGAGATGGGAGCAGACAGGGCGACGGGTC
CTGCCGGTGCCACATGGGGTACCAGGGCCCGCTGTGCACTGACTGCATGGACGGCTACTTCA
GCTCGCTCCGGAACGAGACCCACAGCATCTGCACAGCCTGTGACGAGTCCTGCAAGACGTGC
TCGGGCCTGACCAACAGAGACTGCGGCGAGTGTGAAGTGGGCTGGGTGCTGGACGAGGGCGC
CTGTGTGGATGTGGACGAGTGTGCGGCCGAGCCGCCTCCCTGCAGCGCTGCGCAGT
TCTGTAAGAACGCCAACGGCTCCTACACGTGCGAAGAGTGTGACTCCAGCTGTGTGGGCTGC
ACAGGGGAAGGCCCAGGAAACTGTAAAGAGTGTATCTCTGGCTACGCGAGGGAGCACGGACA
GTGTGCAGATGTGGACGAGTGCTCACTAGCAGAAAAAACCTGTGTGAGGAAAAACGAAAACT
GCTACAATACTCCAGGGAGCTACGTCTGTGTGTGTCCTGACGGCTTCGAAGAAACGGAAGAT
GCCTGTGTGCCGCCGGCAGAGGCTGAAGCCACAGAAGGAGAAAGCCCGACACAGCTGCCCTC
CCGCGAAGACCTGTAATGTGCCGGACTTACCCTTTAAATTATTCAGAAGGATGTCCCGTGGA
AAATGTGGCCCTGAGGATGCCGTCTCCTGCAGTGGACAGCGGCGGGGAGAGGCTGCCTGCTC
TCTAACGGTTGATTCTCATTTGTCCCTTAAACAGCTGCATTTCTTGGTTGTTCTTAAACAGA
CTTGTATATTTTGATACAGTTCTTTGTAATAAAATTGACCATTGTAGGTAATCAGGAGGAAA
AAAAAA

FIGURE 2

Met Arg Leu Pro Arg Arg Ala Ala Leu Gly Leu Leu Pro Leu Leu Leu
Leu Leu Pro Pro Ala Pro Glu Ala Ala Lys Lys Pro Thr Pro Cys His
Arg Cys Arg Gly Leu Val Asp Lys Phe Asn Gln Gly Met Val Asp Thr
Ala Lys Lys Asn Phe Gly Gly Gly Asn Thr Ala Trp Glu Glu Lys Thr
Leu Ser Lys Tyr Glu Ser Ser Glu Ile Arg Leu Leu Glu Ile Leu Glu
Gly Leu Cys Glu Ser Ser Asp Phe Glu Cys Asn Gln Met Leu Glu Ala
Gln Glu Glu His Leu Glu Ala Trp Trp Leu Gln Leu Lys Ser Glu Tyr
Pro Asp Leu Phe Glu Trp Phe Cys Val Lys Thr Leu Lys Val Cys Cys
Ser Pro Gly Thr Tyr Gly Pro Asp Cys Leu Ala Cys Gln Gly Gly Ser
Gln Arg Pro Cys Ser Gly Asn Gly His Cys Ser Gly Asp Gly Ser Arg
Gln Gly Asp Gly Ser Cys Arg Cys His Met Gly Tyr Gln Gly Pro Leu
Cys Thr Asp Cys Met Asp Gly Tyr Phe Ser Ser Leu Arg Asn Glu Thr
His Ser Ile Cys Thr Ala Cys Asp Glu Ser Cys Lys Thr Cys Ser Gly
Leu Thr Asn Arg Asp Cys Gly Glu Cys Glu Val Gly Trp Val Leu Asp
Glu Gly Ala Cys Val Asp Val Asp Glu Cys Ala Ala Glu Pro Pro Pro
Cys Ser Ala Ala Gln Phe Cys Lys Asn Ala Asn Gly Ser Tyr Thr Cys
Glu Glu Cys Asp Ser Ser Cys Val Gly Cys Thr Gly Glu Gly Pro Gly
Asn Cys Lys Glu Cys Ile Ser Gly Tyr Ala Arg Glu His Gly Gln Cys
Ala Asp Val Asp Glu Cys Ser Leu Ala Glu Lys Thr Cys Val Arg Lys
Asn Glu Asn Cys Tyr Asn Thr Pro Gly Ser Tyr Val Cys Val Cys Pro
Asp Gly Phe Glu Glu Thr Glu Asp Ala Cys Val Pro Pro Ala Glu Ala
Glu Ala Thr Glu Gly Glu Ser Pro Thr Gln Leu Pro Ser Arg Glu Asp
Leu

FIGURE 3

CCAGGCCGGGAGGCGACGCGCCCAGCCGTCTAAACGGGAACAGCCCTGGCTGAGGGAGCTGC
AGCGCAGCAGAGTATCTGACGGCGCCAGGTTGCGTAGGTGCGGCACGAGGAGTTTTCCCGGC
AGCGAGGAGGTCCTGAGCAGCATGGCCCGGAGGAGCGCCTTCCCTGCCGCCGCGCTCTGGCT
CTGGAGCATCCTCCTGTGCCTGCTGGCACTGCGGGCGGAGGCCGGGCCGCCGCAGGAGGAGA
GCCTGTACCTATGGATCGATGCTCACCAGGCAAGAGTACTCATAGGATTTGAAGAAGATATC
CTGATTGTTTCAGAGGGGAAAATGGCACCTTTTACACATGATTTCAGAAAAGCGCAACAGAG
AATGCCAGCTATTCCTGTCAATATCCATTCCATGAATTTTACCTGGCAAGCTGCAGGGCAGG
CAGAATACTTCTATGAATTCCTGTCCTTGCGCTCCCTGGATAAAGGCATCATGGCAGATCCA
ACCGTCAATGTCCCTCTGCTGGGAACAGTGCCTCACAAGGCATCAGTTGTTCAAGTTGGTTT
CCCATGTCTTGGAAAACAGGATGGGGTGGCAGCATTTGAAGTGGATGTGATTGTTATGAATT
CTGAAGGCAACACCATTCTCCAAACACCTCAAAATGCTATCTTCTTTAAAACATGTCAACAA
GCTGAGTGCCCAGGCGGGTGCCGAAATGGAGGCTTTTGTAATGAAAGACGCATCTGCGAGTG
TCCTGATGGGTTCCACGGACCTCACTGTGAGAAAGCCCTTTGTACCCCACGATGTATGAATG
GTGGACTTTGTGTGACTCCTGGTTTCTGCATCTGCCCACCTGGATTCTATGGAGTGAACTGT
GACAAAGCAAACTGCTCAACCACCTGCTTTAATGGAGGGACCTGTTTCTACCCTGGAAAATG
TATTTGCCCTCCAGGACTAGAGGGAGAGCAGTGTGAAATCAGCAAATGCCCACAACCCTGTC
GAAATGGAGGTAAATGCATTGGTAAAAGCAAATGTAAGTGTTCCAAAGGTTACCAGGGAGAC
CTCTGTTCAAAGCCTGTCTGCGAGCCTGGCTGTGGTGCACATGGAACCTGCCATGAACCCAA
CAAATGCCAATGTCAAGAAGGTTGGCATGGAAGACACTGCAATAAAAGGTACGAAGCCAGCC
TCATACATGCCCTGAGGCCAGCAGGCGCCCAGCTCAGGCAGCACACGCCTTCACTTAAAAAG
GCCGAGGAGCGGCGGGATCCACCTGAATCCAATTACATCTGGTGAACTCCGACATCTGAAAC
GTTTTAAGTTACACCAAGTTCATAGCCTTTGTTAACCTTTCATGTGTTGAATGTTCAAATAA
TGTTCATTACACTTAAGAATACTGGCCTGAATTTTATTAGCTTCATTATAAATCACTGAGCT
GATATTTACTCTTCCTTTTAAGTTTTCTAAGTACGTCTGTAGCATGATGGTATAGATTTTCT
TGTTTCAGTGCTTTGGGACAGATTTTATATTATGTCAATTGATCAGGTTAAAATTTTCAGTG
TGTAGTTGGCAGATATTTTCAAAATTACAATGCATTTATGGTGTCTGGGGGCAGGGGAACAT
CAGAAAGGTTAAATTGGGCAAAAATGCGTAAGTCACAAGAATTTGGATGGTGCAGTTAATGT
TGAAGTTACAGCATTTCAGATTTTATTGTCAGATATTTAGATGTTTGTTACATTTTTAAAAA
TTGCTCTTAATTTTTAAACTCTCAATACAATATATTTTGACCTTACCATTATTCCAGAGATT
CAGTATTAAAAAAAAAAAAATTACACTGTGGTAGTGGCATTTAAACAATATAATATATTCTA
AACACAATGAAATAGGGAATATAATGTATGAACTTTTTGCATTGGCTTGAAGCAATATAATA
TATTGTAAACAAAACACAGCTCTTACCTAATAAACATTTTATACTGTTTGTATGTATAAAAT
AAAGGTGCTGCTTTAGTTTTTTGGAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 4

Met Ala Arg Arg Ser Ala Phe Pro Ala Ala Ala Leu Trp Leu Trp Ser
Ile Leu Leu Cys Leu Leu Ala Leu Arg Ala Glu Ala Gly ProPro Gln
Glu Glu Ser Leu Tyr Leu Trp Ile Asp Ala His Gln Ala Arg Val Leu
Ile Gly Phe Glu Glu Asp Ile Leu Ile Val Ser Glu Gly Lys Met Ala
Pro Phe Thr His Asp Phe Arg Lys Ala Gln Gln Arg Met Pro Ala Ile
Pro Val Asn Ile His Ser Met Asn Phe Thr Trp Gln Ala Ala Gly Gln
Ala Glu Tyr Phe Tyr Glu Phe Leu Ser Leu Arg Ser Leu Asp Lys Gly
Ile Met Ala Asp Pro Thr Val Asn Val Pro Leu Leu Gly Thr Val Pro
His Lys Ala Ser Val Val Gln Val Gly Phe Pro Cys Leu Gly Lys Gln
Asp Gly Val Ala Ala Phe Glu Val Asp Val Ile Val Met Asn Ser Glu
Gly Asn Thr Ile Leu Gln Thr Pro Gln Asn Ala Ile Phe Phe Lys Thr
Cys Gln Gln Ala Glu Cys Pro Gly Gly Cys Arg Asn Gly Gly Phe Cys
Asn Glu Arg Arg Ile Cys Glu Cys Pro Asp Gly Phe His Gly Pro His
Cys Glu Lys Ala Leu Cys Thr Pro Arg Cys Met Asn Gly Gly Leu Cys
Val Thr Pro Gly Phe Cys Ile Cys Pro Pro Gly Phe Tyr Gly Val Asn
Cys Asp Lys Ala Asn Cys Ser Thr Thr Cys Phe Asn Gly Gly Thr Cys
Phe Tyr Pro Gly Lys Cys Ile Cys Pro Pro Gly Leu Glu Gly Glu Gln
Cys Glu Ile Ser Lys Cys Pro Gln Pro Cys Arg Asn Gly Gly Lys Cys
Ile Gly Lys Ser Lys Cys Lys Cys Ser Lys Gly Tyr Gln Gly Asp Leu
Cys Ser Lys Pro Val Cys Glu Pro Gly Cys Gly Ala His Gly Thr Cys
His Glu Pro Asn Lys Cys Gln Cys Gln Glu Gly Trp His Gly Arg His
Cys Asn Lys Arg Tyr Glu Ala Ser Leu Ile His Ala Leu Arg Pro Ala
Gly Ala Gln Leu Arg Gln His Thr Pro Ser Leu Lys Lys Ala Glu Glu
Arg Arg Asp Pro Pro Glu Ser Asn Tyr Ile Trp

FIGURE 5

CGGACGCGTGGGCGTCCGGCGGTCGCAGAGCCAGGAGGCGGAGGCGCGCGGGCCAGCCTGGG
CCCCAGCCCACACCTTCACCAGGGCCCAGGAGCCACCATGTGGCGATGTCCACTGGGGCTAC
TGCTGTTGCTGCCGCTGGCTGGCCACTTGGCTCTGGGTGCCCAGCAGGGTCGTGGGCGCCGG
GAGCTAGCACCGGGTCTGCACCTGCGGGGCATCCGGGACGCGGGAGGCCGGTACTGCCAGGA
GCAGGACCTGTGCTGCCGCGGCCGTGCCGACGACTGTGCCCTGCCCTACCTGGGCGCCATCT
GTTACTGTGACCTCTTCTGCAACCGCACGGTCTCCGACTGCTGCCCTGACTTCTGGGACTTC
TGCCTCGGCGTGCCACCCCCTTTTCCCCCGATCCAAGGATGTATGCATGGAGGTCGTATCTA
TCCAGTCTTGGGAACGTACTGGGACAACTGTAACCGTTGCACCTGCCAGGAGAACAGGCAGT
GGCAGTGTGACCAAGAACCATGCCTGGTGGATCCAGACATGATCAAAGCCATCAACCAGGGC
AACTATGGCTGGCAGGCTGGGAACCACAGCGCCTTCTGGGGCATGACCCTGGATGAGGGCAT
TCGCTACCGCCTGGGCACCATCCGCCCATCTTCCTCGGTCATGAACATGCATGAAATTTATA
CAGTGCTGAACCCAGGGGAGGTGCTTCCCACAGCCTTCGAGGCCTCTGAGAAGTGGCCCAAC
CTGATTCATGAGCCTCTTGACCAAGGCAACTGTGCAGGCTCCTGGGCCTTCTCCACAGCAGC
TGTGGCATCCGATCGTGTCTCAATCCATTCTCTGGGACACATGACGCCTGTCCTGTCGCCCC
AGAACCTGCTGTCTTGTGACACCCACCAGCAGCAGGGCTGCCGCGGTGGGCGTCTCGATGGT
GCCTGGTGGTTCCTGCGTCGCCGAGGGGTGGTGTCTGACCACTGCTACCCCTTCTCGGGCCG
TGAACGAGACGAGGCTGGCCCTGCGCCCCCCTGTATGATGCACAGCCGAGCCATGGGTCGGG
GCAAGCGCCAGGCCACTGCCCACTGCCCCAACAGCTATGTTAATAACAATGACATCTACCAGGTC
ACTCCTGTCTACCGCCTCGGCTCCAACGACAAGGAGATCATGAAGGAGCTGATGGAGAATGG
CCCTGTCCAAGCCCTCATGGAGGTGCATGAGGACTTCTTCCTATACAAGGGAGGCATCTACA
GCCACACGCCAGTGAGCCTTGGGAGGCCAGAGAGATACCGCCGGCATGGGACCCACTCAGTC
AAGATCACAGGATGGGGAGAGGAGACGCTGCCAGATGGAAGGACGCTCAAATACTGGACTGC
GGCCAACTCCTGGGGCCCAGCCTGGGGCGAGAGGGGCCACTTCCGCATCGTGCGCGGCGTCA
ATGAGTGCGACATCGAGAGCTTCGTGCTGGGCGTCTGGGGCCGCGTGGGCATGGAGGACATG
GGTCATCACTGAGGCTGCGGGCACCACGCGGGGTCCGGCCTGGGATCCAGGCTAAGGGCCGG
CGGAAGAGGCCCCAATGGGGCGGTGACCCCAGCCTCGCCCGACAGAGCCCGGGGCGCAGGCG
GGCGCCAGGGCGCTAATCCCGGCGCGGGTTCCGCTGACGCAGCGCCCCGCCTGGGAGCCGCG
GGCAGGCGAGACTGGCGGAGCCCCCAGACCTCCCAGTGGGGACGGGGCAGGGCCTGGCCTGG
GAAGAGCACAGCTGCAGATCCCAGGCCTCTGGCGCCCCCACTCAAGACTACCAAAGCCAGGA
CACCTCAAGTCTCCAGCCCCAATACCCCACCCCAATCCCGTATTCTTTTTTTTTTTTTTTTAGAC
AGGGTCTTGCTCCGTTGCCCAGGTTGGAGTGCAGTGGCCCATCAGGGCTCACTGTAACCTCC
GACTCCTGGGTTCAAGTGACCCTCCCACCTCAGCCTCTCAAGTAGCTGGGACTACAGGTGCA
CCACCACACCTGGCTAATTTTTTGTATTTTTTTGTAAAGAGGGGGGTCTCACTGTGTTGCCCAG
GCTGGTTTCGAACTCCTGGGCTCAAGCGGTCCACCTGCCTCCGCCTCCCAAAGTGCTGGGAT
TGCAGGCATGAGCCACTGCACCCAGCCCTGTATTCTTATTCTTCAGATATTTATTTTTCTTT
TCACTGTTTTAAAATAAAACCAAAGTATTGATAAAAAAAAA

FIGURE 6

><ss.DNA33223

><subunit 1 of 1, 467 aa, 1 stop

><MW: 52387, pI: 6.95, NX(S/T): 2

MWRCPLGLLLLLPLAGHLALGAQQGRGRRELAPGLHLRGIRDAGGRYCQEQDLCCRGRAD
DCALPYLGAICYCDLFCNRTVSDCCPDFWDFCLGVPPPFPPIQGCMHGGRIYPVLGTYWD
NCNRCTCQENRQWQCDQEPCLVDPDMIKAINQGNYGWQAGNHSAFWGMTLDEGIRYRLGT
IRPSSSVMNMHEIYTVLNPGEVLPTAFEASEKWPNLIHEPLDQGNCAGSWAFSTAAVASD
RVSIHSLGHMTPVLSPQNLLSCDTHQQQGCRGGRLDGAWWFLRRRGVVSDHCYPFSGRER
DEAGPAPPCMMHSRAMGRGKRQATAHCPNSYVNNNDIYQVTPVYRLGSNDKEIMKELMEN
GPVQALMEVHEDFFLYKGGIYSHTPVSLGRPERYRRHGTHSVKITGWGEETLPDGRTLKY
WTAANSWGPAWGERGHFRIVRGVNECDIESFVLGVWGRVGMEDMGHH

FIGURE 7

AGGCTCCTTGGCCCTTTTTCCACAGCAAGCTTNTGCNATCCCGATTCGTTGTCTCAAATC
CAATTCTCTTGGGACACATNACGCCTGTCCTTTNGCCCCAGAACCTGCTGTCTTGTACAC
CCACCAGCAGCAGGGCTGCCGCGNTGGGCGTCTCGATGGTGCCTGGTGGTTCCTGCGTCG
CCGAGGGNTGGTGTCTGACCACTGCTACCCCTTCTCGGGCCGTGAACGAGACGAGGCTGG
CCCTGCGCCCCCCTGTATGATGCACAGCCGAGCCATGGGTCGGGGCAAGCGCCAGGCCAC
TGCCCACTGCCCCAACAGCTATGTTAATAACAATGACATCTACCAGGTCACTCCTGTCTA
CCGCCTCGGCTCCAACGACAAGGAGATCATGAAGGAGCTGATGGAGAATGGCCCTGTCCA
AGCCCTCATGGAGGTGCATGAGGACTTCTTCCTATACAAGGGAGGCATCTACAGCCACAC
GCCAGTGAGCCTTGGGAGGCCAGAGAGATACCGCCGGCATGGGACCCACTCAG

FIGURE 8

GCTGCTTGCCCTGTTGATGGCAGGCTTGGCCCTGCAGCCAGGCACTGCCCTGCTGTGCTACT
CCTGCAAAGCCCAGGTGAGCAACGAGGACTGCCTGCAGGTGGAGAACTGCACCCAGCTGGGG
GAGCAGTGCTGGACCGCGCGCATCCGCGCAGTTGGCCTCCTGACCGTCATCAGCAAAGGCTG
CAGCTTGAACTGCGTGGATGACTCACAGGACTACTACGTGGGCAAGAAGAACATCACGTGCT
GTGACACCGACTTGTGCAACGCCAGCGGGGCCCATGCCCTGCAGCCGGCTGCCGCCATCCTT
GCGCTGCTCCCTGCACTCGGCCTGCTGCTCTGGGGACCCGGCCAGCTATAGGCTCTGGGGGG
CCCCGCTGCAGCCCACACTGGGTGTGGTGCCCCAGGCCTCTGTGCCACTCCTCACAGACCTG
GCCCAGTGGGAGCCTGTCCTGGTTCCTGAGGCACATCCTAACGCAAGTCTGACCATGTATGT
CTGCACCCCTGTCCCCCACCCTGACCCTCCCATGGCCCTCTCCAGGACTCCCACCCGGCAGA
TCAGCTCTAGTGACACAGATCCGCCTGCAGATGGCCCCTCCAACCCTCTCTGCTGCTGTTTCCAT
GGCCCAGCATTCTCCACCCTTAACCCTGTGCTCAGGCACCTCTTCCCCCAGGAAGCCTTCCC
TGCCCACCCCATCTATGACTTGAGCCAGGTCTGGTCCGTGGTGTCCCCCGCACCCAGCAGGG
GACAGGCACTCAGGAGGGCCCAGTAAAGGCTGAGATGAAGTGGACTGAGTAGAACTGGAGGA
CAAGAGTCGACGTGAGTTCCTGGGAGTCTCCAGAGATGGGGCCTGGAGGCCTGGAGGAAGGG
GCCAGGCCTCACATTCGTGGGGCTCCCTGAATGGCAGCCTGAGCACAGCGTAGGCCCTTAAT
AAACACCTGTTGGATAAGCCAAAAAAA

FIGURE 9

MAGLALQPGTALLCYSCKAQVSNEDCLQVENCTQLGEQCWTARIRAVGLLTVISKGCSLNCV
DDSQDYYVGKKNITCCDTDLCNASGAHALQPAAAILALLPALGLLLWGPGQL

FIGURE 10

ATGGGAGCCGCCCGCCTGCTGCCCAACCTCACTCTGTGCTTACAGCTGCTGATTCTCTGCTG
TCAAACTCAGTACGTGAGGGACCAGGGCGCCATGACCGACCAGCTGAGCAGGCGGCAGATCC
GCGAGTACCAACTCTACAGCAGGACCAGTGGCAAGCACGTGCAGGTCACCGGGCGTCGCATC
TCCGCCACCGCCGAGGACGGCAACAAGTTTGCCAAGCTCATAGTGGAGACGGACACGTTTGG
CAGCCGGGTTCGCATCAAAGGGGCTGAGAGTGAGAAGTACATCTGTATGAACAAGAGGGGCA
AGCTCATCGGGAAGCCCAGCGGGAAGAGCAAAGACTGCGTGTTCACGGAGATCGTGCTGGAG
AACAACTATACGGCCTTCCAGAACGCCCGGCACGAGGGCTGGTTCATGGCCTTCACGCGGCA
GGGGCGGCCCCGCCAGGCTTCCCGCAGCCGCCAGAACCAGCGCGAGGCCCACTTCATCAAGC
GCCTCTACCAAGGCCAGCTGCCCTTCCCCAACCACGCCGAGAAGCAGAAGCAGTTCGAGTTT
GTGGGCTCCGCCCCCACCCGCCGGACCAAGCGCACACGGCGGCCCCAGCCCCTCACGTAGTC
TGGGAGGCAGGGGGCAGCAGCCCCTGGGCCGCCTCCCCACCCCTTTCCCTTCTTAATCCAAG
GACTGGGCTGGGGTGGCGGGAGGGGAGCCAGATCCCCGAGGGAGGACCCTGAGGGCCGCGAA
GCATCCGAGCCCCCAGCTGGGAAGGGGCAGGCCGGTGCCCCAGGGGCGGCTGGCACAGTGCC
CCCTTCCCGGACGGGTGGCAGGCCCTGGAGAGGAACTGAGTGTCACCCTGATCTCAGGCCAC
CAGCCTTTGCCGGCCTCCCAGCCGGGCTCCTGAAGCCCGCTGAAAGGTCAGCGACTGAAGGC
CTTGCAGACAACCGTCTGGAGGTGGCTGTCCTCAAAATCTGCTTCTCGGATCTCCCTCAGTC
TGCCCCCAGCCCCCAAACTCCTCCTGGCTAGACTGTAGGAAGGGACTTTTGTTTGTTTGTTT
GTTTCAGGAAAAAAGAAAGGGAGAGAGAGGAAAATAGAGGGTTGTCCACTCCTCACATTCCA
CGACCCAGGCCTGCACCCCACCCCCAACTCCCAGCCCCGGAATAAAACCATTTTCCTGC

FIGURE 11

Met Gly Ala Ala Arg Leu Leu Pro Asn Leu Thr Leu Cys Leu Gln
Leu Leu Ile Leu Cys Cys Gln Thr Gln Tyr Val Arg Asp Gln Gly
Ala Met Thr Asp Gln Leu Ser Arg Arg Gln Ile Arg Glu Tyr Gln
Leu Tyr Ser Arg Thr Ser Gly Lys His Val Gln Val Thr Gly Arg
Arg Ile Ser Ala Thr Ala Glu Asp Gly Asn Lys Phe Ala Lys Leu
Ile Val Glu Thr Asp Thr Phe Gly Ser Arg Val Arg Ile Lys Gly
Ala Glu Ser Glu Lys Tyr Ile Cys Met Asn Lys Arg Gly Lys Leu
Ile Gly Lys Pro Ser Gly Lys Ser Lys Asp Cys Val Phe Thr Glu
Ile Val Leu Glu Asn Asn Tyr Thr Ala Phe Gln Asn Ala Arg His
Glu Gly Trp Phe Met Ala Phe Thr Arg Gln Gly Arg Pro Arg Gln
Ala Ser Arg Ser Arg Gln Asn Gln Arg Glu Ala His Phe Ile Lys
Arg Leu Tyr Gln Gly Gln Leu Pro Phe Pro Asn His Ala Glu Lys
Gln Lys Gln Phe Glu Phe Val Gly Ser Ala Pro Thr Arg Arg Thr
Lys Arg Thr Arg Arg Pro Gln Pro Leu Thr

FIGURE 12

ACTTGCCATCACCTGTTGCCAGTGTGGAAAAATTCTCCCTGTTGAATTTTTTGCACATGGAG
GACAGCAGCAAAGAGGGCAACACAGGCTGATAAGACCAGAGACAGCAGGGAGATTATTTTAC
CATACGCCCTCAGGACGTTCCCTCTAGCTGGAGTTCTGGACTTCAACAGAACCCCATCCAGT
CATTTTGATTTTGCTGTTTATTTTTTTTTTCTTTTTCTTTTTTCCCACCACATTGTATTTTAT
TTCCGTACTTCAGAAATGGGCCTACAGACCACAAAGTGGCCCAGCCATGGGGCTTTTTTCCT
GAAGTCTTGGCTTATCATTTCCCTGGGGCTCTACTCACAGGTGTCCAAACTCCTGGCCTGCC
CTAGTGTGTGCCGCTGCGACAGGAACTTTGTCTACTGTAATGAGCGAAGCTTGACCTCAGTG
CCTCTTGGGATCCCGGAGGGCGTAACCGTACTCTACCTCCACAACAACCAAATTAATAATGC
TGGATTTCCTGCAGAACTGCACAATGTACAGTCGGTGCACACGGTCTACCTGTATGGCAACC
AACTGGACGAATTCCCCATGAACCTTCCCAAGAATGTCAGAGTTCTCCATTTGCAGGAAAACAAT
ATTCAGACCATTTCACGGGCTGCTCTTGCCCAGCTCTTGAAGCTTGAAGAGCTGCACCTGGA
TGACAACTCCATATCCACAGTGGGGGTGGAAGACGGGGCCTTCCGGGAGGCTATTAGCCTCA
AATTGTTGTTTTTGTCTAAGAATCACCTGAGCAGTGTGCCTGTTGGGCTTCCTGTGGACTTG
CAAGAGCTGAGAGTGGATGAAAATCGAATTGCTGTCATATCCGACATGGCCTTCCAGAATCT
CACGAGCTTGGAGCGTCTTATTGTGGACGGGAACCTCCTGACCAACAAGGGTATCGCCGAGG
GCACCTTCAGCCATCTCACCAAGCTCAAGGAATTTTCAATTGTACGTAATTCGCTGTCCCAC
CCTCCTCCCGATCTCCCAGGTACGCATCTGATCAGGCTCTATTTGCAGGACAACCAGATAAA
CCACATTCCTTTGACAGCCTTCTCAAATCTGCGTAAGCTGGAACGGCTGGATATATCCAACA
ACCAACTGCGGATGCTGACTCAAGGGGTTTTTGATAATCTCTCCAACCTGAAGCAGCTCACT
GCTCGGAATAACCCTTGGTTTTGTGACTGCAGTATTAAATGGGTCACAGAATGGCTCAAATA
TATCCCTTCATCTCTCAACGTGCGGGGTTTCATGTGCCAAGGTCCTGAACAAGTCCGGGGGA
TGGCCGTCAGGGAATTAAATATGAATCTTTTGTCCTGTCCCACCACGACCCCCGGCCTGCCT
CTCTTCACCCCAGCCCCAAGTACAGCTTCTCCGACCACTCAGCCTCCCACCCTCTCTATTCC
AAACCCTAGCAGAAGCTACACGCCTCCAACTCCTACCACATCGAAACTTCCCACGATTCCTG
ACTGGGATGGCAGAGAAAGAGTGACCCCACCTATTTCTGAACGGATCCAGCTCTCTATCCAT
TTTGTGAATGATACTTCCATTCAAGTCAGCTGGCTCTCTCTTCACCGTGATGGCATACAA
ACTCACATGGGTGAAAATGGGCCACAGTTTAGTAGGGGGCATCGTTCAGGAGCGCATAGTCA
GCGGTGAGAAGCAACACCTGAGCCTGGTTAACTTAGAGCCCCGATCCACCTATCGGATTTGT
TTAGTGCCACTGGATGCTTTTAACTACCGCGCGGTAGAAGACACCATTTGTTCAGAGGCCAC
CACCCATGCCTCCTATCTGAACAACGGCAGCAACACAGCGTCCAGCCATGAGCAGACGACGT
CCCACAGCATGGGCTCCCCCTTTCTGCTGGCGGGCTTGATCGGGGGCGCGGTGATATTTGTG
CTGGTGGTCTTGCTCAGCGTCTTTTGCTGGCATATGCACAAAAAGGGGCGCTACACCTCCCA
GAAGTGGAAATACAACCGGGGCCGGCGGAAAGATGATTATTGCGAGGCAGGCACCAAGAAGG
ACAACTCCATCCTGGAGATGACAGAAACCAGTTTTCAGATCGTCTCCTTAAATAACGATCAA
CTCCTTAAAGGAGATTTCAGACTGCAGCCCATTTACACCCCAAATGGGGGCATTAATTACAC
AGACTGCCATATCCCCAACAACATGCGATACTGCAACAGCAGCGTGCCAGACCTGGAGCACT
GCCATACGTGACAGCCAGAGGCCCAGCGTTATCAAGGCGGACAATTAGACTCTTGAGAACAC
ACTCGTGTGTGCACATAAAGACACGCAGATTACATTTGATAAATGTTACACAGATGCATTTG
TGCATTTGAATACTCTGTAATTTATACGGTGTACTATATAATGGGATTTAAAAAAAGTGCTA
TCTTTTCTATTTCAAGTTAATTACAAACAGTTTTGTAACTCTTTGCTTTTTAAATCTT

FIGURE 13

>><1158/ss.DNA36350

>><subunit 1 of 1, 660 aa, 1 stop

>><MW: 74049, pI: 8.09, NX(S/T): 7

MGLQTTKWPSHGAFFLKSWLIISLGLYSQVSKLLACPSVCRCDRNFVYCNERSLTSVPLG
IPEGVTVLYLHNNQINNAGFPAELHNVQSVHTVYLYGNQLDEFPMNLPKNVRVLHLQENN
IQTISRAALAQLLKLEELHLDDNSISTVGVEDGAFREAISLKLLFLSKNHLSSVPVGLPV
DLQELRVDENRIAVISDMAFQNLTSLERLIVDGNLLTNKGIAEGTFSHLTKLKEFSIVRN
SLSHPPPDLPGTHLIRLYLQDNQINHIPLTAFSNLRKLERLDISNNQLRMLTQGVFDNLS
NLKQLTARNNPWFCDCSIKWVTEWLKYIPSSLNVRGFMCQGPEQVRGMAVRELNMNLLSC
PTTTPGLPLFTPAPSTASPTTQPPTLSIPNPSRSYTPPTPTTSKLPTIPDWDGRERVTPP
ISERIQLSIHFVNDTSIQVSWLSLFTVMAYKLTWVKMGHSLVGGIVQERIVSGEKQHLSL
VNLEPRSTYRICLVPLDAFNYRAVEDTICSEATTHASYLNNGSNTASSHEQTTSHSMGSP
FLLAGLIGGAVIFVLVVLLSVFCWHMHKKGRYTSQKWKYNRGRRKDDYCEAGTKKDNSIL
EMTETSFQIVSLNNDQLLKGDFRLQPIYTPNGGINYTDCHIPNNMRYCNSSVPDLEHCHT

FIGURE 14A

ACTTGGAGCAAGCGGCGGCGGCGGAGACAGAGGCAGAGGCAGAAGCTGGGGCTCCGTCCTCG
CCTCCCACGAGCGATCCCCGAGGAGAGCCGCGGCCCTCGGCGAGGCGAAGAGGCCGACGAGG
AAGACCCGGGTGGCTGCGCCCCTGCCTCGCTTCCCAGGCGCCGGCGGCTGCAGCCTTGCCCC
TCTTGCTCGCCTTGAAAATGGAAAAGATGCTCGCAGGCTGCTTTCTGCTGATCCTCGGACAG
ATCGTCCTCCTCCCTGCCGAGGCCAGGGAGCGGTCACGTGGGAGGTCCATCTCTAGGGGCAG
ACACGCTCGGACCCACCCGCAGACGGCCCTTCTGGAGAGTTCCTGTGAGAACAAGCGGGCAG
ACCTGGTTTTCATCATTGACAGCTCTCGCAGTGTCAACACCCATGACTATGCAAAGGTCAAG
GAGTTCATCGTGGACATCTTGCAATTCTTGGACATTGGTCCTGATGTCACCCGAGTGGGCCT
GCTCCAATATGGCAGCACTGTCAAGAATGAGTTCTCCCTCAAGACCTTCAAGAGGAAGTCCG
AGGTGGAGCGTGCTGTCAAGAGGATGCGGCATCTGTCCACGGGCACCATGACTGGGCTGGCC
ATCCAGTATGCCCTGAACATCGCATTCTCAGAAGCAGAGGGGGCCCGGCCCCTGAGGGAGAA
TGTGCCACGGGTCATAATGATCGTGACAGATGGGAGACCTCAGGACTCCGTGGCCGAGGTGG
CTGCTAAGGCACGGGACACGGGCATCCTAATCTTTGCCATTGGTGTGGGCCAGGTAGACTTC
AACACCTTGAAGTCCATTGGGAGTGAGCCCCATGAGGACCATGTCTTCCTTGTGGCCAATTT
CAGCCAGATTGAGACGCTGACCTCCGTGTTCCAGAAGAAGTTGTGCACGGCCCACATGTGCA
GCACCCTGGAGCATAACTGTGCCCACTTCTGCATCAACATCCCTGGCTCATACGTCTGCAGG
TGCAAACAAGGCTACATTCTCAACTCGGATCAGACGACTTGCAGAATCCAGGATCTGTGTGC
CATGGAGGACCACAACTGTGAGCAGCTCTGTGTGAATGTGCCGGGCTCCTTCGTCTGCCAGT
GCTACAGTGGCTACGCCCTGGCTGAGGATGGGAAGAGGTGTGTGGCTGTGGACTACTGTGCC
TCAGAAAACCACGGATGTGAACATGAGTGTGTAAATGCTGATGGCTCCTACCTTTGCCAGTG
CCATGAAGGATTTGCTCTTAACCCAGATGAAAAAACGTGCACAAGGATCAACTACTGTGCAC
TGAACAAACCGGGCTGTGAGCATGAGTGCGTCAACATGGAGGAGAGCTACTACTGCCGCTGC
CACCGTGGCTACACTCTGGACCCCAATGGCAAAACCTGCAGCCGAGTGGACCACTGTGCACA
GCAGGACCATGGCTGTGAGCAGCTGTGTCTGAACACGGAGGATTCCTTCGTCTGCCAGTGCT
CAGAAGGCTTCCTCATCAACGAGGACCTCAAGACCTGCTCCCGGGTGGATTACTGCCTGCTG
AGTGACCATGGTTGTGAATACTCCTGTGTCAACATGGACAGATCCTTTGCCTGTCAGTGTCC
TGAGGGACACGTGCTCCGCAGCGATGGGAAGACGTGTGCAAAATTGGACTCTTGTGCTCTGG
GGGACCACGGTTGTGAACATTCGTGTGTAAGCAGTGAAGATTCGTTTGTGTGCCAGTGCTTT
GAAGGTTATATACTCCGTGAAGATGGAAAAACCTGCAGAAGGAAAGATGTCTGCCAAGCTAT
AGACCATGGCTGTGAACACATTTGTGTGAACAGTGACGACTCATACACGTGCGAGTGCTTGG
AGGGATTCCGGCTCGCTGAGGATGGGAAACGCTGCCGAAGGAAGGATGTCTGCAAATCAACC
CACCATGGCTGCGAACACATTTGTGTTAATAATGGGAATTCCTACATCTGCAAATGCTCAGA
GGGATTTGTTCTAGCTGAGGACGGAAGACGGTGCAAGAAATGCACTGAAGGCCCAATTGACC
TGGTCTTTGTGATCGATGGATCCAAGAGTCTTGGAGAAGAGAATTTTGAGGTCGTGAAGCAG
TTTGTCACTGGAATTATAGATTCCTTGACAATTTCCCCCAAAGCCGCTCGAGTGGGGCTGCT
CCAGTATTCCACACAGGTCCACACAGAGTTCACTCTGAGAAACTTCAACTCAGCCAAAGACA
TGAAAAAAGCCGTGGCCCACATGAAATACATGGGAAAGGGCTCTATGACTGGGCTGGCCCTG
AAACACATGTTTGAGAGAAGTTTTACCCAAGGAGAAGGGGCCAGGCCCCTTTCCACAAGGGT
GCCCAGAGCAGCCATTGTGTTCACCGACGGACGGGCTCAGGATGACGTCTCCGAGTGGGCCA
GTAAAGCCAAGGCCAATGGTATCACTATGTATGCTGTTGGGGTAGGAAAAGCCATTGAGGAG
GAACTACAAGAGATTGCCTCTGAGCCCACAAACAAGCATCTCTTCTATGCCGAAGACTTCAG
CACAATGGATGAGATAAGTGAAAAACTCAAGAAAGGCATCTGTGAAGCTCTAGAAGACTCCG
ATGGAAGACAGGACTCTCCAGCAGGGGAACTGCCAAAAACGGTCCAACAGCCAACAGAATCT
GAGCCAGTCACCATAAATATCCAAGACCTACTTTCCTGTTCTAATTTTGCAGTGCAACACAG
ATATCTGTTTGAAGAAGACAATCTTTTACGGTCTACACAAAAGCTTTCCCATTCAACAAAAC
CTTCAGGAAGCCCTTTGGAAGAAAAACACGATCAATGCAAATGTGAAAACCTTATAATGTTC
CAGAACCTTGCAAACGAAGAAGTAAGAAAATTAACACAGCGCTTAGAAGAAATGACACAGAG
AATGGAAGCCCTGGAAAATCGCCTGAGATACAGATGAAGATTAGAAATCGCGACACATTTGT
AGTCATTGTATCACGGATTACAATGAACGCAGTGCAGAGCCCCAAAGCTCAGGCTATTGTTA
AATCAATAATGTTGTGAAGTAAAACAATCAGTACTGAGAAACCTGGTTTGCCACAGAACAAA
GACAAGAAGTATACACTAACTTGTATAAATTTATCTAGGAAAAAAATCCTTCAGAATTCTAA
GATGAATTTACCAGGTGAGAATGAATAAGCTATGCAAGGTATTTTGTAATATACTGTGGACA
CAACTTGCTTCTGCCTCATCCTGCCTTAGTGTGCAATCTCATTTGACTATACGATAAAGTTT

FIGURE 14B

GCACAGTCTTACTTCTGTAGAACACTGGCCATAGGAAATGCTGTTTTTTTTGTACTGGACTTT
ACCTTGATATATGTATATGGATGTATGCATAAAATCATAGGACATATGTACTTGTGGAACAA
GTTGGATTTTTTATACAATATTAAAATTCACCACTTCAG

FIGURE 15

>1164/ss.DNA32290

>subunit 1 of 1, 915 aa, 1 stop

>MW: 102233, pI: 6.02, NX(S/T): 1

MEKMLAGCFLLILGQIVLLPAEARERSRGRSISRGRHARTHPQTALLESSCENKRADLVF
IIDSSRSVNTHDYAKVKEFIVDILQFLDIGPDVTRVGLLQYGSTVKNEFSLKTFKRKSEV
ERAVKRMRHLSTGTMTGLAIQYALNIAFSEAEGARPLRENVPRVIMIVTDGRPQDSVAEV
AAKARDTGILIFAIGVGQVDFNTLKSIGSEPHEDHVFLVANFSQIETLTSVFQKKLCTAH
MCSTLEHNCAHFCINIPGSYVCRCKQGYILNSDQTTCRIQDLCAMEDHNCEQLCVNVPGS
FVCQCYSGYALAEDGKRCVAVDYCASENHGCEHECVNADGSYLCQCHEGFALNPDEKTCT
RINYCALNKPGCEHECVNMEESYYCRCHRGYTLDPNGKTCSRVDHCAQQDHGCEQLCLNT
EDSFVCQCSEGFLINEDLKTCSRVDYCLLSDHGCEYSCVNMDRSFACQCPEGHVLRSDGK
TCAKLDSCALGDHGCEHSCVSSEDSFVCQCFEGYILREDGKTCRRKDVCQAIDHGCEHIC
VNSDDSYTCECLEGFRLAEDGKRCRRKDVCKSTHHGCEHICVNNGNSYICKCSEGFVLAE
DGRRCKKCTEGPIDLVFVIDGSKSLGEENFEVVKQFVTGIIDSLTISPKAARVGLLQYST
QVHTEFTLRNFNSAKDMKKAVAHMKYMGKGSMTGLALKHMFERSFTQGEGARPLSTRVPR
AAIVFTDGRAQDDVSEWASKAKANGITMYAVGVGKAIEEELQEIASEPTNKHLFYAEDFS
TMDEISEKLKKGICEALEDSDGRQDSPAGELPKTVQQPTESEPVTINIQDLLSCSNFAVQ
HRYLFEEDNLLRSTQKLSHSTKPSGSPLEEKHDQCKCENLIMFQNLANEEVRKLTQRLEE
MTQRMEALENRLRYR

FIGURE 16

GGAGCCGCCCTGGGTGTCAGCGGCTCGGCTCCCGCGCACGCTCCGGCCGTCGCGCAGCCTCG
GCACCTGCAGGTCCGTGCGTCCCGCGGCTGGCGCCCCTGACTCCGTCCCGGCCAGGGAGGGC
CATGATTTCCCTCCCGGGGCCCCTGGTGACCAACTTGCTGCGGTTTTTGTTCCTGGGGCTGA
GTGCCCTCGCGCCCCCCTCGCGGGCCCAGCTGCAACTGCACTTGCCCGCCAACCGGTTGCAG
GCGGTGGAGGGAGGGGAAGTGGTGCTTCCAGCGTGGTACACCTTGCACGGGGAGGTGTCTTC
ATCCCAGCCATGGGAGGTGCCCTTTGTGATGTGGTTCTTCAAACAGAAAGAAAAGGAGGATC
AGGTGTTGTCCTACATCAATGGGGTCACAACAAGCAAACCTGGAGTATCCTTGGTCTACTCC
ATGCCCTCCCGGAACCTGTCCCTGCGGCTGGAGGGTCTCCAGGAGAAAGACTCTGGCCCCTA
CAGCTGCTCCGTGAATGTGCAAGACAAACAAGGCAAATCTAGGGGCCACAGCATCAAAACCT
TAGAACTCAATGTACTGGTTCCTCCAGCTCCTCCATCCTGCCGTCTCCAGGGTGTGCCCCAT
GTGGGGGCAAACGTGACCCTGAGCTGCCAGTCTCCAAGGAGTAAGCCCGCTGTCCAATACCA
GTGGGATCGGCAGCTTCCATCCTTCCAGACTTTCTTTGCACCAGCATTAGATGTCATCCGTG
GGTCTTTAAGCCTCACCAACCTTTCGTCTTCCATGGCTGGAGTCTATGTCTGCAAGGCCCAC
AATGAGGTGGGCACTGCCCAATGTAATGTGACGCTGGAAGTGAGCACAGGGCCTGGAGCTGC
AGTGGTTGCTGGAGCTGTTGTGGGTACCCTGGTTGGACTGGGGTTGCTGGCTGGGCTGGTCC
TCTTGTACCACCGCCGGGGCAAGGCCCTGGAGGAGCCAGCCAATGATATCAAGGAGGATGCC
ATTGCTCCCCGGACCCTGCCCTGGCCCAAGAGCTCAGACACAATCTCCAAGAATGGGACCCT
TTCCTCTGTCACCTCCGCACGAGCCCTCCGGCCACCCCATGGCCCTCCCAGGCCTGGTGCAT
TGACCCCCACGCCCAGTCTCTCCAGCCAGGCCCTGCCCTCACCAAGACTGCCCACGACAGAT
GGGGCCCACCCTCAACCAATATCCCCCATCCCTGGTGGGGTTTCTTCCTCTGGCTTGAGCCG
CATGGGTGCTGTGCCTGTGATGGTGCCTGCCCAGAGTCAAGCTGGCTCTCTGGTATGATGAC
CCCACCACTCATTGGCTAAAGGATTTGGGGTCTCTCCTTCCTATAAGGGTCACCTCTAGCAC
AGAGGCCTGAGTCATGGGAAAGAGTCACACTCCTGACCCTTAGTACTCTGCCCCCACCTCTC
TTTACTGTGGGAAAACCATCTCAGTAAGACCTAAGTGTCCAGGAGACAGAAGGAGAAGAGGA
AGTGGATCTGGAATTGGGAGGAGCCTCCACCCACCCCTGACTCCTCCTTATGAAGCCAGCTG
CTGAAATTAGCTACTCACCAAGAGTGAGGGGCAGAGACTTCCAGTCACTGAGTCTCCCAGGC
CCCCTTGATCTGTACCCCACCCCTATCTAACACCACCCTTGGCTCCCACTCCAGCTCCCTGT
ATTGATATAACCTGTCAGGCTGGCTTGGTTAGGTTTTACTGGGGCAGAGGATAGGGAATCTC
TTATTAAAACTAACATGAAATATGTGTTGTTTTCATTTGCAAATTTAAATAAAGATACATAA
TGTTTGTATGAAAAA

FIGURE 17

><1172/ss.DNA35639

><subunit 1 of 1, 390 aa, 1 stop

><MW: 41176, pI: 9.61, NX(S/T): 5

MISLPGPLVTNLLRFLFLGLSALAPPSRAQLQLHLPANRLQAVEGGEVVLPAWYTLHGEV
SSSQPWEVPFVMWFFKQKEKEDQVLSYINGVTTSKPGVSLVYSMPSRNLSLRLEGLQEKD
SGPYSCSVNVQDKQGKSRGHSIKTLELNVLVPPAPPSCRLQGVPHVGANVTLSCQSPRSK
PAVQYQWDRQLPSFQTFFAPALDVIRGSLSLTNLSSSMAGVYVCKAHNEVGTAQCNVTLE
VSTGPGAAVVAGAVVGTLVGLGLLAGLVLLYHRRGKALEEPANDIKEDAIAPRTLPWPKS
SDTISKNGTLSSVTSARALRPPHGPPRPGALTPTPSLSSQALPSPRLPTTDGAHPQPISP
IPGGVSSSGLSRMGAVPVMVPAQSQAGSLV

FIGURE 18

CGCCACCACTGCGGCCACCGCCA
><MET {trans=1-s, dir=f, res=1}>
ATGAAACGCCTCCCGCTCCTAGTGGTTTTTTCCACTTTGTTGAATTGTTCCTATACTCAA
AATTGCACCAAGACACCTTGTCTCCCAAATGCAAAATGTGAAATACGCAATGGAATTGAA
GCCTGCTATTGCAACATGGGATTTTCAGGAAATGGTGTCACAATTTGTGAAGATGATAAT
GAATGTGGAAATTTAACTCAGTCCTGTGGCGAAAATGCTAATTGCACTAACACAGAAGGA
AGTTATTATTGTATGTGTGTACCTGGCTTCAGATCCAGCAGTAACCAAGACAGGTTTATC
ACTAATGATGGAACCGTCTGTATAGAAAATGTGAATGCAAACTGCCATTTAGATAATGTC
TGTATAGCTGCAAATATTAATAAAACTTTAACAAAAATCAGATCCATAAAAGAACCTGTG
GCTTTGCTACAAGAAGTCTATAGAAATTCTGTGACAGATCTTTCACCAACAGATATAATT
ACATATATAGAAATATTAGCTGAATCATCTTCATTACTAGGTTACAAGAACAACACTATC
TCAGCCAAGGACACCCTTTCTAACTCAACTCTTACTGAATTTGTAAAAACCGTGAATAAT
TTTGTTCAAAGGGATACATTTGTAGTTTGGGACAAGTTATCTGTGAATCATAGGAGAACA
CATCTTACAAAACTCATGCACACTGTTGAACAAGCTACTTTAAGGATATCCCAGAGCTTC
CAAAAGACCACAGAGTTTGATACAAATTCAACGGATATAGCTCTCAAAGTTTTCTTTTTT
GATTCATATAACATGAAACATATTCATCCTCATATGAATATGGATGGAGACTACATAAAT
ATATTTCCAAAGAGAAAAGCTGCATATGATTCAAATGGCAATGTTGCAGTTGCATTTTTA
TATTATAAGAGTATTGGTCCTTTGCTTTCATCATCTGACAACTTCTTATTGAAACCTCAA
AATTATGATAATTCTGAAGAGGAGGAAAGAGTCATATCTTCAGTAATTTCAGTCTCAATG
AGCTCAAACCCACCCACATTATATGAACTTGAAAAAATAACATTTACATTAAGTCATCGA
AAGGTCACAGATAGGTATAGGAGTCTATGTGCATTTTGGAATTACTCACCTGATACCATG
AATGGCAGCTGGTCTTCAGAGGGCTGTGAGCTGACATACTCAAATGAGACCCACACCTCA
TGCCGCTGTAATCACCTGACACATTTTGCAATTTTGATGTCCTCTGGTCCTTCCATTGGT
ATTAAAGATTATAATATTCTTACAAGGATCACTCAACTAGGAATAATTATTTCACTGATT
TGTCTTGCCATATGCATTTTTACCTTCTGGTTCTTCAGTGAAATTCAAAGCACCAGGACA
ACAATTCACAAAAATCTTTGCTGTAGCCTATTTCTTGCTGAACTTGTTTTTCTTGTTGGG
ATCAATACAAATACTAATAAGCTCTTCTGTTCAATCATTGCCGGACTGCTACACTACTTC
TTTTTAGCTGCTTTTGCATGGATGTGCATTGAAGGCATACATCTCTATCTCATTGTTGTG
GGTGTCATCTACAACAAGGGATTTTTGCACAAGAATTTTTATATCTTTGGCTATCTAAGC
CCAGCCGTGGTAGTTGGATTTTCGGCAGCACTAGGATACAGATATTATGGCACAACCAAA
GTATGTTGGCTTAGCACCGAAAACAACTTTATTTGGAGTTTTATAGGACCAGCATGCCTA
ATCATTCTTGTTAATCTCTTGGCTTTTGGAGTCATCATATACAAAGTTTTTCGTCACACT
GCAGGGTTGAAACCAGAAGTTAGTTGCTTTGAGAACATAAGGTCTTGTGCAAGAGGAGCC
CTCGCTCTTCTGTTCCTTCTCGGCACCACCTGGATCTTTGGGGTTCTCCATGTTGTGCAC
GCATCAGTGGTTACAGCTTACCTCTTCACAGTCAGCAATGCTTTCCAGGGGATGTTCATT
TTTTTATTCCTGTGTGTTTTATCTAGAAAGATTCAAGAAGAATATTACAGATTGTTCAAA
AATGTCCCCTGTTGTTTTGGATGTTTAAGGTAAACATAGAGAATGGTGGATAATTACAAC
TGCACAAAAATAAAAATTCCAAGCTGTGGATGACCAATGTATAAAAATGACTCATCAAAT
TATCCAATTATTAACTACTAGACAAAAAGTATTTTAAATCAGTTTTTCTGTTTATGCTAT
AGGAACTGTAGATAATAAGGTAAAATTATGTATCATATAGATATACTATGTTTTTCTATG
TGAAATAGTTCTGTCAAAAATAGTATTGCAGATATTTGGAAAGTAATTGGTTTCTCAGGA
GTGATATCACTGCACCCAAGGAAAGATTTTCTTTCTAACACGAGAAGTATATGAATGTCC
TGAAGGAAACCACTGGCTTGATATTTCTGTGACTCGTGTTGCCTTTGAAACTAGTCCCCT
ACCACCTCGGTAATGAGCTCCATTACAGAAAGTGGAACATAAGAGAATGAAGGGGCAGAA
TATCAAACAGTGAAAAGGGAATGATAAGATGTATTTTGAATGAACTGTTTTTTCTGTAGA
CTAGCTGAGAAATTGTTGACATAAAATAAAGAATTGAAGAAACACATTTTACCATTTTGT
GAATTGTTCTGAACTTAAATGTCCACTAAAACAACTTAGACTTCTGTTTGCTAAATCTGT
TTCTTTTTTCTAATATTCTAAAAAAAAAAAAAAAGGTTTACCTCCACAAATTGAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 19

>></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA33092
><subunit 1 of 1, 690 aa, 0 stop
><MW: 77825, pI: 7.96, NX(S/T): 11
MKRLPLLVVFSTLLNCSYTQNCTKTPCLPNAKCEIRNGIEACYCNMGFSGNGVTICEDDN
ECGNLTQSCGENANCTNTEGSYYCMCVPGFRSSSNQDRFITNDGTVCIENVNANCHLDNV
CIAANINKTLTKIRSIKEPVALLQEVYRNSVTDLSPTDIITYIEILAESSSLLGYKNNTI
SAKDTLSNSTLTEFVKTVNNFVQRDTFVVWDKLSVNHRRTHLTKLMHTVEQATLRISQSF
QKTTEFDTNSTDIALKVFFFDSYNMKHIHPHMNMDGDYINIFPKRKAAYDSNGNVAVAFL
YYKSIGPLLSSSDNFLLKPQNYDNSEEEERVISSVISVSMSSNPPTLYELEKITFTLSHR
KVTDRYRSLCAFWNYSPDTMNGSWSSEGCELTYSNETHTSCRCNHLTHFAILMSSGPSIG
IKDYNILTRITQLGIIISLICLAICIFTFWFFSEIQSTRTTIHKNLCCSLFLAELVFLVG
INTNTNKLFCSIIAGLLHYFFLAAFAWMCIEGIHLYLIVVGVIYNKGFLHKNFYIFGYLS
PAVVVGFSAALGYRYYGTTKVCWLSTENNFIWSFIGPACLIILVNLLAFGVIIYKVFRHT
AGLKPEVSCFENIRSCARGALALLFLLGTTWIFGVLHVVHASVVTAYLFTVSNAFQGMFI
FLFLCVLSRKIQEEYYRLFKNVPCCFGCLR

FIGURE 20

ATAGGAGTCTATGTGGCATTTTGGAATACTCACCTGATACCATGAATGGCAGCTGGTCTTCA
GAGGGCTGTGAGCTGACATACTCAAATGAGACCCACACCTCATGCCGCTGTAATCACCTGAC
ACATTTTGCAATTTTGATGTCCTCTGGTCCTTCCATTGGTATTAAAGATTATAATATTCTTA
CAAGGATCACTCAACTAGGAATAATTATTTCACTGATTTGTCTTGCCATATGCATTTTTACC
TTCTGGTTCTTCAGTGAAATTCAAAGCACCAGGA

FIGURE 21

CTCCCAGCCAAGAACCTCGGGGCCGCTGCGCGGTGGGGAGGAGTTCCCCGAAACCCGGCCGC
TAAGCGAGGCCTCCTCCTCCCGCAGATCCGAACGGCCTGGGCGGGGTCACCCCGGCTGGGAC
AAGAAGCCGCCGCCTGCCTGCCCGGGCCCGGGGAGGGGGCTGGGGCTGGGGCCGGAGGCGGG
GTGTGAGTGGGTGTGTGCGGGGGGCGGAGGCTTGATGCAGTCCCGATAAGAAATGCTCGGGT
GTCTTGGGCACCTACCCGTGGGGGCCGTAAGGCGCTACTATATAACGCTGCCGGCCCTGAGC
CGCCGAGCCGTCCGAGCAGGAGCGCTGCGTCCAGGATCTAGGGCACGACCATCCCAACCCGG
CATTCACAGCCCCGCAGCGCATCCGGTCGCCGGCCAGCTTCCGCACCCCATCGCCGGAGCTG
CGCCGAGAGCCCCAGGGAGGTGCCATGCGGAGCGGGTGTGTGGTGGTCCACGTATGGATCCT
GGCCGGCCTCTGGCTGGCCGTGGCCGGGCGCCCCCTCGCCTTCTCGGACGCGGGGCCCCACG
TGCACTACGGCTGGGGCGACCCCATCCGCCTGCGGCACCTGTACACCTCCGGCCCCCACGGG
CTCTCCAGCTGCTTCCTGCGCATCCGTGCCGACGGCGTCGTGGACTGCGCGCGGGGCCAGAG
CGCGCACAGTTTGCTGGAGATCAAGGCAGTCGCTCTGCGGACCGTGGCCATCAAGGGCGTGC
ACAGCGTGCGGTACCTCTGCATGGGCGCCGACGGCAAGATGCAGGGGCTGCTTCAGTACTCG
GAGGAAGACTGTGCTTTCGAGGAGGAGATCCGCCCAGATGGCTACAATGTGTACCGATCCGA
GAAGCACCGCCTCCCGGTCTCCCTGAGCAGTGCCAAACAGCGGCAGCTGTACAAGAACAGAG
GCTTTCTTCCACTCTCTCATTTCCTGCCCATGCTGCCCATGGTCCCAGAGGAGCCTGAGGAC
CTCAGGGGCCACTTGGAATCTGACATGTTCTCTTCGCCCCTGGAGACCGACAGCATGGACCC
ATTTGGGCTTGTCACCGGACTGGAGGCCGTGAGGAGTCCCAGCTTTGAGAAGTAACTGAGAC
CATGCCCGGGCCTCTTCACTGCTGCCAGGGGCTGTGGTACCTGCAGCGTGGGGGACGTGCTT
CTACAAGAACAGTCCTGAGTCCACGTTCTGTTTAGCTTTAGGAAGAAACATCTAGAAGTTGT
ACATATTCAGAGTTTTCCATTGGCAGTGCCAGTTTCTAGCCAATAGACTTGTCTGATCATAA
CATTGGAAGCCTTGTACTTGGCCCAGCTGTTGCCTGGGCCCCCCATTCTGCTCCCTCGAGGT
TGCTGGACAAGCTGCTGCACTGTCTCAGTTCTGCTTGAATACCTCCATCGATGGGGAACTCA
CTTCCTTTGGAAAAATTCTTATGTCAAGCTGAAATTCTCTAATTTTTCTCATCACTTCCCCA
GGAGCAGCCAGAAGACAGGCAGTAGTTTTAATTTCAGGAACAGGTGATCCACTCTGTAAAAC
AGCAGGTAAATTTCACTCAACCCCATGTGGGAATTGATCTATATCTCTACTTCCAGGGACCA
TTTGCCCTTCCCAAATCCCTCCAGGCCAGAACTGACTGGAGCAGGCATGGCCCACCAGGCTT
CAGAAGTAGGGGAAGCCTGGAGCCCCACTCCAGCCCTGGGACAACTTGAGAATTCCCCCTGA
GGCCAGTTCTGTCATGGATGCTGTCCTGAGAATAACTTGCTGTCCCGGTGTCACCTGCTTCC
ATCTCCCAGCCCACCAGCCCTCTGCCCACCTCACATGCCTCCCCATGGATTGGGGCCTCCCA
GGCCCCCCCACCTTATGTCAACCTGCACTTCTTGTTCAAAAATCAGGAAAAGAAAAGATTTGA
AGACCCCAAGTCTTGTCAATAACTTGCTGTGTGGAAGCAGCGGGGGAAGACCTAGAACCCTT
TCCCCAGCACTTGGTTTTCCAACATGATATTTATGAGTAATTTATTTTGATATGTACATCTC
TTATTTTCTTACATTATTTATGCCCCCAAATTATATTTATGTATGTAAGTGAGGTTTGTTTT
GTATATTAAAATGGAGTTTGTTTGT

FIGURE 22

Met Arg Ser Gly Cys Val Val Val His Val Trp Ile Leu Ala Gly
Leu Trp Leu Ala Val Ala Gly Arg Pro Leu Ala Phe Ser Asp Ala
Gly Pro His Val His Tyr Gly Trp Gly Asp Pro Ile Arg Leu Arg
His Leu Tyr Thr Ser Gly Pro His Gly Leu Ser Ser Cys Phe Leu
Arg Ile Arg Ala Asp Gly Val Val Asp Cys Ala Arg Gly Gln Ser
Ala His Ser Leu Leu Glu Ile Lys Ala Val Ala Leu Arg Thr Val
Ala Ile Lys Gly Val His Ser Val Arg Tyr Leu Cys Met Gly Ala
Asp Gly Lys Met Gln Gly Leu Leu Gln Tyr Ser Glu Glu Asp Cys
Ala Phe Glu Glu Glu Ile Arg Pro Asp Gly Tyr Asn Val Tyr Arg
Ser Glu Lys His Arg Leu Pro Val Ser Leu Ser Ser Ala Lys Gln
Arg Gln Leu Tyr Lys Asn Arg Gly Phe Leu Pro Leu Ser His Phe
Leu Pro Met Leu Pro Met Val Pro Glu Glu Pro Glu Asp Leu Arg
Gly His Leu Glu Ser Asp Met Phe Ser Ser Pro Leu Glu Thr Asp
Ser Met Asp Pro Phe Gly Leu Val Thr Gly Leu Glu Ala Val Arg
Ser Pro Ser Phe Glu Lys

FIGURE 23

CCCAGAAGTTCAAGGGCCCCCGGCCTCCTGCGCTCCTGCCGCCGGGACCCTCGACCTCCTCA
GAGCAGCCGGCTGCCGCCCCGGGAAGATGGCGAGGAGGAGCCGCCACCGCCTCCTCCTGCTG
CTGCTGCGCTACCTGGTGGTCGCCCTGGGCTATCATAAGGCCTATGGGTTTTCTGCCCCAAA
AGACCAACAAGTAGTCACAGCAGTAGAGTACCAAGAGGCTATTTTAGCCTGCAAAACCCCAA
AGAAGACTGTTTCCTCCAGATTAGAGTGGAAGAAACTGGGTCGGAGTGTCTCCTTTGTCTAC
TATCAACAGACTCTTCAAGGTGATTTTAAAAATCGAGCTGAGATGATAGATTTCAATATCCG
GATCAAAAATGTGACAAGAAGTGATGCGGGGAAATATCGTTGTGAAGTTAGTGCCCCATCTG
AGCAAGGCCAAAACCTGGAAGAGGATACAGTCACTCTGGAAGTATTAGTGGCTCCAGCAGTT
CCATCATGTGAAGTACCCTCTTCTGCTCTGAGTGGAACTGTGGTAGAGCTACGATGTCAAGA
CAAAGAAGGGAATCCAGCTCCTGAATACACATGGTTTAAGGATGGCATCCGTTTGCTAGAAA
ATCCCAGACTTGGCTCCCAAAGCACCAACAGCTCATACACAATGAATACAAAAACTGGAACT
CTGCAATTTAATACTGTTTCCAAACTGGACACTGGAGAATATTCCTGTGAAGCCCGCAATTC
TGTTGGATATCGCAGGTGTCCTGGGAAACGAATGCAAGTAGATGATCTCAACATAAGTGGCA
TCATAGCAGCCGTAGTAGTTGTGGCCTTAGTGATTTCCGTTTGTGGCCTTGGTGTATGCTAT
GCTCAGAGGAAAGGCTACTTTTCAAAAGAAACCTCCTTCCAGAAGAGTAATTCTTCATCTAA
AGCCACGACAATGAGTGAAAATGTGCAGTGGCTCACGCCTGTAATCCCAGCACTTTGGAAGG
CCGCGGCGGGCGGATCACGAGGTCAGGAGTTCTAGACCAGTCTGGCCAATATGGTGAAACCC
CATCTCTACTAAAATACAAAAATTAGCTGGGCATGGTGGCATGTGCCTGCAGTTCCAGCTGC
TTGGGAGACAGGAGAATCACTTGAACCCGGGAGGCGGAGGTTGCAGTGAGCTGAGATCACGC
CACTGCAGTCCAGCCTGGGTAACAGAGCAAGATTCCATCTCAAAAAATAAAATAAATAAATA
AATAAATACTGGTTTTTACCTGTAGAATTCTTACAATAAATATAGCTTGATATTC

FIGURE 24

>><ss.DNA35638

><subunit 1 of 1, 312 aa, 1 stop

><MW: 34554, pI: 9.39, NX(S/T): 4

MARRSRHRLLLLLLRYLVVALGYHKAYGFSAPKDQQVVTAVEYQEAILACKTPKKTVSSR
LEWKKLGRSVSFVYYQQTLQGDFKNRAEMIDFNIRIKNVTRSDAGKYRCEVSAPSEQGQN
LEEDTVTLEVLVAPAVPSCEVPSSALSGTVVELRCQDKEGNPAPEYTWFKDGIRLLENPR
LGSQSTNSSYTMNTKTGTLQFNTVSKLDTGEYSCEARNSVGYRRCPGKRMQVDDLNISGI
IAAVVVVALVISVCGLGVCYAQRKGYFSKETSFQKSNSSSKATTMSENVQWLTPVIPALW
KAAAGGSRGQEF

FIGURE 25

GACATCGGAGGTGGGCTAGCACTGAAACTGCTTTTCAAGACGAGGAAGAGGAGGAGAAAGAG
AAAGAAGAGGAAGATGTTGGGCAACATTTATTTAACATGCTCCACAGCCCGGACCCTGGCAT
CATGCTGCTATTCCTGCAAATACTGAAGAAGCATGGGATTTAAATATTTTACTTCTAAATAA
ATGAATTACTCAATCTCCTATGACCATCTATACATACTCCACCTTCAAAAAGTACATCAATA
TTATATCATTAAGGAAATAGTAACCTTCTCTTCTCCAATATGCATGACATTTTTGGACAATG
CAATTGTGGCACTGGCACTTATTTCAGTGAAGAAAAACTTTGTGGTTCTATGGCATTCATCA
TTTGACAAATGCAAGCATCTTCCTTATCAATCAGCTCCTATTGAACTTACTAGCACTGACTG
TGGAATCCTTAAGGGCCCATTACATTTCTGAAGAAGAAAGCTAAGATGAAGGACATGCCACT
CCGAATTCATGTGCTACTTGGCCTAGCTATCACTACACTAGTACAAGCTGTAGATAAAAAAG
TGGATTGTCCACGGTTATGTACGTGTGAAATCAGGCCTTGGTTTACACCCAGATCCATTTAT
ATGGAAGCATCTACAGTGGATTGTAATGATTTAGGTCTTTTAACTTTCCCAGCCAGATTGCC
AGCTAACACACAGATTCTTCTCCTACAGACTAACAATATTGCAAAAATTGAATACTCCACAG
ACTTTCCAGTAAACCTTACTGGCCTGGATTTATCTCAAAACAATTTATCTTCAGTCACCAAT
ATTAATGTAAAAAAGATGCCTCAGCTCCTTTCTGTGTACCTAGAGGAAAACAAACTTACTGA
ACTGCCTGAAAAATGTCTGTCCGAACTGAGCAACTTACAAGAACTCTATATTAATCACAACT
TGCTTTCTACAATTTCACCTGGAGCCTTTATTGGCCTACATAATCTTCTTCGACTTCATCTC
AATTCAAATAGATTGCAGATGATCAACAGTAAGTGGTTTGATGCTCTTCCAAATCTAGAGAT
TCTGATGATTGGGGAAAATCCAATTATCAGAATCAAAGACATGAACTTTAAGCCTCTTATCA
ATCTTCGCAGCCTGGTTATAGCTGGTATAAACCTCACAGAAATACCAGATAACGCCTTGGTT
GGACTGGAAAACTTAGAAAGCATCTCTTTTTACGATAACAGGCTTATTAAAGTACCCCATGT
TGCTCTTCAAAAAGTTGTAAATCTCAAATTTTTGGATCTAAATAAAAATCCTATTAATAGAA
TACGAAGGGGTGATTTTAGCAATATGCTACACTTAAAAGAGTTGGGGATAAATAATATGCCT
GAGCTGATTTCCATCGATAGTCTTGCTGTGGATAACCTGCCAGATTTAAGAAAAATAGAAGC
TACTAACAACCCTAGATTGTCTTACATTCACCCCAATGCATTTTTCAGACTCCCCAAGCTGG
AATCACTCATGCTGAACAGCAATGCTCTCAGTGCCCTGTACCATGGTACCATTGAGTCTCTG
CCAAACCTCAAGGAAATCAGCATACACAGTAACCCCATCAGGTGTGACTGTGTCATCCGTTG
GATGAACATGAACAAAACCAACATTCGATTCATGGAGCCAGATTCACTGTTTTGCGTGGACC
CACCTGAATTCCAAGGTCAGAATGTTCGGCAAGTGCATTTCAGGGACATGATGGAAATTTGT
CTCCCTCTTATAGCTCCTGAGAGCTTTCCTTCTAATCTAAATGTAGAAGCTGGGAGCTATGT
TTCCTTTCACTGTAGAGCTACTGCAGAACCACAGCCTGAAATCTACTGGATAACACCTTCTG
GTCAAAAACTCTTGCCTAATACCCTGACAGACAAGTTCTATGTCCATTCTGAGGGAACACTA
GATATAAATGGCGTAACTCCCAAAGAAGGGGGTTTATATACTTGTATAGCAACTAACCTAGT
TGGCGCTGACTTGAAGTCTGTTATGATCAAAGTGGATGGATCTTTTCCACAAGATAACAATG
GCTCTTTGAATATTAAAATAAGAGATATTCAGGCCAATTCAGTTTTGGTGTCCTGGAAAGCA
AGTTCTAAAATTCTCAAATCTAGTGTTAAATGGACAGCCTTTGTCAAGACTGAAAATTCTCA
TGCTGCGCAAAGTGCTCGAATACCATCTGATGTCAAGGTATATAATCTTACTCATCTGAATC
CATCAACTGAGTATAAAATTTGTATTGATATTCCCACCATCTATCAGAAAAACAGAAAAAAA
TGTGTAAATGTCACCACCAAAGGTTTGCACCCTGATCAAAAAGAGTATGAAAAGAATAATAC
CACAACACTTATGGCCTGTCTTGGAGGCCTTCTGGGGATTATTGGTGTGATATGTCTTATCA
GCTGCCTCTCTCCAGAAATGAACTGTGATGGTGGACACAGCTATGTGAGGAATTACTTACAG
AAACCAACCTTTGCATTAGGTGAGCTTTATCCTCCTCTGATAAATCTCTGGGAAGCAGGAAA
AGAAAAAAGTACATCACTGAAAGTAAAAGCAACTGTTATAGGTTTACCAACAAATATGTCCTAA
AAACCACCAAGGAAACCTACTCCAAAAATGAAC

FIGURE 26

MKDMPLRIHVLLGLAITTLVQAVDKKVDCPRLCTCEIRPWFTPRSIYMEASTVDCNDLG
LLTFPARLPANTQILLLQTNNIAKIEYSTDFPVNLTGLDLSQNNLSSVTNINVKKMPQL
LSVYLEENKLTELPEKCLSELSNLQELYINHNLLSTISPGAFIGLHNLLRLHLNSNRLQ
MINSKWFDALPNLEILMIGENPIIRIKDMNFKPLINLRSLVIAGINLTEIPDNALVGLE
NLESISFYDNRLIKVPHVALQKVVNLKFLDLNKNPINRIRRGDFSNMLHLKELGINNMP
ELISIDSLAVDNLPDLRKIEATNNPRLSYIHPNAFFRLPKLESLMLNSNALSALYHGTI
ESLPNLKEISIHSNPIRCDCVIRWMNMNKTNIRFMEPDSLFCVDPPEFQGQNVRQVHFR
DMMEICLPLIAPESFPSNLNVEAGSYVSFHCRATAEPQPEIYWITPSGQKLLPNTLTDK
FYVHSEGTLDINGVTPKEGGLYTCIATNLVGADLKSVMIKVDGSFPQDNNGSLNIKIRD
IQANSVLVSWKASSKILKSSVKWTAFVKTENSHAAQSARIPSDVKVYNLTHLNPSTEYK
ICIDIPTIYQKNRKKCVNVTTKGLHPDQKEYEKNNTTTLMACLGGLLGIIGVICLISCL
SPEMNCDGGHSYVRNYLQKPTFALGELYPPLINLWEAGKEKSTSLKVKATVIGLPTNMS

FIGURE 27

GCCCGGGACTGGCGCAAGGTGCCCAAGCAAGGAAAGAAATAATGAAGAGACACATGTGTTAG
CTGCAGCCTTTTGAAACACGCAAGAAGGAAATCAATAGTGTGGACAGGGCTGGAACCTTTAC
CACGCTTGTTGGAGTAGATGAGGAATGGGCTCGTGATTATGCTGACATTCCAGCATGAATCT
GGTAGACCTGTGGTTAACCCGTTCCCTCTCCATGTGTCTCCTCCTACAAAGTTTTGTTCTTA
TGATACTGTGCTTTCATTCTGCCAGTATGTGTCCCAAGGGCTGTCTTTGTTCTTCCTCTGGG
GGTTTAAATGTCACCTGTAGCAATGCAAATCTCAAGGAAATACCTAGAGATCTTCCTCCTGA
AACAGTCTTACTGTATCTGGACTCCAATCAGATCACATCTATTCCCAATGAAATTTTTAAGG
ACCTCCATCAACTGAGAGTTCTCAACCTGTCCAAAAATGGCATTGAGTTTATCGATGAGCAT
GCCTTCAAAGGAGTAGCTGAAACCTTGCAGACTCTGGACTTGTCCGACAATCGGATTCAAAG
TGTGCACAAAAATGCCTTCAATAACCTGAAGGCCAGGGCCAGAATTGCCAACAACCCCTGGC
ACTGCGACTGTACTCTACAGCAAGTTCTGAGGAGCATGGCGTCCAATCATGAGACAGCCCAC
AACGTGATCTGTAAAACGTCCGTGTTGGATGAACATGCTGGCAGACCATTCCTCAATGCTGC
CAACGACGCTGACCTTTGTAACCTCCCTAAAAAAACTACCGATTATGCCATGCTGGTCACCA
TGTTTGGCTGGTTCACTATGGTGATCTCATATGTGGTATATTATGTGAGGCAAAATCAGGAG
GATGCCCGGAGACACCTCGAATACTTGAAATCCCTGCCAAGCAGGCAGAAGAAAGCAGATGA
ACCTGATGATATTAGCACTGTGGTATAGTGTCCAAACTGACTGTCATTGAGAAAGAAAGAAA
GTAGTTTGCGATTGCAGTAGAAATAAGTGGTTTACTTCTCCCATCCATTGTAAACATTTGAA
ACTTTGTATTTCAGTTTTTTTTGAATTATGCCACTGCTGAACTTTTAACAAACACTACAACA
TAAATAATTTGAGTTTAGGTGATCCACCCCTTAATTGTACCCCCGATGGTATATTTCTGAGT
AAGCTACTATCTGAACATTAGTTAGATCCATCTCACTATTTAATAATGAAATTTATTTTTTT
AATTTAAAAGCAAATAAAAGCTTAACTTTGAACCATGGGAAAAAAAAAAAAAAAAAAAAAACA

FIGURE 28

><ss.DNA33089

><subunit 1 of 1, 259 aa, 1 stop

><MW: 29275, pI: 6.92, NX(S/T): 2

MNLVDLWLTRSLSMCLLLQSFVLMILCFHSASMCPKGCLCSSSGGLNVTCSNANLKEIPR
DLPPETVLLYLDSNQITSIPNEIFKDLHQLRVLNLSKNGIEFIDEHAFKGVAETLQTLDL
SDNRIQSVHKNAFNNLKARARIANNPWHCDCTLQQVLRSMASNHETAHNVICKTSVLDEH
AGRPFLNAANDADLCNLPKKTTDYAMLVTMFGWFTMVISYVVYYVRQNQEDARRHLEYLK
SLPSRQKKADEPDDISTVV

FIGURE 29

ACCGAGCCGAGCGGACCGAAGGCGCGCCCGAGATGCAGGTGAGCAAGAGGATGCTGGCGGGG
GGCGTGAGGAGCATGCCCAGCCCCCTCCTGGCCTGCTGGCAGCCCATCCTCCTGCTGGTGCT
GGGCTCAGTGCTGTCAGGCTCGGCCACGGGCTGCCCGCCCCGCTGCGAGTGCTCCGCCCAGG
ACCGCGCTGTGCTGTGCCACCGCAAGTGCTTTGTGGCAGTCCCCGAGGGCATCCCCACCGAG
ACGCGCCTGCTGGACCTAGGCAAGAACCGCATCAAAACGCTCAACCAGGACGAGTTCGCCAG
CTTCCCGCACCTGGAGGAGCTGGAGCTCAACGAGAACATCGTGAGCGCCGTGGAGCCCGGCG
CCTTCAACAACCTCTTCAACCTCCGGACGCTGGGTCTCCGCAGCAACCGCCTGAAGCTCATC
CCGCTAGGCGTCTTCACTGGCCTCAGCAACCTGACCAAGCAGGACATCAGCGAGAACAAGAT
CGTTATCCTACTGGACTACATGTTTCAGGACCTGTACAACCTCAAGTCACTGGAGGTTGGCG
ACAATGACCTCGTCTACATCTCTCACCGCGCCTTCAGCGGCCTCAACAGCCTGGAGCAGCTG
ACGCTGGAGAAATGCAACCTGACCTCCATCCCCACCGAGGCGCTGTCCCACCTGCACGGCCT
CATCGTCCTGAGGCTCCGGCACCTCAACATCAATGCCATCCGGGACTACTCCTTCAAGAGGC
TGTACCGACTCAAGGTCTTGGAGATCTCCCACTGGCCCTACTTGGACACCATGACACCCAAC
TGCCTCTACGGCCTCAACCTGACGTCCCTGTCCATCACACACTGCAATCTGACCGCTGTGCC
CTACCTGGCCGTCCGCCACCTAGTCTATCTCCGCTTCCTCAACCTCTCCTACAACCCCATCA
GCACCATTGAGGGCTCCATGTTGCATGAGCTGCTCCGGCTGCAGGAGATCCAGCTGGTGGGC
GGGCAGCTGGCCGTGGTGGAGCCCTATGCCTTCCGCGGCCTCAACTACCTGCGCGTGCTCAA
TGTCTCTGGCAACCAGCTGACCACACTGGAGGAATCAGTCTTCCACTCGGTGGGCAACCTGG
AGACACTCATCCTGGACTCCAACCCGCTGGCCTGCGACTGTCGGCTCCTGTGGGTGTTCCGG
CGCCGCTGGCGGCTCAACTTCAACCGGCAGCAGCCCACGTGCGCCACGCCCGAGTTTGTCCA
GGGCAAGGAGTTCAAGGACTTCCCTGATGTGCTACTGCCCAACTACTTCACCTGCCGCCGCG
CCCGCATCCGGGACCGCAAGGCCCAGCAGGTGTTTGTGGACGAGGGCCACACGGTGCAGTTT
GTGTGCCGGGCCGATGGCGACCCGCCGCCCGCCATCCTCTGGCTCTCACCCCGAAAGCACCT
GGTCTCAGCCAAGAGCAATGGGCGGCTCACAGTCTTCCCTGATGGCACGCTGGAGGTGCGCT
ACGCCCAGGTACAGGACAACGGCACGTACCTGTGCATCGCGGCCAACGCGGGCGGCAACGAC
TCCATGCCCCGCCCACCTGCATGTGCGCAGCTACTCGCCCGACTGGCCCCATCAGCCCAACAA
GACCTTCGCTTTCATCTCCAACCAGCCGGGCGAGGGAGAGGCCAACAGCACCCGCGCCACTG
TGCCTTTCCCCTTCGACATCAAGACCCTCATCATCGCCACCACCATGGGCTTCATCTCTTTC
CTGGGCGTCGTCCTCTTCTGCCTGGTGCTGCTGTTTCTCTGGAGCCGGGGCAAGGGCAACAC
AAAGCACAACATCGAGATCGAGTATGTGCCCCGAAAGTCGGACGCAGGCATCAGCTCCGCCG
ACGCGCCCCGCAAGTTCAACATGAAGATGATATGAGGCCGGGGCGGGGGGCAGGGACCCCCG
GGCGGCCGGGCAGGGGAAGGGGCCTGGTCGCCACCTGCTCACTCTCCAGTCCTTCCCACCTC
CTCCCTACCCTTCTACACACGTTCTCTTTCTCCCTCCCGCCTCCGTCCCCTGCTGCCCCCCG
CCAGCCCTCACCACCTGCCCTCCTTCTACCAGGACCTCAGAAGCCCAGACCTGGGGACCCCA
CCTACACAGGGGCATTGACAGACTGGAGTTGAAAGCCGACGAACCGACACGCGGCAGAGTCA
ATAATTCAATAAAAAAGTTACGAACTTTCTCTGTAACTTGGGTTTCAATAATTATGGATTTT
TATGAAAACTTGAAATAATAAAAAGAGAAAAAAACTAAAAAAAAAAAAAAAAAAAAAA

FIGURE 30

><ss.DNA33786

><subunit 1 of 1, 620 aa, 1 stop

><MW: 69838, pI: 8.84, NX(S/T): 10

MQVSKRMLAGGVRSMPSPLLACWQPILLLVLGSVLSGSATGCPPRCECSAQDRAVLCHRK
CFVAVPEGIPTETRLLDLGKNRIKTLNQDEFASFPHLEELELNENIVSAVEPGAFNNLFN
LRTLGLRSNRLKLIPLGVFTGLSNLTKQDISENKIVILLDYMFQDLYNLKSLEVGDNDLV
YISHRAFSGLNSLEQLTLEKCNLTSIPTEALSHLHGLIVLRLRHLNINAIRDYSFKRLYR
LKVLEISHWPYLDTMTPNCLYGLNLTSLSITHCNLTAVPYLAVRHLVYLRFLNLSYNPIS
TIEGSMLHELLRLQEIQLVGGQLAVVEPYAFRGLNYLRVLNVSGNQLTTLEESVFHSVGN
LETLILDSNPLACDCRLLWVFRRRWRLNFNRQQPTCATPEFVQGKEFKDFPDVLLPNYFT
CRRARIRDRKAQQVFVDEGHTVQFVCRADGDPPPAILWLSPRKHLVSAKSNGRLTVFPDG
TLEVRYAQVQDNGTYLCIAANAGGNDSMPAHLHVRSYSPDWPHQPNKTFAFISNQPGEGE
ANSTRATVPFPFDIKTLIIATTMGFISFLGVVLFCLVLLFLWSRGKGNTKHNIEIEYVPR
KSDAGISSADAPRKFNMKMI

FIGURE 31

CCCACGCGTCCGCACCTCGGCCCCGGGCTCCGAAGCGGCTCGGGGGCGCCCTTTCGGTCAAC
ATCGTAGTCCACCCCCTCCCCATCCCCAGCCCCCGGGGATTCAGGCTCGCCAGCGCCCAGCC
AGGGAGCCGGCCGGGAAGCGCGATGGGGGCCCCAGCCGCCTCGCTCCTGCTCCTGCTCCTGC
TGTTCGCCTGCTGCTGGGCGCCCGGCGGGGCCAACCTCTCCCAGGACGACAGCCAGCCCTGG
ACATCTGATGAAACAGTGGTGGCTGGTGGCACCGTGGTGCTCAAGTGCCAAGTGAAAGATCA
CGAGGACTCATCCCTGCAATGGTCTAACCCTGCTCAGCAGACTCTCTACTTTGGGGAGAAGA
GAGCCCTTCGAGATAATCGAATTCAGCTGGTTACCTCTACGCCCCACGAGCTCAGCATCAGC
ATCAGCAATGTGGCCCTGGCAGACGAGGGCGAGTACACCTGCTCAATCTTCACTATGCCTGT
GCGAACTGCCAAGTCCCTCGTCACTGTGCTAGGAATTCCACAGAAGCCCATCATCACTGGTT
ATAAATCTTCATTACGGGAAAAAGACACAGCCACCCTAAACTGTCAGTCTTCTGGGAGCAAG
CCTGCAGCCCGGCTCACCTGGAGAAAGGGTGACCAAGAACTCCACGGAGAACCAACCCGCAT
ACAGGAAGATCCCAATGGTAAAACCTTCACTGTCAGCAGCTCGGTGACATTCCAGGTTACCC
GGGAGGATGATGGGGCGAGCATCGTGTGCTCTGTGAACCATGAATCTCTAAAGGGAGCTGAC
AGATCCACCTCTCAACGCATTGAAGTTTTATACACACCAACTGCGATGATTAGGCCAGACCC
TCCCCATCCTCGTGAGGGCCAGAAGCTGTTGCTACACTGTGAGGGTCGCGGCAATCCAGTCC
CCCAGCAGTACCTATGGGAGAAGGAGGGCAGTGTGCCACCCCTGAAGATGACCCAGGAGAGT
GCCCTGATCTTCCCTTTCCTCAACAAGAGTGACAGTGGCACCTACGGCTGCACAGCCACCAG
CAACATGGGCAGCTACAAGGCCTACTACACCCTCAATGTTAATGACCCCAGTCCGGTGCCCT
CCTCCTCCAGCACCTACCACGCCATCATCGGTGGGATCGTGGCTTTCATTGTCTTCCTGCTG
CTCATCATGCTCATCTTCCTTGGCCACTACTTGATCCGGCACAAAGGAACCTACCTGACACA
TGAGGCAAAAGGCTCCGACGATGCTCCAGACGCGGACACGGCCATCATCAATGCAGAAGGCG
GGCAGTCAGGAGGGGACGACAAGAAGGAATATTTCATCTAGAGGCGCCTGCCCACTTCCTGC
GCCCCCCAGGGGCCCTGTGGGGACTGCTGGGGCCGTCACCAACCCGGACTTGTACAGAGCAA
CCGCAGGGCCGCCCCTCCCGCTTGCTCCCCAGCCCACCCACCCCCCTGTACAGAATGTCTGC
TTTGGGTGCGGTTTTGTACTCGGTTTGGAATGGGGAGGGAGGAGGGCGGGGGGAGGGGAGGG
TTGCCCTCAGCCCTTTCCGTGGCTTCTCTGCATTTGGGTTATTATTATTTTTGTAACAATCC
CAAATCAAATCTGTCTCCAGGCTGGAGAGGCAGGAGCCCTGGGGTGAGAAAAGCAAAAAACA
AACAAAAAACA

FIGURE 32

MGAPAASLLLLLLLLFACCWAPGGANLSQDDSQPWTSDETVVAGGTVVLKCQVKDHEDSS
LQWSNPAQQTLYFGEKRALRDNRIQLVTSTPHELSISISNVALADEGEYTCSIFTMPVR
TAKSLVTVLGIPQKPIITGYKSSLREKDTATLNCQSSGSKPAARLTWRKGDQELHGEPT
RIQEDPNGKTFTVSSSVTFQVTREDDGASIVCSVNHESLKGADRSTSQRIEVLYTPTAM
IRPDPPHPREGQKLLLHCEGRGNPVPQQYLWEKEGSVPPLKMTQESALIFPFLNKSDSG
TYGCTATSNMGSYKAYYTLNVNDPSPVPSSSSTYHAIIGGIVAFIVFLLLIMLIFLGHY
LIRHKGTYLTHEAKGSDDAPDADTAIINAEGGQSGGDDKKEYFI

FIGURE 33

GGGGGTTAGGGAGGAAGGAATCCACCCCCACCCCCCCAAACCCTTTTCTTCTCCTTTCCTGG
CTTCGGACATTGGAGCACTAAATGAACTTGAATTGTGTCTGTGGCGAGCAGGATGGTCGCTG
TTACTTTGTGATGAGATCGGGGATGAATTGCTCGCTTTAAAAATGCTGCTTTGGATTCTGTT
GCTGGAGACGTCTCTTTGTTTTGCCGCTGGAAACGTTACAGGGGACGTTTGCAAAGAGAAGA
TCTGTTCCTGCAATGAGATAGAAGGGGACCTACACGTAGACTGTGAAAAAAAGGGCTTCACA
AGTCTGCAGCGTTTCACTGCCCCGACTTCCCAGTTTTACCATTTATTTCTGCATGGCAATTC
CCTCACTCGACTTTTCCCTAATGAGTTCGCTAACTTTTATAATGCGGTTAGTTTGCACATGG
AAAACAATGGCTTGCATGAAATCGTTCCGGGGGCTTTTCTGGGGCTGCAGCTGGTGAAAAGG
CTGCACATCAACAACAACAAGATCAAGTCTTTTCGAAAGCAGACTTTTCTGGGGCTGGACGA
TCTGGAATATCTCCAGGCTGATTTTAATTTATTACGAGATATAGACCCGGGGGCCTTCCAGG
ACTTGAACAAGCTGGAGGTGCTCATTTTAAATGACAATCTCATCAGCACCCTACCTGCCAAC
GTGTTCCAGTATGTGCCCATCACCCACCTCGACCTCCGGGGTAACAGGCTGAAAACGCTGCC
CTATGAGGAGGTCTTGGAGCAAATCCCTGGTATTGCGGAGATCCTGCTAGAGGATAACCCTT
GGGACTGCACCTGTGATCTGCTCTCCCTGAAAGAATGGCTGGAAAACATTCCCAAGAATGCC
CTGATCGGCCGAGTGGTCTGCGAAGCCCCCACCAGACTGCAGGGTAAAGACCTCAATGAAAC
CACCGAACAGGACTTGTGTCCTTTGAAAAACCGAGTGGATTCTAGTCTCCCGGCGCCCCCTG
CCCAAGAAGAGACCTTTGCTCCTGGACCCCTGCCAACTCCTTTCAAGACAAATGGGCAAGAG
GATCATGCCACACCAGGGTCTGCTCCAAACGGAGGTACAAAGATCCCAGGCAACTGGCAGAT
CAAAATCAGACCCACAGCAGCGATAGCGACGGGTAGCTCCAGGAACAAACCCTTAGCTAACA
GTTTACCCTGCCCTGGGGGCTGCAGCTGCGACCACATCCCAGGGTCGGGTTTAAAGATGAAC
TGCAACAACAGGAACGTGAGCAGCTTGGCTGATTTGAAGCCCAAGCTCTCTAACGTGCAGGA
GCTTTTCCTACGAGATAACAAGATCCACAGCATCCGAAAATCGCACTTTGTGGATTACAAGA
ACCTCATTCTGTTGGATCTGGGCAACAATAACATCGCTACTGTAGAGAACAACACTTTCAAG
AACCTTTTGGACCTCAGGTGGCTATACATGGATAGCAATTACCTGGACACGCTGTCCCGGGA
GAAATTCGCGGGGCTGCAAAACCTAGAGTACCTGAACGTGGAGTACAACGCTATCCAGCTCA
TCCTCCCGGGCACTTTCAATGCCATGCCCAAACTGAGGATCCTCATTCTCAACAACAACCTG
CTGAGGTCCCTGCCTGTGGACGTGTTCGCTGGGGTCTCGCTCTCTAAACTCAGCCTGCACAA
CAATTACTTCATGTACCTCCCGGTGGCAGGGGTGCTGGACCAGTTAACCTCCATCATCCAGA
TAGACCTCCACGGAAACCCCTGGGAGTGCTCCTGCACAATTGTGCCTTTCAAGCAGTGGGCA
GAACGCTTGGGTTCCGAAGTGCTGATGAGCGACCTCAAGTGTGAGACGCCGGTGAACTTCTT
TAGAAAGGATTTCATGCTCCTCTCCAATGACGAGATCTGCCCTCAGCTGTACGCTAGGATCT
CGCCCACGTTAACTTCGCACAGTAAAAACAGCACTGGGTTGGCGGAGACCGGGACGCACTCC
AACTCCTACCTAGACACCAGCAGGGTGTCCATCTCGGTGTTGGTCCCGGGACTGCTGCTGGT
GTTTGTCACCTCCGCCTTCACCGTGGTGGGCATGCTCGTGTTTATCCTGAGGAACCGAAAGC
GGTCCAAGAGACGAGATGCCAACTCCTCCGCGTCCGAGATTAATTCCCTACAGACAGTCTGT
GACTCTTCCTACTGGCACAATGGGCCTTACAACGCAGATGGGGCCCACAGAGTGTATGACTG
TGGCTCTCACTCGCTCTCAGACTAAGACCCCAACCCCAATAGGGGAGGGCAGAGGGAAGGCG
ATACATCCTTCCCCACCGCAGGCACCCCGGGGGCTGGAGGGGCGTGTACCCAAATCCCCGCG
CCATCAGCCTGGATGGGCATAAGTAGATAAATAACTGTGAGCTCGCACAACCGAAAGGGCCT
GACCCCTTACTTAGCTCCCTCCTTGAAACAAAGAGCAGACTGTGGAGAGCTGGGAGAGCGCA
GCCAGCTCGCTCTTTGCTGAGAGCCCCTTTTGACAGAAAGCCCAGCACGACCCTGCTGGAAG
AACTGACAGTGCCCTCGCCCTCGGCCCCGGGGCCTGTGGGGTTGGATGCCGCGGTTCTATAC
ATATATACATATATCCACATCTATATAGAGAGATAGATATCTATTTTTCCCCTGTGGATTAG
CCCCGTGATGGCTCCCTGTTGGCTACGCAGGGATGGGCAGTTGCACGAAGGCATGAATGTAT
TGTAAATAAGTAACTTTGACTTCTGAC

FIGURE 34

MLLWILLLETSLCFAAGNVTGDVCKEKICSCNEIEGDLHVDCEKKGFTSLQRFTAPTSQ
FYHLFLHGNSLTRLFPNEFANFYNAVSLHMENNGLHEIVPGAFLGLQLVKRLHINNNKI
KSFRKQTFLGLDDLEYLQADFNLLRDIDPGAFQDLNKLEVLILNDNLISTLPANVFQYV
PITHLDLRGNRLKTLPYEEVLEQIPGIAEILLEDNPWDCTCDLLSLKEWLENIPKNALI
GRVVCEAPTRLQGKDLNETTEQDLCPLKNRVDSSLPAPPAQEETFAPGPLPTPFKTNGQ
EDHATPGSAPNGGTKIPGNWQIKIRPTAAIATGSSRNKPLANSLPCPGGCSCDHIPGSG
LKMNCNNRNVSSLADLKPKLSNVQELFLRDNKIHSIRKSHFVDYKNLILLDLGNNNIAT
VENNTFKNLLDLRWLYMDSNYLDTLSREKFAGLQNLEYLNVEYNAIQLILPGTFNAMPK
LRILILNNNLLRSLPVDVFAGVSLSKLSLHNNYFMYLPVAGVLDQLTSIIQIDLHGNPW
ECSCTIVPFKQWAERLGSEVLMSDLKCETPVNFFRKDFMLLSNDEICPQLYARISPTLT
SHSKNSTGLAETGTHSNSYLDTSRVSISVLVPGLLLVFVTSAFTVVGMLVFILRNRKRS
KRRDANSSASEINSLQTVCDSSYWHNGPYNADGAHRVYDCGSHSLSD

FIGURE 35

AGTCGACTGCGTCCCCTGTACCCGGCGCCAGCTGTGTTCCTGACCCCAGAATAACTCAGG
GCTGCACCGGGCCTGGCAGCGCTCCGCACACATTTCCTGTCGCGGCCTAAGGGAAACTGT
TGGCCGCTGGGCCCGCGGGGGGATTCTTGGCAGTTGGGGGGTCCGTCGGGAGCGAGGGCG
GAGGGGAAGGGAGGGGGAACCGGGTTGGGGAAGCCAGCTGTAGAGGGCGGTGACCGCGCT
CCAGACACAGCTCTGCGTCCTCGAGCGGGACAGATCCAAGTTGGGAGCAGCTCTGCGTGC
GGGGCCTCAGAGA
><MET {trans=1-s, dir=f, res=1}
ATGAGGCCGGCGTTCGCCCTGTGCCTCCTCTGGCAGGCGCTCTGGCCCGGGCCGGGCGGC
GGCGAACACCCCACTGCCGACCGTGCTGGCTGCTCGGCCTCGGGGGCCTGCTACAGCCTG
CACCACGCTACCATGAAGCGGCAGGCGGCCGAGGAGGCCTGCATCCTGCGAGGTGGGGCG
CTCAGCACCGTGCGTGCGGGCGCCGAGCTGCGCGCTGTGCTCGCGCTCCTGCGGGCAGGC
CCAGGGCCCGGAGGGGGCTCCAAAGACCTGCTGTTCTGGGTCGCACTGGAGCGCAGGCGT
TCCCACTGCACCCTGGAGAACGAGCCTTTGCGGGGTTTCTCCTGGCTGTCCTCCGACCCC
GGCGGTCTCGAAAGCGACACGCTGCAGTGGGTGGAGGAGCCCCAACGCTCCTGCACCGCG
CGGAGATGCGCGGTACTCCAGGCCACCGGTGGGGTCGAGCCCGCAGGCTGGAAGGAGATG
CGATGCCACCTGCGCGCCAACGGCTACCTGTGCAAGTACCAGTTTGAGGTCTTGTGTCCT
GCGCCGCGCCCCGGGGCCGCCTCTAACTTGAGCTATCGCGCGCCCTTCCAGCTGCACAGC
GCCGCTCTGGACTTCAGTCCACCTGGGACCGAGGTGAGTGCGCTCTGCCGGGGACAGCTC
CCGATCTCAGTTACTTGCATCGCGGACGAAATCGGCGCTCGCTGGGACAAACTCTCGGGC
GATGTGTTGTGTCCCTGCCCCGGGAGGTACCTCCGTGCTGGCAAATGCGCAGAGCTCCCT
AACTGCCTAGACGACTTGGGAGGCTTTGCCTGCGAATGTGCTACGGGCTTCGAGCTGGGG
AAGGACGGCCGCTCTTGTGTGACCAGTGGGGAAGGACAGCCGACCCTTGGGGGGACCGGG
GTGCCCACCAGGCGCCCGCCGGCCACTGCAACCAGCCCCGTGCCGCAGAGAACATGGCCA
ATCAGGGTCGACGAGAAGCTGGGAGAGACACCACTTGTCCCTGAACAAGACAATTCAGTA
ACATCTATTCCTGAGATTCCTCGATGGGGATCACAGAGCACGATGTCTACCCTTCAAATG
TCCCTTCAAGCCGAGTCAAAGGCCACTATCACCCCATCAGGGAGCGTGATTTCCAAGTTT
AATTCTACGACTTCCTCTGCCACTCCTCAGGCTTTCGACTCCTCCTCTGCCGTGGTCTTC
ATATTTGTGAGCACAGCAGTAGTAGTGTTGGTGATCTTGACCATGACAGTACTGGGGCTT
GTCAAGCTCTGCTTTCACGAAAGCCCCTCTTCCCAGCCAAGGAAGGAGTCTATGGGCCCG
CCGGGCCTGGAGAGTGATCCTGAGCCCGCTGCTTTGGGCTCCAGTTCTGCACATTGCACA
AACAATGGGGTGAAAGTCGGGGACTGTGATCTGCGGGACAGAGCAGAGGGTGCCTTGCTG
GCGGAGTCCCCTCTTGGCTCTAGTGATGCATAGGGAAACAGGGGACATGGGCACTCCTGT
GAACAGTTTTTCACTTTTGATGAAACGGGGAACCAAGAGGAACTTACTTGTGTAACTGAC
AATTTCTGCAGAAATCCCCCTTCCTCTAAATTCCCTTTACTCCACTGAGGAGCTAAATCA
GAACTGCACACTCCTTCCCTGATGATAGAGGAAGTGGAAGTGCCTTTAGGATGGTGATAC
TGGGGGACCGGGTAGTGCTGGGGAGAGATATTTTCTTATGTTTATTCGGAGAATTTGGAG
AAGTGATTGAACTTTTCAAGACATTGGAAACAAATAGAACACAATATAATTTACATTAAA
AAATAATTTCTACCAAAATGGAAAGGAAATGTTCTATGTTGTTCAGGCTAGGAGTATATT
GGTTCGAAATCCCAGGGAAAAAAATAAAAATAAAAAATTAAAGGATTGTTGAT

FIGURE 36

MRPAFALCLLWQALWPGPGGGEHPTADRAGCSASGACYSLHHATMKRQAAEEACILRGGA
LSTVRAGAELRAVLALLRAGPGPGGGSKDLLFWVALERRRSHCTLENEPLRGFSWLSSDP
GGLESDTLQWVEEPQRSCTARRCAVLQATGGVEPAGWKEMRCHLRANGYLCKYQFEVLCP
APRPGAASNLSYRAPFQLHSAALDFSPPGTEVSALCRGQLPISVTCIADEIGARWDKLSG
DVLCPCPGRYLRAGKCAELPNCLDDLGGFACECATGFELGKDGRSCVTSGEGQPTLGGTG
VPTRRPPATATSPVPQRTWPIRVDEKLGETPLVPEQDNSVTSIPEIPRWGSQSTMSTLQM
SLQAESKATITPSGSVISKFNSTTSSATPQAFDSSSAVVFIFVSTAVVVLVILTMTVLGL
VKLCFHESPSSQPRKESMGPPGLESDPEPAALGSSSAHCTNNGVKVGDCDLRDRAEGALL
AESPLGSSDA

FIGURE 37

CGGACGCGTGGGATTCAGCAGTGGCCTGTGGCTGCCAGAGCAGCTCCTCAGGGGAAACTA
AGCGTCGAGTCAGACGGCACCATAATCGCCTTTAAAAGTGCCTCCGCCCTGCCGGCCGCG
TATCCCCCGGCTACCTGGGCCGCCCCGCGGCGGTGCGCGCGTGAGAGGGAGCGCGCGGGC
AGCCGAGCGCCGGTGTGAGCCAGCGCTGCTGCCAGTGTGAGCGGCGGTGTGAGCGCGGTG
GGTGCGGAGGGGCGTGTGTGCCGGCGCGCGCGCCGTGGGGTGCAAACCCCGAGCGTCTAC
GCTGCC
><MET {trans=1-s, dir=f, res=1}
ATGAGGGGCGCGAACGCCTGGGCGCCACTCTGCCTGCTGCTGGCTGCCGCCACCCAGCTC
TCGCGGCAGCAGTCCCCAGAGAGACCTGTTTTCACATGTGGTGGCATTCTTACTGGAGAG
TCTGGATTTATTGGCAGTGAAGGTTTTCCTGGAGTGTACCCTCCAAATAGCAAATGTACT
TGGAAAATCACAGTTCCCGAAGGAAAAGTAGTCGTTCTCAATTTCCGATTCATAGACCTC
GAGAGTGACAACCTGTGCCGCTATGACTTTGTGGATGTGTACAATGGCCATGCCAATGGC
CAGCGCATTGGCCGCTTCTGTGGCACTTTCCGGCCTGGAGCCCTTGTGTCCAGTGGCAAC
AAGATGATGGTGCAGATGATTTCTGATGCCAACACAGCTGGCAATGGCTTCATGGCCATG
TTCTCCGCTGCTGAACCAAACGAAAGAGGGGATCAGTATTGTGGAGGACTCCTTGACAGA
CCTTCCGGCTCTTTTAAAACCCCCAACTGGCCAGACCGGGATTACCCTGCAGGAGTCACT
TGTGTGTGGCACATTGTAGCCCCAAAGAATCAGCTTATAGAATTAAAGTTTGAGAAGTTT
GATGTGGAGCGAGATAACTACTGCCGATATGATTATGTGGCTGTGTTTAATGGCGGGGAA
GTCAACGATGCTAGAAGAATTGGAAAGTATTGTGGTGATAGTCCACCTGCGCCAATTGTG
TCTGAGAGAAATGAACTTCTTATTCAGTTTTTATCAGACTTAAGTTTAACTGCAGATGGG
TTTATTGGTCACTACATATTCAGGCCAAAAAAACTGCCTACAACTACAGAACAGCCTGTC
ACCACCACATTCCCTGTAACCACGGGTTTAAAACCCACCGTGGCCTTGTGTCAACAAAAG
TGTAGACGGACGGGGACTCTGGAGGGCAATTATTGTTCAAGTGACTTTGTATTAGCCGGC
ACTGTTATCACAACCATCACTCGCGATGGGAGTTTGCACGCCACAGTCTCGATCATCAAC
ATCTACAAAGAGGGAAATTTGGCGATTCAGCAGGCGGGCAAGAACATGAGTGCCAGGCTG
ACTGTCGTCTGCAAGCAGTGCCCTCTCCTCAGAAGAGGTCTAAATTACATTATTATGGGC
CAAGTAGGTGAAGATGGGCGAGGCAAAATCATGCCAAACAGCTTTATCATGATGTTCAAG
ACCAAGAATCAGAAGCTCCTGGATGCCTTAAAAAATAAGCAATGTTAACAGTGAACTGTG
TCCATTTAAGCTGTATTCTGCCATTGCCTTTGAAAGATCTATGTTCTCTCAGTAGAAAAA
AAAATACTTATAAAATTACATATTCTGAAAGAGGATTCCGAAAGATGGGACTGGTTGACT
CTTCACATGATGGAGGTATGAGGCCTCCGAGATAGCTGAGGGAAGTTCTTTGCCTGCTGT
CAGAGGAGCAGCTATCTGATTGGAAACCTGCCGACTTAGTGCGGTGATAGGAAGCTAAAA
GTGTCAAGCGTTGACAGCTTGGAAGCGTTTATTTATACATCTCTGTAAAAGGATATTTTA
GAATTGAGTTGTGTGAAGATGTCAAAAAAAGATTTTAGAAGTGCAATATTTATAGTGTTA
TTTGTTTCACCTTCAAGCCTTTGCCCTGAGGTGTTACAATCTTGTCTTGCGTTTTCTAAA
TCAATGCTTAATAAAATATTTTTAAAGGAAAAAAAAAAAA

FIGURE 38

MRGANAWAPLCLLLAAATQLSRQQSPERPVFTCGGILTGESGFIGSEGFPGVYPPNSKCT
WKITVPEGKVVVLNFRFIDLESDNLCRYDFVDVYNGHANGQRIGRFCGTFRPGALVSSGN
KMMVQMISDANTAGNGFMAMFSAAEPNERGDQYCGGLLDRPSGSFKTPNWPDRDYPAGVT
CVWHIVAPKNQLIELKFEKFDVERDNYCRYDYVAVFNGGEVNDARRIGKYCGDSPPAPIV
SERNELLIQFLSDLSLTADGFIGHYIFRPKKLPTTTEQPVTTTFPVTTGLKPTVALCQQK
CRRTGTLEGNYCSSDFVLAGTVITTITRDGSLHATVSIINIYKEGNLAIQQAGKNMSARL
TVVCKQCPLLRRGLNYIIMGQVGEDGRGKIMPNSFIMMFKTKNQKLLDALKNKQC

FIGURE 39

CGGACGCGTGGGCGGACGCGTGGGCGGCCCACGGCGCCCGCGGGCTGGGGCGGTCGCTTCTT
CCTTCTCCGTGGCCTACGAGGGTCCCCAGCCTGGGTAAAGATGGCCCCATGGCCCCCGAAGG
GCCTAGTCCCAGCTGTGCTCTGGGGCCTCAGCCTCTTCCTCAACCTCCCAGGACCTATCTGG
CTCCAGCCCTCTCCACCTCCCCAGTCTTCTCCCCCGCCTCAGCCCCATCCGTGTCATACCTG
CCGGGGACTGGTTGACAGCTTTAACAAGGGCCTGGAGAGAACCATCCGGGACAACTTTGGAG
GTGGAAACACTGCCTGGGAGGAAGAGAATTTGTCCAAATACAAAGACAGTGAGACCCGCCTG
GTAGAGGTGCTGGAGGGTGTGTGCAGCAAGTCAGACTTCGAGTGCCACCGCCTGCTGGAGCT
GAGTGAGGAGCTGGTGGAGAGCTGGTGGTTTCACAAGCAGCAGGAGGCCCCGGACCTCTTCC
AGTGGCTGTGCTCAGATTCCCTGAAGCTCTGCTGCCCCGCAGGCACCTTCGGGCCCTCCTGC
CTTCCCTGTCCTGGGGGAACAGAGAGGCCCTGCGGTGGCTACGGGCAGTGTGAAGGAGAAGG
GACACGAGGGGGCAGCGGGCACTGTGACTGCCAAGCCGGCTACGGGGGTGAGGCCTGTGGCC
AGTGTGGCCTTGGCTACTTTGAGGCAGAACGCAACGCCAGCCATCTGGTATGTTCGGCTTGT
TTTGGCCCCTGTGCCCGATGCTCAGGACCTGAGGAATCAAACTGTTTGCAATGCAAGAAGGG
CTGGGCCCTGCATCACCTCAAGTGTGTAGACATTGATGAGTGTGGCACAGAGGGAGCCAACT
GTGGAGCTGACCAATTCTGCGTGAACACTGAGGGCTCCTATGAGTGCCGAGACTGTGCCAAG
GCCTGCCTAGGCTGCATGGGGGCAGGGCCAGGTCGCTGTAAGAAGTGTAGCCCTGGCTATCA
GCAGGTGGGCTCCAAGTGTCTCGATGTGGATGAGTGTGAGACAGAGGTGTGTCCGGGAGAGA
ACAAGCAGTGTGAAAACACCGAGGGCGGTTATCGCTGCATCTGTGCCGAGGGCTACAAGCAG
ATGGAAGGCATCTGTGTGAAGGAGCAGATCCCAGAGTCAGCAGGCTTCTTCTCAGAGATGAC
AGAAGACGAGTTGGTGGTGCTGCAGCAGATGTTCTTTGGCATCATCATCTGTGCACTGGCCA
CGCTGGCTGCTAAGGGCGACTTGGTGTTCACCGCCATCTTCATTGGGGCTGTGGCGGCCATG
ACTGGCTACTGGTTGTCAGAGCGCAGTGACCGTGTGCTGGAGGGCTTCATCAAGGGCAGATA
ATCGCGGCCACCACCTGTAGGACCTCCTCCCACCCACGCTGCCCCCAGAGCTTGGGCTGCCC
TCCTGCTGGACACTCAGGACAGCTTGGTTTATTTTTGAGAGTGGGGTAAGCACCCCTACCTG
CCTTACAGAGCAGCCCAGGTACCCAGGCCCGGGCAGACAAGGCCCCTGGGGTAAAAAGTAGC
CCTGAAGGTGGATACCATGAGCTCTTCACCTGGCGGGGACTGGCAGGCTTCACAATGTGTGA
ATTTCAAAAGTTTTTCCTTAATGGTGGCTGCTAGAGCTTTGGCCCCTGCTTAGGATTAGGTG
GTCCTCACAGGGGTGGGGCCATCACAGCTCCCTCCTGCCAGCTGCATGCTGCCAGTTCCTGT
TCTGTGTTCACCACATCCCCACACCCCATTGCCACTTATTTATTCATCTCAGGAAATAAAGA
AAGGTCTTGGAAAGTTAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 40

Met Ala Pro Trp Pro Pro Lys Gly Leu Val Pro Ala Val Leu Trp Gly
Leu Ser Leu Phe Leu Asn Leu Pro Gly Pro Ile Trp Leu Gln Pro Ser
Pro Pro Pro Gln Ser Ser Pro Pro Pro Gln Pro His Pro Cys His Thr
Cys Arg Gly Leu Val Asp Ser Phe Asn Lys Gly Leu Glu Arg Thr Ile
Arg Asp Asn Phe Gly Gly Gly Asn Thr Ala Trp Glu Glu Glu Asn Leu
Ser Lys Tyr Lys Asp Ser Glu Thr Arg Leu Val Glu Val Leu Glu Gly
Val Cys Ser Lys Ser Asp Phe Glu Cys His Arg Leu Leu Glu Leu Ser
Glu Glu Leu Val Glu Ser Trp Trp Phe His Lys Gln Gln Glu Ala Pro
Asp Leu Phe Gln Trp Leu Cys Ser Asp Ser Leu Lys Leu Cys Cys Pro
Ala Gly Thr Phe Gly Pro Ser Cys Leu Pro Cys Pro Gly Gly Thr Glu
Arg Pro Cys Gly Gly Tyr Gly Gln Cys Glu Gly Glu Gly Thr Arg Gly
Gly Ser Gly His Cys Asp Cys Gln Ala Gly Tyr Gly Gly Glu Ala Cys
Gly Gln Cys Gly Leu Gly Tyr Phe Glu Ala Glu Arg Asn Ala Ser His
Leu Val Cys Ser Ala Cys Phe Gly Pro Cys Ala Arg Cys Ser Gly Pro
Glu Glu Ser Asn Cys Leu Gln Cys Lys Lys Gly Trp Ala Leu His His
Leu Lys Cys Val Asp Ile Asp Glu Cys Gly Thr Glu Gly Ala Asn Cys
Gly Ala Asp Gln Phe Cys Val Asn Thr Glu Gly Ser Tyr Glu Cys Arg
Asp Cys Ala Lys Ala Cys Leu Gly Cys Met Gly Ala Gly Pro Gly Arg
Cys Lys Lys Cys Ser Pro Gly Tyr Gln Gln Val Gly Ser Lys Cys Leu
Asp Val Asp Glu Cys Glu Thr Glu Val Cys Pro Gly Glu Asn Lys Gln
Cys Glu Asn Thr Glu Gly Gly Tyr Arg Cys Ile Cys Ala Glu Gly Tyr
Lys Gln Met Glu Gly Ile Cys Val Lys Glu Gln Ile Pro Glu Ser Ala
Gly Phe Phe Ser Glu Met Thr Glu Asp Glu Leu Val Val Leu Gln Gln
Met Phe Phe Gly Ile Ile Ile Cys Ala Leu Ala Thr Leu Ala Ala Lys
Gly Asp Leu Val Phe Thr Ala Ile Phe Ile Gly Ala Val Ala Ala Met
Thr Gly Tyr Trp Leu Ser Glu Arg Ser Asp Arg Val Leu Glu Gly Phe
Ile Lys Gly Arg

FIGURE 41

TGAGACCCTCCTGCAGCCTTCTCAAGGGACAGCCCCACTCTGCCTCTTGCTCCTCCAGGG
CAGCACC
><MET {trans=1-s, dir=f, res=1}
ATGCAGCCCCTGTGGCTCTGCTGGGCACTCTGGGTGTTGCCCCTGGCCAGCCCCGGGGCC
GCCCTGACCGGGGAGCAGCTCCTGGGCAGCCTGCTGCGGCAGCTGCAGCTCAAAGAGGTG
CCCACCCTGGACAGGGCCGACATGGAGGAGCTGGTCATCCCCACCCACGTGAGGGCCCAG
TACGTGGCCCTGCTGCAGCGCAGCCACGGGGACCGCTCCCGCGGAAAGAGGTTCAGCCAG
AGCTTCCGAGAGGTGGCCGGCAGGTTCCTGGCGTTGGAGGCCAGCACACACCTGCTGGTG
TTCGGCATGGAGCAGCGGCTGCCGCCCAACAGCGAGCTGGTGCAGGCCGTGCTGCGGCTC
TTCCAGGAGCCGGTCCCCAAGGCCGCGCTGCACAGGCACGGGCGGCTGTCCCCGCGCAGC
GCCCGGGCCCGGGTGACCGTCGAGTGGCTGCGCGTCCGCGACGACGGCTCCAACCGCACC
TCCCTCATCGACTCCAGGCTGGTGTCCGTCCACGAGAGCGGCTGGAAGGCCTTCGACGTG
ACCGAGGCCGTGAACTTCTGGCAGCAGCTGAGCCGGCCCCGGCAGCCGCTGCTGCTACAG
GTGTCGGTGCAGAGGGAGCATCTGGGCCCGCTGGCGTCCGGCGCCCACAAGCTGGTCCGC
TTTGCCTCGCAGGGGGCGCCAGCCGGGCTTGGGGAGCCCCAGCTGGAGCTGCACACCCTG
GACCTTGGGGACTATGGAGCTCAGGGCGACTGTGACCCTGAAGCACCAATGACCGAGGGC
ACCCGCTGCTGCCGCCAGGAGATGTACATTGACCTGCAGGGGATGAAGTGGGCCGAGAAC
TGGGTGCTGGAGCCCCCGGGCTTCCTGGCTTATGAGTGTGTGGGCACCTGCCGGCAGCCC
CCGGAGGCCCTGGCCTTCAAGTGGCCGTTTCTGGGGCCTCGACAGTGCATCGCCTCGGAG
ACTGACTCGCTGCCCATGATCGTCAGCATCAAGGAGGGAGGCAGGACCAGGCCCCAGGTG
GTCAGCCTGCCCAACATGAGGGTGCAGAAGTGCAGCTGTGCCTCGGATGGTGCGCTCGTG
CCAAGGAGGCTCCAGCCATAGGCGCCTAGTGTAGCCATCGAGGGACTTGACTTGTGTGTG
TTTCTGAAGTGTTCGAGGGTACCAGGAGAGCTGGCGATGACTGAACTGCTGATGGACAAA
TGCTCTGTGCTCTCTAGTGAGCCCTGAATTTGCTTCCTCTGACAAGTTACCTCACCTAAT
TTTTGCTTCTCAGGAATGAGAATCTTTGGCCACTGGAGAGCCCTTGCTCAGTTTTCTCTA
TTCTTATTATTCACTGCACTATATTCTAAGCACTTACATGTGGAGATACTGTAACCTGAG
GGCAGAAAGCCCANTGTGTCATTGTTTACTTGTCCTGTCACTGGATCTGGGCTAAAGTCC
TCCACCACCACTCTGGACCTAAGACCTGGGGTTAAGTGTGGGTTGTGCATCCCCAATCCA
GATAATAAAGACTTTGTAAAACATGAATAAAACACATTTTATTCTAAAA

FIGURE 42

MQPLWLCWALWVLPLASPGAALTGEQLLGSLLRQLQLKEVPTLDRADMEELVIPTHVRAQ
YVALLQRSHGDRSRGKRFSQSFREVAGRFLALEASTHLLVFGMEQRLPPNSELVQAVLRL
FQEPVPKAALHRHGRLSPRSARARVTVEWLRVRDDGSNRTSLIDSRLVSVHESGWKAFDV
TEAVNFWQQLSRPRQPLLLQVSVQREHLGPLASGAHKLVRFASQGAPAGLGEPQLELHTL
DLGDYGAQGDCDPEAPMTEGTRCCRQEMYIDLQGMKWAENWVLEPPGFLAYECVGTCRQP
PEALAFKWPFLGPRQCIASETDSLPMIVSIKEGGRTRPQVVSLPNMRVQKCSCASDGALV
PRRLQP

FIGURE 43

GTCTGTTCCCAGGAGTCCTTCGGCGGCTGTTGTGTCAGTGGCCTGATCGCGATGGGGACAAA
GGCGCAAGTCGAGAGGAAACTGTTGTGCCTCTTCATATTGGCGATCCTGTTGTGCTCCCTGG
CATTGGGCAGTGTTACAGTGCACTCTTCTGAACCTGAAGTCAGAATTCCTGAGAATAATCCT
GTGAAGTTGTCCTGTGCCTACTCGGGCTTTTCTTCTCCCCGTGTGGAGTGGAAGTTTGACCA
AGGAGACACCACCAGACTCGTTTGCTATAATAACAAGATCACAGCTTCCTATGAGGACCGGG
TGACCTTCTTGCCAACTGGTATCACCTTCAAGTCCGTGACACGGGAAGACACTGGGACATAC
ACTTGTATGGTCTCTGAGGAAGGCGGCAACAGCTATGGGGAGGTCAAGGTCAAGCTCATCGT
GCTTGTGCCTCCATCCAAGCCTACAGTTAACATCCCCTCCTCTGCCACCATTGGGAACCGGG
CAGTGCTGACATGCTCAGAACAAGATGGTTCCCCACCTTCTGAATACACCTGGTTCAAAGAT
GGGATAGTGATGCCTACGAATCCCAAAAGCACCCGTGCCTTCAGCAACTCTTCCTATGTCCT
GAATCCCACAACAGGAGAGCTGGTCTTTGATCCCCTGTCAGCCTCTGATACTGGAGAATACA
GCTGTGAGGCACGGAATGGGTATGGGACACCCATGACTTCAAATGCTGTGCGCATGGAAGCT
GTGGAGCGGAATGTGGGGGTCATCGTGGCAGCCGTCCTTGTAACCCTGATTCTCCTGGGAAT
CTTGGTTTTTGGCATCTGGTTTGCCTATAGCCGAGGCCACTTTGACAGAACAAAGAAAGGGA
CTTCGAGTAAGAAGGTGATTTACAGCCAGCCTAGTGCCCGAAGTGAAGGAGAATTCAAACAG
ACCTCGTCATTCCTGGTGTGAGCCTGGTCGGCTCACCGCCTATCATCTGCATTTGCCTTACT
CAGGTGCTACCGGACTCTGGCCCCTGATGTCTGTAGTTTCACAGGATGCCTTATTTGTCTTC
TACACCCCACAGGGCCCCCTACTTCTTCGGATGTGTTTTTAATAATGTCAGCTATGTGCCCC
ATCCTCCTTCATGCCCTCCCTCCCTTTCCTACCACTGCTGAGTGGCCTGGAACTTGTTTAAA
GTGTTTATTCCCCATTTCTTTGAGGGATCAGGAAGGAATCCTGGGTATGCCATTGACTTCCC
TTCTAAGTAGACAGCAAAAATGGCGGGGGTCGCAGGAATCTGCACTCAACTGCCCACCTGGC
TGGCAGGGATCTTTGAATAGGTATCTTGAGCTTGGTTCTGGGCTCTTTCCTTGTGTACTGAC
GACCAGGGCCAGCTGTTCTAGAGCGGGAATTAGAGGCTAGAGCGGCTGAAATGGTTGTTTGG
TGATGACACTGGGGTCCTTCCATCTCTGGGGCCCACTCTCTTCTGTCTTCCCATGGGAAGTG
CCACTGGGATCCCTCTGCCCTGTCCTCCTGAATACAAGCTGACTGACATTGACTGTGTCTGT
GGAAAATGGGAGCTCTTGTTGTGGAGAGCATAGTAAATTTTCAGAGAACTTGAAGCCAAAAG
GATTTAAAACCGCTGCTCTAAAGAAAAGAAAACTGGAGGCTGGGCGCAGTGGCTCACGCCTG
TAATCCCAGAGGCTGAGGCAGGCGGATCACCTGAGGTCGGGAGTTCGGGATCAGCCTGACCA
ACATGGAGAAACCCTACTGGAAATACAAAGTTAGCCAGGCATGGTGGTGCATGCCTGTAGTC
CCAGCTGCTCAGGAGCCTGGCAACAAGAGCAAAACTCCAGCTCA

FIGURE 44

Met Gly Thr Lys Ala Gln Val Glu Arg Lys Leu Leu Cys Leu Phe Ile
Leu Ala Ile Leu Leu Cys Ser Leu Ala Leu Gly Ser Val Thr Val His
Ser Ser Glu Pro Glu Val Arg Ile Pro Glu Asn Asn Pro Val Lys Leu
Ser Cys Ala Tyr Ser Gly Phe Ser Ser Pro Arg Val Glu Trp Lys Phe
Asp Gln Gly Asp Thr Thr Arg Leu Val Cys Tyr Asn Asn Lys Ile Thr
Ala Ser Tyr Glu Asp Arg Val Thr Phe Leu Pro Thr Gly Ile Thr Phe
Lys Ser Val Thr Arg Glu Asp Thr Gly Thr Tyr Thr Cys Met Val Ser
Glu Glu Gly Gly Asn Ser Tyr Gly Glu Val Lys Val Lys Leu Ile Val
Leu Val Pro Pro Ser Lys Pro Thr Val Asn Ile Pro Ser Ser Ala Thr
Ile Gly Asn Arg Ala Val Leu Thr Cys Ser Glu Gln Asp Gly Ser Pro
Pro Ser Glu Tyr Thr Trp Phe Lys Asp Gly Ile Val Met Pro Thr Asn
Pro Lys Ser Thr Arg Ala Phe Ser Asn Ser Ser Tyr Val Leu Asn Pro
Thr Thr Gly Glu Leu Val Phe Asp Pro Leu Ser Ala Ser Asp Thr Gly
Glu Tyr Ser Cys Glu Ala Arg Asn Gly Tyr Gly Thr Pro Met Thr Ser
Asn Ala Val Arg Met Glu Ala Val Glu Arg Asn Val Gly Val Ile Val
Ala Ala Val Leu Val Thr Leu Ile Leu Leu Gly Ile Leu Val Phe Gly
Ile Trp Phe Ala Tyr Ser Arg Gly His Phe Asp Arg Thr Lys Lys Gly
Thr Ser Ser Lys Lys Val Ile Tyr Ser Gln Pro Ser Ala Arg Ser Glu
Gly Glu Phe Lys Gln Thr Ser Ser Phe Leu Val

FIGURE 45

CAGCGCGTGGCCGGCGCCGCTGTGGGGACAGCATGAGCGGCGGTTGGATGGCGCAGGTTGGA
GCGTGGCGAACAGGGGCTCTGGGCCTGGCGCTGCTGCTGCTGCTCGGCCTCGGACTAGGCCT
GGAGGCCGCCGCGAGCCCGCTTTCCACCCCGACCTCTGCCCAGGCCGCAGGCCCCAGCTCAG
GCTCGTGCCCACCCACCAAGTTCCAGTGCCGCACCAGTGGCTTATGCGTGCCCCTCACCTGG
CGCTGCGACAGGGACTTGGACTGCAGCGATGGCAGCGATGAGGAGGAGTGCAGGATTGAGCC
ATGTACCCAGAAAGGGCAATGCCCACCGCCCCCTGGCCTCCCCTGCCCCTGCACCGGCGTCA
GTGACTGCTCTGGGGGAACTGACAAGAAACTGCGCAACTGCAGCCGCCTGGCCTGCCTAGCA
GGCGAGCTCCGTTGCACGCTGAGCGATGACTGCATTCCACTCACGTGGCGCTGCGACGGCCA
CCCAGACTGTCCCGACTCCAGCGACGAGCTCGGCTGTGGAACCAATGAGATCCTCCCGGAAG
GGGATGCCACAACCATGGGGCCCCCTGTGACCCTGGAGAGTGTCACCTCTCTCAGGAATGCC
ACAACCATGGGGCCCCCTGTGACCCTGGAGAGTGTCCCCTCTGTCGGGAATGCCACATCCTC
CTCTGCCGGAGACCAGTCTGGAAGCCCAACTGCCTATGGGGTTATTGCAGCTGCTGCGGTGC
TCAGTGCAAGCCTGGTCACCGCCACCCTCCTCCTTTTGTCCTGGCTCCGAGCCCAGGAGCGC
CTCCGCCCACTGGGGTTACTGGTGGCCATGAAGGAGTCCCTGCTGCTGTCAGAACAGAAGAC
CTCGCTGCCCTGAGGACAAGCACTTGCCACCACCGTCACTCAGCCCTGGGCGTAGCCGGACA
GGAGGAGAGCAGTGATGCGGATGGGTACCCGGGCACACCAGCCCTCAGAGACCTGAGTTCTT
CTGGCCACGTGGAACCTCGAACCCGAGCTCCTGCAGAAGTGGCCCTGGAGATTGAGGGTCCC
TGGACACTCCCTATGGAGATCCGGGGAGCTAGGATGGGGAACCTGCCACAGCCAGAACTGAG
GGGCTGGCCCCAGGCAGCTCCCAGGGGGTAGAACGGCCCTGTGCTTAAGACACTCCCTGCTG
CCCCGTCTGAGGGTGGCGATTAAAGTTGCTTC

FIGURE 46

><ss.DNA33221

><subunit 1 of 1, 282 aa, 1 stop

><MW: 28991, pI: 4.62, NX(S/T): 3

MSGGWMAQVGAWRTGALGLALLLLLGLGLGLEAAASPLSTPTSAQAAGPSSGSCPPTKFQ
CRTSGLCVPLTWRCDRDLDCSDGSDEEECRIEPCTQKGQCPPPPGLPCPCTGVSDCSGGT
DKKLRNCSRLACLAGELRCTLSDDCIPLTWRCDGHPDCPDSSDELGCGTNEILPEGDATT
MGPPVTLESVTSLRNATTMGPPVTLESVPSVGNATSSSAGDQSGSPTAYGVIAAAAVLSA
SLVTATLLLLLSWLRAQERLRPLGLLVAMKESLLLSEQKTSLP

FIGURE 47

CCCACGCGTCCGGTCTCGCTCGCTCGCGCAGCGGCGGCAGCAGAGGTCGCGCACAGATGCGG
GTTAGACTGGCGGGGGGAGGAGGCGGAGGAGGGAAGGAAGCTGCATGCATGAGACCCACAGA
CTCTTGCAAGCTGGATGCCCTCTGTGGATGAAAGATGTATCATGGAATGAACCCGAGCAATG
GAGATGGATTTCTAGAGCAGCAGCAGCAGCAGCAACCTCAGTCCCCCCAGAGACTCTTG
GCCGTGATCCTGTGGTTTCAGCTGGCGCTGTGCTTCGGCCCTGCACAGCTCACGGGCGGGTT
CGATGACCTTCAAGTGTGTGCTGACCCCGGCATTCCCGAGAATGGCTTCAGGACCCCCAGCG
GAGGGGTTTTCTTTGAAGGCTCTGTAGCCCGATTTCACTGCCAAGACGGATTCAAGCTGAAG
GGCGCTACAAAGAGACTGTGTTTGAAGCATTTTAATGGAACCCTAGGCTGGATCCCAAGTGA
TAATTCCATCTGTGTGCAAGAAGATTGCCGTATCCCTCAAATCGAAGATGCTGAGATTCATA
ACAAGACATATAGACATGGAGAGAAGCTAATCATCACTTGTCATGAAGGATTCAAGATCCGG
TACCCCGACCTACACAATATGGTTTCATTATGTCGCGATGATGGAACGTGGAATAATCTGCC
CATCTGTCAAGGCTGCCTGAGACCTCTAGCCTCTTCTAATGGCTATGTAAACATCTCTGAGC
TCCAGACCTCCTTCCCGGTGGGGACTGTGATCTCCTATCGCTGCTTTCCCGGATTTAAACTT
GATGGGTCTGCGTATCTTGAGTGCTTACAAAACCTTATCTGGTCGTCCAGCCCACCCCGGTG
CCTTGCTCTGGAAGCCCAAGTCTGTCCACTACCTCCAATGGTGAGTCACGGAGATTTCGTCT
GCCACCCGCGGCCTTGTGAGCGCTACAACCACGGAACTGTGGTGGAGTTTTACTGCGATCCT
GGCTACAGCCTCACCAGCGACTACAAGTACATCACCTGCCAGTATGGAGAGTGGTTTCCTTC
TTATCAAGTCTACTGCATCAAATCAGAGCAAACGTGGCCCAGCACCCATGAGACCCTCCTGA
CCACGTGGAAGATTGTGGCGTTCACGGCAACCAGTGTGCTGCTGGTGCTGCTGCTCGTCATC
CTGGCCAGGATGTTCCAGACCAAGTTCAAGGCCCACTTTCCCCCCAGGGGGCCTCCCCGGAG
TTCCAGCAGTGACCCTGACTTTGTGGTGGTAGACGGCGTGCCCGTCATGCTCCCGTCCTATG
ACGAAGCTGTGAGTGGCGGCTTGAGTGCCTTAGGCCCCGGGTACATGGCCTCTGTGGGCCAG
GGCTGCCCCTTACCCGTGGACGACCAGAGCCCCCCAGCATACCCCGGCTCAGGGGACACGGA
CACAGGCCCAGGGGAGTCAGAAACCTGTGACAGCGTCTCAGGCTCTTCTGAGCTGCTCCAAA
GTCTGTATTCACCTCCCAGGTGCCAAGAGAGCACCCACCCTGCTTCGGACAACCCTGACATA
ATTGCCAGCACGGCAGAGGAGGTGGCATCCACCAGCCCAGGCATCCATCATGCCCACTGGGT
GTTGTTCCTAAGAAACTGATTGATTAAAAAATTTCCCAAAGTGTCCTGAAGTGTCTCTTCAA
ATACATGTTGATCTGTGGAGTTGATTCCTTTCCTTCTCTTGGTTTTAGACAAATGTAAACAA
AGCTCTGATCCTTAAAATTGCTATGCTGATAGAGTGGTGAGGGCTGGAAGCTTGATCAAGTC
CTGTTTCTTCTTGACACAGACTGATTAAAAATTAAAAGNAAAAAA

FIGURE 48

><ss.DNA33107

><subunit 1 of 1, 490 aa, 1 stop

><MW: 53920, pI: 5.41, NX(S/T): 4

MYHGMNPSNGDGFLEQQQQQQQPQSPQRLLAVILWFQLALCFGPAQLTGGFDDLQVCADP
GIPENGFRTPSGGVFFEGSVARFHCQDGFKLKGATKRLCLKHFNGTLGWIPSDNSICVQE
DCRIPQIEDAEIHNKTYRHGEKLIITCHEGFKIRYPDLHNMVSLCRDDGTWNNLPICQGC
LRPLASSNGYVNISELQTSFPVGTVISYRCFPGFKLDGSAYLECLQNLIWSSSPPRCLAL
EAQVCPLPPMVSHGDFVCHPRPCERYNHGTVVEFYCDPGYSLTSDYKYITCQYGEWFPSY
QVYCIKSEQTWPSTHETLLTTWKIVAFTATSVLLVLLLVILARMFQTKFKAHFPPRGPPR
SSSSDPDFVVVDGVPVMLPSYDEAVSGGLSALGPGYMASVGQGCPLPVDDQSPPAYPGSG
DTDTGPGESETCDSVSGSSELLQSLYSPPRCQESTHPASDNPDIIASTAEEVASTSPGIH
HAHWVLFLRN

FIGURE 49

CCCACGCGTCCGCTCCGCGCCCTCCCCCCCCGCCTCCCGTGCGGTCCGTCGGTGGCCTAGAGA
TGCTGCTGCCGCGGTTGCAGTTGTCGCGCACGCCTCTGCCCGCCAGCCCGCTCCACCGCCGT
AGCGCCCGAGTGTCGGGGGGCGCACCCGAGTCGGGCCATGAGGCCGGGAACCGCGCTACAGG
CCGTGCTGCTGGCCGTGCTGCTGGTGGGGCTGCGGGCCGCGACGGGTCGCCTGCTGAGTGCC
TCGGATTTGGACCTCAGAGGAGGGCAGCCAGTCTGCCGGGGAGGGACACAGAGGCCTTGTTA
TAAAGTCATTTACTTCCATGATACTTCTCGAAGACTGAACTTTGAGGAAGCCAAAGAAGCCT
GCAGGAGGGATGGAGGCCAGCTAGTCAGCATCGAGTCTGAAGATGAACAGAAACTGATAGAA
AAGTTCATTGAAAACCTCTTGCCATCTGATGGTGACTTCTGGATTGGGCTCAGGAGGCGTGA
GGAGAAACAAAGCAATAGCACAGCCTGCCAGGACCTTTATGCTTGGACTGATGGCAGCATAT
CACAATTTAGGAACTGGTATGTGGATGAGCCGTCCTGCGGCAGCGAGGTCTGCGTGGTCATG
TACCATCAGCCATCGGCACCCGCTGGCATCGGAGGCCCCTACATGTTCCAGTGGAATGATGA
CCGGTGCAACATGAAGAACAATTTCATTTGCAAATATTCTGATGAGAAACCAGCAGTTCCTT
CTAGAGAAGCTGAAGGTGAGGAAACAGAGCTGACAACACCTGTACTTCCAGAAGAAACACAG
GAAGAAGATGCCAAAAAAACATTTAAAGAAAGTAGAGAAGCTGCCTTGAATCTGGCCTACAT
CCTAATCCCCAGCATTCCCCTTCTCCTCCTCCTTGTGGTCACCACAGTTGTATGTTGGGTTT
GGATCTGTAGAAAAAGAAAACGGGAGCAGCCAGACCCTAGCACAAAGAAGCAACACACCATC
TGGCCCTCTCCTCACCAGGGAAACAGCCCGGACCTAGAGGTCTACAATGTCATAAGAAAACA
AAGCGAAGCTGACTTAGCTGAGACCCGGCCAGACCTGAAGAATATTTCATTCCGAGTGTGTT
CGGGAGAAGCCACTCCCGATGACATGTCTTGTGACTATGACAACATGGCTGTGAACCCATCA
GAAAGTGGGTTTGTGACTCTGGTGAGCGTGGAGAGTGGATTTGTGACCAATGACATTTATGA
GTTCTCCCCAGACCAAATGGGGAGGAGTAAGGAGTCTGGATGGGTGGAAAATGAAATATATG
GTTATTAGGACATATAAAAAACTGAAACTGACAACAATGGAAAAGAAATGATAAGCAAAATC
CTCTTATTTTCTATAAGGAAAATACACAGAAGGTCTATGAACAAGCTTAGATCAGGTCCTGT
GGATGAGCATGTGGTCCCCACGACCTCCTGTTGGACCCCCACGTTTTGGCTGTATCCTTTAT
CCCAGCCAGTCATCCAGCTCGACCTTATGAGAAGGTACCTTGCCCAGGTCTGGCACATAGTA
GAGTCTCAATAAATGTCACTTGGTTGGTTGTATCTAACTTTTAAGGGACAGAGCTTTACCTG
GCAGTGATAAAGATGGGCTGTGGAGCTTGGAAAACCACCTCTGTTTTCCTTGCTCTATACAG
CAGCACATATTATCATACAGACAGAAAATCCAGAATCTTTTCAAAGCCCACATATGGTAGCA
CAGGTTGGCCTGTGCATCGGCAATTCTCATATCTGTTTTTTTCAAAGAATAAAATCAAATAA
AGAGCAGGAAAAAAAAA

FIGURE 50

Met Arg Pro Gly Thr Ala Leu Gln Ala Val Leu Leu Ala Val Leu Leu
Val Gly Leu Arg Ala Ala Thr Gly Arg Leu Leu Ser Ala Ser Asp Leu
Asp Leu Arg Gly Gly Gln Pro Val Cys Arg Gly Gly Thr Gln Arg Pro
Cys Tyr Lys Val Ile Tyr Phe His Asp Thr Ser Arg Arg Leu Asn Phe
Glu Glu Ala Lys Glu Ala Cys Arg Arg Asp Gly Gly Gln Leu Val Ser
Ile Glu Ser Glu Asp Glu Gln Lys Leu Ile Glu Lys Phe Ile Glu Asn
Leu Leu Pro Ser Asp Gly Asp Phe Trp Ile Gly Leu Arg Arg Arg Glu
Glu Lys Gln Ser Asn Ser Thr Ala Cys Gln Asp Leu Tyr Ala Trp Thr
Asp Gly Ser Ile Ser Gln Phe Arg Asn Trp Tyr Val Asp Glu Pro Ser
Cys Gly Ser Glu Val Cys Val Val Met Tyr His Gln Pro Ser Ala Pro
Ala Gly Ile Gly Gly Pro Tyr Met Phe Gln Trp Asn Asp Asp Arg Cys
Asn Met Lys Asn Asn Phe Ile Cys Lys Tyr Ser Asp Glu Lys Pro Ala
Val Pro Ser Arg Glu Ala Glu Gly Glu Glu Thr Glu Leu Thr Thr Pro
Val Leu Pro Glu Glu Thr Gln Glu Glu Asp Ala Lys Lys Thr Phe Lys
Glu Ser Arg Glu Ala Ala Leu Asn Leu Ala Tyr Ile Leu Ile Pro Ser
Ile Pro Leu Leu Leu Leu Leu Val Val Thr Thr Val Val Cys Trp Val
Trp Ile Cys Arg Lys Arg Lys Arg Glu Gln Pro Asp Pro Ser Thr Lys
Lys Gln His Thr Ile Trp Pro Ser Pro His Gln Gly Asn Ser Pro Asp
Leu Glu Val Tyr Asn Val Ile Arg Lys Gln Ser Glu Ala Asp Leu Ala
Glu Thr Arg Pro Asp Leu Lys Asn Ile Ser Phe Arg Val Cys Ser Gly
Glu Ala Thr Pro Asp Asp Met Ser Cys Asp Tyr Asp Asn Met Ala Val
Asn Pro Ser Glu Ser Gly Phe Val Thr Leu Val Ser Val Glu Ser Gly
Phe Val Thr Asn Asp Ile Tyr Glu Phe Ser Pro Asp Gln Met Gly Arg
Ser Lys Glu Ser Gly Trp Val Glu Asn Glu Ile Tyr Gly Tyr

FIGURE 51

GGGGTCTCCCTCAGGGCCGGGAGGCACAGCGGTCCCTGCTTGCTGAAGGGCTGGATGTACGC
ATCCGCAGGTTCCCGCGGACTTGGGGGCGCCCGCTGAGCCCCGGCGCCCGCAGAAGACTTGT
GTTTGCCTCCTGCAGCCTCAACCCGGAGGGCAGCGAGGGCCTACCACCATGATCACTGGTGT
GTTCAGCATGCGCTTGTGGACCCCAGTGGGCGTCCTGACCTCGCTGGCGTACTGCCTGCACC
AGCGGCGGGTGGCCCTGGCCGAGCTGCAGGAGGCCGATGGCCAGTGTCCGGTCGACCGCAGC
CTGCTGAAGTTGAAAATGGTGCAGGTCGTGTTTCGACACGGGGCTCGGAGTCCTCTCAAGCC
GCTCCCGCTGGAGGAGCAGGTAGAGTGGAACCCCCAGCTATTAGAGGTCCCACCCCAAACTC
AGTTTGATTACACAGTCACCAATCTAGCTGGTGGTCCGAAACCATATTCTCCTTACGACTCT
CAATACCATGAGACCACCCTGAAGGGGGGCATGTTTGCTGGGCAGCTGACCAAGGTGGGCAT
GCAGCAAATGTTTGCCTTGGGAGAGAGACTGAGGAAGAACTATGTGGAAGACATTCCCTTTC
TTTCACCAACCTTCAACCCACAGGAGGTCTTTATTCGTTCCACTAACATTTTTCGGAATCTG
GAGTCCACCCGTTGTTTGCTGGCTGGGCTTTTCCAGTGTCAGAAAGAAGGACCCATCATCAT
CCACACTGATGAAGCAGATTCAGAAGTCTTGTATCCCAACTACCAAAGCTGCTGGAGCCTGA
GGCAGAGAACCAGAGGCCGGAGGCAGACTGCCTCTTTACAGCCAGGAATCTCAGAGGATTTG
AAAAAGGTGAAGGACAGGATGGGCATTGACAGTAGTGATAAAGTGGACTTCTTCATCCTCCT
GGACAACGTGGCTGCCGAGCAGGCACACAACCTCCCAAGCTGCCCCATGCTGAAGAGATTTG
CACGGATGATCGAACAGAGAGCTGTGGACACATCCTTGTACATACTGCCCAAGGAAGACAGG
GAAAGTCTTCAGATGGCAGTAGGCCCATTCCTCCACATCCTAGAGAGCAACCTGCTGAAAGC
CATGGACTCTGCCACTGCCCCCGACAAGATCAGAAAGCTGTATCTCTATGCGGCTCATGATG
TGACCTTCATACCGCTCTTAATGACCCTGGGGATTTTTGACCACAAATGGCCACCGTTTGCT
GTTGACCTGACCATGGAACTTTACCAGCACCTGGAATCTAAGGAGTGGTTTGTGCAGCTCTA
TTACCACGGGAAGGAGCAGGTGCCGAGAGGTTGCCCTGATGGGCTCTGCCCGCTGGACATGT
TCTTGAATGCCATGTCAGTTTATACCTTAAGCCCAGAAAAATACCATGCACTCTGCTCTCAA
ACTCAGGTGATGGAAGTTGGAAATGAAGAGTAACTGATTTATAAAAGCAGGATGTGTTGATT
TTAAAATAAAGTGCCTTTATACAATG

FIGURE 52

><ss.DNA34434

><subunit 1 of 1, 428 aa, 1 stop

><MW: 48886, pI: 6.39, NX(S/T): 0

MITGVFSMRLWTPVGVLTSLAYCLHQRRVALAELQEADGQCPVDRSLLKLKMVQVVFRHG
ARSPLKPLPLEEQVEWNPQLLEVPPQTQFDYTVTNLAGGPKPYSPYDSQYHETTLKGGMF
AGQLTKVGMQQMFALGERLRKNYVEDIPFLSPTFNPQEVFIRSTNIFRNLESTRCLLAGL
FQCQKEGPIIIHTDEADSEVLYPNYQSCWSLRQRTRGRRQTASLQPGISEDLKKVKDRMG
IDSSDKVDFFILLDNVAAEQAHNLPSCPMLKRFARMIEQRAVDTSLYILPKEDRESLQMA
VGPFLHILESNLLKAMDSATAPDKIRKLYLYAAHDVTFIPLLMTLGIFDHKWPPFAVDLT
MELYQHLESKEWFVQLYYHGKEQVPRGCPDGLCPLDMFLNAMSVYTLSPEKYHALCSQTQ
VMEVGNEE

FIGURE 53

CTCCTCTTAACATACTTGCAGCTAAAACTAAATATTGCTGCTTGGGGACCTCCTTCTAGCCT
TAAATTTCAGCTCATCACCTTCACCTGCCTTGGTCATGGCTCTGCTATTCTCCTTGATCCTT
GCCATTTGCACCAGACCTGGATTCCTAGCGTCTCCATCTGGAGTGCGGCTGGTGGGGGGCCT
CCACCGCTGTGAAGGGCGGGTGGAGGTGGAACAGAAAGGCCAGTGGGGCACCGTGTGTGATG
ACGGCTGGGACATTAAGGACGTGGCTGTGTTGTGCCGGGAGCTGGGCTGTGGAGCTGCCAGC
GGAACCCCTAGTGGTATTTTGTATGAGCCACCAGCAGAAAAAGAGCAAAAGGTCCTCATCCA
ATCAGTCAGTTGCACAGGAACAGAAGATACATTGGCTCAGTGTGAGCAAGAAGAAGTTTATG
ATTGTTCACATGATGAAGATGCTGGGGCATCGTGTGAGAACCCAGAGAGCTCTTTCTCCCCA
GTCCCAGAGGGTGTCAGGCTGGCTGACGGCCCTGGGCATTGCAAGGGACGCGTGGAAGTGAA
GCACCAGAACCAGTGGTATACCGTGTGCCAGACAGGCTGGAGCCTCCGGGCCGCAAAGGTGG
TGTGCCGGCAGCTGGGATGTGGGAGGGCTGTACTGACTCAAAAACGCTGCAACAAGCATGCC
TATGGCCGAAAACCCATCTGGCTGAGCCAGATGTCATGCTCAGGACGAGAAGCAACCCTTCA
GGATTGCCCTTCTGGGCCTTGGGGGAAGAACACCTGCAACCATGATGAAGACACGTGGGTCG
AATGTGAAGATCCCTTTGACTTGAGACTAGTAGGAGGAGACAACCTCTGCTCTGGGCGACTG
GAGGTGCTGCACAAGGGCGTATGGGGCTCTGTCTGTGATGACAACTGGGGAGAAAAGGAGGA
CCAGGTGGTATGCAAGCAACTGGGCTGTGGGAAGTCCCTCTCTCCCTCCTTCAGAGACCGGA
AATGCTATGGCCCTGGGGTTGGCCGCATCTGGCTGGATAATGTTCGTTGCTCAGGGGAGGAG
CAGTCCCTGGAGCAGTGCCAGCACAGATTTTGGGGGTTTCACGACTGCACCCACCAGGAAGA
TGTGGCTGTCATCTGCTCAGTGTAGGTGGGCATCATCTAATCTGTTGAGTGCCTGAATAGAA
GAAAAACACAGAAGAAGGGAGCATTTACTGTCTACATGACTGCATGGGATGAACACTGATCT
TCTTCTGCCCTTGGACTGGGACTTATACTTGGTGCCCCTGATTCTCAGGCCTTCAGAGTTGG
ATCAGAACTTACAACATCAGGTCTAGTTCTCAGGCCATCAGACATAGTTTGGAACTACATCA
CCACCTTTCCTATGTCTCCACATTGCACACAGCAGATTCCCAGCCTCCATAATTGTGTGTAT
CAACTACTTAAATACATTCTCACACACACACACACACACACACACACACACACACACATA
CACCATTTGTCCTGTTTCTCTGAAGAACTCTGACAAAATACAGATTTTGGTACTGAAAGAGA
TTCTAGAGGAACGGAATTTTAAGGATAAATTTTCTGAATTGGTTATGGGGTTTCTGAAATTG
GCTCTATAATCTAATTAGATATAAAATTCTGGTAACTTTATTTACAATAATAAAGATAGCAC
TATGTGTTCAAA

FIGURE 54

><ss.DNA33100

><subunit 1 of 1, 347 aa, 1 stop

><MW: 38130, pI: 5.40, NX(S/T): 0

MALLFSLILAICTRPGFLASPSGVRLVGGLHRCEGRVEVEQKGQWGTVCDDGWDIKDVAV
LCRELGCGAASGTPSGILYEPPAEKEQKVLIQSVSCTGTEDTLAQCEQEEVYDCSHDEDA
GASCENPESSFSPVPEGVRLADGPGHCKGRVEVKHQNQWYTVCQTGWSLRAAKVVCRQLG
CGRAVLTQKRCNKHAYGRKPIWLSQMSCSGREATLQDCPSGPWGKNTCNHDEDTWVECED
PFDLRLVGGDNLCSGRLEVLHKGVWGSVCDDNWGEKEDQVVCKQLGCGKSLSPSFRDRKC
YGPGVGRIWLDNVRCSGEEQSLEQCQHRFWGFHDCTHQEDVAVICSV

FIGURE 55

ACTGCACTCGGTTCTATCGATTGAATTCCCCGGGGATCCTCTAGAGATCCCTCGACCTCGAC
CCACGCGTCCGCGGACGCGTGGGCGGACGCGTGGGCCGGCTACCAGGAAGAGTCTGCCGAAG
GTGAAGGCCATGGACTTCATCACCTCCACAGCCATCCTGCCCCTGCTGTTCGGCTGCCTGGG
CGTCTTCGGCCTCTTCCGGCTGCTGCAGTGGGTGCGCGGGAAGGCCTACCTGCGGAATGCTG
TGGTGGTGATCACAGGCGCCACCTCAGGGCTGGGCAAAGAATGTGCAAAAGTCTTCTATGCT
GCGGGTGCTAAACTGGTGCTCTGTGGCCGGAATGGTGGGGCCCTAGAAGAGCTCATCAGAGA
ACTTACCGCTTCTCATGCCACCAAGGTGCAGACACACAAGCCTTACTTGGTGACCTTCGACC
TCACAGACTCTGGGGCCATAGTTGCAGCAGCAGCTGAGATCCTGCAGTGCTTTGGCTATGTC
GACATACTTGTCAACAATGCTGGGATCAGCTACCGTGGTACCATCATGGACACCACAGTGGA
TGTGGACAAGAGGGTCATGGAGACAAACTACTTTGGCCCAGTTGCTCTAACGAAAGCACTCC
TGCCCTCCATGATCAAGAGGAGGCAAGGCCACATTGTCGCCATCAGCAGCATCCAGGGCAAG
ATGAGCATTCCTTTTCGATCAGCATATGCAGCCTCCAAGCACGCAACCCAGGCTTTCTTTGA
CTGTCTGCGTGCCGAGATGGAACAGTATGAAATTGAGGTGACCGTCATCAGCCCCGGCTACA
TCCACACCAACCTCTCTGTAAATGCCATCACCGCGGATGGATCTAGGTATGGAGTTATGGAC
ACCACCACAGCCCAGGGCCGAAGCCCTGTGGAGGTGGCCCAGGATGTTCTTGCTGCTGTGGG
GAAGAAGAAGAAAGATGTGATCCTGGCTGACTTACTGCCTTCCTTGGCTGTTTATCTTCGAA
CTCTGGCTCCTGGGCTCTTCTTCAGCCTCATGGCCTCCAGGGCCAGAAAAGAGCGGAAATCC
AAGAACTCCTAGTACTCTGACCAGCCAGGGCCAGGGCAGAGAAGCAGCACTCTTAGGCTTGC
TTACTCTACAAGGGACAGTTGCATTTGTTGAGACTTTAATGGAGATTTGTCTCACAAGTGGG
AAAGACTGAAGAAACACATCTCGTGCAGATCTGCTGGCAGAGGACAATCAAAAACGACAACA
AGCTTCTTCCCAGGGTGAGGGGAAACACTTAAGGAATAAATATGGAGCTGGGGTTTAACACT
AAAAACTAGAAATAAACATCTCAAACAGTAAAAAAAAAAAAAAAAGGGCGGCCGCGACTCTAG
AGTCGACCTGCAGAAGCTTGGCCGCCATGGCCCAACTTGTTTATTGCAGCTTATAATGGTTAC

FIGURE 56

><ss.DNA35600

><subunit 1 of 1, 310 aa, 1 stop

><MW: 33524, pI: 9.55, NX(S/T): 1

MDFITSTAILPLLFGCLGVFGLFRLLQWVRGKAYLRNAVVVITGATSGLGKECAKVFYAA
GAKLVLCGRNGGALEELIRELTASHATKVQTHKPYLVTFDLTDSGAIVAAAAEILQCFGY
VDILVNNAGISYRGTIMDTTVDVDKRVMETNYFGPVALTKALLPSMIKRRQGHIVAISSI
QGKMSIPFRSAYAASKHATQAFFDCLRAEMEQYEIEVTVISPGYIHTNLSVNAITADGSR
YGVMDTTTAQGRSPVEVAQDVLAAVGKKKKDVILADLLPSLAVYLRTLAPGLFFSLMASR
ARKERKSKNS

FIGURE 57

CCCACGCGTCCGCTGGTGTTAGATCGAGCAACCCTCTAAAAGCAGTTTAGAGTGGTAAAAAA
AAAAAAAAACACACCAAACGCTCGCAGCCACAAAAGGG
><MET {trans=1-s, dir=f, res=1}>
ATGAAATTTCTTCTGGACATCCTCCTGCTTCTCCCGTTACTGATCGTCTGCTCCCTAGAGTC
CTTCGTGAAGCTTTTTATTCCTAAGAGGAGAAAATCAGTCACCGGCGAAATCGTGCTGATTA
CAGGAGCTGGGCATGGAATTGGGAGACTGACTGCCTATGAATTTGCTAAACTTAAAAGCAAG
CTGGTTCTCTGGGATATAAATAAGCATGGACTGGAGGAAACAGCTGCCAAATGCAAGGGACT
GGGTGCCAAGGTTCATACCTTTGTGGTAGACTGCAGCAACCGAGAAGATATTTACAGCTCTG
CAAAGAAGGTGAAGGCAGAAATTGGAGATGTTAGTATTTTAGTAAATAATGCTGGTGTAGTC
TATACATCAGATTTGTTTGCTACACAAGATCCTCAGATTGAAAAGACTTTTGAAGTTAATGT
ACTTGCACATTTCTGGACTACAAAGGCATTTCTTCCTGCAATGACGAAGAATAACCATGGCC
ATATTGTCACTGTGGCTTCGGCAGCTGGACATGTCTCGGTCCCCTTCTTACTGGCTTACTGT
TCAAGCAAGTTTGCTGCTGTTGGATTTCATAAAACTTTGACAGATGAACTGGCTGCCTTACA
AATAACTGGAGTCAAAACAACATGTCTGTGTCCTAATTTCGTAAACACTGGCTTCATCAAAA
ATCCAAGTACAAGTTTGGGACCCACTCTGGAACCTGAGGAAGTGGTAAACAGGCTGATGCAT
GGGATTCTGACTGAGCAGAAGATGATTTTTATTCCATCTTCTATAGCTTTTTTTAACAACATT
GGAAAGGATCCTTCCTGAGCGTTTCCTGGCAGTTTTAAAACGAAAAATCAGTGTTAAGTTTG
ATGCAGTTATTGGATATAAAATGAAAGCGCAATAAGCACCTAGTTTTCTGAAAACTGATTTA
CCAGGTTTAGGTTGATGTCATCTAATAGTGCCAGAATTTTAATGTTTGAACTTCTGTTTTTT
CTAATTATCCCCATTTCTTCAATATCATTTTTGAGGCTTTGGCAGTCTTCATTTACTACCAC
TTGTTCTTTAGCCAAAAGCTGATTACATATGATATAAACAGAGAAATACCTTTAGAGGTGAC
TTTAAGGAAAATGAAGAAAAAGAACCAAAATGACTTTATTAAAATAATTTCCAAGATTATTT
GTGGCTCACCTGAAGGCTTTGCAAAATTTGTACCATAACCGTTTATTTAACATATATTTTTA
TTTTTGATTGCACTTAAATTTTGTATAATTTGTGTTTCTTTTTCTGTTCTACATAAAATCAG
AAACTTCAAGCTCTCTAAATAAAATGAAGGACTATATCTAGTGGTATTTCACAATGAATATC
ATGAACTCTCAATGGGTAGGTTTCATCCTACCCATTGCCACTCTGTTTCCTGAGAGATACCT
CACATTCCAATGCCAAACATTTCTGCACAGGGAAGCTAGAGGTGGATACACGTGTTGCAAGT
ATAAAAGCATCACTGGGATTTAAGGAGAATTGAGAGAATGTACCCACAAATGGCAGCAATAA
TAAATGGATCACACTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 58

><subunit 1 of 1, 300 aa, 1 stop

><MW: 32964, pI: 9.52

><signal peptide>

MKFLLDILLLLPLLIVCSL

><start mature protein>

ESFVKLFIPKRRKSVTGEIVLITGAGHGIGRLTAYEFAKLKSKLVLWDINKHGLEETAAK

CKGLGAKVHTFVVDCSNREDIYSSAKKVKAEIGDVSILVNNAGVVYTSDLFATQDPQIEK

TFEVNVLAHFWTTKAFLPAMTKNNHGHIVTVASAAGHVSVPFLLA

><putative oxidoreductase active site, by similarity to Y00P_MYCTU and BUDC_KLETE>

YCSSKFAAVGFHKTLTDELAALQITGVKTTCLCPNFVNTGFIKNPSTSLGPTLEPEEVVN

RLMHGILTEQKMIFIPSSIAFLTTLERILPERFLAVLKRKISVKFDAVIGYKMKAQ

FIGURE 59

CCCACGCGTCCGCGGACGCGTGGGTCGACTAGTTCTAGATCGCGAGCGGCCGCCCGCGGCTC
AGGGAGGAGCACCGACTGCGCCGCACCCTGAGAGATGGTTGGTGCCATGTGGAAGGTGATTG
TTTCGCTGGTCCTGTTGATGCCTGGCCCCTGTGATGGGCTGTTTCGCTCCCTATACAGAAGT
GTTTCCATGCCACCTAAGGGAGACTCAGGACAGCCATTATTTCTCACCCCTTACATTGAAGC
TGGGAAGATCCAAAAAGGAAGAGAATTGAGTTTGGTCGGCCCTTTCCCAGGACTGAACATGA
AGAGTTATGCCGGCTTCCTCACCGTGAATAAGACTTACAACAGCAACCTCTTCTTCTGGTTC
TTCCCAGCTCAGATACAGCCAGAAGATGCCCCAGTAGTTCTCTGGCTACAGGGTGGGCCGGG
AGGTTCATCCATGTTTGGACTCTTTGTGGAACATGGGCCTTATGTTGTCACAAGTAACATGA
CCTTGCGTGACAGAGACTTCCCCTGGACCACAACGCTCTCCATGCTTTACATTGACAATCCA
GTGGGCACAGGCTTCAGTTTTACTGATGATACCCACGGATATGCAGTCAATGAGGACGATGT
AGCACGGGATTTATACAGTGCACTAATTCAGTTTTTCCAGATATTTCCTGAATATAAAAATA
ATGACTTTTATGTCACTGGGGAGTCTTATGCAGGGAAATATGTGCCAGCCATTGCACACCTC
ATCCATTCCCTCAACCCTGTGAGAGAGGTGAAGATCAACCTGAACGGAATTGCTATTGGAGA
TGGATATTCTGATCCCGAATCAATTATAGGGGGCTATGCAGAATTCCTGTACCAAATTGGCT
TGTTGGATGAGAAGCAAAAAAAGTACTTCCAGAAGCAGTGCCATGAATGCATAGAACACATC
AGGAAGCAGAACTGGTTTGAGGCCTTTGAAATACTGGATAAACTACTAGATGGCGACTTAAC
AAGTGATCCTTCTTACTTCCAGAATGTTACAGGATGTAGTAATTACTATAACTTTTTGCGGT
GCACGGAACCTGAGGATCAGCTTTACTATGTGAAATTTTTGTCACTCCCAGAGGTGAGACAA
GCCATCCACGTGGGGAATCAGACTTTTAATGATGGAACTATAGTTGAAAAGTACTTGCGAGA
AGATACAGTACAGTCAGTTAAGCCATGGTTAACTGAAATCATGAATAATTATAAGGTTCTGA
TCTACAATGGCCAACTGGACATCATCGTGGCAGCTGCCCTGACAGAGCGCTCCTTGATGGGC
ATGGACTGGAAAGGATCCCAGGAATACAAGAAGGCAGAAAAAAAAAGTTTGGAAGATCTTTAA
ATCTGACAGTGAAGTGGCTGGTTACATCCGGCAAGCGGGTGACTTCCATCAGGTAATTATTC
GAGGTGGAGGACATATTTTACCCTATGACCAGCCTCTGAGAGCTTTTGACATGATTAATCGA
TTCATTTATGGAAAAGGATGGGATCCTTATGTTGGATAAACTACCTTCCCAAAAGAGAACAT
CAGAGGTTTTCATTGCTGAAAAGAAAATCGTAAAAACAGAAAATGTCATAGGAATAAAAAAA
TTATCTTTTCATATCTGCAAGATTTTTTTCATCAATAAAAATTATCCTTGAAACAAGTGAGC
TTTTGTTTTTGGGGGGAGATGTTTACTACAAAATTAACATGAGTACATGAGTAAGAATTACA
TTATTTAACTTAAAGGATGAAAGGTATGGATGATGTGACACTGAGACAAGATGTATAAATGA
AATTTTAGGGTCTTGAATAGGAAGTTTTAATTTCTTCTAAGAGTAAGTGAAAAGTGCAGTTG
TAACAAACAAAGCTGTAACATCTTTTTCTGCCAATAACAGAAGTTTGGCATGCCGTGAAGGT
GTTTGGAAATATTATTGGATAAGAATAGCTCAATTATCCCAAATAAATGGATGAAGCTATAA
TAGTTTTGGGGAAAAGATTCTCAAATGTATAAAGTCTTAGAACAAAAGAATTCTTTGAAATA
AAAATATTATATATAAAAGTAAAAAAAAAA

FIGURE 60

><ss.DNA33206

><subunit 1 of 1, 476 aa, 1 stop

><MW: 54164, pI: 5.52, NX(S/T): 4

MVGAMWKVIVSLVLLMPGPCDGLFRSLYRSVSMPPKGDSGQPLFLTPYIEAGKIQKGREL
SLVGPFPGLNMKSYAGFLTVNKTYNSNLFFWFFPAQIQPEDAPVVLWLQGGPGGSSMFGL
FVEHGPYVVTSNMTLRDRDFPWTTTLSMLYIDNPVGTGFSFTDDTHGYAVNEDDVARDLY
SALIQFFQIFPEYKNNDFYVTGESYAGKYVPAIAHLIHSLNPVREVKINLNGIAIGDGYS
DPESIIGGYAEFLYQIGLLDEKQKKYFQKQCHECIEHIRKQNWFEAFEILDKLLDGDLTS
DPSYFQNVTGCSNYYNFLRCTEPEDQLYYVKFLSLPEVRQAIHVGNQTFNDGTIVEKYLR
EDTVQSVKPWLTEIMNNYKVLIYNGQLDIIVAAALTERSLMGMDWKGSQEYKKAEKKVWK
IFKSDSEVAGYIRQAGDFHQVIIRGGGHILPYDQPLRAFDMINRFIYGKGWDPYVG

FIGURE 61

CGAGGGCTTTTCCGGCTCCGGAATGGCACATGTGGGAATCCCAGTCTTGTTGGCTACAACAT
TTTTCCCTTTCCTAACAAGTTCTAACAGCTGTTCTAACAGCTAGTGATCAGGGGTTCTTCTT
GCTGGAGAAGAAAGGGCTGAGGGCAGAGCAGGGCACTCTCACTCAGGGTGACCAGCTCCTTG
CCTCTCTGTGGATAACAGAGCATGAGAAAGTGAAGAGATGCAGCGGAGTGAGGTGATGGAAG
TCTAAAATAGGAAGGAATTTTGTGTGCAATATCAGACTCTGGGAGCAGTTGACCTGGAGAGC
CTGGGGGAGGGCCTGCCTAACAAGCTTTCAAAAAACAGGAGCGACTTCCACTGGGCTGGGAT
AAGACGTGCCGGTAGGATAGGGAAGACTGGGTTTAGTCCTAATATCAAATTGACTGGCTGGG
TGAACTTCAACAGCCTTTTAACCTCTCTGGGAGATGAAAACGATGGCTTAAGGGGCCAGAAA
TAGAGATGCTTTGTAAAATAAAATTTTAAAAAAAGCAAGTATTTTATAGCATAAAGGCTAGA
GACCAAAATAGATAACAGGATTCCCTGAACATTCCTAAGAGGGAGAAAGTATGTTAAAAATA
GAAAAACCAAAATGCAGAAGGAGGAGACTCACAGAGCTAAACCAGGATGGGGACCCTGGGTC
AGGCCAGCCTCTTTGCTCCTCCCGGAAATTATTTTTGGTCTGACCACTCTGCCTTGTGTTTT
GCAGAATCATGTGAGGGCCAACCGGGGAAGGTGGAGCAGATGAGCACACACAGGAGCCGTCT
CCTCACCGCCGCCCCTCTCAGCATGGAACAGAGGCAGCCCTGGCCCCGGGCCCTGGAGGTGG
ACAGCCGCTCTGTGGTCCTGCTCTCAGTGGTCTGGGTGCTGCTGGCCCCCCCAGCAGCCGGC
ATGCCTCAGTTCAGCACCTTCCACTCTGAGAATCGTGACTGGACCTTCAACCACTTGACCGT
CCACCAAGGGACGGGGGCCGTCTATGTGGGGGCCATCAACCGGGTCTATAAGCTGACAGGCA
ACCTGACCATCCAGGTGGCTCATAAGACAGGGCCAGAAGAGGACAACAAGTCTCGTTACCCG
CCCCTCATCGTGCAGCCCTGCAGCGAAGTGCTCACCCTCACCAACAATGTCAACAAGCTGCT
CATCATTGACTACTCTGAGAACCGCCTGCTGGCCTGTGGGAGCCTCTACCAGGGGGTCTGCA
AGCTGCTGCGGCTGGATGACCTCTTCATCCTGGTGGAGCCATCCCACAAGAAGGAGCACTAC
CTGTCCAGTGTCAACAAGACGGGCACCATGTACGGGGTGATTGTGCGCTCTGAGGGTGAGGA
TGGCAAGCTCTTCATCGGCACGGCTGTGGATGGGAAGCAGGATTACTTCCCGACCCTGTCCA
GCCGGAAGCTGCCCCGAGACCCTGAGTCCTCAGCCATGCTCGACTATGAGCTACACAGCGAT
TTTGTCTCCTCTCTCATCAAGATCCCTTCAGACACCCTGGCCCTGGTCTCCCACTTTGACAT
CTTCTACATCTACGGCTTTGCTAGTGGGGGCTTTGTCTACTTTCTCACTGTCCAGCCCGAGA
CCCCTGAGGGTGTGGCCATCAACTCCGCTGGAGACCTCTTCTACACCTCACGCATCGTGCGG
CTCTGCAAGGATGACCCCAAGTTCCACTCATACGTGTCCCTGCCCTTCGGCTGCACCCGGGC
CGGGGTGGAATACCGCCTCCTGCAGGCTGCTTACCTGGCCAAGCCTGGGGACTCACTGGCCC
AGGCCTTCAATATCACCAGCCAGGACGATGTACTCTTTGCCATCTTCTCCAAAGGGCAGAAG
CAGTATCACCACCCGCCCGATGACTCTGCCCTGTGTGCCTTCCCTATCCGGGCCATCAACTT
GCAGATCAAGGAGCGCCTGCAGTCCTGCTACCAGGGCGAGGGCAACCTGGAGCTCAACTGGC
TGCTGGGGAAGGACGTCCAGTGCACGAAGGCGCCTGTCCCCATCGATGATAACTTCTGTGGA
CTGGACATCAACCAGCCCCTGGGAGGCTCAACTCCAGTGGAGGGCCTGACCCTGTACACCAC
CAGCAGGGACCGCATGACCTCTGTGGCCTCCTACGTTTACAACGGCTACAGCGTGGTTTTTG
TGGGGACTAAGAGTGGCAAGCTGAAAAAGGTAAGAGTCTATGAGTTCAGATGCTCCAATGCC
ATTCACCTCCTCAGCAAAGAGTCCCTCTTGGAAGGTAGCTATTGGTGGAGATTTAACTATAG
GCAACTTTATTTTCTTGGGGAACAAAGGTGAAATGGGGAGGTAAGAAGGGGTTAATTTTGTG
ACTTAGCTTCTAGCTACTTCCTCCAGCCATCAGTCATTGGGTATGTAAGGAATGCAAGCGTA
TTTCAATATTTCCCAAACTTTAAGAAAAAACTTTAAGAAGGTACATCTGCAAAAGCAAA

FIGURE 62

><ss.DNA35558

><subunit 1 of 1, 552 aa, 1 stop

><MW: 61674, pI: 6.95, NX(S/T): 4

MGTLGQASLFAPPGNYFWSDHSALCFAESCEGQPGKVEQMSTHRSRLLTAAPLSMEQRQP
WPRALEVDSRSVVLLSVVWVLLAPPAAGMPQFSTFHSENRDWTFNHLTVHQGTGAVYVGA
INRVYKLTGNLTIQVAHKTGPEEDNKSRYPPLIVQPCSEVLTLTNNVNKLLIIDYSENRL
LACGSLYQGVCKLLRLDDLFILVEPSHKKEHYLSSVNKTGTMYGVIVRSEGEDGKLFIGT
AVDGKQDYFPTLSSRKLPRDPESSAMLDYELHSDFVSSLIKIPSDTLALVSHFDIFYIYG
FASGGFVYFLTVQPETPEGVAINSAGDLFYTSRIVRLCKDDPKFHSYVSLPFGCTRAGVE
YRLLQAAYLAKPGDSLAQAFNITSQDDVLFAIFSKGQKQYHHPPDDSALCAFPIRAINLQ
IKERLQSCYQGEGNLELNWLLGKDVQCTKAPVPIDDNFCGLDINQPLGGSTPVEGLTLYT
TSRDRMTSVASYVYNGYSVVFVGTKSGKLKKVRVYEFRCSNAIHLLSKESLLEGSYWWRF
NYRQLYFLGEQR

FIGURE 63

AGGCTCCCGCGCGCGGCTGAGTGCGGACTGGAGTGGGAACCCGGGTCCCCGCGCTTAGAGAA
CACGCGATGACCACGTGGAGCCTCCGGCGGAGGCCGGCCCGCACGCTGGGACTCCTGCTGCT
GGTCGTCTTGGGCTTCCTGGTGCTCCGCAGGCTGGACTGGAGCACCCTGGTCCCTCTGCGGC
TCCGCCATCGACAGCTGGGGCTGCAGGCCAAGGGCTGGAACTTCATGCTGGAGGATTCCACC
TTCTGGATCTTCGGGGGCTCCATCCACTATTTCCGTGTGCCCAGGGAGTACTGGAGGGACCG
CCTGCTGAAGATGAAGGCCTGTGGCTTGAACACCCTCACCACCTATGTTCCGTGGAACCTGC
ATGAGCCAGAAAGAGGCAAATTTGACTTCTCTGGGAACCTGGACCTGGAGGCCTTCGTCCTG
ATGGCCGCAGAGATCGGGCTGTGGGTGATTCTGCGTCCAGGCCCCTACATCTGCAGTGAGAT
GGACCTCGGGGGCTTGCCCAGCTGGCTACTCCAAGACCCTGGCATGAGGCTGAGGACAACTT
ACAAGGGCTTCACCGAAGCAGTGGACCTTTATTTTGACCACCTGATGTCCAGGGTGGTGCCA
CTCCAGTACAAGCGTGGGGGACCTATCATTGCCGTGCAGGTGGAGAATGAATATGGTTCCTA
TAATAAAGACCCCGCATACATGCCCTACGTCAAGAAGGCACTGGAGGACCGTGGCATTGTGG
AACTGCTCCTGACTTCAGACAACAAGGATGGGCTGAGCAAGGGGATTGTCCAGGGAGTCTTG
GCCACCATCAACTTGCAGTCAACACACGAGCTGCAGCTACTGACCACCTTTCTCTTCAACGT
CCAGGGGACTCAGCCCAAGATGGTGATGGAGTACTGGACGGGGTGGTTTGACTCGTGGGGAG
GCCCTCACAATATCTTGGATTCTTCTGAGGTTTTGAAAACCGTGTCTGCCATTGTGGACGCC
GGCTCCTCCATCAACCTCTACATGTTCCACGGAGGCACCAACTTTGGCTTCATGAATGGAGC
CATGCACTTCCATGACTACAAGTCAGATGTCACCAGCTATGACTATGATGCTGTGCTGACAG
AAGCCGGCGATTACACGGCCAAGTACATGAAGCTTCGAGACTTCTTCGGCTCCATCTCAGGC
ATCCCTCTCCCTCCCCCACCTGACCTTCTTCCCAAGATGCCGTATGAGCCCTTAACGCCAGT
CTTGTACCTGTCTCTGTGGGACGCCCTCAAGTACCTGGGGGAGCCAATCAAGTCTGAAAAGC
CCATCAACATGGAGAACCTGCCAGTCAATGGGGGAAATGGACAGTCCTTCGGGTACATTCTC
TATGAGACCAGCATCACCTCGTCTGGCATCCTCAGTGGCCACGTGCATGATCGGGGGCAGGT
GTTTGTGAACACAGTATCCATAGGATTCTTGGACTACAAGACAACGAAGATTGCTGTCCCCC
TGATCCAGGGTTACACCGTGCTGAGGATCTTGGTGGAGAATCGTGGGCGAGTCAACTATGGG
GAGAATATTGATGACCAGCGCAAAGGCTTAATTGGAAATCTCTATCTGAATGATTCACCCCT
GAAAAACTTCAGAATCTATAGCCTGGATATGAAGAAGAGCTTCTTTCAGAGGTTCGGCCTGG
ACAAATGGNGTTCCCTCCCAGAAACACCCACATTACCTGCTTTCTTCTTGGGTAGCTTGTCC
ATCAGCTCCACGCCTTGTGACACCTTTCTGAAGCTGGAGGGCTGGGAGAAGGGGGTTGTATT
CATCAATGGCCAGAACCTTGGACGTTACTGGAACATTGGACCCCAGAAGACGCTTTACCTCC
CAGGTCCCTGGTTGAGCAGCGGAATCAACCAGGTCATCGTTTTTGAGGAGACGATGGCGGGC
CCTGCATTACAGTTCACGGAAACCCCCCACCTGGGCAGGAACCAGTACATTAAGTGAGCGGT
GGCACCCCCTCCTGCTGGTGCCAGTGGGAGACTGCCGCCTCCTCTTGACCTGAAGCCTGGTG
GCTGCTGCCCCACCCCTCACTGCAAAAGCATCTCCTTAAGTAGCAACCTCAGGGACTGGGGG
CTACAGTCTGCCCCTGTCTCAGCTCAAAACCCTAAGCCTGCAGGGAAAGGTGGGATGGCTCT
GGGCCTGGCTTTGTTGATGATGGCTTTCCTACAGCCCTGCTCTTGTGCCGAGGCTGTCGGGC
TGTCTCTAGGGTGGGAGCAGCTAATCAGATCGCCCAGCCTTTGGCCCTCAGAAAAAGTGCTG
AAACGTGCCCTTGCACCGGACGTCACAGCCCTGCGAGCATCTGCTGGACTCAGGCGTGCTCT
TTGCTGGTTCCTGGGAGGCTTGGCCACATCCCTCATGGCCCCATTTTATCCCCGAAATCCTG
GGTGTGTCACCAGTGTAGAGGGTGGGGAAGGGGTGTCTCACCTGAGCTGACTTTGTTCTTCC
TTCACAACCTTCTGAGCCTTCTTTGGGATTCTGGAAGGAACTCGGCGTGAGAAACATGTGAC
TTCCCCTTTCCCTTCCCACTCGCTGCTTCCCACAGGGTGACAGGCTGGGCTGGAGAAACAGA
AATCCTCACCCTGCGTCTTCCCAAGTTAGCAGGTGTCTCTGGTGTTCAGTGAGGAGGACATG
TGAGTCCTGGCAGAAGCCATGGCCCATGTCTGCACATCCAGGGAGGAGGACAGAAGGCCCAG
CTCACATGTGAGTCCTGGCAGAAGCCATGGCCCATGTCTGCACATCCAGGGAGGAGGACAGA
AGGCCCAGCTCACATGTGAGTCCTGGCAGAAGCCATGGCCCATGTCTGCACATCCAGGGAGG
AGGACAGAAGGCCCAGCTCACATGTGAGTCCTGGCAGAAGCCATGGCCCATGTCTGCACATC
CAGGGAGGAGGACAGAAGGCCCAGCTCAGTGGCCCCCGCTCCCCACCCCCCACGCCCGAACA
GCAGGGGCAGAGCAGCCCTCCTTCGAAGTGTGTCCAAGTCCGCATTTGAGCCTTGTTCTGGG
GCCCAGCCCAACACCTGGCTTGGGCTCACTGTCCTGAGTTGCAGTAAAGCTATAACCTTGAA
TCACAA

FIGURE 64

MTTWSLRRRPARTLGLLLLVVLGFLVLRRLDWSTLVPLRLRHRQLGLQAKGWNFMLEDS
TFWIFGGSIHYFRVPREYWRDRLLKMKACGLNTLTTYVPWNLHEPERGKFDFSGNLDLE
AFVLMAAEIGLWVILRPGPYICSEMDLGGLPSWLLQDPGMRLRTTYKGFTEAVDLYFDH
LMSRVVPLQYKRGGPIIAVQVENEYGSYNKDPAYMPYVKKALEDRGIVELLLTSDNKDG
LSKGIVQGVLATINLQSTHELQLLTTFLFNVQGTQPKMVMEYWTGWFDSWGGPHNILDS
SEVLKTVSAIVDAGSSINLYMFHGGTNFGFMNGAMHFHDYKSDVTSYDYDAVLTEAGDY
TAKYMKLRDFFGSISGIPLPPPPDLLPKMPYEPLTPVLYLSLWDALKYLGEPIKSEKPI
NMENLPVNGGNGQSFGYILYETSITSSGILSGHVHDRGQVFVNTVSIGFLDYKTTKIAV
PLIQGYTVLRILVENRGRVNYGENIDDQRKGLIGNLYLNDSPLKNFRIYSLDMKKSFFQ
RFGLDKWXSLPETPTLPAFFLGSLSISSTPCDTFLKLEGWEKGVVFINGQNLGRYWNIG
PQKTLYLPGPWLSSGINQVIVFEETMAGPALQFTETPHLGRNQYIK

FIGURE 65

GGGGACGCGGAGCTGAGAGGCTCCGGGCTAGCTAGGTGTAGGGGTGGACGGGTCCCAGGACC
CTGGTGAGGGTTCTCTACTTGGCCTTCGGTGGGGGTCAAGACGCAGGCACCTACGCCAAAGG
GGAGCAAAGCCGGGCTCGGCCCGAGGCCCCCAGGACCTCCATCTCCCAATGTTGGAGGAATC
CGACACGTGACGGTCTGTCCGCCGTCTCAGACTAGAGGAGCGCTGTAAACGCCATGGCTCCC
AAGAAGCTGTCCTGCCTTCGTTCCCTGCTGCTGCCGCTCAGCCTGACGCTACTGCTGCCCCA
GGCAGACACTCGGTCGTTCGTAGTGGATAGGGGTCATGACCGGTTTCTCCTAGACGGGGCCC
CGTTCCGCTATGTGTCTGGCAGCCTGCACTACTTTCGGGTACCGCGGGTGCTTTGGGCCGAC
CGGCTTTTGAAGATGCGATGGAGCGGCCTCAACGCCATACAGTTTTATGTGCCCTGGAACTA
CCACGAGCCACAGCCTGGGGTCTATAACTTTAATGGCAGCCGGGACCTCATTGCCTTTCTGA
ATGAGGCAGCTCTAGCGAACCTGTTGGTCATACTGAGACCAGGACCTTACATCTGTGCAGAG
TGGGAGATGGGGGGTCTCCCATCCTGGTTGCTTCGAAAACCTGAAATTCATCTAAGAACCTC
AGATCCAGACTTCCTTGCCGCAGTGGACTCCTGGTTCAAGGTCTTGCTGCCCAAGATATATC
CATGGCTTTATCACAATGGGGGCAACATCATTAGCATTCAGGTGGAGAATGAATATGGTAGC
TACAGAGCCTGTGACTTCAGCTACATGAGGCACTTGGCTGGGCTCTTCCGTGCACTGCTAGG
AGAAAAGATCTTGCTCTTCACCACAGATGGGCCTGAAGGACTCAAGTGTGGCTCCCTCCGGG
GACTCTATACCACTGTAGATTTTGGCCCAGCTGACAACATGACCAAAATCTTTACCCTGCTT
CGGAAGTATGAACCCCATGGGCCATTGGTAAACTCTGAGTACTACACAGGCTGGCTGGATTA
CTGGGGCCAGAATCACTCCACACGGTCTGTGTCAGCTGTAACCAAAGGACTAGAGAACATGC
TCAAGTTGGGAGCCAGTGTGAACATGTACATGTTCCATGGAGGTACCAACTTTGGATATTGG
AATGGTGCCGATAAGAAGGGACGCTTCCTTCCGATTACTACCAGCTATGACTATGATGCACC
TATATCTGAAGCAGGGGACCCCACACCTAAGCTTTTTGCTCTTCGAGATGTCATCAGCAAGT
TCCAGGAAGTTCCTTTGGGACCTTTACCTCCCCCGAGCCCCAAGATGATGCTTGGACCTGTG
ACTCTGCACCTGGTTGGGCATTTACTGGCTTTCCTAGACTTGCTTTGCCCCCGTGGGCCCAT
TCATTCAATCTTGCCAATGACCTTTGAGGCTGTCAAGCAGGACCATGGCTTCATGTTGTACC
GAACCTATATGACCCATACCATTTTTGAGCCAACACCATTCTGGGTGCCAAATAATGGAGTC
CATGACCGTGCCTATGTGATGGTGGATGGGGTGTTCCAGGGTGTTGTGGAGCGAAATATGAG
AGACAAACTATTTTTGACGGGGAAACTGGGGTCCAAACTGGATATCTTGGTGGAGAACATGG
GGAGGCTCAGCTTTGGGTCTAACAGCAGTGACTTCAAGGGCCTGTTGAAGCCACCAATTCTG
GGGCAAACAATCCTTACCCAGTGGATGATGTTCCCTCTGAAAATTGATAACCTTGTGAAGTG
GTGGTTTCCCCTCCAGTTGCCAAAATGGCCATATCCTCAAGCTCCTTCTGGCCCCACATTCT
ACTCCAAAACATTTCCAATTTTAGGCTCAGTTGGGGACACATTTCTATATCTACCTGGATGG
ACCAAGGGCCAAGTCTGGATCAATGGGTTTAACTTGGGCCGGTACTGGACAAAGCAGGGGCC
ACAACAGACCCTCTACGTGCCAAGATTCCTGCTGTTTCCTAGGGGAGCCCTCAACAAAATTA
CATTGCTGGAACTAGAAGATGTACCTCTCCAGCCCCAAGTCCAATTTTTGGATAAGCCTATC
CTCAATAGCACTAGTACTTTGCACAGGACACATATCAATTCCCTTTCAGCTGATACACTGAG
TGCCTCTGAACCAATGGAGTTAAGTGGGCACTGAAAGGTAGGCCGGGCATGGTGGCTCATGC
CTGTAATCCCAGCACTTTGGGAGGCTGAGACGGGTGGATTACCTGAGGTCAGGACTTCAAGA
CCAGCCTGGCCAACATGGTGAAACCCCGTCTCCACTAAAAATACAAAAATTAGCCGGGCGTG
ATGGTGGGCACCTCTAATCCCAGCTACTTGGGAGGCTGAGGGCAGGAGAATTGCTTGAATCC
AGGAGGCAGAGGTTGCAGTGAGTGGAGGTTGTACCACTGCACTCCAGCCTGGCTGACAGTGA
GACACTCCATCTCAAAAAAAAAAAA

FIGURE 66

MRWSGLNAIQFYVPWNYHEPQPGVYNFNGSRDLIAFLNEAALANLLVILRPGPYICAEW
EMGGLPSWLLRKPEIHLRTSDPDFLAAVDSWFKVLLPKIYPWLYHNGGNIISIQVENEY
GSYRACDFSYMRHLAGLFRALLGEKILLFTTDGPEGLKCGSLRGLYTTVDFGPADNMTK
IFTLLRKYEPHGPLVNSEYYTGWLDYWGQNHSTRSVSAVTKGLENMLKLGASVNMYMFH
GGTNFGYWNGADKKGRFLPITTSYDYDAPISEAGDPTPKLFALRDVISKFQEVPLGPLP
PPSPKMMLGPVTLHLVGHLLAFLDLLCPRGPIHSILPMTFEAVKQDHGFMLYRTYMTHT
IFEPTPFWVPNNGVHDRAYVMVDGVFQGVVERNMRDKLFLTGKLGSKLDILVENMGRLS
FGSNSSDFKGLLKPPILGQTILTQWMMFPLKIDNLVKWWFPLQLPKWPYPQAPSGPTFY
SKTFPILGSVGDTFLYLPGWTKGQVWINGFNLGRYWTKQGPQQTLYVPRFLLFPRGALN
KITLLELEDVPLQPQVQFLDKPILNSTSTLHRTHINSLSADTLSASEPMELSGH

FIGURE 67

GCTTTGAACACGTCTGCAAGCCCAAAGTTGAGCATCTGATTGGTTATGAGGTATTTGAGTGC
ACCCACAATATGGCTTACATGTTGAAAAAGCTTCTCATCAGTTACATATCCATTATTTGTGT
TTATGGCTTTATCTGCCTCTACACTCTCTTCTGGTTATTCAGGATACCTTTGAAGGAATATT
CTTTCGAAAAAGTCAGAGAAGAGAGCAGTTTTAGTGACATTCCAGATGTCAAAAACGATTTT
GCGTTCCTTCTTCACATGGTAGACCAGTATGACCAGCTATATTCCAAGCGTTTTGGTGTGTT
CTTGTCAGAAGTTAGTGAAAATAAACTTAGGGAAATTAGTTTGAACCATGAGTGGACATTTG
AAAAACTCAGGCAGCACATTTCACGCAACGCCCAGGACAAGCAGGAGTTGCATCTGTTCATG
CTGTCGGGGGTGCCCGATGCTGTCTTTGACCTCACAGACCTGGATGTGCTAAAGCTTGAACT
AATTCCAGAAGCTAAAATTCCTGCTAAGATTTCTCAAATGACTAACCTCCAAGAGCTCCACC
TCTGCCACTGCCCTGCAAAAGTTGAACAGACTGCTTTTAGCTTTCTTCGCGATCACTTGAGA
TGCCTTCACGTGAAGTTCACTGATGTGGCTGAAATTCCTGCCTGGGTGTATTTGCTCAAAAA
CCTTCGAGAGTTGTACTTAATAGGCAATTTGAACTCTGAAAACAATAAGATGATAGGACTTG
AATCTCTCCGAGAGTTGCGGCACCTTAAGATTCTCCACGTGAAGAGCAATTTGACCAAAGTT
CCCTCCAACATTACAGATGTGGCTCCACATCTTACAAAGTTAGTCATTCATAATGACGGCAC
TAAACTCTTGGTACTGAACAGCCTTAAGAAAATGATGAATGTCGCTGAGCTGGAACTCCAGA
ACTGTGAGCTAGAGAGAATCCCACATGCTATTTTCAGCCTCTCTAATTTACAGGAACTGGAT
TTAAAGTCCAATAACATTCGCACAATTGAGGAAATCATCAGTTTCCAGCATTTAAAACGACT
GACTTGTTTAAAATTATGGCATAACAAAATTGTTACTATTCCTCCCTCTATTACCCATGTCA
AAAACTTGGAGTCACTTTATTTCTCTAACAACAAGCTCGAATCCTTACCAGTGGCAGTATTT
AGTTTACAGAAACTCAGATGCTTAGATGTGAGCTACAACAACATTTCAATGATTCCAATAGA
AATAGGATTGCTTCAGAACCTGCAGCATTTGCATATCACTGGGAACAAAGTGGACATTCTGC
CAAAACAATTGTTTAAATGCATAAAGTTGAGGACTTTGAATCTGGGACAGAACTGCATCACC
TCACTCCCAGAGAAAGTTGGTCAGCTCTCCCAGCTCACTCAGCTGGAGCTGAAGGGGAACTG
CTTGGACCGCCTGCCAGCCCAGCTGGGCCAGTGTCGGATGCTCAAGAAAAGCGGGCTTGTTG
TGGAAGATCACCTTTTTGATACCCTGCCACTCGAAGTCAAAGAGGCATTGAATCAAGACATA
AATATTCCCTTTGCAAATGGGATTTAAACTAAGATAATATATGCACAGTGATGTGCAGGAAC
AACTTCCTAGATTGCAAGTGCTCACGTACAAGTTATTACAAGATAATGCATTTTAGGAGTAG
ATACATCTTTTAAAATAAAACAGAGAGGATGCATAGAAGGCTGATAGAAGACATAACTGAAT
GTTCAATGTTTGTAGGGTTTTAAGTCATTCATTTCCAAATCATTTTTTTTTTTCTTTTGGGG
AAAGGGAAGGAAAAATTATAATCACTAATCTTGGTTCTTTTTAAATTGTTTGTAACTTGGAT
GCTGCCGCTACTGAATGTTTACAAATTGCTTGCCTGCTAAAGTAAATGATTAAATTGACATT
TTCTTACTAAAAAAAAAAAAAAAAA

FIGURE 68

><ss.DNA34407

><subunit 1 of 1, 501 aa, 1 stop

><MW: 57819, pI: 8.15, NX(S/T): 3

MAYMLKKLLISYISIICVYGFICLYTLFWLFRIPLKEYSFEKVREESSFSDIPDVKNDFA
FLLHMVDQYDQLYSKRFGVFLSEVSENKLREISLNHEWTFEKLRQHISRNAQDKQELHLF
MLSGVPDAVFDLTDLDVLKLELIPEAKIPAKISQMTNLQELHLCHCPAKVEQTAFSFLRD
HLRCLHVKFTDVAEIPAWVYLLKNLRELYLIGNLNSENNKMIGLESLRELRHLKILHVKS
NLTKVPSNITDVAPHLTKLVIHNDGTKLLVLNSLKKMMNVAELELQNCELERIPHAIFSL
SNLQELDLKSNNIRTIEEIISFQHLKRLTCLKLWHNKIVTIPPSITHVKNLESLYFSNNK
LESLPVAVFSLQKLRCLDVSYNNISMIPIEIGLLQNLQHLHITGNKVDILPKQLFKCIKL
RTLNLGQNCITSLPEKVGQLSQLTQLELKGNCLDRLPAQLGQCRMLKKSGLVVEDHLFDT
LPLEVKEALNQDINIPFANGI

FIGURE 69

CCCACGCGTCCGGCCTTCTCTCTGGACTTTGCATTTCCATTCCTTTTCATTGACAAACTGAC
TTTTTTTATTTCTTTTTTTTCCATCTCTGGGCCAGCTTGGGATCCTAGGCCGCCCTGGGAAGA
CATTTGTGTTTTACACACATAAGGATCTGTGTTTGGGGTTTCTTCTTCCTCCCCTGACATTG
GCATTGCTTAGTGGTTGTGTGGGGAGGGAGACCACGTGGGCTCAGTGCTTGCTTGCACTTAT
CTGCCTAGGTACATCGAAGTCTTTTGACCTCCATACAGTGATTATGCCTGTCATCGCTGGTG
GTATCCTGGCGGCCTTGCTCCTGCTGATAGTTGTCGTGCTCTGTCTTTACTTCAAAATACAC
AACGCGCTAAAAGCTGCAAAGGAACCTGAAGCTGTGGCTGTAAAAAATCACAACCCAGACAA
GGTGTGGTGGGCCAAGAACAGCCAGGCCAAAACCATTGCCACGGAGTCTTGTCCTGCCCTGC
AGTGCTGTGAAGGATATAGAATGTGTGCCAGTTTTGATTCCCTGCCACCTTGCTGTTGCGAC
ATAAATGAGGGCCTCTGAGTTAGGAAAGGCTCCCTTCTCAAAGCAGAGCCCTGAAGACTTCA
ATGATGTCAATGAGGCCACCTGTTTGTGATGTGCAGGCACAGAAGAAAGGCACAGCTCCCCA
TCAGTTTCATGGAAAATAACTCAGTGCCTGCTGGGAACCAGCTGCTGGAGATCCCTACAGAG
AGCTTCCACTGGGGGCAACCCTTCCAGGAAGGAGTTGGGGAGAGAGAACCCTCACTGTGGGG
AATGCTGATAAACCAGTCACACAGCTGCTCTATTCTCACACAAATCTACCCCTTGCGTGGCT
GGAACTGACGTTTCCCTGGAGGTGTCCAGAAAGCTGATGTAACACAGAGCCTATAAAAGCTG
TCGGTCCTTAAGGCTGCCCAGCGCCTTGCCAAAATGGAGCTTGTAAGAAGGCTCATGCCATT
GACCCTCTTAATTCTCTCCTGTTTGGCGGAGCTGACAATGGCGGAGGCTGAAGGCAATGCAA
GCTGCACAGTCAGTCTAGGGGGTGCCAATATGGCAGAGACCCACAAAGCCATGATCCTGCAA
CTCAATCCCAGTGAGAACTGCACCTGGACAATAGAAAGACCAGAAAACAAAAGCATCAGAAT
TATCTTTTCCTATGTCCAGCTTGATCCAGATGGAAGCTGTGAAAGTGAAAACATTAAAGTCT
TTGACGGAACCTCCAGCAATGGGCCTCTGCTAGGGCAAGTCTGCAGTAAAAACGACTATGTT
CCTGTATTTGAATCATCATCCAGTACATTGACGTTTCAAATAGTTACTGACTCAGCAAGAAT
TCAAAGAACTGTCTTTGTCTTCTACTACTTCTTCTCTCCTAACATCTCTATTCCAAACTGTG
GCGGTTACCTGGATACCTTGGAAGGATCCTTCACCAGCCCCAATTACCCAAAGCCGCATCCT
GAGCTGGCTTATTGTGTGTGGCACATACAAGTGGAGAAAGATTACAAGATAAAACTAAACTT
CAAAGAGATTTTCCTAGAAATAGACAAACAGTGCAAATTTGATTTTCTTGCCATCTATGATG
GCCCCTCCACCAACTCTGGCCTGATTGGACAAGTCTGTGGCCGTGTGACTCCCACCTTCGAA
TCGTCATCAAACTCTCTGACTGTCGTGTTGTCTACAGATTATGCCAATTCTTACCGGGGATT
TTCTGCTTCCTACACCTCAATTTATGCAGAAAACATCAACACTACATCTTTAACTTGCTCTT
CTGACAGGATGAGAGTTATTATAAGCAAATCCTACCTAGAGGCTTTTAACTCTAATGGGAAT
AACTTGCAACTAAAAGACCCAACTTGCAGACCAAAATTATCAAATGTTGTGGAATTTTCTGT
CCCTCTTAATGGATGTGGTACAATCAGAAAGGTAGAAGATCAGTCAATTACTTACACCAATA
TAATCACCTTTTCTGCATCCTCAACTTCTGAAGTGATCACCCGTCAGAAACAACTCCAGATT
ATTGTGAAGTGTGAAATGGGACATAATTCTACAGTGGAGATAATATACATAACAGAAGATGA
TGTAATACAAAGTCAAAATGCACTGGGCAAATATAACACCAGCATGGCTCTTTTTGAATCCA
ATTCATTTGAAAAGACTATACTTGAATCACCATATTATGTGGATTTGAACCAAACTCTTTTT
GTTCAAGTTAGTCTGCACACCTCAGATCCAAATTTGGTGGTGTTTCTTGATACCTGTAGAGC
CTCTCCCACCTCTGACTTTGCATCTCCAACCTACGACCTAATCAAGAGTGGATGTAGTCGAG
ATGAAACTTGTAAGGTGTATCCCTTATTTGGACACTATGGGAGATTCCAGTTTAATGCCTTT
AAATTCTTGAGAAGTATGAGCTCTGTGTATCTGCAGTGTAAAGTTTTGATATGTGATAGCAG
TGACCACCAGTCTCGCTGCAATCAAGGTTGTGTCTCCAGAAGCAAACGAGACATTTCTTCAT
ATAAATGGAAAACAGATTCCATCATAGGACCCATTCGTCTGAAAAGGGATCGAAGTGCAAGT
GGCAATTCAGGATTTCAGCATGAAACACATGCGGAAGAAACTCCAAACCAGCCTTTCAACAG
TGTGCATCTGTTTTCCTTCATGGTTCTAGCTCTGAATGTGGTGACTGTAGCGACAATCACAG
TGAGGCATTTTGTAAATCAACGGGCAGACTACAAATACCAGAAGCTGCAGAACTATTAACTA
ACAGGTCCAACCCTAAGTGAGACATGTTTCTCCAGGATGCCAAAGGAAATGCTACCTCGTGG
CTACACATATTATGAATAAATGAGGAAGGGCCTGAAAGTGACACACAGGCCTGCATGTAAAAAAA

FIGURE 70

><ss.DNA35841

><subunit 1 of 1, 607 aa, 1 stop

><MW: 68153, pI: 6.39, NX(S/T): 9

MELVRRLMPLTLLILSCLAELTMAEAEGNASCTVSLGGANMAETHKAMILQLNPSENCTW
TIERPENKSIRIIFSYVQLDPDGSCESENIKVFDGTSSNGPLLGQVCSKNDYVPVFESSS
STLTFQIVTDSARIQRTVFVFYYFFSPNISIPNCGGYLDTLEGSFTSPNYPKPHPELAYC
VWHIQVEKDYKIKLNFKEIFLEIDKQCKFDFLAIYDGPSTNSGLIGQVCGRVTPTFESSS
NSLTVVLSTDYANSYRGFSASYTSIYAENINTTSLTCSSDRMRVIISKSYLEAFNSNGNN
LQLKDPTCRPKLSNVVEFSVPLNGCGTIRKVEDQSITYTNIITFSASSTSEVITRQKQLQ
IIVKCEMGHNSTVEIIYITEDDVIQSQNALGKYNTSMALFESNSFEKTILESPYYVDLNQ
TLFVQVSLHTSDPNLVVFLDTCRASPTSDFASPTYDLIKSGCSRDETCKVYPLFGHYGRF
QFNAFKFLRSMSSVYLQCKVLICDSSDHQSRCNQGCVSRSKRDISSYKWKTDSIIGPIRL
KRDRSASGNSGFQHETHAEETPNQPFNSVHLFSFMVLALNVVTVATITVRHFVNQRADYK
YQKLQNY

FIGURE 71

GACGGAAGAACAGCGCTCCCGAGGCCGCGGGAGCCTGCAGAGAGGACAGCCGGCCTGCGCCG
GGACATGCGGCCCCAGGAGCTCCCCAGGCTCGCGTTCCCGTTGCTGCTGTTGCTGTTGCTGC
TGCTGCCGCCGCCGCCGTGCCCTGCCCACAGCGCCACGCGCTTCGACCCCACCTGGGAGTCC
CTGGACGCCCGCCAGCTGCCCGCGTGGTTTGACCAGGCCAAGTTCGGCATCTTCATCCACTG
GGGAGTGTTTTCCGTGCCCAGCTTCGGTAGCGAGTGGTTCTGGTGGTATTGGCAAAAGGAAA
AGATACCGAAGTATGTGGAATTTATGAAAGATAATTACCCTCCTAGTTTCAAATATGAAGAT
TTTGGACCACTATTTACAGCAAAATTTTTTAATGCCAACCAGTGGGCAGATATTTTTCAGGC
CTCTGGTGCCAAATACATTGTCTTAACTTCCAAACATCATGAAGGCTTTACCTTGTGGGGGT
CAGAATATTCGTGGAACTGGAATGCCATAGATGAGGGGCCCAAGAGGGACATTGTCAAGGAA
CTTGAGGTAGCCATTAGGAACAGAACTGACCTGCGTTTTGGACTGTACTATTCCCTTTTTGA
ATGGTTTCATCCGCTCTTCCTTGAGGATGAATCCAGTTCATTCCATAAGCGGCAATTTCCAG
TTTCTAAGACATTGCCAGAGCTCTATGAGTTAGTGAACAACTATCAGCCTGAGGTTCTGTGG
TCGGATGGTGACGGAGGAGCACCGGATCAATACTGGAACAGCACAGGCTTCTTGGCCTGGTT
ATATAATGAAAGCCCAGTTCGGGGCACAGTAGTCACCAATGATCGTTGGGGAGCTGGTAGCA
TCTGTAAGCATGGTGGCTTCTATACCTGCAGTGATCGTTATAACCCAGGACATCTTTTGCCA
CATAAATGGGAAAACTGCATGACAATAGACAAACTGTCCTGGGGCTATAGGAGGGAAGCTGG
AATCTCTGACTATCTTACAATTGAAGAATTGGTGAAGCAACTTGTAGAGACAGTTTCATGTG
GAGGAAATCTTTTGATGAATATTGGGCCCACACTAGATGGCACCATTTCTGTAGTTTTTGAG
GAGCGACTGAGGCAAGTGGGGTCCTGGCTAAAAGTCAATGGAGAAGCTATTTATGAAACCTA
TACCTGGCGATCCCAGAATGACACTGTCACCCCAGATGTGTGGTACACATCCAAGCCTAAAG
AAAAATTAGTCTATGCCATTTTTCTTAAATGGCCCACATCAGGACAGCTGTTCCTTGGCCAT
CCCAAAGCTATTCTGGGGGCAACAGAGGTGAAACTACTGGGCCATGGACAGCCACTTAACTG
GATTTCTTTGGAGCAAAATGGCATTATGGTAGAACTGCCACAGCTAACCATTCATCAGATGC
CGTGTAAATGGGGCTGGGCTCTAGCCCTAACTAATGTGATCTAAAGTGCAGCAGAGTGGCTG
ATGCTGCAAGTTATGTCTAAGGCTAGGAACTATCAGGTGTCTATAATTGTAGCACATGGAGA
AAGCAATGTAAACTGGATAAGAAAATTATTTGGCAGTTCAGCCCTTTCCCTTTTTCCCACTA
AATTTTTCTTAAATTACCCATGTAACCATTTTAACTCTCCAGTGCACTTTGCCATTAAAGTC
TCTTCACATTGATTTGTTTCCATGTGTGACTCAGAGGTGAGAATTTTTTCACATTATAGTAG
CAAGGAATTGGTGGTATTATGGACCGAACTGAAAATTTTATGTTGAAGCCATATCCCCCATG
ATTATATAGTTATGCATCACTTAATATGGGGATATTTTCTGGGAAATGCATTGCTAGTCAAT
TTTTTTTTGTGCCAACATCATAGAGTGTATTTACAAAATCCTAGATGGCATAGCCTACTACA
CACCTAATGTGTATGGTATAGACTGTTGCTCCTAGGCTACAGACATATACAGCATGTTACTG
AATACTGTAGGCAATAGTAACAGTGGTATTTGTATATCGAAACATATGGAAACATAGAGAAG
GTACAGTAAAAATACTGTAAAATAAATGGTGCACCTGTATAGGGCACTTACCACGAATGGAG
CTTACAGGACTGGAAGTTGCTCTGGGTGAGTCAGTGAGTGAATGTGAAGGCCTAGGACATTA
TTGAACACTGCCAGACGTTATAAATACTGTATGCTTAGGCTACACTACATTTATAAAAAAAA
GTTTTTCTTTCTTCAATTATAAATTAACATAAGTGTACTGTAACTTTACAAACGTTTTAATT
TTTAAAACCTTTTTGGCTCTTTTGTAATAACACTTAGCTTAAAACATAAACTCATTGTGCAA
ATGTAA

FIGURE 72

MRPQELPRLAFPLLLLLLLLLPPPPCPAHSATRFDPTWESLDARQLPAWFDQAKFGIFI
HWGVFSVPSFGSEWFWWYWQKEKIPKYVEFMKDNYPPSFKYEDFGPLFTAKFFNANQWA
DIFQASGAKYIVLTSKHHEGFTLWGSEYSWNWNAIDEGPKRDIVKELEVAIRNRTDLRF
GLYYSLFEWFHPLFLEDESSSFHKRQFPVSKTLPELYELVNNYQPEVLWSDGDGGAPDQ
YWNSTGFLAWLYNESPVRGTVVTNDRWGAGSICKHGGFYTCSDRYNPGHLLPHKWENCM
TIDKLSWGYRREAGISDYLTIEELVKQLVETVSCGGNLLMNIGPTLDGTISVVFEERLR
QVGSWLKVNGEAIYETYTWRSQNDTVTPDVWYTSKPKEKLVYAIFLKWPTSGQLFLGHP
KAILGATEVKLLGHGQPLNWISLEQNGIMVELPQLTIHQMPCKWGWALALTNVI

FIGURE 73

AGCAGGGAAATCCGGATGTCTCGGTTATGAAGTGGAGCAGTGAGTGTGAGCCTCAACATAGT
TCCAGAACTCTCCATCCGGACTAGTTATTGAGCATCTGCCTCTCATATCACCAGTGGCCATC
TGAGGTGTTTCCCTGGCTCTGAAGGGGTAGGCACGATGGCCAGGTGCTTCAGCCTGGTGTTG
CTTCTCACTTCCATCTGGACCACGAGGCTCCTGGTCCAAGGCTCTTTGCGTGCAGAAGAGCT
TTCCATCCAGGTGTCATGCAGAATTATGGGGATCACCCTTGTGAGCAAAAAGGCGAACCAGC
AGCTGAATTTCACAGAAGCTAAGGAGGCCTGTAGGCTGCTGGGACTAAGTTTGGCCGGCAAG
GACCAAGTTGAAACAGCCTTGAAAGCTAGCTTTGAAACTTGCAGCTATGGCTGGGTTGGAGA
TGGATTCGTGGTCATCTCTAGGATTAGCCCAAACCCCAAGTGTGGGAAAAATGGGGTGGGTG
TCCTGATTTGGAAGGTTCCAGTGAGCCGACAGTTTGCAGCCTATTGTTACAACTCATCTGAT
ACTTGGACTAACTCGTGCATTCCAGAAATTATCACCACCAAAGATCCCATATTCAACACTCA
AACTGCAACACAAACAACAGAATTTATTGTCAGTGACAGTACCTACTCGGTGGCATCCCCTT
ACTCTACAATACCTGCCCCTACTACTACTCCTCCTGCTCCAGCTTCCACTTCTATTCCACGG
AGAAAAAAATTGATTTGTGTCACAGAAGTTTTTATGGAAACTAGCACCATGTCTACAGAAAC
TGAACCATTTGTTGAAAATAAAGCAGCATTCAAGAATGAAGCTGCTGGGTTTGGAGGTGTCC
CCACGGCTCTGCTAGTGCTTGCTCTCCTCTTCTTTGGTGCTGCAGCTGGTCTTGGATTTTGC
TATGTCAAAAGGTATGTGAAGGCCTTCCCTTTTACAAACAAGAATCAGCAGAAGGAAATGAT
CGAAACCAAAGTAGTAAAGGAGGAGAAGGCCAATGATAGCAACCCTAATGAGGAATCAAAGA
AAACTGATAAAAACCCAGAAGAGTCCAAGAGTCCAAGCAAAACTACCGTGCGATGCCTGGAA
GCTGAAGTTTAGATGAGACAGAAATGAGGAGACACACCTGAGGCTGGTTTCTTTCATGCTCC
TTACCCTGCCCCAGCTGGGGAAATCAAAAGGGCCAAAGAACCAAAGAAGAAAGTCCACCCTT
GGTTCCTAACTGGAATCAGCTCAGGACTGCCATTGGACTATGGAGTGCACCAAAGAGAATGC
CCTTCTCCTTATTGTAACCCTGTCTGGATCCTATCCTCCTACCTCCAAAGCTTCCCACGGCC
TTTCTAGCCTGGCTATGTCCTAATAATATCCCACTGGGAGAAAGGAGTTTTGCAAAGTGCAA
GGACCTAAAACATCTCATCAGTATCCAGTGGTAAAAAGGCCTCCTGGCTGTCTGAGGCTAGG
TGGGTTGAAAGCCAAGGAGTCACTGAGACCAAGGCTTTCTCTACTGATTCCGCAGCTCAGAC
CCTTTCTTCAGCTCTGAAAGAGAAACACGTATCCCACCTGACATGTCCTTCTGAGCCCGGTA
AGAGCAAAAGAATGGCAGAAAAGTTTAGCCCCTGAAAGCCATGGAGATTCTCATAACTTGAG
ACCTAATCTCTGTAAAGCTAAAATAAAGAAATAGAACAAGGCTGAGGATACGACAGTACACT
GTCAGCAGGGACTGTAAACACAGACAGGGTCAAAGTGTTTTCTCTGAACACATTGAGTTGGA
ATCACTGTTTAGAACACACACACTTACTTTTTCTGGTCTCTACCACTGCTGATATTTTCTCT
AGGAAATATACTTTTACAAGTAACAAAAATAAAAACTCTTATAAATTTCTATTTTTATCTGA
GTTACAGAAATGATTACTAAGGAAGATTACTCAGTAATTTGTTTAAAAAGTAATAAAATTCA
ACAAACATTTGCTGAATAGCTACTATATGTCAAGTGCTGTGCAAGGTATTACACTCTGTAAT
TGAATATTATTCCTCAAAAAATTGCACATAGTAGAACGCTATCTGGGAAGCTATTTTTTTCA
GTTTTGATATTTCTAGCTTATCTACTTCCAAACTAATTTTTATTTTTGCTGAGACTAATCTT
ATTCATTTTCTCTAATATGGCAACCATTATAACCTTAATTTATTATTAACATACCTAAGAAG
TACATTGTTACCTCTATATACCAAAGCACATTTTAAAAGTGCCATTAACAAATGTATCACTA
GCCCTCCTTTTTTCCAACAAGAAGGGACTGAGAGATGCAGAAATATTTGTGACAAAAAATTAA
AGCATTTAGAAAACTT

FIGURE 74

><ss.DNA34431

><subunit 1 of 1, 322 aa, 1 stop

><MW: 35213, pI: 8.71, NX(S/T): 3

MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEA
CRLLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPV
SRQFAAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPA
PTTTPPAPASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTAL
LVLALLFFGAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKT
DKNPEESKSPSKTTVRCLEAEV

FIGURE 75

AG

><MET {trans=1-s, dir=f, res=1}>

ATGGCGGTCTTGGCACCTCTAATTGCTCTCGTGTATTCGGTGCCGCGACTTTCACGATGG

CTCGCCCAACCTTACTACCTTCTGTCGGCCCTGCTCTCTGCTGCCTTCCTACTCGTGAGG

AAACTGCCGCCGCTCTGCCACGGTCTGCCCACCCAACGCGAAGACGGTAACCCGTGTGAC

TTTGACTGGAGAGAAGTGGAGATCCTGATGTTTCTCAGTGCCATTGTGATGATGAAGAAC

CGCAGATCCATCACTGTGGAGCAACATATAGGCAACATTTTCATGTTTAGTAAAGTGGCC

AACACAATTCTTTTCTTCCGCTTGGATATTCGCATGGGCCTACTTTACATCACACTCTGC

ATAGTGTTCCTGATGACGTGCAAACCCCCCCTATATATGGGCCCTGAGTATATCAAGTAC

TTCAATGATAAAACCATTGATGAGGAACTAGAACGGGACAAGAGGGTCACTTGGATTGTG

GAGTTCTTTGCCAATTGGTCTAATGACTGCCAATCATTTGCCCCTATCTATGCTGACCTC

TCCCTTAAATACAACTGTACAGGGCTAAATTTTGGGAAGGTGGATGTTGGACGCTATACT

GATGTTAGTACGCGGTACAAAGTGAGCACATCACCCCTCACCAAGCAACTCCCTACCCTG

ATCCTGTTCCAAGGTGGCAAGGAGGCAATGCGGCGGCCACAGATTGACAAGAAAGGACGG

GCTGTCTCATGGACCTTCTCTGAGGAGAATGTGATCCGAGAATTTAACTTAAATGAGCTA

TACCAGCGGGCCAAGAAACTATCAAAGGCTGGAGACAATATCCCTGAGGAGCAGCCTGTG

GCTTCAACCCCCACCACAGTGTCAGATGGGGAAAACAAGAAGGATAAATAAGATCCTCAC

TTTGGCAGTGCTTCCTCTCCTGTCAATTCCAGGCTCTTTCCATAACCACAAGCCTGAGGC

TGCAGCCTTTNATTNATGTTTTCCCTTTGGCTGNGACTGGNTGGGGCAGCATGCAGCTTC

TGATTTTAAAGAGGCATCTAGGGAATTGTCAGGCACCCTACAGGAAGGCCTGCCATGCTG

TGGCCAACTGTTTCACTGGAGCAAGAAAGAGATCTCATAGGACGGAGGGGGAAATGGTTT

CCCTCCAAGCTTGGGTCAGTGTGTTAACTGCTTATCAGCTATTCAGACATCTCCATGGTT

TCTCCATGAAACTCTGTGGTTTCATCATTCCTTCTTAGTTGACCTGCACAGCTTGGTTAG

ACCTAGATTTAACCCTAAGGTAAGATGCTGGGGTATAGAACGCTAAGAATTTTCCCCCAA

GGACTCTTGCTTCCTTAAGCCCTTCTGGCTTCGTTTATGGTCTTCATTAAAAGTATAAGC

CTAACTTTGTCGCTAGTCCTAAGGAGAAACCTTTAACCACAAAGTTTTTATCATTGAAGA

CAATATTGAACAACCCCCTATTTTGTGGGGATTGAGAAGGGGTGAATAGAGGCTTGAGAC

TTTCCTTTGTGTGGTAGGACTTGGAGGAGAAATCCCCTGGACTTTCACTAACCCTCTGAC

ATACTCCCCACACCCAGTTGATGGCTTTCCGTAATAAAAAGATTGGGATTTCCTTTTG

FIGURE 76

MAVLAPLIALVYSVPRLSRWLAQPYYLLSALLSAAFLLVRKLPPLCHGLPTQREDGNPCD
FDWREVEILMFLSAIVMMKNRRSITVEQHIGNIFMFSKVANTILFFRLDIRMGLLYITLC
IVFLMTCKPPLYMGPEYIKYFNDKTIDEELERDKRVTWIVEFFANWSNDCQSFAPIYADL
SLKYNCTGLNFGKVDVGRYTDVSTRYKVSTSPLTKQLPTLILFQGGKEAMRRPQIDKKGR
AVSWTFSEENVIREFNLNELYQRAKKLSKAGDNIPEEQPVASTPTTVSDGENKKDK

FIGURE 77

GGACAGCTCGCGGCCCCCGAGAGCTCTAGCCGTCGAGGAGCTGCCTGGGGACGTTTGCCCTG
GGGCCCCAGCCTGGCCCGGGTCACCCTGGCATGAGGAGATGGGCCTGTTGCTCCTGGTCCCA
TTGCTCCTGCTGCCCGGCTCCTACGGACTGCCCTTCTACAACGGCTTCTACTACTCCAACAG
CGCCAACGACCAGAACCTAGGCAACGGTCATGGCAAAGACCTCCTTAATGGAGTGAAGCTGG
TGGTGGAGACACCCGAGGAGACCCTGTTCACCTACCAAGGGGCCAGTGTGATCCTGCCCTGC
CGCTACCGCTACGAGCCGGCCCTGGTCTCCCCGCGGCGTGTGCGTGTCAAATGGTGGAAGCT
GTCGGAGAACGGGGCCCCAGAGAAGGACGTGCTGGTGGCCATCGGGCTGAGGCACCGCTCCT
TTGGGGACTACCAAGGCCGCGTGCACCTGCGGCAGGACAAAGAGCATGACGTCTCGCTGGAG
ATCCAGGATCTGCGGCTGGAGGACTATGGGCGTTACCGCTGTGAGGTCATTGACGGGCTGGA
GGATGAAAGCGGTCTGGTGGAGCTGGAGCTGCGGGGTGTGGTCTTTCCTTACCAGTCCCCCA
ACGGGCGCTACCAGTTCAACTTCCACGAGGGCCAGCAGGTCTGTGCAGAGCAGGCTGCGGTG
GTGGCCTCCTTTGAGCAGCTCTTCCGGGCCTGGGAGGAGGGCCTGGACTGGTGCAACGCGGG
CTGGCTGCAGGATGCTACGGTGCAGTACCCCATCATGTTGCCCCGGCAGCCCTGCGGTGGCC
CAGGCCTGGCACCTGGCGTGCGAAGCTACGGCCCCCGCCACCGCCGCCTGCACCGCTATGAT
GTATTCTGCTTCGCTACTGCCCTCAAGGGGCGGGTGTACTACCTGGAGCACCCTGAGAAGCT
GACGCTGACAGAGGCAAGGGAGGCCTGCCAGGAAGATGATGCCACGATCGCCAAGGTGGGAC
AGCTCTTTGCCGCCTGGAAGTTCCATGGCCTGGACCGCTGCGACGCTGGCTGGCTGGCAGAT
GGCAGCGTCCGCTACCCTGTGGTTCACCCGCATCCTAACTGTGGGCCCCCAGAGCCTGGGGT
CCGAAGCTTTGGCTTCCCCGACCCGCAGAGCCGCTTGTACGGTGTTTACTGCTACCGCCAGC
ACTAGGACCTGGGGCCCTCCCCTGCCGCATTCCCTCACTGGCTGTGTATTTATTGAGTGGTT
CGTTTTCCCTTGTGGGTTGGAGCCATTTTAACTGTTTTTATACTTCTCAATTTAAATTTTCT
TTAAACATTTTTTTACTATTTTTTGTAAAGCAAACAGAACCCAATGCCTCCCTTTGCTCCTG
GATGCCCCACTCCAGGAATCATGCTTGCTCCCCTGGGCCATTTGCGGTTTTGTGGGCTTCTG
GAGGGTTCCCCGCCATCCAGGCTGGTCTCCCTCCCTTAAGGAGGTTGGTGCCCAGAGTGGGC
GGTGGCCTGTCTAGAATGCCGCCGGGAGTCCGGGCATGGTGGGCACAGTTCTCCCTGCCCCT
CAGCCTGGGGGAAGAAGAGGGCCTCGGGGGCCTCCGGAGCTGGGCTTTGGGCCTCTCCTGCC
CACCTCTACTTCTCTGTGAAGCCGCTGACCCCAGTCTGCCCACTGAGGGGCTAGGGCTGGAA
GCCAGTTCTAGGCTTCCAGGCGAAATCTGAGGGAAGGAAGAAACTCCCCTCCCCGTTCCCCT
TCCCCTCTCGGTTCCAAAGAATCTGTTTTGTTGTCATTTGTTTCTCCTGTTTCCCTGTGTGG
GGAGGGGCCCTCAGGTGTGTGTACTTTGGACAATAAATGGTGCTATGACTGCCTTCCGCCAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 78

>><ss.DNA39423

>><subunit 1 of 1, 360 aa, 1 stop

>><MW: 40894, pI: 6.44, NX(S/T): 0

MGLLLLVPLLLLPGSYGLPFYNGFYYSNSANDQNLGNGHGKDLLNGVKLVVETPEETLFT
YQGASVILPCRYRYEPALVSPRRVRVKWWKLSENGAPEKDVLVAIGLRHRSFGDYQGRVH
LRQDKEHDVSLEIQDLRLEDYGRYRCEVIDGLEDESGLVELELRGVVFPYQSPNGRYQFN
FHEGQQVCAEQAAVVASFEQLFRAWEEGLDWCNAGWLQDATVQYPIMLPRQPCGGPGLAP
GVRSYGPRHRRLHRYDVFCFATALKGRVYYLEHPEKLTLTEAREACQEDDATIAKVGQLF
AAWKFHGLDRCDAGWLADGSVRYPVVHPHPNCGPPEPGVRSFGFPDPQSRLYGVYCYRQH

FIGURE 79

GGAGAGCGGAGCGAAGCTGGATAACAGGGGACCG

><MET {trans=1-s, dir=f, res=1}

ATGATGTGGCGACCATCAGTTCTGCTGCTTCTGTTGCTACTGAGGCACGGGGCCCAGGGG

AAGCCATCCCCAGACGCAGGCCCTCATGGCCAGGGGAGGGTGCACCAGGCGGCCCCCCTG

AGCGACGCTCCCCATGATGACGCCCACGGGAACTTCCAGTACGACCATGAGGCTTTCCTG

GGACGGGAAGTGGCCAAGGAATTCGACCAACTCACCCCAGAGGAAAGCCAGGCCCGTCTG

GGGCGGATCGTGGACCGCATGGACCGCGCGGGGGACGGCGACGGCTGGGTGTCGCTGGCC

GAGCTTCGCGCGTGGATCGCGCACACGCAGCAGCGGCACATACGGGACTCGGTGAGCGCG

GCCTGGGACACGTACGACACGGACCGCGACGGGCGTGTGGGTTGGGAGGAGCTGCGCAAC

GCCACCTATGGCCACTACGCGCCCGGTGAAGAATTTCATGACGTGGAGGATGCAGAGACC

TACAAAAAGATGCTGGCTCGGGACGAGCGGCGTTTCCGGGTGGCCGACCAGGATGGGGAC

TCGATGGCCACTCGAGAGGAGCTGACAGCCTTCCTGCACCCCGAGGAGTTCCCTCACATG

CGGGACATCGTGATTGCTGAAACCCTGGAGGACCTGGACAGAAACAAAGATGGCTATGTC

CAGGTGGAGGAGTACATCGCGGATCTGTACTCAGCCGAGCCTGGGGAGGAGGAGCCGGCG

TGGGTGCAGACGGAGAGGCAGCAGTTCCGGGACTTCCGGGATCTGAACAAGGATGGGCAC

CTGGATGGGAGTGAGGTGGGCCACTGGGTGCTGCCCCCTGCCCAGGACCAGCCCCTGGTG

GAAGCCAACCACCTGCTGCACGAGAGCGACACGGACAAGGATGGGCGGCTGAGCAAAGCG

GAAATCCTGGGTAATTGGAACATGTTTGTGGGCAGTCAGGCCACCAACTATGGCGAGGAC

CTGACCCCGGCACCACGATGAGCTGTGAGCACCGCGCACCTGCCACAGCCTCAGAGGCCCG

CACAATGACCGGAGGAGGGGCCGCTGTGGTCTGGCCCCCTCCCTGTCCAGGCCCCGCAGG

AGGCAGATGCAGTCCCAGGCATCCTCCTGCCCCTGGGCTCTCAGGGACCCCCTGGGTCGG

CTTCTGTCCCTGTCACACCCCCAACCCCAGGGAGGGGCTGTCATAGTCCCAGAGGATAAG

CAATACCTATTTCTGACTGAGTCTCCCAGCCCAGACCCAGGGACCCTTGGCCCCAAGCTC

AGCTCTAAGAACCGCCCCAACCCCTCCAGCTCCAAATCTGAGCCTCCACCACATAGACTG

AAACTCCCCTGGCCCCAGCCCTCTCCTGCCTGGCCTGGCCTGGGACACCTCCTCTCTGCC

AGGAGGCAATAAAAGCCAGCGCCGGGACCTTGAAAAAAAAAAAAAAAAAAAAAAAAAAA

AAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 80

>></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA40620

><subunit 1 of 1, 328 aa, 0 stop

><MW: 37493, pI: 4.77, NX(S/T): 1

MMWRPSVLLLLLLLLRHGAQGKPSPDAGPHGQGRVHQAAPLSDAPHDDAHGNFQYDHEAFL
GREVAKEFDQLTPEESQARLGRIVDRMDRAGDGDGWVSLAELRAWIAHTQQRHIRDSVSA
AWDTYDTDRDGRVGWEELRNATYGHYAPGEEFHDVEDAETYKKMLARDERRFRVADQDGD
SMATREELTAFLHPEEFPHMRDIVIAETLEDLDRNKDGYVQVEEYIADLYSAEPGEEEPA
WVQTERQQFRDFRDLNKDGHLDGSEVGHWVLPPAQDQPLVEANHLLHESDTDKDGRLSKA
EILGNWNMFVGSQATNYGEDLTRHHDEL

FIGURE 81

GGGGCCTTGCCTTCCGCACTCGGGCGCAGCCGGGTGGATCTCGAGCAGGTGCGGAGCCCC
GGGCGGCGGGCGCGGGTGCGAGGGATCCCTGACGCCTCTGTCCCTGTTTCTTTGTCGCTC
CCAGCCTGTCTGTCGTCGTTTTGGCGCCCCCGCCTCCCCGCGGTGCGGGGTTGCACACCG
ATCCTGGGCTTCGCTCGATTTGCCGCCGAGGCGCCTCCCAGACCTAGAGGGGCGCTGGCC
TGGAGCAGCGGGTCGTCTGTGTCCTCTCTCCTCTGCGCCGCGCCCGGGGATCCGAAGGGT
GCGGGGCTCTGAGGAGGTGACGCGCGGGGCCTCCCGCACCCTGGCCTTGCCCCGCATTCTC
CCTCTCTCCCAGGTGTGAGCAGCCTATCAGTCACC
><MET {trans=1-s, dir=f, res=1}
ATGTCCGCAGCCTGGATCCCGGCTCTCGGCCTCGGTGTGTGTCTGCTGCTGCTGCCGGGG
CCCGCGGGCAGCGAGGGAGCCGCTCCCATTGCTATCACATGTTTTACCAGAGGCTTGGAC
ATCAGGAAAGAGAAAGCAGATGTCCTCTGCCCAGGGGGCTGCCCTCTTGAGGAATTCTCT
GTGTATGGGAACATAGTATATGCTTCTGTATCGAGCATATGTGGGGCTGCTGTCCACAGG
GGAGTAATCAGCAACTCAGGGGGACCTGTACGAGTCTATAGCCTACCTGGTCGAGAAAAC
TATTCCTCAGTAGATGCCAATGGCATCCAGTCTCAAATGCTTTCTAGATGGTCTGCTTCT
TTCACAGTAACTAAAGGCAAAAGTAGTACACAGGAGGCCACAGGACAAGCAGTGTCCACA
GCACATCCACCAACAGGTAAACGACTAAAGAAAACACCCGAGAAGAAAACTGGCAATAAA
GATTGTAAAGCAGACATTGCATTTCTGATTGATGGAAGCTTTAATATTGGGCAGCGCCGA
TTTAATTTACAGAAGAATTTTGTTGGAAAAGTGGCTCTAATGTTGGGAATTGGAACAGAA
GGACCACATGTGGGCCTTGTTCAAGCCAGTGAACATCCCAAAATAGAATTTTACTTGAAA
AACTTTACATCAGCCAAAGATGTTTTGTTTGCCATAAAGGAAGTAGGTTTCAGAGGGGGT
AATTCCAATACAGGAAAAGCCTTGAAGCATACTGCTCAGAAATTCTTCACGGTAGATGCT
GGAGTAAGAAAAGGGATCCCCAAAGTGGTGGTGGTATTTATTGATGGTTGGCCTTCTGAT
GACATCGAGGAAGCAGGCATTGTGGCCAGAGAGTTTGGTGTCAATGTATTTATAGTTTCT
GTGGCCAAGCCTATCCCTGAAGAACTGGGGATGGTTCAGGATGTCACATTTGTTGACAAG
GCTGTCTGTCGGAATAATGGCTTCTTCTCTTACCACATGCCCAACTGGTTTGGCACCACA
AAATACGTAAAGCCTCTGGTACAGAAGCTGTGCACTCATGAACAAATGATGTGCAGCAAG
ACCTGTTATAACTCAGTGAACATTGCCTTTCTAATTGATGGCTCCAGCAGTGTTGGAGAT
AGCAATTTCCGCCTCATGCTTGAATTTGTTTCCAACATAGCCAAGACTTTTGAAATCTCG
GACATTGGTGCCAAGATAGCTGCTGTACAGTTTACTTATGATCAGCGCACGGAGTTCAGT
TTCACTGACTATAGCACCAAAGAGAATGTCCTAGCTGTCATCAGAAACATCCGCTATATG
AGTGGTGGAACAGCTACTGGTGATGCCATTTCCTTCACTGTTAGAAATGTGTTTGGCCCT
ATAAGGGAGAGCCCCAACAAGAACTTCCTAGTAATTGTCACAGATGGGCAGTCCTATGAT
GATGTCCAAGGCCCTGCAGCTGCTGCACATGATGCAGGAATCACTATCTTCTCTGTTGGT
GTGGCTTGGGCACCTCTGGATGACCTGAAAGATATGGCTTCTAAACCGAAGGAGTCTCAC
GCTTTCTTCACAAGAGAGTTCACAGGATTAGAACCAATTGTTTCTGATGTCATCAGAGGC
ATTTGTAGAGATTTCTTAGAATCCCAGCAATAATGGTAACATTTTGACAACTGAAAGAAA
AAGTACAAGGGGATCCAGTGTGTAAATTGTATTCTCATAATACTGAAATGCTTTAGCATA
CTAGAATCAGATACAAAACTATTAAGTATGTCAACAGCCATTTAGGCAAATAAGCACTCC
TTTAAAGCCGCTGCCTTCTGGTTACAATTTACAGTGTACTTTGTTAAAAACACTGCTGAG
GCTTCATAATCATGGCTCTTAGAAACTCAGGAAAGAGGAGATAATGTGGATTAAAACCTT
AAGAGTTCTAACCATGCCTACTAAATGTACAGATATGCAAATTCCATAGCTCAATAAAAG
AATCTGATACTTAGACCAAAAAAAAAAA

FIGURE 82

>></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA40604
><subunit 1 of 1, 550 aa, 0 stop
><MW: 59483, pI: 8.34, NX(S/T): 2
MSAAWIPALGLGVCLLLLPGPAGSEGAAPIAITCFTRGLDIRKEKADVLCPGGCPLEEFS
VYGNIVYASVSSICGAAVHRGVISNSGGPVRVYSLPGRENYSSVDANGIQSQMLSRWSAS
FTVTKGKSSTQEATGQAVSTAHPPTGKRLKKTPEKKTGNKDCKADIAFLIDGSFNIGQRR
FNLQKNFVGKVALMLGIGTEGPHVGLVQASEHPKIEFYLKNFTSAKDVLFAIKEVGFRGG
NSNTGKALKHTAQKFFTVDAGVRKGIPKVVVVFIDGWPSDDIEEAGIVAREFGVNVFIVS
VAKPIPEELGMVQDVTFVDKAVCRNNGFFSYHMPNWFGTTKYVKPLVQKLCTHEQMMCSK
TCYNSVNIAFLIDGSSSVGDSNFRLMLEFVSNIAKTFEISDIGAKIAAVQFTYDQRTEFS
FTDYSTKENVLAVIRNIRYMSGGTATGDAISFTVRNVFGPIRESPNKNFLVIVTDGQSYD
DVQGPAAAAHDAGITIFSVGVAWAPLDDLKDMASKPKESHAFFTREFTGLEPIVSDVIRG
ICRDFLESQQ

FIGURE 83

CGCCGCGCTCCCGCACCCGCGGCCCGCCCACCGCGCCGCTCCCGCATCTGCACCCGCAGC
CCGGCGGCCTCCCGGCGGGAGCGAGCAGATCCAGTCCGGCCCGCAGCGCAACTCGGTCCA
GTCGGGGCGGCGGCTGCGGGCGCAGAGCGGAG
><MET {trans=1-s, dir=f, res=1}
ATGCAGCGGCTTGGGGCCACCCTGCTGTGCCTGCTGCTGGCGGCGGCGGTCCCCACGGCC
CCCGCGCCCGCTCCGACGGCGACCTCGGCTCCAGTCAAGCCCGGCCCGGCTCTCAGCTAC
CCGCAGGAGGAGGCCACCCTCAATGAGATGTTCCGCGAGGTTGAGGAACTGATGGAGGAC
ACGCAGCACAAATTGCGCAGCGCGGTGGAAGAGATGGAGGCAGAAGAAGCTGCTGCTAAA
GCATCATCAGAAGTGAACCTGGCAAACTTACCTCCCAGCTATCACAATGAGACCAACACA
GACACGAAGGTTGGAAATAATACCATCCATGTGCACCGAGAAATTCACAAGATAACCAAC
AACCAGACTGGACAAATGGTCTTTTCAGAGACAGTTATCACATCTGTGGGAGACGAAGAA
GGCAGAAGGAGCCACGAGTGCATCATCGACGAGGACTGTGGGCCCAGCATGTACTGCCAG
TTTGCCAGCTTCCAGTACACCTGCCAGCCATGCCGGGGCCAGAGGATGCTCTGCACCCGG
GACAGTGAGTGCTGTGGAGACCAGCTGTGTGTCTGGGGTCACTGCACCAAAATGGCCACC
AGGGGCAGCAATGGGACCATCTGTGACAACCAGAGGGACTGCCAGCCGGGGCTGTGCTGT
GCCTTCCAGAGAGGCCTGCTGTTCCCTGTGTGCACACCCCTGCCCGTGGAGGGCGAGCTT
TGCCATGACCCCGCCAGCCGGCTTCTGGACCTCATCACCTGGGAGCTAGAGCCTGATGGA
GCCTTGGACCGATGCCCTTGTGCCAGTGGCCTCCTCTGCCAGCCCCACAGCCACAGCCTG
GTGTATGTGTGCAAGCCGACCTTCGTGGGGAGCCGTGACCAAGATGGGGAGATCCTGCTG
CCCAGAGAGGTCCCCGATGAGTATGAAGTTGGCAGCTTCATGGAGGAGGTGCGCCAGGAG
CTGGAGGACCTGGAGAGGAGCCTGACTGAAGAGATGGCGCTGGGGGAGCCTGCGGCTGCC
GCCGCTGCACTGCTGGGAGGGGAAGAGATTTAGATCTGGACCAGGCTGTGGGTAGATGTG
CAATAGAAATAGCTAATTTATTTCCCCAGGTGTGTGCTTTAGGCGTGGGCTGACCAGGCT
TCTTCCTACATCTTCTTCCCAGTAAGTTTCCCCTCTGGCTTGACAGCATGAGGTGTTGTG
CATTTGTTCAGCTCCCCCAGGCTGTTCTCCAGGCTTCACAGTCTGGTGCTTGGGAGAGTC
AGGCAGGGTTAAACTGCAGGAGCAGTTTGCCACCCCTGTCCAGATTATTGGCTGCTTTGC
CTCTACCAGTTGGCAGACAGCCGTTTGTTCTACATGGCTTTGATAATTGTTTGAGGGGAG
GAGATGGAAACAATGTGGAGTCTCCCTCTGATTGGTTTTGGGGAAATGTGGAGAAGAGTG
CCCTGCTTTGCAAACATCAACCTGGCAAAAATGCAACAAATGAATTTTCCACGCAGTTCT
TTCCATGGGCATAGGTAAGCTGTGCCTTCAGCTGTTGCAGATGAAATGTTCTGTTCACCC
TGCATTACATGTGTTTATTCATCCAGCAGTGTTGCTCAGCTCCTACCTCTGTGCCAGGGC
AGCATTTTCATATCCAAGATCAATTCCCTCTCTCAGCACAGCCTGGGGAGGGGGTCATTG
TTCTCCTCGTCCATCAGGGATCTCAGAGGCTCAGAGACTGCAAGCTGCTTGCCCAAGTCA
CACAGCTAGTGAAGACCAGAGCAGTTTCATCTGGTTGTGACTCTAAGCTCAGTGCTCTCT
CCACTACCCCACACCAGCCTTGGTGCCACCAAAAGTGCTCCCCAAAAGGAAGGAGAATGG
GATTTTTCTTGAGGCATGCACATCTGGAATTAAGGTCAAACTAATTCTCACATCCCTCTA
AAAGTAAACTACTGTTAGGAACAGCAGTGTTCTCACAGTGTGGGGCAGCCGTCCTTCTAA
TGAAGACAATGATATTGACACTGTCCCTCTTTGGCAGTTGCATTAGTAACTTTGAAAGGT
ATATGACTGAGCGTAGCATACAGGTTAACCTGCAGAAACAGTACTTAGGTAATTGTAGGG
CGAGGATTATAAATGAAATTTGCAAAATCACTTAGCAGCAACTGAAGACAATTATCAACC
ACGTGGAGAAAATCAAACCGAGCAGGGCTGTGTGAAACATGGTTGTAATATGCGACTGCG
AACACTGAACTCTACGCCACTCCACAAATGATGTTTTCAGGTGTCATGGACTGTTGCCAC
CATGTATTCATCCAGAGTTCTTAAAGTTTAAAGTTGCACATGATTGTATAAGCATGCTTT
CTTTGAGTTTTAAATTATGTATAAACATAAGTTGCATTTAGAAATCAAGCATAAATCACT
TCAACTGCAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 84

MQRLGATLLCLLLAAAVPTAPAPAPTATSAPVKPGPALSYPQEEATLNEMFREVEELMED
TQHKLRSAVEEMEAEEAAAKASSEVNLANLPPSYHNETNTDTKVGNNTIHVHREIHKITN
NQTGQMVFSETVITSVGDEEGRRSHECIIDEDCGPSMYCQFASFQYTCQPCRGQRMLCTR
DSECCGDQLCVWGHCTKMATRGSNGTICDNQRDCQPGLCCAFQRGLLFPVCTPLPVEGEL
CHDPASRLLDLITWELEPDGALDRCPCASGLLCQPHSHSLVYVCKPTFVGSRDQDGEILL
PREVPDEYEVGSFMEEVRQELEDLERSLTEEMALGEPAAAAAALLGGEEI

FIGURE 85A

AAGGAGGCTGGGAGGAAAGAGGTAAGAAAGGTTAGAGAACCTACCTCACATCTCTCTGGGCT
CAGAAGGACTCTGAAGATAACAATAATTTCAGCCCATCCACTCTCCTTCCCTCCCAAACACA
CATGTGCATGTACACACACACATACACACACATACACCTTCCTCTCCTTCACTGAAGACTCA
CAGTCACTCACTCTGTGAGCAGGTCATAGAAAAGGACACTAAAGCCTTAAGGACAGGCCTGG
CCATTACCTCTGCAGCTCCTTTGGCTTGTTGAGTCAAAAAACATGGGAGGGGCCAGGCACGG
TGACTCACACCTGTAATCCCAGCATTTTGGGAGACCGAGGTGAGCAGATCACTTGAGGTCAG
GAGTTCGAGACCAGCCTGGCCAACATGGAGAAACCCCCATCTCTACTAAAAATACAAAAATT
AGCCAGGAGTGGTGGCAGGTGCCTGTAATCCCAGCTACTCAGGTGGCTGAGCCAGGAGAATC
GCTTGAATCCAGGAGGCGGAGGATGCAGTCAGCTGAGTGCACCGCTGCACTCCAGCCTGGGT
GACAGAATGAGACTCTGTCTCAAACAAACAAACACGGGAGGAGGGGTAGATACTGCTTCTCT
GCAACCTCCTTAACTCTGCATCCTCTTCTTCCAGGGCTGCCCCTGATGGGGCCTGGCAATGA
CTGAGCAGGCCCAGCCCCAGAGGACAAGGAAGAGAAGGCATATTGAGGAGGGCAAGAAGTGA
CGCCCGGTGTAGAATGACTGCCCTGGGAGGGTGGTTCCTTGGGCCCTGGCAGGGTTGCTGAC
CCTTACCCTGCAAAACACAAAGAGCAGGACTCCAGACTCTCCTTGTGAATGGTCCCCTGCCC
TGCAGCTCCACCATGAGGCTTCTCGTGGCCCCACTCTTGCTAGCTTGGGTGGCTGGTGCCAC
TGCCACTGTGCCCGTGGTACCCTGGCATGTTCCCTGCCCCCCTCAGTGTGCCTGCCAGATCC
GGCCCTGGTATACGCCCCGCTCGTCCTACCGCGAGGCTACCACTGTGGACTGCAATGACCTA
TTCCTGACGGCAGTCCCCCCGGCACTCCCCGCAGGCACACAGACCCTGCTCCTGCAGAGCAA
CAGCATTGTCCGTGTGGACCAGAGTGAGCTGGGCTACCTGGCCAATCTCACAGAGCTGGACC
TGTCCCAGAACAGCTTTTCGGATGCCCGAGACTGTGATTTCCATGCCCTGCCCCAGCTGCTG
AGCCTGCACCTAGAGGAGAACCAGCTGACCCGGCTGGAGGACCACAGCTTTGCAGGGCTGGC
CAGCCTACAGGAACTCTATCTCAACCACAACCAGCTCTACCGCATCGCCCCCAGGGCCTTTT
CTGGCCTCAGCAACTTGCTGCGGCTGCACCTCAACTCCAACCTCCTGAGGGCCATTGACAGC
CGCTGGTTTGAAATGCTGCCCAACTTGGAGATACTCATGATTGGCGGCAACAAGGTAGATGC
CATCCTGGACATGAACTTCCGGCCCCTGGCCAACCTGCGTAGCCTGGTGCTAGCAGGCATGA
ACCTGCGGGAGATCTCCGACTATGCCCTGGAGGGGCTGCAAAGCCTGGAGAGCCTCTCCTTC
TATGACAACCAGCTGGCCCGGGTGCCCAGGCGGGCACTGGAACAGGTGCCCGGGCTCAAGTT
CCTAGACCTCAACAAGAACCCGCTCCAGCGGGTAGGGCCGGGGGACTTTGCCAACATGCTGC
ACCTTAAGGAGCTGGGACTGAACAACATGGAGGAGCTGGTCTCCATCGACAAGTTTGCCCTG
GTGAACCTCCCCGAGCTGACCAAGCTGGACATCACCAATAACCCACGGCTGTCCTTCATCCA
CCCCCGCGCCTTCCACCACCTGCCCCAGATGGAGACCCTCATGCTCAACAACAACGCTCTCA
GTGCCTTGCACCAGCAGACGGTGGAGTCCCTGCCCAACCTGCAGGAGGTAGGTCTCCACGGC
AACCCCATCCGCTGTGACTGTGTCATCCGCTGGGCCAATGCCACGGGCACCCGTGTCCGCTT
CATCGAGCCGCAATCCACCCTGTGTGCGGAGCCTCCGGACCTCCAGCGCCTCCCGGTCCGTG
AGGTGCCCTTCCGGGAGATGACGGACCACTGTTTGCCCCTCATCTCCCCACGAAGCTTCCCC
CCAAGCCTCCAGGTAGCCAGTGGAGAGAGCATGGTGCTGCATTGCCGGGCACTGGCCGAACC
CGAACCCGAGATCTACTGGGTCACTCCAGCTGGGCTTCGACTGACACCTGCCCATGCAGGCA
GGAGGTACCGGGTGTACCCCGAGGGGACCCTGGAGCTGCGGAGGGTGACAGCAGAAGAGGCA
GGGCTATACACCTGTGTGGCCCAGAACCTGGTGGGGGCTGACACTAAGACGGTTAGTGTGGT
TGTGGGCCGTGCTCTCCTCCAGCCAGGCAGGGACGAAGGACAGGGGCTGGAGCTCCGGGTGC
AGGAGACCCACCCCTATCACATCCTGCTATCTTGGGTCACCCCACCCAACACAGTGTCCACC
AACCTCACCTGGTCCAGTGCCTCCTCCCTCCGGGGCCAGGGGGCCACAGCTCTGGCCCCGCCT
GCCTCGGGGAACCCACAGCTACAACATTACCCGCCTCCTTCAGGCCACGGAGTACTGGGCCT
GCCTGCAAGTGGCCTTTGCTGATGCCCACACCCAGTTGGCTTGTGTATGGGCCAGGACCAAA
GAGGCCACTTCTTGCCACAGAGCCTTAGGGGATCGTCCTGGGCTCATTGCCATCCTGGCTCT
CGCTGTCCTTCTCCTGGCAGCTGGGCTAGCGGCCCACCTTGGCACAGGCCAACCCAGGAAGG
GTGTGGGTGGGAGGCGGCCTCTCCCTCCAGCCTGGGCTTTCTGGGGCTGGAGTGCCCCTTCT
GTCCGGGTTGTGTCTGCTCCCCTCGTCCTGCCCTGGAATCCAGGGAGGAAGCTGCCCAGATC
CTCAGAAGGGGAGACACTGTTGCCACCATTGTCTCAAAATTCTTGAAGCTCAGCCTGTTCTC
AGCAGTAGAGAAATCACTAGGACTACTTTTTACCAAAAGAGAAGCAGTCTGGGCCAGATGCC
CTGCCAGGAAAGGGACATGGACCCACGTGCTTGAGGCCTGGCAGCTGGGCCAAGACAGATGG
GGCTTTGTGGCCCTGGGGGTGCTTCTGCAGCCTTGAAAAAGTTGCCCTTACCTCCTAGGGTC
ACCTCTGCTGCCATTCTGAGGAACATCTCCAAGGAACAGGAGGGACTTTGGCTAGAGCCTCC

FIGURE 85B

TGCCTCCCCATCTTCTCTCTGCCCAGAGGCTCCTGGGCCTGGCTTGGCTGTCCCCTACCTGT
GTCCCCGGGCTGCACCCCTTCCTCTTCTCTTTCTCTGTACAGTCTCAGTTGCTTGCTCTTGT
GCCTCCTGGGCAAGGGCTGAAGGAGGCCACTCCATCTCACCTCGGGGGGCTGCCCTCAATGT
GGGAGTGACCCCAGCCAGATCTGAAGGACATTTGGGAGAGGGATGCCCAGGAACGCCTCATC
TCAGCAGCCTGGGCTCGGCATTCCGAAGCTGACTTTCTATAGGCAATTTTGTACCTTTGTGG
AGAAATGTGTCACCTCCCCCAACCCGATTCACTCTTTTCTCCTGTTTTGTAAAAAATAAAAA
TAAATAATAACAATAAAAAAA

FIGURE 86

MRLLVAPLLLAWVAGATATVPVVPWHVPCPPQCACQIRPWYTPRSSYREATTVDCNDLF
LTAVPPALPAGTQTLLLQSNSIVRVDQSELGYLANLTELDLSQNSFSDARDCDFHALPQ
LLSLHLEENQLTRLEDHSFAGLASLQELYLNHNQLYRIAPRAFSGLSNLLRLHLNSNLL
RAIDSRWFEMLPNLEILMIGGNKVDAILDMNFRPLANLRSLVLAGMNLREISDYALEGL
QSLESLSFYDNQLARVPRRALEQVPGLKFLDLNKNPLQRVGPGDFANMLHLKELGLNNM
EELVSIDKFALVNLPELTKLDITNNPRLSFIHPRAFHHLPQMETLMLNNNALSALHQQT
VESLPNLQEVGLHGNPIRCDCVIRWANATGTRVRFIEPQSTLCAEPPDLQRLPVREVPF
REMTDHCLPLISPRSFPPSLQVASGESMVLHCRALAEPEPEIYWVTPAGLRLTPAHAGR
RYRVYPEGTLELRRVTAEEAGLYTCVAQNLVGADTKTVSVVVGRALLQPGRDEGQGLEL
RVQETHPYHILLSWVTPPNTVSTNLTWSSASSLRGQGATALARLPRGTHSYNITRLLQA
TEYWACLQVAFADAHTQLACVWARTKEATSCHRALGDRPGLIAILALAVLLLAAGLAAH
LGTGQPRKGVGGRRPLPPAWAFWGWSAPSVRVVSAPLVLPWNPGRKLPRSSEGETLLPP
LSQNS

FIGURE 87A

GCAAGCCAAGGCGCTGTTTGAGAAGGTGAAGAAGTTCCGGACCCATGTGGAGGAGGGGGA
CATTGTGTACCGCCTCTAC
><MET {trans=1-s, dir=f, res=1}>
ATGCGGCAGACCATCATCAAGGTGATCAAGTTCATCCTCATCATCTGCTACACCGTCTAC
TACGTGCACAACATCAAGTTCGACGTGGACTGCACCGTGGACATTGAGAGCCTGACGGGC
TACCGCACCTACCGCTGTGCCCACCCCCTGGCCACACTCTTCAAGATCCTGGCGTCCTTC
TACATCAGCCTAGTCATCTTCTACGGCCTCATCTGCATGTACACACTGTGGTGGATGCTA
CGGCGCTCCCTCAAGAAGTACTCGTTTGAGTCGATCCGTGAGGAGAGCAGCTACAGCGAC
ATCCCCGACGTCAAGAACGACTTCGCCTTCATGCTGCACCTCATTGACCAATACGACCCG
CTCTACTCCAAGCGCTTCGCCGTCTTCCTGTCGGAGGTGAGTGAGAACAAGCTGCGGCAG
CTGAACCTCAACAACGAGTGGACGCTGGACAAGCTCCGGCAGCGGCTCACCAAGAACGCG
CAGGACAAGCTGGAGCTGCACCTGTTCATGCTCAGTGGCATCCCTGACACTGTGTTTGAC
CTGGTGGAGCTGGAGGTCCTCAAGCTGGAGCTGATCCCCGACGTGACCATCCCGCCCAGC
ATTGCCCAGCTCACGGGCCTCAAGGAGCTGTGGCTCTACCACACAGCGGCCAAGATTGAA
GCGCCTGCGCTGGCCTTCCTGCGCGAGAACCTGCGGGCGCTGCACATCAAGTTCACCGAC
ATCAAGGAGATCCCGCTGTGGATCTATAGCCTGAAGACACTGGAGGAGCTGCACCTGACG
GGCAACCTGAGCGCGGAGAACAACCGCTACATCGTCATCGACGGGCTGCGGGAGCTCAAA
CGCCTCAAGGTGCTGCGGCTCAAGAGCAACCTAAGCAAGCTGCCACAGGTGGTCACAGAT
GTGGGCGTGCACCTGCAGAAGCTGTCCATCAACAATGAGGGCACCAAGCTCATCGTCCTC
AACAGCCTCAAGAAGATGGCGAACCTGACTGAGCTGGAGCTGATCCGCTGCGACCTGGAG
CGCATCCCCCACTCCATCTTCAGCCTCCACAACCTGCAGGAGATTGACCTCAAGGACAAC
AACCTCAAGACCATCGAGGAGATCATCAGCTTCCAGCACCTGCACCGCCTCACCTGCCTT
AAGCTGTGGTACAACCACATCGCCTACATCCCCATCCAGATCGGCAACCTCACCAACCTG
GAGCGCCTCTACCTGAACCGCAACAAGATCGAGAAGATCCCCACCCAGCTCTTCTACTGC
CGCAAGCTGCGCTACCTGGACCTCAGCCACAACAACCTGACCTTCCTCCCTGCCGACATC
GGCCTCCTGCAGAACCTCCAGAACCTAGCCATCACGGCCAACCGGATCGAGACGCTCCCT
CCGGAGCTCTTCCAGTGCCGGAAGCTGCGGGCCCTGCACCTGGGCAACAACGTGCTGCAG
TCACTGCCCTCCAGGGTGGGCGAGCTGACCAACCTGACGCAGATCGAGCTGCGGGGCAAC
CGGCTGGAGTGCCTGCCTGTGGAGCTGGGCGAGTGCCCACTGCTCAAGCGCAGCGGCTTG
GTGGTGGAGGAGGACCTGTTCAACACACTGCCACCCGAGGTGAAGGAGCGGCTGTGGAGG
GCTGACAAGGAGCAGGCCTGAGCGAGGCCGGCCCAGCACAGCAAGCAGCAGGACCGCTGC
CCAGTCCTCAGGCCCGGAGGGGCAGGCCTAGCTTCTCCCAGAACTCCCGGACAGCCAGGA
CAGCCTCGCGGCTGGGCAGGAGCCTGGGGCCGCTTGTGAGTCAGGCCAGAGCGAGAGGAC
AGTATCTGTGGGGCTGGCCCCTTTTCTCCCTCTGAGACTCACGTCCCCCAGGGCAAGTGC
TTGTGGAGGAGAGCAAGTCTCAAGAGCGCAGTATTTGGATAATCAGGGTCTCCTCCCTGG
AGGCCAGCTCTGCCCCAGGGGCTGAGCTGCCACCAGAGGTCCTGGGACCCTCACTTTAGT
TCTTGGTATTTATTTTTCTCCATCTCCCACCTCCTTCATCCAGATAACTTATACATTCCC
AAGAAAGTTCAGCCCAGATGGAAGGTGTTCAGGGAAAGGTGGGCTGCCTTTTCCCCTTGT
CCTTATTTAGCGATGCCGCCGGGCATTTAACACCCACCTGGACTTCAGCAGAGTGGTCCG
GGGCGAACCAGCCATGGGACGGTCACCCAGCAGTGCCGGGCTGGGCTCTGCGGTGCGGTC
CACGGGAGAGCAGGCCTCCAGCTGGAAAGGCCAGGCCTGGAGCTTGCCTCTTCAGTTTTT
GTGGCAGTTTTAGTTTTTTGTTTTTTTTTTTTTTAATCAAAAAACAATTTTTTTTAAAAA
AAAGCTTTGAAAATGGATGGTTTGGGTATTAAAAAGAAAAAAAAAACTTAAAAAAAAAAA
GACACTAACGGCCAGTGAGTTGGAGTCTCAGGGCAGGGTGGCAGTTTCCCTTGAGCAAAG
CAGCCAGACGTTGAACTGTGTTTCCTTTCCCTGGGCGCAGGGTGCAGGGTGTCTTCCGGA
TCTGGTGTGACCTTGGTCCAGGAGTTCTATTTGTTCCTGGGGAGGGAGGTTTTTTTGTTT
GTTTTTTGGGTTTTTTTGGTGTCTTGTTTTCTTTCTCCTCCATGTGTCTTGGCAGGCACT
CATTTCTGTGGCTGTCGGCCAGAGGGAATGTTCTGGAGCTGCCAAGGAGGGAGGAGACTC
GGGTTGGCTAATCCCCGGATGAACGGTGCTCCATTCGCACCTCCCCTCCTCGTGCCTGCC
CTGCCTCTCCACGCACAGTGTTAAGGAGCCAAGAGGAGCCACTTCGCCCAGACTTTGTTT
CCCCACCTCCTGCGGCATGGGTGTGTCCAGTGCCACCGCTGGCCTCCGCTGCTTCCATCA
GCCCTGTCGCCACCTGGTCCTTCATGAAGAGCAGACACTTAGAGGCTGGTCGGGAATGGG
GAGGTCGCCCCTGGGAGGGCAGGCGTTGGTTCCAAGCCGGTTCCCGTCCCTGGCGCCTGG

FIGURE 87B

AGTGCACACAGCCCAGTCGGCACCTGGTGGCTGGAAGCCAACCTGCTTTAGATCACTCGG
GTCCCCACCTTAGAAGGGTCCCCGCCTTAGATCAATCACGTGGACACTAAGGCACGTTTT
AGAGTCTCTTGTCTTAATGATTATGTCCATCCGTCTGTCCGTCCATTTGTGTTTTCTGCG
TCGTGTCATTGGATATAATCCTCAGAAATAATGCACACTAGCCTCTGACAACCATGAAGC
AAAAATCCGTTACATGTGGGTCTGAACTTGTAGACTCGGTCACAGTATCAAATAAAATCT
ATAACAGAAAAAAAAAAAAAAA

FIGURE 88

>></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA35673

><subunit 1 of 1, 546 aa, 0 stop

><MW: 63742, pI: 8.62, NX(S/T): 6

MRQTIIKVIKFILIICYTVYYVHNIKFDVDCTVDIESLTGYRTYRCAHPLATLFKILASF
YISLVIFYGLICMYTLWWMLRRSLKKYSFESIREESSYSDIPDVKNDFAFMLHLIDQYDP
LYSKRFAVFLSEVSENKLRQLNLNNEWTLDKLRQRLTKNAQDKLELHLFMLSGIPDTVFD
LVELEVLKLELIPDVTIPPSIAQLTGLKELWLYHTAAKIEAPALAFLRENLRALHIKFTD
IKEIPLWIYSLKTLEELHLTGNLSAENNRYIVIDGLRELKRLKVLRLKSNLSKLPQVVTD
VGVHLQKLSINNEGTKLIVLNSLKKMANLTELELIRCDLERIPHSIFSLHNLQEIDLKDN
NLKTIEEIISFQHLHRLTCLKLWYNHIAYIPIQIGNLTNLERLYLNRNKIEKIPTQLFYC
RKLRYLDLSHNNLTFLPADIGLLQNLQNLAITANRIETLPPELFQCRKLRALHLGNNVLQ
SLPSRVGELTNLTQIELRGNRLECLPVELGECPLLKRSGLVVEEDLFNTLPPEVKERLWR
ADKEQA

FIGURE 89

GCCTGTTGCTGATGCTGCCGTGCGGTACTTGTC

><MET {trans=1-s, dir=f, res=1}

ATGGAGCTGGCACTGCGGCGCTCTCCCGTCCCGCGGTGGTTGCTGCTGCTGCCGCTGCTG

CTGGGCCTGAACGCAGGAGCTGTCATTGACTGGCCCACAGAGGAGGGCAAGGAAGTATGG

GATTATGTGACGGTCCGCAAGGATGCCTACATGTTCTGGTGGCTCTATTATGCCACCAAC

TCCTGCAAGAACTTCTCAGAACTGCCCCTGGTCATGTGGCTTCAGGGCGGTCCAGGCGGT

TCTAGCACTGGATTTGGAAACTTTGAGGAAATTGGGCCCCTTGACAGTGATCTCAAACCA

CGGAAAACCACCTGGCTCCAGGCTGCCAGTCTCCTATTTGTGGATAATCCCGTGGGCACT

GGGTTCAGTTATGTGAATGGTAGTGGTGCCTATGCCAAGGACCTGGCTATGGTGGCTTCA

GACATGATGGTTCTCCTGAAGACCTTCTTCAGTTGCCACAAAGAATTCCAGACAGTTCCA

TTCTACATTTTCTCAGAGTCCTATGGAGGAAAAATGGCAGCTGGCATTGGTCTAGAGCTT

TATAAGGCCATTCAGCGAGGGACCATCAAGTGCAACTTTGCGGGGGTTGCCTTGGGTGAT

TCCTGGATCTCCCCTGTTGATTCGGTGCTCTCCTGGGGACCTTACCTGTACAGCATGTCT

CTTCTCGAAGACAAAGGTCTGGCAGAGGTGTCTAAGGTTGCAGAGCAAGTACTGAATGCC

GTAAATAAGGGGCTCTACAGAGAGGCCACAGAGCTGTGGGGGAAAGCAGAAATGATCATT

GAACAGAACACAGATGGGGTGAACTTCTATAACATCTTAACTAAAAGCACTCCCACGTCT

ACAATGGAGTCGAGTCTAGAATTCACACAGAGCCACCTAGTTTGTCTTTGTCAGCGCCAC

GTGAGACACCTACAACGAGATGCCTTAAGCCAGCTCATGAATGGCCCCATCAGAAAGAAG

CTCAAAATTATTCCTGAGGATCAATCCTGGGGAGGCCAGGCTACCAACGTCTTTGTGAAC

ATGGAGGAGGACTTCATGAAGCCAGTCATTAGCATTGTGGACGAGTTGCTGGAGGCAGGG

ATCAACGTGACGGTGTATAATGGACAGCTGGATCTCATCGTAGATACCATGGGTCAGGAG

GCCTGGGTGCGGAAACTGAAGTGGCCAGAACTGCCTAAATTCAGTCAGCTGAAGTGGAAG

GCCCTGTACAGTGACCCTAAATCTTTGGAAACATCTGCTTTTGTCAAGTCCTACAAGAAC

CTTGCTTTCTACTGGATTCTGAAAGCTGGTCATATGGTTCCTTCTGACCAAGGGGACATG

GCTCTGAAGATGATGAGACTGGTGACTCAGCAAGAATAGGATGGATGGGGCTGGAGATGA

GCTGGTTTGGCCTTGGGGCACAGAGCTGAGCTGAGGCCGCTGAAGCTGTAGGAAGCGCCA

TTCTTCCCTGTATCTAACTGGGGCTGTGATCAAGAAGGTTCTGACCAGCTTCTGCAGAGG

ATAAAATCATTGTCTCTGGAGGCAATTTGGAAATTATTTCTGCTTCTTAAAAAAACCTAA

GATTTTTTAAAAAATTGATTTGTTTTGATCAAAATAAAGGATGATAATAGATATTAA

FIGURE 90

><signal peptide>

MELALRRSPVPRWLLLLPLLLGLNA

><start mature protein>

GAVIDWPTEEGKEVW

><homology to peptidases 40-end>

DYVTVRKDAYMFWWLYYATNSCK

><potential N-glycosylation site>

NFSELPLVMWLQGGPGGSSTGFGNFEEIGPLDSDLKPRKTTWLQAASLLFVDNPVGT

GFSYV

><potential N-glycosylation site>

NGSGAYAKDLAMVASDMMVLLKTFFSCHKEFQTVPFYIFSESYGGKMAAGIGLELY

KAIQRGTIKCNFAGVALGDSWISPVDSVLSWGPYLYSMSLLEDKGLAEVSKVAEQVL

NAVNKGLYREATELWGKAEMIEQNTDGVNFYNILTKSTPTSTMESSLEFTQSHLVCL

CQRHVRHLQRDALSQLMNGPIRKKLKIIPEDQSWGGQATNVFVNMEEDFMKPVISIV

DELLEAGI

><potential N-glycosylation site>

NVTVYNGQLDLIVDTMGQEAWVRKLKWPELPKFSQLKWKALYSDPKSLETSAFVKS

YKNLAFYWILKAGHMVPSDQGDMALKMMRLVTQQE

FIGURE 91

GGCCGCGGGAGAGGAGGCC

><MET {trans=1-s, dir=f, res=1}

ATGGGCGCGCCGCGGGGCGCTGCTGCTGGCGCTGCTGCTGGCTCGGGCTGGACTCAGGAAG
CCGGAGTCGCAGGAGGCGGCGCCGTTATCAGGACCATGCGGCCGACGGGTCATCACGTCG
CGCATCGTGGGTGGAGAGGACGCCGAACTCGGGCGTTGGCCGTGGCAGGGGAGCCTGCGC
CTGTGGGATTCCCACGTATGCGGAGTGAGCCTGCTCAGCCACCGCTGGGCACTCACGGCG
GCGCACTGCTTTGAAACCTATAGTGACCTTAGTGATCCCTCCGGGTGGATGGTCCAGTTT
GGCCAGCTGACTTCCATGCCATCCTTCTGGAGCCTGCAGGCCTACTACACCCGTTACTTC
GTATCGAATATCTATCTGAGCCCTCGCTACCTGGGGAATTCACCCTATGACATTGCCTTG
GTGAAGCTGTCTGCACCTGTCACCTACACTAAACACATCCAGCCCATCTGTCTCCAGGCC
TCCACATTTGAGTTTGAGAACCGGACAGACTGCTGGGTGACTGGCTGGGGGGTACATCAAA
GAGGATGAGGCACTGCCATCTCCCCACACCCTCCAGGAAGTTCAGGTCGCCATCATAAAC
AACTCTATGTGCAACCACCTCTTCCTCAAGTACAGTTTCCGCAAGGACATCTTTGGAGAC
ATGGTTTGTGCTGGCAACGCCCAAGGCGGGAAGGATGCCTGCTTCGGTGACTCAGGTGGA
CCCTTGGCCTGTAACAAGAATGGACTGTGGTATCAGATTGGAGTCGTGAGCTGGGGAGTG
GGCTGTGGTCGGCCCAATCGGCCCGGTGTCTACACCAATATCAGCCACCACTTTGAGTGG
ATCCAGAAGCTGATGGCCCAGAGTGGCATGTCCCAGCCAGACCCCTCCTGGCCACTACTC
TTTTTCCCTCTTCTCTGGGCTCTCCCACTCCTGGGGCCGGTCTGAGCCTACCTGAGCCCA
TGCAGCCTGGGGCCACTGCCAAGTCAGGCCCTGGTTCTCTTCTGTCTTGTTTGGTAATAA
ACACATTCCAGTTGATGCCTTGCAGGGCATTCTTCAAAAAAAAAAAAAAAAAAAAAAAAA
A

FIGURE 92

><signal peptide>

MGARGALLLALLLARAGL

><start mature protein>

RKPESQEAAPLSGPCGRRVITSRIVGGEDAELGRWPWQGSLRLWDSHVCGVSLLSHRWA

><Serine proteases, trypsin family, histidine active site 'LTAAHC'>

LTAAHCFETYSDLSDPSGWMVQFGQLTSMPSFWSLQAYYTRYFVSNIYLSPRYLGNS

PYDIALVKLSAPVTYTKHIQPICLQASTFEFE

><potential N-glycosylation site>

NRTDCWVTGWGYIKEDEALPSPHTLQEVQVAII

><potential N-glycosylation site>

NNSMCNHLFLKYSFRKDIFGDMVCAGNAQGGKDACFGDSGGPLACNKNGLWYQIG

VVSWGVGCGRPNRPGVYT

><potential N-glycosylation site>

NISHHFEWIQKLMAQSGMSQPDPSWPLLFFPLLWALPLLGPV

FIGURE 93

CCCACGCGTCCGCGGACGCGTGGGAAGGGCAGA
<MET {trans=1-s, dir=f, res=1}
ATGGGACTCCAAGCCTGCCTCCTAGGGCTCTTTGCCCTCATCCTCTCTGGCAAATGCAGT
TACAGCCCGGAGCCCGACCAGCGGAGGACGCTGCCCCCAGGCTGGGTGTCCCTGGGCCGT
GCGGACCCTGAGGAAGAGCTGAGTCTCACCTTTGCCCTGAGACAGCAGAATGTGGAAAGA
CTCTCGGAGCTGGTGCAGGCTGTGTCGGATCCCAGCTCTCCTCAATACGGAAAATACCTG
ACCCTAGAGAATGTGGCTGATCTGGTGAGGCCATCCCCACTGACCCTCCACACGGTGCAA
AAATGGCTCTTGGCAGCCGGAGCCCAGAAGTGCCATTCTGTGATCACACAGGACTTTCTG
ACTTGCTGGCTGAGCATCCGACAAGCAGAGCTGCTGCTCCCTGGGGCTGAGTTTCATCAC
TATGTGGGAGGACCTACGGAAACCCATGTTGTAAGGTCCCCACATCCCTACCAGCTTCCA
CAGGCCTTGGCCCCCCATGTGGACTTTGTGGGGGGACTGCACCGTTTTCCCCCAACATCA
TCCCTGAGGCAACGTCCTGAGCCGCAGGTGACAGGGACTGTAGGCCTGCATCTGGGGGTA
ACCCCCTCTGTGATCCGTAAGCGATACAACTTGACCTCACAAGACGTGGGCTCTGGCACC
AGCAATAACAGCCAAGCCTGTGCCCAGTTCCTGGAGCAGTATTTCCATGACTCAGACCTG
GCTCAGTTCATGCGCCTCTTCGGTGGCAACTTTGCACATCAGGCATCAGTAGCCCGTGTG
GTTGGACAACAGGGCCGGGGCCGGGCCGGGATTGAGGCCAGTCTAGATGTGCAGTACCTG
ATGAGTGCTGGTGCCAACATCTCCACCTGGGTCTACAGTAGCCCTGGCCGGCATGAGGGA
CAGGAGCCCTTCCTGCAGTGGCTCATGCTGCTCAGTAATGAGTCAGCCCTGCCACATGTG
CATACTGTGAGCTATGGAGATGATGAGGACTCCCTCAGCAGCGCCTACATCCAGCGGGTC
AACACTGAGCTCATGAAGGCTGCCGCTCGGGGTCTCACCCTGCTCTTCGCCTCAGGTGAC
AGTGGGGCCGGGTGTTGGTCTGTCTCTGGAAGACACCAGTTCCGCCCTACCTTCCCTGCC
TCCAGCCCCTATGTCACCACAGTGGGAGGCACATCCTTCCAGGAACCTTTCCTCATCACA
AATGAAATTGTTGACTATATCAGTGGTGGTGGCTTCAGCAATGTGTTCCCACGGCCTTCA
TACCAGGAGGAAGCTGTAACGAAGTTCCTGAGCTCTAGCCCCCACCTGCCACCATCCAGT
TACTTCAATGCCAGTGGCCGTGCCTACCCAGATGTGGCTGCACTTTCTGATGGCTACTGG
GTGGTCAGCAACAGAGTGCCCATTCCATGGGTGTCCGGAACCTCGGCCTCTACTCCAGTG
TTTGGGGGGATCCTATCCTTGATCAATGAGCACAGGATCCTTAGTGGCCGCCCCCCTCTT
GGCTTTCTCAACCCAAGGCTCTACCAGCAGCATGGGGCAGGTCTCTTTGATGTAACCCGT
GGCTGCCATGAGTCCTGTCTGGATGAAGAGGTAGAGGGCCAGGGTTTCTGCTCTGGTCCT
GGCTGGGATCCTGTAACAGGCTGGGGAACACCAACTTCCCAGCTTTGCTGAAGACTCTAC
TCAACCCCTGACCCTTTCCTATCAGGAGAGATGGCTTGTCCCCTGCCCTGAAGCTGGCAG
TTCAGTCCCTTATTCTGCCCTGTTGGAAGCCCTGCTGAACCCTCAACTATTGACTGCTGC
AGACAGCTTATCTCCCTAACCCTGAAATGCTGTGAGCTTGACTTGACTCCCAACCCTACC
ATGCTCCATCATACTCAGGTCTCCCTACTCCTGCCTTAGATTCCTCAATAAGATGCTGTA
ACTAGCATTTTTTGAATGCCTCTCCCTCCGCATCTCATCTTTCTCTTTTCAATCAGGCTT
TTCCAAAGGGTTGTATACAGACTCTGTGCACTATTTCACTTGATATTCATTCCCCAATTC
ACTGCAAGGAGACCTCTACTGTCACCGTTTACTCTTTCCTACCCTGACATCCAGAAACAA
TGGCCTCCAGTGCATACTTCTCAATCTTTGCTTTATGGCCTTTCCATCATAGTTGCCCAC
TCCCTCTCCTTACTTAGCTTCCAGGTCTTAACTTCTCTGACTACTCTTGTCTTCCTCTCT
CATCAATTTCTGCTTCTTCATGGAATGCTGACCTTCATTGCTCCATTTGTAGATTTTTGC
TCTTCTCAGTTTACTCATTGTCCCCTGGAACAAATCACTGACATCTACAACCATTACCAT
CTCACTAAATAAGACTTTCTATCCAATAATGATTGATACCTCAAATGTAAAAAA

FIGURE 94

><signal peptide>

MGLQACLLGLFALILS

><start mature protein>

GKCSYSPEPDQRRTLPPGWVSLGRADPEEELSLTFALRQQNVERLSELVQAVSDPSSP

QYGKYLTLENVADLVRPSPLTLHTVQKWLLAAGAQKCHSVITQDFLTCWLSIRQAEL

LLPGAEFHHYVGGPTETHVVRSPHPYQLPQALAPHVDFVGGLHRFPPTSSLRQRPEPQ

VTGTVGLHLGVTPSVIRKRY

><potential N-glycosylation site>

NLTSQDVGSGTS

><potential N-glycosylation site>

NNSQACAQFLEQYFHDSDLAQFMRLFGGNFAHQASVARVVGQQGRGRAGIEASLDV

QYLMSAGA

><potential N-glycosylation site>

NISTWVYSSPGRHEGQEPFLQWLMLLS

><potential N-glycosylation site>

NESALPHVHTVSYGDDEDSLSSAYIQRVNTELMKAAARGLTLLFASGDSGAGCWSVS

GRHQFRPTFPASSPYVTTVGGTSFQEPFLITNEIVDYISGGGFSNVFPRPSYQEEAVTKF

LSSSPHLPPSSYF

><potential N-glycosylation site>

NASGRAYPDVAALSDGYWVVSNRVPIPWVSGTSASTPVFGGILSLINEHRILSGRPPL

GFLNPRLYQQHGAGLFDVTRGCHESCLDEEVEGQGFCSGPGWDPVTGWGTPTSQLC

FIGURE 95

GCCGCGCGCTCTCTCCCGGCGCCCACACCTGTCTGAGCGGCGCAGCGAGCCGCGGCCCGG
GCGGGCTGCTCGGCGCGGAACAGTGCTCGGC
><MET {trans=1-s, dir=f, res=1}>
ATGGCAGGGATTCCAGGGCTCCTCTTCCTTCTCTTCTTTCTGCTCTGTGCTGTTGGGCAA
GTGAGCCCTTACAGTGCCCCCTGGAAACCCACTTGGCCTGCATACCGCCTCCCTGTCGTC
TTGCCCCAGTCTACCCTCAATTTAGCCAAGCCAGACTTTGGAGCCGAAGCCAAATTAGAA
GTATCTTCTTCATGTGGACCCCAGTGTCATAAGGGAACTCCACTGCCCACTTACGAAGAG
GCCAAGCAATATCTGTCTTATGAAACGCTCTATGCCAATGGCAGCCGCACAGAGACGCAG
GTGGGCATCTACATCCTCAGCAGTAGTGGAGATGGGGCCCAACACCGAGACTCAGGGTCT
TCAGGAAAGTCTCGAAGGAAGCGGCAGATTTATGGCTATGACAGCAGGTTCAGCATTTTT
GGGAAGGACTTCCTGCTCAACTACCCTTTCTCAACATCAGTGAAGTTATCCACGGGCTGC
ACCGGCACCCTGGTGGCAGAGAAGCATGTCCTCACAGCTGCCCACTGCATACACGATGGA
AAAACCTATGTGAAAGGAACCCAGAAGCTTCGAGTGGGCTTCCTAAAGCCCAAGTTTAAA
GATGGTGGTCGAGGGGCCAACGACTCCACTTCAGCCATGCCCGAGCAGATGAAATTTCAG
TGGATCCGGGTGAAACGCACCCATGTGCCCAAGGGTTGGATCAAGGGCAATGCCAATGAC
ATCGGCATGGATTATGATTATGCCCTCCTGGAACTCAAAAAGCCCCACAAGAGAAAATTT
ATGAAGATTGGGGTGAGCCCTCCTGCTAAGCAGCTGCCAGGGGGCAGAATTCACTTCTCT
GGTTATGACAATGACCGACCAGGCAATTTGGTGTATCGCTTCTGTGACGTCAAAGACGAG
ACCTATGACTTGCTCTACCAGCAATGCGATGCCCAGCCAGGGGCCAGCGGGTCTGGGGTC
TATGTGAGGATGTGGAAGAGACAGCAGCAGAAGTGGGAGCGAAAAATTATTGGCATTTTT
TCAGGGCACCAGTGGGTGGACATGAATGGTTCCCCACAGGATTTCAACGTGGCTGTCAGA
ATCACTCCTCTCAAATATGCCCAGATTTGCTATTGGATTAAAGGAAACTACCTGGATTGT
AGGGAGGGGTGACACAGTGTTCCCTCCTGGCAGCAATTAAGGGTCTTCATGTTCTTATTT
TAGGAGAGGCCAAATTGTTTTTTGTCATTGGCGTGCACACGTGTGTGTGTGTGTGTGTGT
GTGTGTAAGGTGTCTTATAATCTTTTACCTATTTCTTACAATTGCAAGATGACTGGCTTT
ACTATTTGAAAACTGGTTTGTGTATCATATCATATATCATTTAAGCAGTTTGAAGGCATA
CTTTTGCATAGAAATAAAAAAAATACTGATTTGGGGCAATGAGGAATATTTGACAATTAA
GTTAATCTTCACGTTTTTGCAAACTTTGATTTTTATTTCATCTGAACTTGTTTCAAAGAT
TTATATTAAATATTTGGCATACAAGAGATATGAAAAAAAAAAAAAAA

FIGURE 96

><signal peptide>

MAGIPGLLFLLFFLLCAVG

><start mature protein>

QVSPYSAPWKPTWPAYRLPVVLPQSTLNLAKPDFGAEAKLEVSSSCGPQCHKGTPLP

TYEEAKQYLSYETLYA

><potential N-glycosylation site>

NGSRTETQVGIYILSSSGDGAQHRDSGSSGKSRRKRQIYGYDSRFSIFGKDFLLNYPFS

TSVKLSTGCTGTLVAEKHV

><serine proteases, trypsin family, histidine active site 'LTAAHC'>

LTAAHCIHDGKTYVKGTQKLRVGFLKPKFKDGGRGA

><potential N-glycosylation site>

NDSTSAMPEQMKFQWIRVKRTHVPKGWIKGNANDIGMDYDYALLELKKPHKRKFM

KIGVSPPAKQLPGGRIHFSGYDNDRPGNLVYRFCDVKDETYDLLYQQCDAQPGASGS

GVYVRMWKRQQQKWERKIIGIFSGHQWVDMNGSPQDFNVAVRITPLKYAQICYWIK

GNYLDCREG

FIGURE 97

GCATCGCCCTGGGTCTCTCGAGCCTGCTGCCTGCTCCCCCGCCCCACCAGCC

><MET {trans=1-s, dir=f, res=1}

ATGGTGGTTTCTGGAGCGCCCCCAGCCCTGGGTGGGGGCTGTCTCGGCACCTTCACCTCC

CTGCTGCTGCTGGCGTCGACAGCCATCCTCAATGCGGCCAGGATACCTGTTCCCCCAGCC

TGTGGGAAGCCCCAGCAGCTGAACCGGGTTGTGGGCGGCGAGGACAGCACTGACAGCGAG

TGGCCCTGGATCGTGAGCATCCAGAAGAATGGGACCCACCACTGCGCAGGTTCTCTGCTC

ACCAGCCGCTGGGTGATCACTGCTGCCCACTGTTTCAAGGACAACCTGAACAAACCATAC

CTGTTCTCTGTGCTGCTGGGGGCCTGGCAGCTGGGGAACCCTGGCTCTCGGTCCCAGAAG

GTGGGTGTTGCCTGGGTGGAGCCCCACCCTGTGTATTCCTGGAAGGAAGGTGCCTGTGCA

GACATTGCCCTGGTGCGTCTCGAGCGCTCCATACAGTTCTCAGAGCGGGTCCTGCCCATC

TGCCTACCTGATGCCTCTATCCACCTCCCTCCAAACACCCACTGCTGGATCTCAGGCTGG

GGGAGCATCCAAGATGGAGTTCCCTTGCCCCACCCTCAGACCCTGCAGAAGCTGAAGGTT

CCTATCATCGACTCGGAAGTCTGCAGCCATCTGTACTGGCGGGGAGCAGGACAGGGACCC

ATCACTGAGGACATGCTGTGTGCCGGCTACTTGGAGGGGGAGCGGGATGCTTGTCTGGGC

GACTCCGGGGGCCCCCTCATGTGCCAGGTGGACGGCGCCTGGCTGCTGGCCGGCATCATC

AGCTGGGGCGAGGGCTGTGCCGAGCGCAACAGGCCCGGGGTCTACATCAGCCTCTCTGCG

CACCGCTCCTGGGTGGAGAAGATCGTGCAAGGGGTGCAGCTCCGCGGGCGCGCTCAGGGG

GGTGGGGCCCTCAGGGCACCGAGCCAGGGCTCTGGGGCCGCCGCGCGCTCCTAGGGCGCA

GCGGGACGCGGGGCTCGGATCTGAAAGGCGGCCAGATCCACATCTGGATCTGGATCTGCG

GCGGCCTCGGGCGGTTTCCCCCGCCGTAAATAGGCTCATCTACCTCTACCTCTGGGGGCC

CGGACGGCTGCTGCGGAAAGGAAACCCCCTCCCCGACCCGCCCGACGGCCTCAGGCCCCC

CTCCAAGGCATCAGGCCCCGCCCAACGGCCTCATGTCCCCGCCCCCACGACTTCCGGCCC

CGCCCCCCGGGCCCCAGCGCTTTTGTGTATATAAATGTTAATGATTTTTATAGGTATTTGT

AACCCTGCCCACATATCTTATTTATTCCTCCAATTTCAATAAATTATTTATTCTCCAAAA

AAAAAA

FIGURE 98

><signal peptide>

MVVSGAPPALGGGCLGTFTSLLLLASTAILNA

><start mature peptide>

ARIPVPPACGKPQQLNRVVGGEDSTDSEWPWIVSIQK

><potential N-glycosylation site>

NGTHHCAGSLLTSRWV

><Serine proteases active site ITAAHC>

ITAAHCFKDNLNKPYLFSVLLGAWQLGNPGSRSQKVGVAWVEPHPVYSWKEGACA

DIALVRLERSIQFSERVLPICLPDASIHLPPNTHCWISGWGSIQDGVPLPHPQTLQKLKV

PIIDSEVCSHLYWRGAGQGPITEDMLCAGYLEGERDACLGDSGGPLMCQVDGAWLL

AGIISWGEGCAERNRPGVYISLSAHRSWVEKIVQGVQLRGRAQGGGALRAPSQGSGA

AARS

FIGURE 99

GACGGCTGGCCACC
><MET {trans=1-s, dir=f, res=1}
ATGCACGGCTCCTGCAGTTTCCTGATGCTTCTGCTGCCGCTACTGCTACTGCTGGTGGCC
ACCACAGGCCCCGTTGGAGCCCTCACAGATGAGGAGAAACGTTTGATGGTGGAGCTGCAC
AACCTCTACCGGGCCCAGGTATCCCCGACGGCCTCAGACATGCTGCACATGAGATGGGAC
GAGGAGCTGGCCGCCTTCGCCAAGGCCTACGCACGGCAGTGCGTGTGGGGCCACAACAAG
GAGCGCGGGCGCCGCGGCGAGAATCTGTTCGCCATCACAGACGAGGGCATGGACGTGCCG
CTGGCCATGGAGGAGTGGCACCACGAGCGTGAGCACTACAACCTCAGCGCCGCCACCTGC
AGCCCAGGCCAGATGTGCGGCCACTACACGCAGGTGGTATGGGCCAAGACAGAGAGGATC
GGCTGTGGTTCCCACTTCTGTGAGAAGCTCCAGGGTGTTGAGGAGACCAACATCGAATTA
CTGGTGTGCAACTATGAGCCTCCGGGGAACGTGAAGGGGAAACGGCCCTACCAGGAGGGG
ACTCCGTGCTCCCAATGTCCCTCTGGCTACCACTGCAAGAACTCCCTCTGTGAACCCATC
GGAAGCCCGGAAGATGCTCAGGATTTGCCTTACCTGGTAACTGAGGCCCCATCCTTCCGG
GCGACTGAAGCATCAGACTCTAGGAAAATGGGTACTCCTTCTTCCCTAGCAACGGGGATT
CCGGCTTTCTTGGTAACAGAGGTCTCAGGCTCCCTGGCAACCAAGGCTCTGCCTGCTGTG
GAAACCCAGGCCCCAACTTCCTTAGCAACGAAAGACCCGCCCTCCATGGCAACAGAGGCT
CCACCTTGCGTAACAACTGAGGTCCCTTCCATTTTGGCAGCTCACAGCCTGCCCTCCTTG
GATGAGGAGCCAGTTACCTTCCCCAAATCGACCCATGTTCCTATCCCAAAATCAGCAGAC
AAAGTGACAGACAAAACAAAAGTGCCCTCTAGGAGCCCAGAGAACTCTCTGGACCCCAAG
ATGTCCCTGACAGGGGCAAGGGAACTCCTACCCCATGCCCAGGAGGAGGCTGAGGCTGAG
GCTGAGTTGCCTCCTTCCAGTGAGGTCTTGGCCTCAGTTTTTCCAGCCCAGGACAAGCCA
GGTGAGCTGCAGGCCACACTGGACCACACGGGGCACACCTCCTCCAAGTCCCTGCCCAAT
TTCCCCAATACCTCTGCCACCGCTAATGCCACGGGTGGGCGTGCCCTGGCTCTGCAGTCG
TCCTTGCCAGGTGCAGAGGGCCCTGACAAGCCTAGCGTTGTGTCAGGGCTGAACTCGGGC
CCTGGTCATGTGTGGGGCCCTCTCCTGGGACTACTGCTCCTGCCTCCTCTGGTGTTGGCT
GGAATCTTCTGAATGGGATACCACTCAAAGGGTGAAGAGGTCAGCTGTCCTCCTGTCATC
TTCCCCACCCTGTCCCCAGCCCCTAAACAAGATACTTCTTGGTTAAGGCCCTCCGGAAGG
GAAAGGCTACGGGGCATGTGCCTCATCACACCATCCATCCTGGAGGCACAAGGCCTGGCT
GGCTGCGAGCTCAGGAGGCCGCCTGAGGACTGCACACCGGGCCCACACCTCTCCTGCCCC
TCCCTCCTGAGTCCTGGGGGTGGGAGGATTTGAGGGAGCTCACTGCCTACCTGGCCTGGG
GCTGTCTGCCCACACAGCATGTGCGCTCTCCCTGAGTGCCTGTGTAGCTGGGGATGGGGA
TTCCTAGGGGCAGATGAAGGACAAGCCCCACTGGAGTGGGGTTCTTTGAGTGGGGGAGGC
AGGGACGAGGGAAGGAAAGTAACTCCTGACTCTCCAATAAAAACCTGTCCAACCTGTGAA
A

FIGURE 100

><signal peptide>

MHGSCSFLMLLLPLLLLLVATT

><start of mature peptide, extracellular domain>

GPVGALTDEEKRLMVELHNLYRAQVSPTASDMLHMRWDEELAAFAKAYARQCVW

GHNKERGRRGENLFAITDEGMDVPLAMEEWHHEREHY

><potential N-glycosylation site>

NLSAATCSPGQMC

><GHYTQVVWAKT Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 signature 1 - CRISP signature>

GHYTQVVWAKTERIGCGSHFCEKLQGVEETNIEL

><LLVCNYEPPGNV Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 signature 2 -CRISP signature>

LVCNYEPPGNVKGKRPYQEGTPCSQCPSGYHCKNSLCEPIGSPEDAQDLPYLVTEAPSFR

ATEASDSRKMGTPSSLATGIPAFLVTEVSGSLATKALPAVETQAPTSLATKDPPSMATEA

PPCVTTEVPSILAAHSLPSLDEEPVTFPKSTHVPIPKSADKVTDKTKVPSRSPENSLDPK

MSLTGARELLPHAQEEAEAEAELPPSSEVLASVFPAQDKPGELQATLDHTGHTSSKSLPN

FP

><potential N-glycosylation site>

NTSATA

><potential N-glycosylation site>

NATGGRALALQSSLPGAEGPDKPSVVSGLNS

><potential glycosylphosphatidylinositol attachment site>

GPGHVWGPLLGLLLLPPLVLAGIF

FIGURE 101A

GTAACTGAAGTCAGGCTTTTCATTTGGGAAGCCCCCTCAACAGAATTCGGTCATTCTCCA
AGTT
><MET {trans=1-s, dir=f, res=1}
ATGGTGGACGTACTTCTGTTGTTCTCCCTCTGCTTGCTTTTTTCACATTAGCAGACCGGAC
TTAAGTCACAACAGATTATCTTTCATCAAGGCAAGTTCCATGAGCCACCTTCAAAGCCTT
CGAGAAGTGAAACTGAACAACAATGAATTGGAGACCATTCCAAATCTGGGACCAGTCTCG
GCAAATATTACACTTCTCTCCTTGGCTGGAAACAGGATTGTTGAAATACTCCCTGAACAT
CTGAAAGAGTTTCAGTCCCTTGAAACTTTGGACCTTAGCAGCAACAATATTTCAGAGCTC
CAAACTGCATTTCCAGCCCTACAGCTCAAATATCTGTATCTCAACAGCAACCGAGTCACA
TCAATGGAACCTGGGTATTTTGACAATTTGGCCAACACACTCCTTGTGTTAAAGCTGAAC
AGGAACCGAATCTCAGCTATCCCACCCAAGATGTTTAAACTGCCCCAACTGCAACATCTC
GAATTGAACCGAAACAAGATTAAAAATGTAGATGGACTGACATTCCAAGGCCTTGGTGCT
CTGAAGTCTCTGAAAATGCAAAGAAATGGAGTAACGAAACTTATGGATGGAGCTTTTTGG
GGGCTGAGCAACATGGAAATTTTGCAGCTGGACCATAACAACCTAACAGAGATTACCAAA
GGCTGGCTTTACGGCTTGCTGATGCTGCAGGAACTTCATCTCAGCCAAAATGCCATCAAC
AGGATCAGCCCTGATGCCTGGGAGTTCTGCCAGAAGCTCAGTGAGCTGGACCTAACTTTC
AATCACTTATCAAGGTTAGATGATTCAAGCTTCCTTGGCCTAAGCTTACTAAATACACTG
CACATTGGGAACAACAGAGTCAGCTACATTGCTGATTGTGCCTTCCGGGGGCTTTCCAGT
TTAAAGACTTTGGATCTGAAGAACAATGAAATTTCCTGGACTATTGAAGACATGAATGGT
GCTTTCTCTGGGCTTGACAAACTGAGGCGACTGATACTCCAAGGAAATCGGATCCGTTCT
ATTACTAAAAAAGCCTTCACTGGTTTGGATGCATTGGAGCATCTAGACCTGAGTGACAAC
GCAATCATGTCTTTACAAGGCAATGCATTTTCACAAATGAAGAAACTGCAACAATTGCAT
TTAAATACATCAAGCCTTTTGTGCGATTGCCAGCTAAAATGGCTCCCACAGTGGGTGGCG
GAAAACAACTTTCAGAGCTTTGTAAATGCCAGTTGTGCCCATCCTCAGCTGCTAAAAGGA
AGAAGCATTTTTGCTGTTAGCCCAGATGGCTTTGTGTGTGATGATTTTCCCAAACCCCAG
ATCACGGTTCAGCCAGAAACACAGTCGGCAATAAAAGGTTCCAATTTGAGTTTCATCTGC
TCAGCTGCCAGCAGCAGTGATTCCCCAATGACTTTTGCTTGGAAAAAAGACAATGAACTA
CTGCATGATGCTGAAATGGAAAATTATGCACACCTCCGGGCCCAAGGTGGCGAGGTGATG
GAGTATACCACCATCCTTCGGCTGCGCGAGGTGGAATTTGCCAGTGAGGGGAAATATCAG
TGTGTCATCTCCAATCACTTTGGTTCATCCTACTCTGTCAAAGCCAAGCTTACAGTAAAT
ATGCTTCCCTCATTCACCAAGACCCCCATGGATCTCACCATCCGAGCTGGGGCCATGGCA
CGCTTGGAGTGTGCTGCTGTGGGGCACCCAGCCCCCCAGATAGCCTGGCAGAAGGATGGG
GGCACAGACTTCCCAGCTGCACGGGAGAGACGCATGCATGTGATGCCCGAGGATGACGTG
TTCTTTATCGTGGATGTGAAGATAGAGGACATTGGGGTATACAGCTGCACAGCTCAGAAC
AGTGCAGGAAGTATTTCAGCAAATGCAACTCTGACTGTCCTAGAAACACCATCATTTTTG
CGGCCACTGTTGGACCGAACTGTAACCAAGGGAGAAACAGCCGTCCTACAGTGCATTGCT
GGAGGAAGCCCTCCCCCTAAACTGAACTGGACCAAAGATGATAGCCCATTGGTGGTAACC
GAGAGGCACTTTTTTGCAGCAGGCAATCAGCTTCTGATTATTGTGGACTCAGATGTCAGT
GATGCTGGGAAATACACATGTGAGATGTCTAACACCCTTGGCACTGAGAGAGGAAACGTG
CGCCTCAGTGTGATCCCCACTCCAACCTGCGACTCCCCTCAGATGACAGCCCCATCGTTA
GACGATGACGGATGGGCCACTGTGGGTGTCGTGATCATAGCCGTGGTTTGCTGTGTGGTG
GGCACGTCACTCGTGTGGGTGGTCATCATATACCACACAAGGCGGAGGAATGAAGATTGC
AGCATTACCAACACAGATGAGACCAACTTGCCAGCAGATATTCCTAGTTATTTGTCATCT
CAGGGAACGTTAGCTGACAGGCAGGATGGGTACGTGTCTTCAGAAAGTGGAAGCCACCAC
CAGTTTGTCACATCTTCAGGTGCTGGATTTTTCTTACCACAACATGACAGTAGTGGGACC
TGCCATATTGACAATAGCAGTGAAGCTGATGTGGAAGCTGCCACAGATCTGTTCCTTTGT
CCGTTTTTGGGATCCACAGGCCCTATGTATTTGAAGGGAAATGTGTATGGCTCAGATCCT
TTTGAAACATATCATACAGGTTGCAGTCCTGACCCAAGAACAGTTTTAATGGACCACTAT
GAGCCCAGTTACATAAAGAAAAAGGAGTGCTACCCATGTTCTCATCCTTCAGAAGAATCC
TGCGAACGGAGCTTCAGTAATATATCGTGGCCTTCACATGTGAGGAAGCTACTTAACACT
AGTTACTCTCACAATGAAGGACCTGGAATGAAAAATCTGTGTCTAAACAAGTCCTCTTTA
GATTTTAGTGCAAATCCAGAGCCAGCGTCGGTTGCCTCGAGTAATTCTTTCATGGGTACC
TTTGGAAAAGCTCTCAGGAGACCTCACCTAGATGCCTATTCAAGCTTTGGACAGCCATCA

FIGURE 101B

GATTGTCAGCCAAGAGCCTTTTATTTGAAAGCTCATTCTTCCCCAGACTTGGACTCTGGG
TCAGAGGAAGATGGGAAAGAAAGGACAGATTTTCAGGAAGAAAATCACATTTGTACCTTT
AAACAGACTTTAGAAAACTACAGGACTCCAAATTTTCAGTCTTATGACTTGGACACATAG
ACTGAATGAGACCAAAGGAAAAGCTTAACATACTACCTCAAGTGAACTTTTATTTAAAAG
AGAGAGAATCTTATGTTTTTTAAATGGAGTTATGAATTTTAAAAGGATAAAAATGCTTTA
TTTATACAGATGAACCAAAATTACAAAAAGTTATGAAAATTTTTATACTGGGAATGATGC
TCATATAAGAATACCTTTTTAAACTATTTTTTAACTTTGTTTTATGCAAAAAAGTATCTT
ACGTAAATTAATGATATAAATCATGATTATTTTATGTATTTTTATAATGCCAGATTTCTT
TTTATGGAAAATGAGTTACTAAAGCATTTTAAATAATACCTGCCTTGTACCATTTTTTAA
ATAGAAGTTACTTCATTATATTTTGCACATTATATTTAATAAAATGTGTCAATTTGAA

FIGURE 102

MVDVLLLFSLCLLFHISRPDLSHNRLSFIKASSMSHLQSLREVKLNNNELETIPNLGPVS
ANITLLSLAGNRIVEILPEHLKEFQSLETLDLSSNNISELQTAFPALQLKYLYLNSNRVT
SMEPGYFDNLANTLLVLKLNRNRISAIPPKMFKLPQLQHLELNRNKIKNVDGLTFQGLGA
LKSLKMQRNGVTKLMDGAFWGLSNMEILQLDHNNLTEITKGWLYGLLMLQELHLSQNAIN
RISPDAWEFCQKLSELDLTFNHLSRLDDSSFLGLSLLNTLHIGNNRVSYIADCAFRGLSS
LKTLDLKNNEISWTIEDMNGAFSGLDKLRRLILQGNRIRSITKKAFTGLDALEHLDLSDN
AIMSLQGNAFSQMKKLQQLHLNTSSLLCDCQLKWLPQWVAENNFQSFVNASCAHPQLLKG
RSIFAVSPDGFVCDDFPKPQITVQPETQSAIKGSNLSFICSAASSSDSPMTFAWKKDNEL
LHDAEMENYAHLRAQGGEVMEYTTILRLREVEFASEGKYQCVISNHFGSSYSVKAKLTVN
MLPSFTKTPMDLTIRAGAMARLECAAVGHPAPQIAWQKDGGTDFPAARERRMHVMPEDDV
FFIVDVKIEDIGVYSCTAQNSAGSISANATLTVLETPSFLRPLLDRTVTKGETAVLQCIA
GGSPPPKLNWTKDDSPLVVTERHFFAAGNQLLIIVDSDVSDAGKYTCEMSNTLGTERGNV
RLSVIPTPTCDSPQMTAPSLDDDGWATVGVVIIAVVCCVVGTSLVWVVIIYHTRRRNEDC
SITNTDETNLPADIPSYLSSQGTLADRQDGYVSSESGSHHQFVTSSGAGFFLPQHDSSGT
CHIDNSSEADVEAATDLFLCPFLGSTGPMYLKGNVYGSDPFETYHTGCSPDPRTVLMDHY
EPSYIKKKECYPCSHPSEESCERSFSNISWPSHVRKLLNTSYSHNEGPGMKNLCLNKSSL
DFSANPEPASVASSNSFMGTFGKALRRPHLDAYSSFGQPSDCQPRAFYLKAHSSPDLDSG
SEEDGKERTDFQEENHICTFKQTLENYRTPNFQSYDLDT

FIGURE 103

GGGGAGAGGAATTGACCATGTAAAAGGAGACTTTTTTTTTTTGGTGGTGGTGGCTGTTGGG
TGCCTTGCAAAAATGAAGGATGCAGGACGCAGCTTTCTCCTGGAACCGAACGCAATGGAT
AAACTGATTGTGCAAGAGAGAAGGAAGAACGAAGCTTTTTCTTGTGAGCCCTGGATCTTA
ACACAAATGTGTATATGTGCACACAGGGAGCATTCAAGAATGAAATAAACCAGAGTTAGA
CCCGCGGGGGTTGGTGTGTTCTGACATAAATAAATAATCTTAAAGCAGCTGTTCCCCTCC
CCACCCCCAAAAAAAAGGATGATTGGAAATGAAGAACCGAGGATTCACAAAGAAAAAAGT
ATGTTCATTTTTCTCTATAAAGGAGAAAGTGAGCCAAGGAGATATTTTTGGAATGAAAAG
TTTGGGGCTTTTTTTAGTAAAGTAAAGAACTGGTGTGGTGGTGTTTTCCTTTCTTTTTGAA
TTTCCCACAAGAGGAGAGGAAATTAATAATACATCTGCAAAGAAATTTCAGAGAAGAAAA
GTTGACCGCGGCAGATTGAGGCATTGATTGGGGGAGAGAAACCAGCAGAGCACAGTTGGA
TTTGTGCCTATGTTGACTAAAATTGACGGATAATTGCAGTTGGATTTTTCTTCATCAACC
TCCTTTTTTTTTAAATTTTTATTCCTTTTGGTATCAAGATCATGCGTTTTCTCTTGTTCTT
AACCACCTGGATTTCCATCTGGATGTTGCTGTGATCAGTCTGAAATACAACTGTTTGAAT
TCCAGAAGGACCAACACCAGATAAATTATGA

><MET {trans=1-s, dir=f, res=1}

ATGTTGAACAAGATGACCTTACATCCACAGCAGATAATGATAGGTCCTAGGTTTAACAGG
GCCCTATTTGACCCCCTGCTTGTGGTGCTGCTGGCTCTTCAACTTCTTGTGGTGGCTGGT
CTGGTGCGGGCTCAGACCTGCCCTTCTGTGTGCTCCTGCAGCAACCAGTTCAGCAAGGTG
ATTTGTGTTCGGAAAAACCTGCGTGAGGTTCCGGATGGCATCTCCACCAACACACGGCTG
CTGAACCTCCATGAGAACCAAATCCAGATCATCAAAGTGAACAGCTTCAAGCACTTGAGG
CACTTGGAAATCCTACAGTTGAGTAGGAACCATATCAGAACCATTGAAATTGGGGCTTTC
AATGGTCTGGCGAACCTCAACACTCTGGAACTCTTTGACAATCGTCTTACTACCATCCCG
AATGGAGCTTTTGTATACTTGTCTAAACTGAAGGAGCTCTGGTTGCGAAACAACCCCATT
GAAAGCATCCCTTCTTATGCTTTTAACAGAATTCCTTCTTTGCGCCGACTAGACTTAGGG
GAATTGAAAAGACTTTCATACATCTCAGAAGGTGCCTTTGAAGGTCTGTCCAACTTGAGG
TATTTGAACCTTGCCATGTGCAACCTTCGGGAAATCCCTAACCTCACACCGCTCATAAAA
CTAGATGAGCTGGATCTTTCTGGGAATCATTTATCTGCCATCAGGCCTGGCTCTTTCCAG
GGTTTGATGCACCTTCAAAAACTGTGGATGATACAGTCCCAGATTCAAGTGATTGAACGG
AATGCCTTTGACAACCTTCAGTCACTAGTGGAGATCAACCTGGCACACAATAATCTAACA
TTACTGCCTCATGACCTCTTCACTCCCTTGCATCATCTAGAGCGGATACATTTACATCAC
AACCCTTGGAACTGTAACTGTGACATACTGTGGCTCAGCTGGTGGATAAAAGACATGGCC
CCCTCGAACACAGCTTGTTGTGCCCGGTGTAACACTCCTCCCAATCTAAAGGGGAGGTAC
ATTGGAGAGCTCGACCAGAATTACTTCACATGCTATGCTCCGGTGATTGTGGAGCCCCCT
GCAGACCTCAATGTCACTGAAGGCATGGCAGCTGAGCTGAAATGTCGGGCCTCCACATCC
CTGACATCTGTATCTTGGATTACTCCAAATGGAACAGTCATGACACATGGGGCGTACAAA
GTGCGGATAGCTGTGCTCAGTGATGGTACGTTAAATTTCACAAATGTAACTGTGCAAGAT
ACAGGCATGTACACATGTATGGTGAGTAATTCCGTTGGGAATACTACTGCTTCAGCCACC
CTGAATGTTACTGCAGCAACCACTACTCCTTTCTCTTACTTTTCAACCGTCACAGTAGAG
ACTATGGAACCGTCTCAGGATGAGGCACGGACCACAGATAACAATGTGGGTCCCACTCCA
GTGGTCGACTGGGAGACCACCAATGTGACCACCTCTCTCACACCACAGAGCACAAGGTCG
ACAGAGAAAACCTTCACCATCCCAGTGACTGATATAAACAGTGGGATCCCAGGAATTGAT
GAGGTCATGAAGACTACCAAAATCATCATTGGGTGTTTTGTGGCCATCACACTCATGGCT
GCAGTGATGCTGGTCATTTTCTACAAGATGAGGAAGCAGCACCATCGGCAAAACCATCAC
GCCCCAACAAGGACTGTTGAAATTATTAATGTGGATGATGAGATTACGGGAGACACACCC
ATGGAAAGCCACCTGCCCATGCCTGCTATCGAGCATGAGCACCTAAATCACTATAACTCA
TACAAATCTCCCTTCAACCACACAACAACAGTTAACACAATAAATTCAATACACAGTTCA
GTGCATGAACCGTTATTGATCCGAATGAACTCTAAAGACAATGTACAAGAGACTCAAATC
TAAAACATTTACAGAGTTACAAAAAACAAACAATCAAAAAAAAAAGACAGTTTATTAAAAA
TGACACAAATGACTGGGCTAAATCTACTGTTTCAAAAAAGTGTCTTTACAAAAAAACAAA
AAAGAAAAGAAATTTATTTATTAAAAATTCTATTGTGATCTAAAGCAGACAAAAA

FIGURE 104

><signal peptide>

MLNKMTLHPQQIMIGPRFNRALFDPLLVVLLALQLLVVAGLVRA

><start mature peptide, extracellular domain>

QTCPSVCSCSNQFSKVICVRKNLREVPDGISTNTRL

><start leucine rich repeat domains>

LNLHENQIQIIKVNSFKHLRHLEILQLSRNHIRTIEIGAFNGLANLNTLELFDNRLTTIPN
GAFVYLSKLKELWLRNNPIESIPSYAFNRIPSLRRLDLGELKRLSYISEGAFEGLSNLRY
LNLAMCNLREIPNLTPLIKLDELDLSGNHLSAIRPGSFQGLMHLQKLWMIQSQIQVIER
NAFDNLQSLVEINLAHNNLTLLPHDLFTPLHHLERIHLHHNPWNCNCDILWLSWWIK
DMAPSNTACCARCNTPPNLKGRYIGELDQNYFTCYAPVIVEPPADLNVTEGMAAELK
CRASTSLTSVSWITPNGTVMTHGAYKVRIAVLSDGTLNFTNVTVQDTGMYTCMVSN
SVGNTTASATLNVTAATTTPFSYFSTVTVETMEPSQDEARTTDNNVGPTPVVDWE

><end leucine rich repeat domains>

TTNVTTSLTPQSTRSTEKTFTIPVTDINSGIPGIDE

><start transmembrane domain>

VMKTTKIIIGCFVAITLMAAVMLVI

><start intracellular domain>

FYKMRKQHHRQNHHAPTRTVEIINVDDEITGDTPMESHLPMPAIEHEHLNHYNSYKS
PFNHTTTVNTINSIHSSVHEPLLIRMNSKDNVQETQI

FIGURE 105A

AGCCGACGCTGCTCAAGCTGCAACTCTGTTGCAGTTGGCAGTTCTTTTCGGTTTCCCTCCTG
CTGTTTGGGGGCATGAAAGGGCTTCGCCGCCGGGAGTAAAAGAAGGAATTGACCGGGCAGCG
CGAGGGAGGAGCGCGCACGCGACCGCGAGGGCGGGCGTGCACCCTCGGCTGGAAGTTTGTGC
CGGGCCCCGAGCGCGCGCCGGCTGGGAGCTTCGGGTAGAGACCTAGGCCGCTGGACCGCGAT
GAGCGCGCCGAGCCTCCGTGCGCGCGCCGCGGGGTTGGGGCTGCTGCTGTGCGCGGTGCTGG
GGCGCGCTGGCCGGTCCGACAGCGGCGGTCGCGGGGAACTCGGGCAGCCCTCTGGGGTAGCC
GCCGAGCGCCCATGCCCCACTACCTGCCGCTGCCTCGGGGACCTGCTGGACTGCAGTCGTAA
GCGGCTAGCGCGTCTTCCCGAGCCACTCCCGTCCTGGGTCGCTCGGCTGGACTTAAGTCACA
ACAGATTATCTTTCATCAAGGCAAGTTCCATGAGCCACCTTCAAAGCCTTCGAGAAGTGAAA
CTGAACAACAATGAATTGGAGACCATTCCAAATCTGGGACCAGTCTCGGCAAATATTACACT
TCTCTCCTTGGCTGGAAACAGGATTGTTGAAATACTCCCTGAACATCTGAAAGAGTTTCAGT
CCCTTGAAACTTTGGACCTTAGCAGCAACAATATTTCAGAGCTCCAAACTGCATTTCCAGCC
CTACAGCTCAAATATCTGTATCTCAACAGCAACCGAGTCACATCAATGGAACCTGGGTATTT
TGACAATTTGGCCAACACACTCCTTGTGTTAAAGCTGAACAGGAACCGAATCTCAGCTATCC
CACCCAAGATGTTTAAACTGCCCCAACTGCAACATCTCGAATTGAACCGAAACAAGATTAAA
AATGTAGATGGACTGACATTCCAAGGCCTTGGTGCTCTGAAGTCTCTGAAAATGCAAAGAAA
TGGAGTAACGAAACTTATGGATGGAGCTTTTTGGGGGCTGAGCAACATGGAAATTTTGCAGC
TGGACCATAACAACCTAACAGAGATTACCAAAGGCTGGCTTTACGGCTTGCTGATGCTGCAG
GAACTTCATCTCAGCCAAAATGCCATCAACAGGATCAGCCCTGATGCCTGGGAGTTCTGCCA
GAAGCTCAGTGAGCTGGACCTAACTTTCAATCACTTATCAAGGTTAGATGATTCAAGCTTCC
TTGGCCTAAGCTTACTAAATACACTGCACATTGGGAACAACAGAGTCAGCTACATTGCTGAT
TGTGCCTTCCGGGGGCTTTCCAGTTTAAAGACTTTGGATCTGAAGAACAATGAAATTTCCTG
GACTATTGAAGACATGAATGGTGCTTTCTCTGGGCTTGACAAACTGAGGCGACTGATACTCC
AAGGAAATCGGATCCGTTCTATTACTAAAAAAGCCTTCACTGGTTTGGATGCATTGGAGCAT
CTAGACCTGAGTGACAACGCAATCATGTCTTTACAAGGCAATGCATTTTCACAAATGAAGAA
ACTGCAACAATTGCATTTAAATACATCAAGCCTTTTGTGCGATTGCCAGCTAAAATGGCTCC
CACAGTGGGTGGCGGAAAACAACTTTCAGAGCTTTGTAAATGCCAGTTGTGCCCATCCTCAG
CTGCTAAAAGGAAGAAGCATTTTTGCTGTTAGCCCAGATGGCTTTGTGTGTGATGATTTTCC
CAAACCCCAGATCACGGTTCAGCCAGAAACACAGTCGGCAATAAAAGGTTCCAATTTGAGTT
TCATCTGCTCAGCTGCCAGCAGCAGTGATTCCCCAATGACTTTTGCTTGGAAAAAAGACAAT
GAACTACTGCATGATGCTGAAATGGAAAATTATGCACACCTCCGGGCCCAAGGTGGCGAGGT
GATGGAGTATACCACCATCCTTCGGCTGCGCGAGGTGGAATTTGCCAGTGAGGGGAAATATC
AGTGTGTCATCTCCAATCACTTTGGTTCATCCTACTCTGTCAAAGCCAAGCTTACAGTAAAT
ATGCTTCCCTCATTCACCAAGACCCCCATGGATCTCACCATCCGAGCTGGGGCCATGGCACG
CTTGGAGTGTGCTGCTGTGGGGCACCCAGCCCCCCAGATAGCCTGGCAGAAGGATGGGGGCA
CAGACTTCCCAGCTGCACGGGAGAGACGCATGCATGTGATGCCCGAGGATGACGTGTTCTTT
ATCGTGGATGTGAAGATAGAGGACATTGGGGTATACAGCTGCACAGCTCAGAACAGTGCAGG
AAGTATTTCAGCAAATGCAACTCTGACTGTCCTAGAAACACCATCATTTTTGCGGCCACTGT
TGGACCGAACTGTAACCAAGGGAGAAACAGCCGTCCTACAGTGCATTGCTGGAGGAAGCCCT
CCCCCTAAACTGAACTGGACCAAAGATGATAGCCCATTGGTGGTAACCGAGAGGCACTTTTT
TGCAGCAGGCAATCAGCTTCTGATTATTGTGGACTCAGATGTCAGTGATGCTGGGAAATACA
CATGTGAGATGTCTAACACCCTTGGCACTGAGAGAGGAAACGTGCGCCTCAGTGTGATCCCC
ACTCCAACCTGCGACTCCCCTCAGATGACAGCCCCATCGTTAGACGATGACGGATGGGCCAC
TGTGGGTGTCGTGATCATAGCCGTGGTTTGCTGTGTGGTGGGCACGTCACTCGTGTGGGTGG
TCATCATATACCACACAAGGCGGAGGAATGAAGATTGCAGCATTACCAACACAGATGAGACC
AACTTGCCAGCAGATATTCCTAGTTATTTGTCATCTCAGGGAACGTTAGCTGACAGGCAGGA
TGGGTACGTGTCTTCAGAAAGTGGAAGCCACCACCAGTTTGTCACATCTTCAGGTGCTGGAT
TTTTCTTACCACAACATGACAGTAGTGGGACCTGCCATATTGACAATAGCAGTGAAGCTGAT
GTGGAAGCTGCCACAGATCTGTTCCTTTGTCCGTTTTTGGGATCCACAGGCCCTATGTATTT
GAAGGGAAATGTGTATGGCTCAGATCCTTTTGAAACATATCATACAGGTTGCAGTCCTGACC
CAAGAACAGTTTTAATGGACCACTATGAGCCCAGTTACATAAAGAAAAAGGAGTGCTACCCA
TGTTCTCATCCTTCAGAAGAATCCTGCGAACGGAGCTTCAGTAATATATCGTGGCCTTCACA
TGTGAGGAAGCTACTTAACACTAGTTACTCTCACAATGAAGGACCTGGAATGAAAAATCTGT
GTCTAAACAAGTCCTCTTTAGATTTTAGTGCAAATCCAGAGCCAGCGTCGGTTGCCTCGAGT
AATTCTTTCATGGGTACCTTTGGAAAAGCTCTCAGGAGACCTCACCTAGATGCCTATTCAAG

FIGURE 105B

CTTTGGACAGCCATCAGATTGTCAGCCAAGAGCCTTTTATTTGAAAGCTCATTCTTCCCCAG
ACTTGGACTCTGGGTCAGAGGAAGATGGGAAAGAAAGGACAGATTTTCAGGAAGAAAATCAC
ATTTGTACCTTTAAACAGACTTTAGAAAACTACAGGACTCCAAATTTTCAGTCTTATGACTT
GGACACATAGACTGAATGAGACCAAAGGAAAAGCTTAACATACTACCTCAAGTGAACTTTTA
TTTAAAAGAGAGAGAATCTTATGTTTTTTAAATGGAGTTATGAATTTTAAAAGGATAAAAAT
GCTTTATTTATACAGATGAACCAAAATTACAAAAAGTTATGAAAATTTTTATACTGGGAATG
ATGCTCATATAAGAATACCTTTTTAAACTATTTTTTAACTTTGTTTTATGCAAAAAAGTATC
TTACGTAAATTAATGATATAAATCATGATTATTTTATGTATTTTTATAATGCCAGATTTCTT
TTTATGGAAAATGAGTTACTAAAGCATTTTAAATAATACCTGCCTTGTACCATTTTTTAAAT
AGAAGTTACTTCATTATATTTTGCACATTATATTTAATAAAATGTGTCAATTTGAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAA

FIGURE 106

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA37140
><subunit 1 of 1, 1119 aa, 1 stop
><MW: 123434, pI: 6.09, NX(S/T): 12
MSAPSLRARAAGLGLLLCAVLGRAGRSDSGGRGELGQPSGVAAERPCPTTCRCLGDLLDC
SRKRLARLPEPLPSWVARLDLSHNRLSFIKASSMSHLQSLREVKLNNNELETIPNLGPVS
ANITLLSLAGNRIVEILPEHLKEFQSLETLDLSSNNISELQTAFPALQLKYLYLNSNRVT
SMEPGYFDNLANTLLVLKLNRNRISAIPPKMFKLPQLQHLELNRNKIKNVDGLTFQGLGA
LKSLKMQRNGVTKLMDGAFWGLSNMEILQLDHNNLTEITKGWLYGLLMLQELHLSQNAIN
RISPDAWEFCQKLSELDLTFNHLSRLDDSSFLGLSLLNTLHIGNNRVSYIADCAFRGLSS
LKTLDLKNNEISWTIEDMNGAFSGLDKLRRLILQGNRIRSITKKAFTGLDALEHLDLSDN
AIMSLQGNAFSQMKKLQQLHLNTSSLLCDCQLKWLPQWVAENNFQSFVNASCAHPQLLKG
RSIFAVSPDGFVCDDFPKPQITVQPETQSAIKGSNLSFICSAASSSDSPMTFAWKKDNEL
LHDAEMENYAHLRAQGGEVMEYTTILRLREVEFASEGKYQCVISNHFGSSYSVKAKLTVN
MLPSFTKTPMDLTIRAGAMARLECAAVGHPAPQIAWQKDGGTDFPAARERRMHVMPEDDV
FFIVDVKIEDIGVYSCTAQNSAGSISANATLTVLETPSFLRPLLDRTVTKGETAVLQCIA
GGSPPPKLNWTKDDSPLVVTERHFFAAGNQLLIIVDSDVSDAGKYTCEMSNTLGTERGNV
RLSVIPTPTCDSPQMTAPSLDDDGWATVGVVIIAVVCCVVGTSLVWVVIIYHTRRRNEDC
SITNTDETNLPADIPSYLSSQGTLADRQDGYVSSESGSHHQFVTSSGAGFFLPQHDSSGT
CHIDNSSEADVEAATDLFLCPFLGSTGPMYLKGNVYGSDPFETYHTGCSPDPRTVLMDHY
EPSYIKKKECYPCSHPSEESCERSFSNISWPSHVRKLLNTSYSHNEGPGMKNLCLNKSSL
DFSANPEPASVASSNSFMGTFGKALRRPHLDAYSSFGQPSDCQPRAFYLKAHSSPDLDSG
SEEDGKERTDFQEENHICTFKQTLENYRTPNFQSYDLDT

FIGURE 107A

CAAAACTTGCGTCGCGGAGAGCGCCCAGCTTGACTTGAATGGAAGGAGCCCGAGCCCGCGGA
GCGCAGCTGAGACTGGGGGAGCGCGTTCGGCCTGTGGGGCGCCGCTCGGCGCCGGGGCGCAG
CAGGGAAGGGGAAGCTGTGGTCTGCCCTGCTCCACGAGGCGCCACTGGTGTGAACCGGGAGA
GCCCCTGGGTGGTCCCGTCCCCTATCCCTCCTTTATATAGAAACCTTCCACACTGGGAAGGC
AGCGGCGAGGCAGGAGGGCTCATGGTGAGCAAGGAGGCCGGCTGATCTGCAGGCGCACAGCA
TTCCGAGTTTACAGATTTTTACAGATACCAAATGGAAGGCGAGGAGGCAGAACAGCCTGCCT
GGTTCCATCAGCCCTGGCGCCCAGGCGCATCTGACTCGGCACCCCCTGCAGGCACCATGGCC
CAGAGCCGGGTGCTGCTGCTCCTGCTGCTGCTGCCGCCACAGCTGCACCTGGGACCTGTGCT
TGCCGTGAGGGCCCCAGGATTTGGCCGAAGTGGCGGCCACAGCCTGAGCCCCGAAGAGAACG
AATTTGCGGAGGAGGAGCCGGTGCTGGTACTGAGCCCTGAGGAGCCCGGGCCTGGCCCAGCC
GCGGTCAGCTGCCCCCGAGACTGTGCCTGTTCCCAGGAGGGCGTCGTGGACTGTGGCGGTAT
TGACCTGCGTGAGTTCCCGGGGGACCTGCCTGAGCACACCAACCACCTATCTCTGCAGAACA
ACCAGCTGGAAAAGATCTACCCTGAGGAGCTCTCCCGGCTGCACCGGCTGGAGACACTGAAC
CTGCAAAACAACCGCCTGACTTCCCGAGGGCTCCCAGAGAAGGCGTTTGAGCATCTGACCAA
CCTCAATTACCTGTACTTGGCCAATAACAAGCTGACCTTGGCACCCCGCTTCCTGCCAAACG
CCCTGATCAGTGTGGACTTTGCTGCCAACTATCTCACCAAGATCTATGGGCTCACCTTTGGC
CAGAAGCCAAACTTGAGGTCTGTGTACCTGCACAACAACAAGCTGGCAGACGCCGGGCTGCC
GGACAACATGTTCAACGGCTCCAGCAACGTCGAGGTCCTCATCCTGTCCAGCAACTTCCTGC
GCCACGTGCCCAAGCACCTGCCGCCTGCCCTGTACAAGCTGCACCTCAAGAACAACAAGCTG
GAGAAGATCCCCCCGGGGGCCTTCAGCGAGCTGAGCAGCCTGCGCGAGCTATACCTGCAGAA
CAACTACCTGACTGACGAGGGCCTGGACAACGAGACCTTCTGGAAGCTCTCCAGCCTGGAGT
ACCTGGATCTGTCCAGCAACAACCTGTCTCGGGTCCCAGCTGGGCTGCCGCGCAGCCTGGTG
CTGCTGCACTTGGAGAAGAACGCCATCCGGAGCGTGGACGCGAATGTGCTGACCCCCATCCG
CAGCCTGGAGTACCTGCTGCTGCACAGCAACCAGCTGCGGGAGCAGGGCATCCACCCACTGG
CCTTCCAGGGCCTCAAGCGGTTGCACACGGTGCACCTGTACAACAACGCGCTGGAGCGCGTG
CCCAGTGGCCTGCCTCGCCGCGTGCGCACCCTCATGATCCTGCACAACCAGATCACAGGCAT
TGGCCGCGAAGACTTTGCCACCACCTACTTCCTGGAGGAGCTCAACCTCAGCTACAACCGCA
TCACCAGCCCACAGGTGCACCGCGACGCCTTCCGCAAGCTGCGCCTGCTGCGCTCGCTGGAC
CTGTCGGGCAACCGGCTGCACACGCTGCCACCTGGGCTGCCTCGAAATGTCCATGTGCTGAA
GGTCAAGCGCAATGAGCTGGCTGCCTTGGCACGAGGGGCGCTGGCGGGCATGGCTCAGCTGC
GTGAGCTGTACCTCACCAGCAACCGACTGCGCAGCCGAGCCCTGGGCCCCCGTGCCTGGGTG
GACCTCGCCCATCTGCAGCTGCTGGACATCGCCGGGAATCAGCTCACAGAGATCCCCGAGGG
GCTCCCCGAGTCACTTGAGTACCTGTACCTGCAGAACAACAAGATTAGTGCGGTGCCCGCCA
ATGCCTTCGACTCCACGCCCAACCTCAAGGGGATCTTTCTCAGGTTTAACAAGCTGGCTGTG
GGCTCCGTGGTGGACAGTGCCTTCCGGAGGCTGAAGCACCTGCAGGTCTTGGACATTGAAGG
CAACTTAGAGTTTGGTGACATTTCCAAGGACCGTGGCCGCTTGGGGAAGGAAAAGGAGGAGG
AGGAAGAGGAGGAGGAGGAGGAAGAGGAAACAAGATAGTGACAAGGTGATGCAGATGTGACC
TAGGATGATGGACCGCCGGACTCTTTTCTGCAGCACACGCCTGTGTGCTGTGAGCCCCCCAC
TCTGCCGTGCTCACACAGACACACCCAGCTGCACACATGAGGCATCCCACATGACACGGGCT
GACACAGTCTCATATCCCCACCCCTTCCCACGGCGTGTCCCACGGCCAGACACATGCACACA
CATCACACCCTCAAACACCCAGCTCAGCCACACACAACTACCCTCCAAACCACCACAGTCTC
TGTCACACCCCCACTACCGCTGCCACGCCCTCTGAATCATGCAGGGAAGGGTCTGCCCCTGC
CCTGGCACACACAGGCACCCATTCCCTCCCCCTGCTGACATGTGTATGCGTATGCATACACA
CCACACACACACACATGCACAAGTCATGTGCGAACAGCCCTCCAAAGCCTATGCCACAGACA
GCTCTTGCCCCCAGCCAGAATCAGCCATAGCAGCTCGCCGTCTGCCCTGTCCATCTGTCCGTC
CGTTCCCTGGAGAAGACACAAGGGTATCCATGCTCTGTGGCCAGGTGCCTGCCACCCTCTGG
AACTCACAAAAGCTGGCTTTTATTCCTTTCCCATCCTATGGGGACAGGAGCCTTCAGGACTG
CTGGCCTGGCCTGGCCCACCCTGCTCCTCCAGGTGCTGGGCAGTCACTCTGCTAAGAGTCCC
TCCCTGCCACGCCCTGGCAGGACACAGGCACTTTTCCAATGGGCAAGCCCAGTGGAGGCAGG
ATGGGAGAGCCCCCTGGGTGCTGCTGGGGCCTTGGGGCAGGAGTGAAGCAGAGGTGATGGGG

FIGURE 107B

CTGGGCTGAGCCAGGGAGGAAGGACCCAGCTGCACCTAGGAGACACCTTTGTTCTTCAGGCC
TGTGGGGGAAGTTCCGGGTGCCTTTATTTTTTATTCTTTTCTAAGGAAAAAAATGATAAAAA
TCTCAAAGCTGATTTTTCTTGTTATAGAAAAACTAATATAAAAGCATTATCCCTATCCCTGC
AAAAAAAAAA

FIGURE 108

Met Glu Gly Glu Glu Ala Glu Gln Pro Ala Trp Phe His Gln Pro Trp
Arg Pro Gly Ala Ser Asp Ser Ala Pro Pro Ala Gly Thr Met Ala Gln
Ser Arg Val Leu Leu Leu Leu Leu Leu Leu Pro Pro Gln Leu His Leu
Gly Pro Val Leu Ala Val Arg Ala Pro Gly Phe Gly Arg Ser Gly Gly
His Ser Leu Ser Pro Glu Glu Asn Glu Phe Ala Glu Glu Glu Pro Val
Leu Val Leu Ser Pro Glu Glu Pro Gly Pro Gly Pro Ala Ala Val Ser
Cys Pro Arg Asp Cys Ala Cys Ser Gln Glu Gly Val Val Asp Cys Gly
Gly Ile Asp Leu Arg Glu Phe Pro Gly Asp Leu Pro Glu His Thr Asn
His Leu Ser Leu Gln Asn Asn Gln Leu Glu Lys Ile Tyr Pro Glu Glu
Leu Ser Arg Leu His Arg Leu Glu Thr Leu Asn Leu Gln Asn Asn Arg
Leu Thr Ser Arg Gly Leu Pro Glu Lys Ala Phe Glu His Leu Thr Asn
Leu Asn Tyr Leu Tyr Leu Ala Asn Asn Lys Leu Thr Leu Ala Pro Arg
Phe Leu Pro Asn Ala Leu Ile Ser Val Asp Phe Ala Ala Asn Tyr Leu
Thr Lys Ile Tyr Gly Leu Thr Phe Gly Gln Lys Pro Asn Leu Arg Ser
Val Tyr Leu His Asn Asn Lys Leu Ala Asp Ala Gly Leu Pro Asp Asn
Met Phe Asn Gly Ser Ser Asn Val Glu Val Leu Ile Leu Ser Ser Asn
Phe Leu Arg His Val Pro Lys His Leu Pro Pro Ala Leu Tyr Lys Leu
His Leu Lys Asn Asn Lys Leu Glu Lys Ile Pro Pro Gly Ala Phe Ser
Glu Leu Ser Ser Leu Arg Glu Leu Tyr Leu Gln Asn Asn Tyr Leu Thr
Asp Glu Gly Leu Asp Asn Glu Thr Phe Trp Lys Leu Ser Ser Leu Glu
Tyr Leu Asp Leu Ser Ser Asn Asn Leu Ser Arg Val Pro Ala Gly Leu
Pro Arg Ser Leu Val Leu Leu His Leu Glu Lys Asn Ala Ile Arg Ser
Val Asp Ala Asn Val Leu Thr Pro Ile Arg Ser Leu Glu Tyr Leu Leu
Leu His Ser Asn Gln Leu Arg Glu Gln Gly Ile His Pro Leu Ala Phe
Gln Gly Leu Lys Arg Leu His Thr Val His Leu Tyr Asn Asn Ala Leu
Glu Arg Val Pro Ser Gly Leu Pro Arg Arg Val Arg Thr Leu Met Ile
Leu His Asn Gln Ile Thr Gly Ile Gly Arg Glu Asp Phe Ala Thr Thr
Tyr Phe Leu Glu Glu Leu Asn Leu Ser Tyr Asn Arg Ile Thr Ser Pro
Gln Val His Arg Asp Ala Phe Arg Lys Leu Arg Leu Leu Arg Ser Leu
Asp Leu Ser Gly Asn Arg Leu His Thr Leu Pro Pro Gly Leu Pro Arg
Asn Val His Val Leu Lys Val Lys Arg Asn Glu Leu Ala Ala Leu Ala
Arg Gly Ala Leu Ala Gly Met Ala Gln Leu Arg Glu Leu Tyr Leu Thr
Ser Asn Arg Leu Arg Ser Arg Ala Leu Gly Pro Arg Ala Trp Val Asp
Leu Ala His Leu Gln Leu Leu Asp Ile Ala Gly Asn Gln Leu Thr Glu
Ile Pro Glu Gly Leu Pro Glu Ser Leu Glu Tyr Leu Tyr Leu Gln Asn
Asn Lys Ile Ser Ala Val Pro Ala Asn Ala Phe Asp Ser Thr Pro Asn
Leu Lys Gly Ile Phe Leu Arg Phe Asn Lys Leu Ala Val Gly Ser Val
Val Asp Ser Ala Phe Arg Arg Leu Lys His Leu Gln Val Leu Asp Ile
Glu Gly Asn Leu Glu Phe Gly Asp Ile Ser Lys Asp Arg Gly Arg Leu
Gly Lys Glu Lys Glu Glu Glu Glu Glu Glu Glu Glu Glu Glu Glu Glu
Thr Arg

FIGURE 109

GGGAGGGGGCTCCGGGCGCCGCGCAGCAGACCTGCTCCGGCCGCGCGCCTCGCCGCTGTC
CTCCGGGAGCGGCAGCAGTAGCCCGGGCGGCGAGGGCTGGGGGTTCCTCGAGACTCTCAG
AGGGGCGCCTCCCATCGGCGCCCACCACCCCAACCTGTTCCTCGCGCGCCACTGCGCTGC
GCCCCAGGACCCGCTGCCCAACATGGATTTTCTCCTGGCGCTGGTGCTGGTATCCTCGCT
CTACCTGCAGGCGGCCGCCGAGTTCGACGGGAGGTGGCCCAGGCAAATAGTGTCATCGAT
TGGCCTATGTCGTTATGGTGGGAGGATTGACTGCTGCTGGGGCTGGGCTCGCCAGTCTTG
GGGACAGTGTCAGCCTGTGTGCCAACCACGATGCAAACATGGTGAATGTATCGGGCCAAA
CAAGTGCAAGTGTCATCCTGGTTATGCTGGAAAAACCTGTAATCAAGATCTAAATGAGTG
TGGCCTGAAGCCCCGGCCCTGTAAGCACAGGTGCATGAACACTTACGGCAGCTACAAGTG
CTACTGTCTCAACGGATATATGCTCATGCCGGATGGTTCCTGCTCAAGTGCCCTGACCTG
CTCCATGGCAAACTGTCAGTATGGCTGTGATGTTGTTAAAGGACAAATACGGTGCCAGTG
CCCATCCCCTGGCCTGCACCTGGCTCCTGATGGGAGGACCTGTGTAGATGTTGATGAATG
TGCTACAGGAAGAGCCTCCTGCCCTAGATTTAGGCAATGTGTCAACACTTTTGGGAGCTA
CATCTGCAAGTGTCATAAAGGCTTCGATCTCATGTATATTGGAGGCAAATATCAATGTCA
TGACATAGACGAATGCTCACTTGGTCAGTATCAGTGCAGCAGCTTTGCTCGATGTTATAA
CGTACGTGGGTCCTACAAGTGCAAATGTAAAGAAGGATACCAGGGTGATGGACTGACTTG
TGTGTATATCCCAAAAGTTATGATTGAACCTTCAGGTCCAATTCATGTACCAAAGGGAAA
TGGTACCATTTTAAAGGGTGACACAGGAAATAATAATTGGATTCCTGATGTTGGAAGTAC
TTGGTGGCCTCCGAAGACACCATATATTCCTCCTATCATTACCAACAGGCCTACTTCTAA
GCCAACAACAAGACCTACACCAAAGCCAACACCAATTCCTACTCCACCACCACCACCACC
CCTGCCAACAGAGCTCAGAACACCTCTACCACCTACAACCCCAGAAAGGCCAACCACCGG
ACTGACAACTATAGCACCAGCTGCCAGTACACCTCCAGGAGGGATTACAGTTGACAACAG
GGTACAGACAGACCCTCAGAAACCCAGAGGAGATGTGTTCAGTGTTCTGGTACACAGTTG
TAATTTTGACCATGGACTTTGTGGATGGATCAGGGAGAAAGACAATGACTTGCACTGGGA
ACCAATCAGGGACCCAGCAGGTGGACAATATCTGACAGTGTCGGCAGCCAAAGCCCCAGG
GGGAAAAGCTGCACGCTTGGTGCTACCTCTCGGCCGCCTCATGCATTCAGGGGACCTGTG
CCTGTCATTCAGGCACAAGGTGACGGGGCTGCACTCTGGCACACTCCAGGTGTTTGTGAG
AAAACACGGTGCCCACGGAGCAGCCCTGTGGGGAAGAAATGGTGGCCATGGCTGGAGGCA
AACACAGATCACCTTGCGAGGGGCTGACATCAAGAGCGAATCACAAAGATGATTAAAGGG
TTGGAAAAAAAGATCTATGATGGAAAATTAAAGGAACTGGGATTATTGAGCCTGGAGAAG
AGAAGACTGAGGGGCAAACCATTGATGGTTTTCAAGTATATGAAGGGTTGGCACAGAGAG
GGTGGCGACCAGCTGTTCTCCATATGCACTAAGAATAGAACAAGAGGAAACTGGCTTAGA
CTAGAGTATAAGGGAGCATTTCTTGGCAGGGGCCATTGTTAGAATACTTCATAAAAAAAG
AAGTGTGAAAATCTCAGTATCTCTCTCTCTTTCTAAAAAATTAGATAAAAATTTGTCTAT
TTAAGATGGTTAAAGATGTTCTTACCCAAGGAAAAGTAACAAATTATAGAATTTCCCAAA
AGATGTTTTGATCCTACTAGTAGTATGCAGTGAAAATCTTTAGAACTAAATAATTTGGAC
AAGGCTTAATTTAGGCATTTCCCTCTTGACCTCCTAATGGAGAGGGATTGAAAGGGGAAG
AGCCCACCAAATGCTGAGCTCACTGAAATATCTCTCCCTTATGGCAATCCTAGCAGTATT
AAAGAAAAAAGGAAACTATTTATTCCAAATGAGAGTATGATGGACAGATATTTTAGTATC
TCAGTAATGTCCTAGTGTGGCGGTGGTTTTCAATGTTTCTTCATGGTAAAGGTATAAGCC
TTTCATTTGTTCAATGGATGATGTTTCAGATTTTTTTTTTTTTAAGAGATCCTTCAAGGA
ACACAGTTCAGAGAGATTTTCATCGGGTGCATTCTCTCTGCTTCGTGTGTGACAAGTTAT
CTTGGCTGCTGAGAAAGAGTGCCCTGCCCCACACCGGCAGACCTTTCCTTCACCTCATCA
GTATGATTCAGTTTCTCTTATCAATTGGACTCTCCCAGGTTCCACAGAACAGTAATATTT
TTTGAACAATAGGTACAATAGAAGGTCTTCTGTCATTTAACCTGGTAAAGGCAGGGCTGG
AGGGGGAAAATAAATCATTAAGCCTTTGAGTAACGGCAGAATATATGGCTGTAGATCCAT
TTTTAATGGTTCATTTCCTTTATGGTCATATAACTGCACAGCTGAAGATGAAAGGGGAAA
ATAAATGAAAATTTTACTTTTCGATGCCAATGATACATTGCACTAAACTGATGGAAGAAG
TTATCCAAAGTACTGTATAACATCTTGTTTATTATTTAATGTTTTCTAAAATAAAAAATG
TTAGTGGTTTTCCAAATGGCCTAATAAAAACAATTATTTGTAAATAAAAACACTGTTAGT
AAT

FIGURE 110

><signal peptide>
MDFLLALVLVSSLYLQA
><start mature protein>
AAEFDGRWPRQIVSSIGLCRYGGRIDCCWGWARQSWGQCQPV
><EGF-like repeats, 60-253>
CQPRCKHGECIGPNKCKCHPGYAGKTCNQDLNECGLKPRPCKHRCMNTYGSYKCYCLNGYML
MPDGSCSSALTCSMANCQYGCDVVKGQIRCQCPSPGLHLAPDGRTCVDVDECATGRASCPRF
RQCVNTFGSYICKCHKGFDLMYIGGKYQCHDIDECSLGQYQCSSFARCYNVRGSYKCKCKEG
YQGDGLTCVYIPKVMIEPSGPIHVPKG
><potential N-glycosylation site>
NGTILKGDTGNNNWIPDVGSTWWPPKTPYIPPIITNRPTSKPTTRPTPKPTPIPTPPPPPPL
PTELRTPLPPTTPERPTTGLTTIAPAASTPPGGITVDNRVQTDPQKPRGDVFSVLVHSCNFD
HGLCGWIREKDNDLHWEPIRDPAGGQYLTVSAAKAPGGKAARLVLPLGRLMHSGDLCLSFRH
KVTGLHSGTLQVFVRKHGAHGAALWGRNGGHGWRQTQITLRGADIKSESQR

FIGURE 111

CTTCTTTGAAAAGGATTATCACCTGATCAGGTTCTCTCTGCATTTGCCCCTTTAGATTGTGA
AATGTGGCTCAAGGTCTTCACAACTTTCCTTTCCTTTGCAACAGGTGCTTGCTCGGGGCTGA
AGGTGACAGTGCCATCACACACTGTCCATGGCGTCAGAGGTCAGGCCCTCTACCTACCCGTC
CACTATGGCTTCCACACTCCAGCATCAGACATCCAGATCATATGGCTATTTGAGAGACCCCA
CACAATGCCCAAATACTTACTGGGCTCTGTGAATAAGTCTGTGGTTCCTGACTTGGAATACC
AACACAAGTTCACCATGATGCCACCCAATGCATCTCTGCTTATCAACCCACTGCAGTTCCCT
GATGAAGGCAATTACATCGTGAAGGTCAACATTCAGGGAAATGGAACTCTATCTGCCAGTCA
GAAGATACAAGTCACGGTTGATGATCCTGTCACAAAGCCAGTGGTGCAGATTCATCCTCCCT
CTGGGGCTGTGGAGTATGTGGGGAACATGACCCTGACATGCCATGTGGAAGGGGGCACTCGG
CTAGCTTACCAATGGCTAAAAAATGGGAGACCTGTCCACACCAGCTCCACCTACTCCTTTTC
TCCCCAAAACAATACCCTTCATATTGCTCCAGTAACCAAGGAAGACATTGGGAATTACAGCT
GCCTGGTGAGGAACCCTGTCAGTGAAATGGAAAGTGATATCATTATGCCCATCATATATTAT
GGACCTTATGGACTTCAAGTGAATTCTGATAAAGGGCTAAAAGTAGGGGAAGTGTTTACTGT
TGACCTTGGAGAGGCCATCCTATTTGATTGTTCTGCTGATTCTCATCCCCCCAACACCTACT
CCTGGATTAGGAGGACTGACAATACTACATATATCATTAAGCATGGGCCTCGCTTAGAAGTT
GCATCTGAGAAAGTAGCCCAGAAGACAATGGACTATGTGTGCTGTGCTTACAACAACATAAC
CGGCAGGCAAGATGAAACTCATTTCACAGTTATCATCACTTCCGTAGGACTGGAGAAGCTTG
CACAGAAAGGAAAATCATTGTCACCTTTAGCAAGTATAACTGGAATATCACTATTTTTGATT
ATATCCATGTGTCTTCTCTTCCTATGGAAAAAATATCAACCCTACAAAGTTATAAAACAGAA
ACTAGAAGGCAGGCCAGAAACAGAATACAGGAAAGCTCAAACATTTTCAGGCCATGAAGATG
CTCTGGATGACTTCGGAATATATGAATTTGTTGCTTTTCCAGATGTTTCTGGTGTTTCCAGG
ATTCCAAGCAGGTCTGTTCCAGCCTCTGATTGTGTATCGGGGCAAGATTTGCACAGTACAGT
GTATGAAGTTATTCAGCACATCCCTGCCCAGCAGCAAGACCATCCAGAGTGAACTTTCATGG
GCTAAACAGTACATTCGAGTGAAATTCTGAAGAAACATTTTAAGGAAAAACAGTGGAAAAGT
ATATTAATCTGGAATCAGTGAAGAAACCAGGACCAACACCTCTTACTCATTATTCCTTTACA
TGCAGAATAGAGGCATTTATGCAAATTGAACTGCAGGTTTTTCAGCATATACACAATGTCTT
GTGCAACAGAAAAACATGTTGGGGAAATATTCCTCAGTGGAGAGTCGTTCTCATGCTGACGG
GGAGAACGAAAGTGACAGGGGTTTCCTCATAAGTTTTGTATGAAATATCTCTACAAACCTCA
ATTAGTTCTACTCTACACTTTCACTATCATCAACACTGAGACTATCCTGTCTCACCTACAAA
TGTGGAAACTTTACATTGTTCGATTTTTCAGCAGACTTTGTTTTATTAAATTTTTATTAGTG
TTAAGAATGCTAAATTTATGTTTCAATTTTATTTCCAAATTTCTATCTTGTTATTTGTACAA
CAAAGTAATAAGGATGGTTGTCACAAAAACAAAACTATGCCTTCTCTTTTTTTTCAATCACC
AGTAGTATTTTTGAGAAGACTTGTGAACACTTAAGGAAATGACTATTAAAGTCTTATTTTTA
TTTTTTTCAAGGAAAGATGGATTCAAATAAATTATTCTGTTTTTGCTTTTAAAAAAAAAAAAA

FIGURE 112

Met Trp Leu Lys Val Phe Thr Thr Phe Leu Ser Phe Ala Thr Gly Ala
Cys Ser Gly Leu Lys Val Thr Val Pro Ser His Thr Val His Gly Val
Arg Gly Gln Ala Leu Tyr Leu Pro Val His Tyr Gly Phe His Thr Pro
Ala Ser Asp Ile Gln Ile Ile Trp Leu Phe Glu Arg Pro His Thr Met
Pro Lys Tyr Leu Leu Gly Ser Val Asn Lys Ser Val Val Pro Asp Leu
Glu Tyr Gln His Lys Phe Thr Met Met Pro Pro Asn Ala Ser Leu Leu
Ile Asn Pro Leu Gln Phe Pro Asp Glu Gly Asn Tyr Ile Val Lys Val
Asn Ile Gln Gly Asn Gly Thr Leu Ser Ala Ser Gln Lys Ile Gln Val
Thr Val Asp Asp Pro Val Thr Lys Pro Val Val Gln Ile His Pro Pro
Ser Gly Ala Val Glu Tyr Val Gly Asn Met Thr Leu Thr Cys His Val
Glu Gly Gly Thr Arg Leu Ala Tyr Gln Trp Leu Lys Asn Gly Arg Pro
Val His Thr Ser Ser Thr Tyr Ser Phe Ser Pro Gln Asn Asn Thr Leu
His Ile Ala Pro Val Thr Lys Glu Asp Ile Gly Asn Tyr Ser Cys Leu
Val Arg Asn Pro Val Ser Glu Met Glu Ser Asp Ile Ile Met Pro Ile
Ile Tyr Tyr Gly Pro Tyr Gly Leu Gln Val Asn Ser Asp Lys Gly Leu
Lys Val Gly Glu Val Phe Thr Val Asp Leu Gly Glu Ala Ile Leu Phe
Asp Cys Ser Ala Asp Ser His Pro Pro Asn Thr Tyr Ser Trp Ile Arg
Arg Thr Asp Asn Thr Thr Tyr Ile Ile Lys His Gly Pro Arg Leu Glu
Val Ala Ser Glu Lys Val Ala Gln Lys Thr Met Asp Tyr Val Cys Cys
Ala Tyr Asn Asn Ile Thr Gly Arg Gln Asp Glu Thr His Phe Thr Val
Ile Ile Thr Ser Val Gly Leu Glu Lys Leu Ala Gln Lys Gly Lys Ser
Leu Ser Pro Leu Ala Ser Ile Thr Gly Ile Ser Leu Phe Leu Ile Ile
Ser Met Cys Leu Leu Phe Leu Trp Lys Lys Tyr Gln Pro Tyr Lys Val
Ile Lys Gln Lys Leu Glu Gly Arg Pro Glu Thr Glu Tyr Arg Lys Ala
Gln Thr Phe Ser Gly His Glu Asp Ala Leu Asp Asp Phe Gly Ile Tyr
Glu Phe Val Ala Phe Pro Asp Val Ser Gly Val Ser Arg Ile Pro Ser
Arg Ser Val Pro Ala Ser Asp Cys Val Ser Gly Gln Asp Leu His Ser
Thr Val Tyr Glu Val Ile Gln His Ile Pro Ala Gln Gln Gln Asp His
Pro Glu

FIGURE 113

GCAAGCGGCGAA
><MET {trans=1-s, dir=f, res=1}
ATGGCGCCCTCCGGGAGTCTTGCAGTTCCCCTGGCAGTCCTGGTGCTGTTGCTTTGGGGT
GCTCCCTGGACGCACGGGCGGCGGAGCAACGTTCGCGTCATCACGGACGAGAACTGGAGA
GAACTGCTGGAAGGAGACTGGATGATAGAATTTTATGCCCCGTGGTGCCCTGCTTGTCAA
AATCTTCAACCGGAATGGGAAAGTTTTGCTGAATGGGGAGAAGATCTTGAGGTTAATATT
GCGAAAGTAGATGTCACAGAGCAGCCAGGACTGAGTGGACGGTTTATCATAACTGCTCTT
CCTACTATTTATCATTGTAAAGATGGTGAATTTAGGCGCTATCAGGGTCCAAGGACTAAG
AAGGACTTCATAAACTTTATAAGTGATAAAGAGTGGAAGAGTATTGAGCCCGTTTCATCA
TGGTTTGGTCCAGGTTCTGTTCTGATGAGTAGTATGTCAGCACTCTTTCAGCTATCTATG
TGGATCAGGACGTGCCATAACTACTTTATTGAAGACCTTGGATTGCCAGTGTGGGGATCA
TATACTGTTTTTGCTTTAGCAACTCTGTTTTCCGGACTGTTATTAGGACTCTGTATGATA
TTTGTGGCAGATTGCCTTTGTCCTTCAAAAAGGCGCAGACCACAGCCATACCCATACCCT
TCAAAAAAATTATTATCAGAATCTGCACAACCTTTGAAAAAAGTGGAGGAGGAACAAGAG
GCGGATGAAGAAGATGTTTCAGAAGAAGAAGCTGAAAGTAAAGAAGGAACAAACAAAGAC
TTTCCACAGAATGCCATAAGACAACGCTCTCTGGGTCCATCATTGGCCACAGATAAATCC
TAGTTAAATTTTATAGTTATCTTAATATTATGATTTTGATAAAAACAGAAGATTGATCAT
TTTGTTTGGTTTGAAGTGAACTGTGACTTTTTTGAATATTGCAGGGTTCAGTCTAGATTG
TCATTAAATTGAAGAGTCTACATTCAGAACATAAAAGCACTAGGTATACAAGTTTGAAAT
ATGATTTAAGCACAGTATGATGGTTTAAATAGTTCTCTAATTTTTGAAAAATCGTGCCAA
GCAATAAGATTTATGTATATTTGTTTAATAATAACCTATTTCAAGTCTGAGTTTTGAAAA
TTTACATTTCCCAAGTATTGCATTATTGAGGTATTTAAGAAGATTATTTTAGAGAAAAAT
ATTTCTCATTTGATATAATTTTTCTCTGTTTCACTGTGTGAAAAAAAGAAGATATTTCCC
ATAAATGGGAAGTTTGCCCATTGTCTCAAGAAATGTGTATTTCAGTGACAATTTCGTGGT
CTTTTTAGAGGTATATTCCAAAATTTCCTTGTATTTTTAGGTTATGCAACTAATAAAAAC
TACCTTACATTAATTAATTACAGTTTTCTACACATGGTAATACAGGATATGCTACTGATT
TAGGAAGTTTTTAAGTTCATGGTATTCTCTTGATTCCAACAAAGTTTGATTTTCTCTTGT
ATTTTTCTTACTTACTATGGGTTACATTTTTTATTTTTCAAATTGGATGATAATTTCTTG
GAAACATTTTTTATGTTTTAGTAAACAGTATTTTTTTGTTGTTTCAAACTGAAGTTTACT
GAGAGATCCATCAAATTGAACAATCTGTTGTAATTTAAAATTTTGGCCACTTTTTTCAGA
TTTTACATCATTCTTGCTGAACTTCAACTTGAAATTGTTTTTTTTTTCCTTTTGGATGTG
AAGGTGAACATTCCTGATTTTTGTCTGATGTGAAAAAGCCTTGGTATTTTACATTTTGAA
AATTCAAAGAAGCTTAATATAAAAGTTTGCATTCTACTCAGGAAAAAGCATCTTCTTGTA
TATGTCTTAAATGTATTTTTGTCCTCATATACAGAAAGTTCTTAATTGATTTTACAGTCT
GTAATGCTTGATGTTTTAAAATAATAACATTTTTATATTTTTTAAAAGACAAACTTCATA
TTATCCTGTGTTCTTTCCTGACTGGTAATATTGTGTGGGATTTCACAGGTAAAAGTCAGT
AGGATGGAACATTTTAGTGTATTTTTACTCCTTAAAGAGCTAGAATACATAGTTTTCACC
TTAAAAGAAGGGGGAAAATCATAAATACAATGAATCAACTGACCATTACGTAGTAGACAA
TTTCTGTAATGTCCCCTTCTTTCTAGGCTCTGTTGCTGTGTGAATCCATTAGATTTACAG
TATCGTAATATACAAGTTTTCTTTAAAGCCCTCTCCTTTAGAATTTAAAATATTGTACCA
TTAAAGAGTTTGGATGTGTAACTTGTGATGCCTTAGAAAAATATCCTAAGCACAAAATAA
ACCTTTCTAACCACTTCATTAAAGCTGAAAAAAAAAAAAAAAAAA

FIGURE 114

MAPSGSLAVPLAVLVLLLWGAPWTHGRRSNVRVITDENWRELLEGDWMIEFYAPWCPACQ
NLQPEWESFAEWGEDLEVNIAKVDVTEQPGLSGRFIITALPTIYHCKDGEFRRYQGPRTK
KDFINFISDKEWKSIEPVSSWFGPGSVLMSSMSALFQLSMWIRTCHNYFIEDLGLPVWGS
YTVFALATLFSGLLLGLCMIFVADCLCPSKRRRPQPYPYPSKKLLSESAQPLKKVEEEQE
ADEEDVSEEEAESKEGTNKDFPQNAIRQRSLGPSLATDKS

FIGURE 115

GCGAGTGTCCAGCTGCGGAGACCCGTGATAATTCGTTAACTAATTCAACAAACGGGACCC
TTCTGTGTGCCAGAAACCGCAAGCAGTTGCTAACCCAGTGGGACAGGCGGATTGGAAGAG
CGGGAAGGTCCTGGCCCAGAGCAGTGTGACACTTCCCTCTGTGACC
><MET {trans=1-s, dir=f, res=1}
ATGAAACTCTGGGTGTCTGCATTGCTGATGGCCTGGTTTGGTGTCCTGAGCTGTGTGCAG
GCCGAATTCTTCACCTCTATTGGGCACATGACTGACCTGATTTATGCAGAGAAAGAGCTG
GTGCAGTCTCTGAAAGAGTACATCCTTGTGGAGGAAGCCAAGCTTTCCAAGATTAAGAGC
TGGGCCAACAAAATGGAAGCCTTGACTAGCAAGTCAGCTGCTGATGCTGAGGGCTACCTG
GCTCACCCTGTGAATGCCTACAAACTGGTGAAGCGGCTAAACACAGACTGGCCTGCGCTG
GAGGACCTTGTCCTGCAGGACTCAGCTGCAGGTTTTATCGCCAACCTCTCTGTGCAGCGG
CAGTTCTTCCCCACTGATGAGGACGAGATAGGAGCTGCCAAAGCCCTGATGAGACTTCAG
GACACATACAGGCTGGACCCAGGCACAATTTCCAGAGGGGAACTTCCAGGAACCAAGTAC
CAGGCAATGCTGAGTGTGGATGACTGCTTTGGGATGGGCCGCTCGGCCTACAATGAAGGG
GACTATTATCATACGGTGTTGTGGATGGAGCAGGTGCTAAAGCAGCTTGATGCCGGGGAG
GAGGCCACCACAACCAAGTCACAGGTGCTGGACTACCTCAGCTATGCTGTCTTCCAGTTG
GGTGATCTGCACCGTGCCCTGGAGCTCACCCGCCGCCTGCTCTCCCTTGACCCAAGCCAC
GAACGAGCTGGAGGGAATCTGCGGTACTTTGAGCAGTTATTGGAGGAAGAGAGAGAAAAA
ACGTTAACAAATCAGACAGAAGCTGAGCTAGCAACCCCAGAAGGCATCTATGAGAGGCCT
GTGGACTACCTGCCTGAGAGGGATGTTTACGAGAGCCTCTGTCGTGGGGAGGGTGTCAAA
CTGACACCCCGTAGACAGAAGAGGCTTTTCTGTAGGTACCACCATGGCAACAGGGCCCCA
CAGCTGCTCATTGCCCCCTTCAAAGAGGAGGACGAGTGGGACAGCCCGCACATCGTCAGG
TACTACGATGTCATGTCTGATGAGGAAATCGAGAGGATCAAGGAGATCGCAAAACCTAAA
CTTGCACGAGCCACCGTTCGTGATCCCAAGACAGGAGTCCTCACTGTCGCCAGCTACCGG
GTTTCCAAAAGCTCCTGGCTAGAGGAAGATGATGACCCTGTTGTGGCCCGAGTAAATCGT
CGGATGCAGCATATCACAGGGTTAACAGTAAAGACTGCAGAATTGTTACAGGTTGCAAAT
TATGGAGTGGGAGGACAGTATGAACCGCACTTCGACTTCTCTAGGCGACCTTTTGACAGC
GGCCTCAAAACAGAGGGGAATAGGTTAGCGACGTTTCTTAACTACATGAGTGATGTAGAA
GCTGGTGGTGCCACCGTCTTCCCTGATCTGGGGGCTGCAATTTGGCCTAAGAAGGGTACA
GCTGTGTTCTGGTACAACCTCTTGCGGAGCGGGGAAGGTGACTACCGAACAAGACATGCT
GCCTGCCCTGTGCTTGTGGGCTGCAAGTGGGTCTCCAATAAGTGGTTCCATGAACGAGGA
CAGGAGTTCTTGAGACCTTGTGGATCAACAGAAGTTGACTGACATCCTTTTCTGTCCTTC
CCCTTCCTGGTCCTTCAGCCCATGTCAACGTGACAGACACCTTTGTATGTTCCTTTGTAT
GTTCCTATCAGGCTGATTTTTGGAGAAATGAATGTTTGTCTGGAGCAGAGGGAGACCATA
CTAGGGCGACTCCTGTGTGACTGAAGTCCCAGCCCTTCCATTCAGCCTGTGCCATCCCTG
GCCCCAAGGCTAGGATCAAAGTGGCTGCAGCAGAGTTAGCTGTCTAGCGCCTAGCAAGGT
GCCTTTGTACCTCAGGTGTTTTAGGTGTGAGATGTTTCAGTGAACCAAAGTTCTGATACC
TTGTTTACATGTTTGTTTTTATGGCATTTCTATCTATTGTGGCTTTACCAAAAAATAAAA
TGTCCCTACCAGAAAAAAAAA

FIGURE 116

><signal peptide>

MKLWVSALLMAWFGVLS

><start mature protein>

CVQAEFFTSIGHMTDLIYAEKELVQSLKEYILVEEAKLSKIKSWANKMEALTSKSAADAE

GYLAHPVNAYKLVKRLNTDWPALEDLVLQDSAAGFIA

><N-glycosylation site.>

NLSVQRQFFPTDEDEIGAAKALMRLQDTYRLDPGTISRGELPGTKYQAMLSVDDCF

GMGRSAYNEGDYYHTVLWMEQVLKQLDAGEEATTTKSQVLDY

><potential Leucine zipper 213-234>

LSYAVFQLGDLHRALELTRRLLSLDPSHERAGGNLRYFEQLLEEEREKTLT

><N-glycosylation site.>

NQTEAELATPEGIYERPVDYLPERDVYESLCRGEGVKLTPRRQKRLFCRYHHGNRAP

QLLIAPFKEEDEWDSPHIVRYYDVMSDEEIERIKEIAKPKLARATVRDPKTGVLTVAS

YRVSKSSWLEEDDDPVVARVNRRMQHITGLTVKTAELLQVANYGVGGQYEPHFDFS

RRPFDSGLKTEGNRLATFLNYMSDVEAGGATVFPDLGAAIWPKKGTAVFWYNLLRS

GEGDYRTRHAACPVLVGCKWVSNKWFHERGQEFLRPCGSTEVD

FIGURE 117

GCAGTATTGAGTTTTACTTCCTCCTCTTTTTAGTGGAAGACAGACCATAATCCCAGTGTGAG
TGAAATTGATTGTTTCATTTATTACCGTTTTGGCTGGGGGTTAGTTCCGACACCTTCACAGT
TGAAGAGCAGGCAGAAGGAGTTGTGAAGACAGGACAATCTTCTTGGGGATGCTGGTCCTGGA
AGCCAGCGGGCCTTGCTCTGTCTTTGGCCTCATTGACCCCAGGTTCTCTGGTTAAAACTGAA
AGCCTACTACTGGCCTGGTGCCCATCAATCCATTGATCCTTGAGGCTGTGCCCCTGGGGCAC
CCACCTGGCAGGGCCTACCACCATGCGACTGAGCTCCCTGTTGGCTCTGCTGCGGCCAGCGC
TTCCCCTCATCTTAGGGCTGTCTCTGGGGTGCAGCCTGAGCCTCCTGCGGGTTTCCTGGATC
CAGGGGGAGGGAGAAGATCCCTGTGTCGAGGCTGTAGGGGAGCGAGGAGGGCCACAGAATCC
AGATTCGAGAGCTCGGCTAGACCAAAGTGATGAAGACTTCAAACCCCGGATTGTCCCCTACT
ACAGGGACCCCAACAAGCCCTACAAGAAGGTGCTCAGGACTCGGTACATCCAGACAGAGCTG
GGCTCCCGTGAGCGGTTGCTGGTGGCTGTCCTGACCTCCCGAGCTACACTGTCCACTTTGGC
CGTGGCTGTGAACCGTACGGTGGCCCATCACTTCCCTCGGTTACTCTACTTCACTGGGCAGC
GGGGGGCCCGGGCTCCAGCAGGGATGCAGGTGGTGTCTCATGGGGATGAGCGGCCCGCCTGG
CTCATGTCAGAGACCCTGCGCCACCTTCACACACACTTTGGGGCCGACTACGACTGGTTCTT
CATCATGCAGGATGACACATATGTGCAGGCCCCCCGCCTGGCAGCCCTTGCTGGCCACCTCA
GCATCAACCAAGACCTGTACTTAGGCCGGGCAGAGGAGTTCATTGGCGCAGGCGAGCAGGCC
CGGTACTGTCATGGGGGCTTTGGCTACCTGTTGTCACGGAGTCTCCTGCTTCGTCTGCGGCC
ACATCTGGATGGCTGCCGAGGAGACATTCTCAGTGCCCGTCCTGACGAGTGGCTTGGACGCT
GCCTCATTGACTCTCTGGGCGTCGGCTGTGTCTCACAGCACCAGGGGCAGCAGTATCGCTCA
TTTGAACTGGCCAAAAATAGGGACCCTGAGAAGGAAGGGAGCTCGGCTTTCCTGAGTGCCTT
CGCCGTGCACCCTGTCTCCGAAGGTACCCTCATGTACCGGCTCCACAAACGCTTCAGCGCTC
TGGAGTTGGAGCGGGCTTACAGTGAAATAGAACAACTGCAGGCTCAGATCCGGAACCTGACC
GTGCTGACCCCCGAAGGGGAGGCAGGGCTGAGCTGGCCCGTTGGGCTCCCTGCTCCTTTCAC
ACCACACTCTCGCTTTGAGGTGCTGGGCTGGGACTACTTCACAGAGCAGCACACCTTCTCCT
GTGCAGATGGGGCTCCCAAGTGCCCACTACAGGGGCTAGCAGGGCGGACGTGGGTGATGCG
TTGGAGACTGCCCTGGAGCAGCTCAATCGGCGCTATCAGCCCCGCCTGCGCTTCCAGAAGCA
GCGACTGCTCAACGGCTATCGGCGCTTCGACCCAGCACGGGGCATGGAGTACACCCTGGACC
TGCTGTTGGAATGTGTGACACAGCGTGGGCACCGGCGGGCCCTGGCTCGCAGGGTCAGCCTG
CTGCGGCCACTGAGCCGGGTGGAAATCCTACCTATGCCCTATGTCACTGAGGCCACCCGAGT
GCAGCTGGTGCTGCCACTCCTGGTGGCTGAAGCTGCTGCAGCCCCGGCTTTCCTCGAGGCGT
TTGCAGCCAATGTCCTGGAGCCACGAGAACATGCATTGCTCACCCTGTTGCTGGTCTACGGG
CCACGAGAAGGTGGCCGTGGAGCTCCAGACCCATTTCTTGGGGTGAAGGCTGCAGCAGCGGA
GTTAGAGCGACGGTACCCTGGGACGAGGCTGGCCTGGCTCGCTGTGCGAGCAGAGGCCCCTT
CCCAGGTGCGACTCATGGACGTGGTCTCGAAGAAGCACCCTGTGGACACTCTCTTCTTCCTT
ACCACCGTGTGGACAAGGCCTGGGCCCGAAGTCCTCAACCGCTGTCGCATGAATGCCATCTC
TGGCTGGCAGGCCTTCTTTCCAGTCCATTTCCAGGAGTTCAATCCTGCCCTGTCACCACAGA
GATCACCCCCAGGGCCCCCGGGGGCTGGCCCTGACCCCCCCTCCCCTCCTGGTGCTGACCCC
TCCCGGGGGGCTCCTATAGGGGGGAGATTTGACCGGCAGGCTTCTGCGGAGGGCTGCTTCTA
CAACGCTGACTACCTGGCGGCCCGAGCCCGGCTGGCAGGTGAACTGGCAGGCCAGGAAGAGG
AGGAAGCCCTGGAGGGGCTGGAGGTGATGGATGTTTTCCTCCGGTTCTCAGGGCTCCACCTC
TTTCGGGCCGTAGAGCCAGGGCTGGTGCAGAAGTTCTCCCTGCGAGACTGCAGCCCACGGCT
CAGTGAAGAACTCTACCACCGCTGCCGCCTCAGCAACCTGGAGGGGCTAGGGGGCCGTGCCC
AGCTGGCTATGGCTCTCTTTGAGCAGGAGCAGGCCAATAGCACTTAGCCCGCCTGGGGGCCC
TAACCTCATTACCTTTCCTTTGTCTGCCTCAGCCCCAGGAAGGGCAAGGCAAGATGGTGGAC
AGATAGAGAATTGTTGCTGTATTTTTTAAATATGAAAATGTTATTAAACATGTCTTCTGCC

FIGURE 118

><signal peptide>

MRLSSLLALLRPALP

><start mature protein>

LILGLSLGCSLSLLRVSWIQGEGEDPCVEAVGERGGPQNPDSRARLDQSDEDFKPRIVPY

YRDPNKPYKKVLRTRYIQTELGSRERLLVAVLTSRATLSTLAVAV

><potential N-glycosylation site>

NRTVAHHFPRLLYFTGQRGARAPAGMQVVSHGDERPAWLMSETLRHLHTHFGADYD

><homology to radical fringe -177-241>

WFFIMQDDTYVQAPRLAALAGHLSINQDLYLGRAEEFIGAGEQARYCHGGFGYLLSRSLLLR

LRPHLDGCRGDILSARPDEWLGRCLIDSLGVGCVSQHQGQQYRSFELAKNRDPEKEGSSAFL

SAFAVHPVSEGTLMYRLHKRFSALELERAYSEIEQLQAQIR

><potential N-glycosylation site>

NLTVLTPEGEAGLSWPVGLPAPFTPHSRFEVLGWDYFTEQHTFSCADGAPKCPLQGASRADV

GDALETALEQLNRRYQPRLRFQKQRLLNGYRRFDPARGMEYTLDLLLECVTQRGHRRALARR

VSLLRPLSRVEILPMPYVTEATRVQLVLPLLVAEAAAAPAFLEAFAANVLEPREHALLTLLL

VYGPREGGRGAPDPFLGVKAAAAELERRYPGTRLAWLAVRAEAPSQVRLMDVVSKKHPVDTL

FFLTTVWTRPGPEVLNRCRMNAISGWQAFFPVHFQEFNPALSPQRSPPGPPGAGPDPPSPPG

ADPSRGAPIGGRFDRQASAEGCFYNADYLAARARLAVNWQARKRRKPLEGLEVMDVFLRFSG

LHLFRAVEPGLVQKFSLRDCSPRLSEELYHRCRLSNLEGLGGRAQLAMALFEQEQANST

FIGURE 119

CGGAGTGGTGCGCCAACGTGAGAGGAAACCCGTGCGCGGCTGCGCTTTCCTGTCCCCAAGCC
GTTCTAGACGCGGGAAAAATGCTTTCTGAAAGCAGCTCCTTTTTGAAGGGTGTGATGCTTGG
AAGCATTTTCTGTGCTTTGATCACTATGCTAGGACACATTAGGATTGGTCATGGAAATAGAA
TGCACCACCATGAGCATCATCACCTACAAGCTCCTAACAAAGAAGATATCTTGAAAATTTCA
GAGGATGAGCGCATGGAGCTCAGTAAGAGCTTTCGAGTATACTGTATTATCCTTGTAAAACC
CAAAGATGTGAGTCTTTGGGCTGCAGTAAAGGAGACTTGGACCAAACACTGTGACAAAGCAG
AGTTCTTCAGTTCTGAAAATGTTAAAGTGTTTGAGTCAATTAATATGGACACAAATGACATG
TGGTTAATGATGAGAAAAGCTTACAAATACGCCTTTGATAAGTATAGAGACCAATACAACTG
GTTCTTCCTTGCACGCCCCACTACGTTTGCTATCATTGAAAACCTAAAGTATTTTTTGTTAA
AAAAGGATCCATCACAGCCTTTCTATCTAGGCCACACTATAAAATCTGGAGACCTTGAATAT
GTGGGTATGGAAGGAGGAATTGTCTTAAGTGTAGAATCAATGAAAAGACTTAACAGCCTTCT
CAATATCCCAGAAAAGTGTCCTGAACAGGGAGGGATGATTTGGAAGATATCTGAAGATAAAC
AGCTAGCAGTTTGCCTGAAATATGCTGGAGTATTTGCAGAAAATGCAGAAGATGCTGATGGA
AAAGATGTATTTAATACCAAATCTGTTGGGCTTTCTATTAAAGAGGCAATGACTTATCACCC
CAACCAGGTAGTAGAAGGCTGTTGTTCAGATATGGCTGTTACTTTTAATGGACTGACTCCAA
ATCAGATGCATGTGATGATGTATGGGGTATACCGCCTTAGGGCATTTGGGCATATTTTCAAT
GATGCATTGGTTTTCTTACCTCCAAATGGTTCTGACAATGACTGAGAAGTGGTAGAAAAGCG
TGAATATGATCTTTGTATAGGACGTGTGTTGTCATTATTTGTAGTAGTAACTACATATCCAA
TACAGCTGTATGTTTCTTTTTCTTTTCTAATTTGGTGGCACTGGTATAACCACACATTAAAG
TCAGTAGTACATTTTTAAATGAGGGTGGTTTTTTTCTTTAAAACACATGAACATTGTAAATG
TGTTGGAAAGAAGTGTTTTAAGAATAATAATTTTGCAAATAAACTATTAATAAATATTATAT
GTGATAAATTCTAAATTATGAACATTAGAAATCTGTGGGGCACATATTTTTGCTGATTGGTT
AAAAAATTTTAACAGGTCTTTAGCGTTCTAAGATATGCAAATGATATCTCTAGTTGTGAATT
TGTGATTAAAGTAAAACTTTTAGCTGTGTGTTCCCTTTACTTCTAATACTGATTTATGTTCT
AAGCCTCCCCAAGTTCCAATGGATTTGCCTTCTCAAAATGTACAACTAAGCAACTAAAGAAA
ATTAAAGTGAAAGTTGAAAAAT

FIGURE 120

><signal peptide>

MLSESSSFLKGVMLGSIFCALITMLGHIRIGHG

><start mature protein>

NRMHHHEHHHLQAPNKEDILKISEDERMELSKSFRVYCIILVKPKDVSLWAAVKETW

TKHCDKAEFFSSENVKVFESINMDTNDMWLMMRKAYKYAFDKYRDQYNWFFLARP

TTFAIIENLKYFLLKKDPSQPFYLGHTIKSGDLEYVGMEGGIVLSVESMKRLNSLLNIP

EKCPEQGGMIWKISEDKQLAVCLKYAGVFAENAEDADGKDVFNTKSVGLSIKEAMT

YHPNQVVEGCCSDMAVTFNGLTPNQMHVMMYGVYRLRAFGHIFNDALVFLPP

><potential N-glycosylation site>

NGSDND

FIGURE 121

CCCACGCGTCCGATCTTACCAACAAAACACTCCTGAGGAGAAAGAAAGAGAGGGAGGGAGAG
AAAAAGAGAGAGAGAGAAACAAAAAACCAAAGAGAGAGAAAAAATGAATTCATCTAAATCAT
CTGAAACACAATGCACAGAGAGAGGATGCTTCTCTTCCCAAATGTTCTTATGGACTGTTGCT
GGGATCCCCATCCTATTTCTCAGTGCCTGTTTCATCACCAGATGTGTTGTGACATTTCGCAT
CTTTCAAACCTGTGATGAGAAAAAGTTTCAGCTACCTGAGAATTTCACAGAGCTCTCCTGCT
ACAATTATGGATCAGGTTCAGTCAAGAATTGTTGTCCATTGAACTGGGAATATTTTCAATCC
AGCTGCTACTTCTTTTCTACTGACACCATTTCCTGGGCGTTAAGTTTAAAGAACTGCTCAGC
CATGGGGGCTCACCTGGTGGTTATCAACTCACAGGAGGAGCAGGAATTCCTTTCCTACAAGA
AACCTAAAATGAGAGAGTTTTTTATTGGACTGTCAGACCAGGTTGTCGAGGGTCAGTGGCAA
TGGGTGGACGGCACACCTTTGACAAAGTCTCTGAGCTTCTGGGATGTAGGGGAGCCCAACAA
CATAGCTACCCTGGAGGACTGTGCCACCATGAGAGACTCTTCAAACCCAAGGCAAAATTGGA
ATGATGTAACCTGTTTCCTCAATTATTTTCGGATTTGTGAAATGGTAGGAATAAATCCTTTG
AACAAAGGAAAATCTCTTTAAGAACAGAAGGCACAACTCAAATGTGTAAAGAAGGAAGAGCA
AGAACATGGCCACACCCACCGCCCCACACGAGAAATTTGTGCGCTGAACTTCAAAGGACTTC
ATAAGTATTTGTTACTCTGATACAAATAAAAATAAGTAGTTTTAAATGTTAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAA

FIGURE 122

><subunit 1 of 1, 219 aa, 0 stop

><MW: 25072, pI: 5.16, NX(S/T): 3

MNSSKSSETQCTERGCFSSQMFLWTVAGIPILFLSACFITRCVVTFRIFQTCDEKKFQLPEN
FTELSCYNYGSGSVKNCCPLNWEYFQSSCYFFSTDTISWALSLKNCSAMGAHLVVINSQEEQ
EFLSYKKPKMREFFIGLSDQVVEGQWQWVDGTPLTKSLSFWDVGEPNNIATLEDCATMRDSS
NPRQNWNDVTCFLNYFRICEMVGINPLNKGKSL

GGCCGGAGCAGCACGGCCGCAGGACCTGGAGCTCCGGCTGCGTCTTCCCGCAGCGCTACCCG
CCATGCGCCTGCCGCGCCGGGCCGCGCTGGGGCTCCTGCCGCTTCTGCTGCTGCTGCCGCCC
GCGCCGGAGGCCGCCAAGAAGCCGACGCCCTGCCACCGGTGCCGGGGGCTGGTGGACAAGTT
TAACCAGGGGATGGTGGACACCGCAAAGAAGAACTTTGGCGGCGGGAACACGGCTTGGGAGG
AAAAGACGCTGTCCAAGTACGAGTCCAGCGAGATTCGCCTGCTGGAGATCCTGGAGGGGCTG
TGCGAGAGCAGCGACTTCGAATGCAATCAGATGCTAGAGGCGCAGGAGGAGCACCTGGAGGC
CTGGTGGCTGCAGCTGAAGAGCGAATATCCTGACTTATTCGAGTGGTTTTGTGTGAAGACAC
TGAAAGTGTGCTGCTCTCCAGGAACCTACGGTCCCGACTGTCTCGCATGCCAGGGCGGATCC
CAGAGGCCCTGCAGCGGGAATGGCCACTGCAGCGGAGATGGGAGCAGACAGGGCGACGGGTC
CTGCCGGTGCCACATGGGGTACCAGGGCCCGCTGTGCACTGACTGCATGGACGGCTACTTCA
GCTCGCTCCGGAACGAGACCCACAGCATCTGCACAGCCTGTGACGAGTCCTGCAAGACGTGC
TCGGGCCTGACCAACAGAGACTGCGGCGAGTGTGAAGTGGGCTGGGTGCTGGACGAGGGCGC
CTGTGTGGATGTGGACGAGTGTGCGGCCGAGCCGCCTCCCTGCAGCGCTGCGCAGT
TCTGTAAGAACGCCAACGGCTCCTACACGTGCGAAGAGTGTGACTCCAGCTGTGTGGGCTGC
ACAGGGGAAGGCCCAGGAAACTGTAAAGAGTGTATCTCTGGCTACGCGAGGGAGCACGGACA
GTGTGCAGATGTGGACGAGTGCTCACTAGCAGAAAAAACCTGTGTGAGGAAAAACGAAAACT
GCTACAATACTCCAGGGAGCTACGTCTGTGTGTGTCCTGACGGCTTCGAAGAAACGGAAGAT
GCCTGTGTGCCGCCGGCAGAGGCTGAAGCCACAGAAGGAGAAAGCCCGACACAGCTGCCCTC
CCGCGAAGACCTGTAATGTGCCGGACTTACCCTTTAAATTATTCAGAAGGATGTCCCGTGGA
AAATGTGGCCCTGAGGATGCCGTCTCCTGCAGTGGACAGCGGCGGGGAGAGGCTGCCTGCTC
TCTAACGGTTGATTCTCATTTGTCCCTTAAACAGCTGCATTTCTTGGTTGTTCTTAAACAGA
CTTGTATATTTTGATACAGTTCTTTGTAATAAAATTGACCATTGTAGGTAATCAGGAGGAAA
AAAAAA