PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C07K 14/47

A2

(11) International Publication Number: **WO 98/41539**

(43) International Publication Date: 24 September 1998 (24.09.98)

(21) International Application Number: PCT/US98/05474

(22) International Filing Date: 19 March 1998 (19.03.98)

(30) Priority Data:

08/820,493	19 March 1997 (19.03.97)	US
09/040,963	18 March 1998 (18.03.98)	US

(71) Applicant: GENETICS INSTITUTE, INC. [US/US]; 87 CambridgePark Drive, Cambridge, MA 02140 (US).

(72) Inventors: JACOBS, Kenneth; 151 Beaumont Avenue, Newton, MA 02160 (US). MCCOY, John, M.; 56 Howard Street, Reading, MA 01867 (US). LAVALLIE, Edward, R.; 90 Green Meadow Drive, Tewksbury, MA 01876 (US). RACIE, Lisa, A.; 124 School Street, Acton, MA 01720 (US). MERBERG, David; 2 Orchard Drive, Acton, MA 01720 (US). TREACY, Maurice; 93 Walcott Road, Chestnut Hill, MA 02167 (US). SPAULDING, Vikki; 11 Meadowbank Road, Billerica, MA 01821 (US). AGOSTINO, Michael, J.; 26 Wolcott Avenue, Andover, MA 01810 (US).

(74) Agent: SPRUNGER, Suzanne, A.; Genetics Institute, Inc., 87 CambridgePark Drive, Cambridge, MA 02140 (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, GW, 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, 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, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published

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

(54) Title: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

(57) Abstract

Polynucleotides and the proteins encoded thereby are disclosed.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

This application is a continuation-in-part of application Ser. No. 60/XXX,XXX (converted to a provisional application from non-provisional application 08/820,493), filed March 19, 1997, which is incorporated by reference herein.

FIELD OF THE INVENTION

The present invention provides novel polynucleotides and proteins encoded by such polynucleotides, along with therapeutic, diagnostic and research utilities for these polynucleotides and proteins.

BACKGROUND OF THE INVENTION

Technology aimed at the discovery of protein factors (including e.g., cytokines, such as lymphokines, interferons, CSFs and interleukins) has matured rapidly over the past decade. The now routine hybridization cloning and expression cloning techniques clone novel polynucleotides "directly" in the sense that they rely on information directly related to the discovered protein (i.e., partial DNA/amino acid sequence of the protein in the case of hybridization cloning; activity of the protein in the case of expression cloning). More recent "indirect" cloning techniques such as signal sequence cloning, which isolates DNA sequences based on the presence of a now well-recognized secretory leader sequence motif, as well as various PCR-based or low stringency hybridization cloning techniques, have advanced the state of the art by making available large numbers of DNA/amino acid sequences for proteins that are known to have biological activity by virtue of their secreted nature in the case of leader sequence cloning, or by virtue of the cell or tissue source in the case of PCR-based techniques. It is to these proteins and the polynucleotides encoding them that the present invention is directed.

SUMMARY OF THE INVENTION

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 463 to nucleotide 606;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 1 to nucleotide 501;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone bd164_7 deposited under accession number ATCC 98364;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone bd164_7 deposited under accession number ATCC 98364;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone bd164_7 deposited under accession number ATCC 98364;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone bd164_7 deposited under accession number ATCC 98364;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:2;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:2 having biological activity, the fragment comprising the amino acid sequence from amino acid 19 to amino acid 28 of SEQ ID NO:2;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:1 from nucleotide 463 to nucleotide 606; the nucleotide sequence of SEQ ID NO:1 from nucleotide 1 to nucleotide 501; the nucleotide sequence of the full-length protein

coding sequence of clone bd164_7 deposited under accession number ATCC 98364; or the nucleotide sequence of a mature protein coding sequence of clone bd164_7 deposited under accession number ATCC 98364. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone bd164_7 deposited under accession number ATCC 98364.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:1.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:2;

(b) fragments of the amino acid sequence of SEQ ID NO:2 comprising the amino acid sequence from amino acid 19 to amino acid 28 of SEQ ID NO:2; and

(c) the amino acid sequence encoded by the cDNA insert of clone bd164_7 deposited under accession number ATCC 98364;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:2.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 202 to nucleotide 849;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 511 to nucleotide 849;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone bi129_2 deposited under accession number ATCC 98364;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone bi129_2 deposited under accession number ATCC 98364;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone bi129_2 deposited under accession number ATCC 98364;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone bi129_2 deposited under accession number ATCC 98364;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:4;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:4 having biological activity, the fragment comprising the amino acid sequence from amino acid 103 to amino acid 112 of SEQ ID NO:4;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:3 from nucleotide 202 to nucleotide 849; the nucleotide sequence of SEQ ID NO:3 from nucleotide 511 to nucleotide 849; the nucleotide sequence of the full-length protein coding sequence of clone bi129_2 deposited under accession number ATCC 98364; or the nucleotide sequence of a mature protein coding sequence of clone bi129_2 deposited under accession number ATCC 98364. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone bi129_2 deposited under accession number ATCC 98364. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:4 from amino acid 88 to amino acid 209.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:3.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:4;

(b) the amino acid sequence of SEQ ID NO:4 from amino acid 88 to amino acid 209;

(c) fragments of the amino acid sequence of SEQ ID NO:4 comprising the amino acid sequence from amino acid 103 to amino acid 112 of SEQ ID NO:4; and

(d) the amino acid sequence encoded by the cDNA insert of clone bi129_2 deposited under accession number ATCC 98364;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:4 or the amino acid sequence of SEQ ID NO:4 from amino acid 88 to amino acid 209.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 51 to nucleotide 356;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 348 to nucleotide 356;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone bk95_3 deposited under accession number ATCC 98364;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone bk95_3 deposited under accession number ATCC 98364;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone bk95_3 deposited under accession number ATCC 98364;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone bk95_3 deposited under accession number ATCC 98364;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:6;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:6 having biological activity, the fragment comprising the amino acid sequence from amino acid 46 to amino acid 55 of SEQ ID NO:6;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:5 from nucleotide 51 to nucleotide 356; the nucleotide sequence of SEQ ID NO:5 from nucleotide 348 to nucleotide 356; the nucleotide sequence of the full-length protein coding sequence of clone bk95_3 deposited under accession number ATCC 98364; or the nucleotide sequence of a mature protein coding sequence of clone bk95_3 deposited under accession number ATCC 98364. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone bk95_3 deposited under accession number ATCC 98364. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:6 from amino acid 2 to amino acid 102.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:5 or SEQ ID NO:7.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:6;

(b) the amino acid sequence of SEQ ID NO:6 from amino acid 2 to amino acid 102;

(c) fragments of the amino acid sequence of SEQ ID NO:6 comprising the amino acid sequence from amino acid 46 to amino acid 55 of SEQ ID NO:6; and

(d) the amino acid sequence encoded by the cDNA insert of clone bk95_3 deposited under accession number ATCC 98364;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:6 or the amino acid sequence of SEQ ID NO:6 from amino acid 2 to amino acid 102.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:8;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:8 from nucleotide 156 to nucleotide 902;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:8 from nucleotide 225 to nucleotide 902;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:8 from nucleotide 237 to nucleotide 654;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone cg160_6 deposited under accession number ATCC 98364;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone cg160_6 deposited under accession number ATCC 98364;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone cg160_6 deposited under accession number ATCC 98364;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone cg160_6 deposited under accession number ATCC 98364;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:9;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:9 having biological activity, the fragment comprising the amino acid sequence from amino acid 119 to amino acid 128 of SEQ ID NO:9;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:8 from nucleotide 156 to nucleotide 902; the nucleotide sequence of SEQ ID NO:8 from nucleotide 225 to nucleotide 902; the nucleotide sequence of SEQ ID NO:8 from nucleotide 237 to nucleotide 654; the nucleotide sequence of the full-length protein coding sequence of clone cg160_6 deposited under accession number ATCC 98364; or the nucleotide sequence of a mature protein coding sequence of clone cg160_6 deposited

under accession number ATCC 98364. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone cg160_6 deposited under accession number ATCC 98364. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:9 from amino acid 28 to amino acid 166.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:8.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:9;

(b) the amino acid sequence of SEQ ID NO:9 from amino acid 28 to amino acid 166;

(c) fragments of the amino acid sequence of SEQ ID NO:9 comprising the amino acid sequence from amino acid 119 to amino acid 128 of SEQ ID NO:9; and

(d) the amino acid sequence encoded by the cDNA insert of clone cg160_6 deposited under accession number ATCC 98364;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:9 or the amino acid sequence of SEQ ID NO:9 from amino acid 28 to amino acid 166.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:10;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:10 from nucleotide 400 to nucleotide 2454;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:10 from nucleotide 1454 to nucleotide 1787;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone cw775_1 deposited under accession number ATCC 98364;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone cw775_1 deposited under accession number ATCC 98364;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone cw775_1 deposited under accession number ATCC 98364;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone cw775_1 deposited under accession number ATCC 98364;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:11;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:11 having biological activity, the fragment comprising the amino acid sequence from amino acid 337 to amino acid 346 of SEQ ID NO:11;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:10 from nucleotide 400 to nucleotide 2454; the nucleotide sequence of SEQ ID NO:10 from nucleotide 1454 to nucleotide 1787; the nucleotide sequence of the full-length protein coding sequence of clone cw775_1 deposited under accession number ATCC 98364; or the nucleotide sequence of a mature protein coding sequence of clone cw775_1 deposited under accession number ATCC 98364. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone cw775_1 deposited under accession number ATCC 98364.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:10.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:11;

(b) fragments of the amino acid sequence of SEQ ID NO:11 comprising the amino acid sequence from amino acid 337 to amino acid 346 of SEQ ID NO:11; and

(c) the amino acid sequence encoded by the cDNA insert of clone cw775_1 deposited under accession number ATCC 98364;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:11.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:12;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:12 from nucleotide 506 to nucleotide 1096;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:12 from nucleotide 656 to nucleotide 1096;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:12 from nucleotide 2 to nucleotide 1078;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone dn740_3 deposited under accession number ATCC 98364;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone dn740_3 deposited under accession number ATCC 98364;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone dn740_3 deposited under accession number ATCC 98364;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone dn740_3 deposited under accession number ATCC 98364;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:13;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:13 having biological activity, the fragment comprising the amino acid sequence from amino acid 93 to amino acid 102 of SEQ ID NO:13;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:12 from nucleotide 506 to nucleotide 1096; the nucleotide sequence of SEQ ID NO:12 from nucleotide 656 to nucleotide 1096; the nucleotide sequence of SEQ ID NO:12 from nucleotide 2 to nucleotide 1078; the nucleotide sequence of the full-length protein coding sequence of clone dn740_3 deposited under accession number ATCC 98364; or the nucleotide sequence of a mature protein coding sequence of clone dn740_3 deposited under accession number ATCC 98364. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone dn740_3 deposited under accession number ATCC 98364. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:13 from amino acid 1 to amino acid 191.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:12.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:13;

(b) the amino acid sequence of SEQ ID NO:13 from amino acid 1 to amino acid 191;

(c) fragments of the amino acid sequence of SEQ ID NO:13 comprising the amino acid sequence from amino acid 93 to amino acid 102 of SEQ ID NO:13; and

(d) the amino acid sequence encoded by the cDNA insert of clone dn740_3 deposited under accession number ATCC 98364;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:13 or the amino acid sequence of SEQ ID NO:13 from amino acid 1 to amino acid 191.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:14;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:14 from nucleotide 1563 to nucleotide 1685;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:14 from nucleotide 1100 to nucleotide 1646;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone dn904_2 deposited under accession number ATCC 98364;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone dn904_2 deposited under accession number ATCC 98364;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone dn904_2 deposited under accession number ATCC 98364;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone dn904_2 deposited under accession number ATCC 98364;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:15;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:15 having biological activity, the fragment comprising the amino acid sequence from amino acid 15 to amino acid 24 of SEQ ID NO:15;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:14 from nucleotide 1563 to nucleotide 1685; the nucleotide sequence of SEQ ID NO:14 from nucleotide 1100 to nucleotide 1646; the nucleotide sequence of the full-length protein coding sequence of clone dn904_2 deposited under accession number ATCC 98364; or the

nucleotide sequence of a mature protein coding sequence of clone dn904_2 deposited under accession number ATCC 98364. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone dn904_2 deposited under accession number ATCC 98364. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:15 from amino acid 1 to amino acid 28.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:14.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:15;

(b) the amino acid sequence of SEQ ID NO:15 from amino acid 1 to amino acid 28;

(c) fragments of the amino acid sequence of SEQ ID NO:15 comprising the amino acid sequence from amino acid 15 to amino acid 24 of SEQ ID NO:15; and

(d) the amino acid sequence encoded by the cDNA insert of clone dn904_2 deposited under accession number ATCC 98364;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:15 or the amino acid sequence of SEQ ID NO:15 from amino acid 1 to amino acid 28.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:16;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:16 from nucleotide 359 to nucleotide 1369;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:16 from nucleotide 1547 to nucleotide 1868;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone do568_11 deposited under accession number ATCC 98364;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone do568_11 deposited under accession number ATCC 98364;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone do568_11 deposited under accession number ATCC 98364;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone do568_11 deposited under accession number ATCC 98364;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:17;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:17 having biological activity, the fragment comprising the amino acid sequence from amino acid 163 to amino acid 172 of SEQ ID NO:17;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:16 from nucleotide 359 to nucleotide 1369; the nucleotide sequence of SEQ ID NO:16 from nucleotide 1547 to nucleotide 1868; the nucleotide sequence of the full-length protein coding sequence of clone do568_11 deposited under accession number ATCC 98364; or the nucleotide sequence of a mature protein coding sequence of clone do568_11 deposited under accession number ATCC 98364. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone do568_11 deposited under accession number ATCC 98364.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:16.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:17;

(b) fragments of the amino acid sequence of SEQ ID NO:17 comprising the amino acid sequence from amino acid 163 to amino acid 172 of SEQ ID NO:17; and

(c) the amino acid sequence encoded by the cDNA insert of clone do568_11 deposited under accession number ATCC 98364;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:17.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:18;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:18 from nucleotide 85 to nucleotide 1263;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:18 from nucleotide 265 to nucleotide 608;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone ek626_3 deposited under accession number ATCC 98364;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone ek626_3 deposited under accession number ATCC 98364;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone ek626_3 deposited under accession number ATCC 98364;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone ek626_3 deposited under accession number ATCC 98364;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:19;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:19 having biological activity, the fragment comprising the amino acid sequence from amino acid 191 to amino acid 200 of SEQ ID NO:19;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:18 from nucleotide 85 to nucleotide 1263; the nucleotide sequence of SEQ ID NO:18 from nucleotide 265 to nucleotide 608; the nucleotide sequence of the full-length protein coding sequence of clone ek626_3 deposited under accession number ATCC 98364; or the nucleotide sequence of a mature protein coding sequence of clone ek626_3 deposited under accession number ATCC 98364. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone ek626_3 deposited under accession number ATCC 98364. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:19 from amino acid 61 to amino acid 175.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:18.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:19;

(b) the amino acid sequence of SEQ ID NO:19 from amino acid 61 to amino acid 175;

(c) fragments of the amino acid sequence of SEQ ID NO:19 comprising the amino acid sequence from amino acid 191 to amino acid 200 of SEQ ID NO:19; and

(d) the amino acid sequence encoded by the cDNA insert of clone ek626_3 deposited under accession number ATCC 98364;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:19 or the amino acid sequence of SEQ ID NO:19 from amino acid 61 to amino acid 175.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:20;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:20 from nucleotide 3746 to nucleotide 4027;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:20 from nucleotide 3815 to nucleotide 4027;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:20 from nucleotide 3640 to nucleotide 3940;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone fe366_1 deposited under accession number ATCC 98364;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone fe366_1 deposited under accession number ATCC 98364;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone fe366_1 deposited under accession number ATCC 98364;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone fe366_1 deposited under accession number ATCC 98364;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:21;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:21 having biological activity, the fragment comprising the amino acid sequence from amino acid 42 to amino acid 51 of SEQ ID NO:21;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:20 from nucleotide 3746 to nucleotide 4027; the nucleotide sequence of SEQ ID NO:20 from nucleotide 3815 to nucleotide 4027; the nucleotide sequence of SEQ ID NO:20 from nucleotide 3640 to nucleotide 3940; the nucleotide sequence of the full-length protein

coding sequence of clone fe366_1 deposited under accession number ATCC 98364; or the nucleotide sequence of a mature protein coding sequence of clone fe366_1 deposited under accession number ATCC 98364. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone fe366_1 deposited under accession number ATCC 98364. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:21 from amino acid 1 to amino acid 65.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:20.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:21;

(b) the amino acid sequence of SEQ ID NO:21 from amino acid 1 to amino acid 65;

(c) fragments of the amino acid sequence of SEQ ID NO:21 comprising the amino acid sequence from amino acid 42 to amino acid 51 of SEQ ID NO:21; and

(d) the amino acid sequence encoded by the cDNA insert of clone fe366_1 deposited under accession number ATCC 98364;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:21 or the amino acid sequence of SEQ ID NO:21 from amino acid 1 to amino acid 65.

In certain preferred embodiments, the polynucleotide is operably linked to an expression control sequence. The invention also provides a host cell, including bacterial, yeast, insect and mammalian cells, transformed with such polynucleotide compositions. Also provided by the present invention are organisms that have enhanced, reduced, or modified expression of the gene(s) corresponding to the polynucleotide sequences disclosed herein.

Processes are also provided for producing a protein, which comprise:

(a) growing a culture of the host cell transformed with such polynucleotide compositions in a suitable culture medium; and

(b) purifying the protein from the culture.

The protein produced according to such methods is also provided by the present invention. Preferred embodiments include those in which the protein produced by such process is a mature form of the protein.

Protein compositions of the present invention may further comprise a pharmaceutically acceptable carrier. Compositions comprising an antibody which specifically reacts with such protein are also provided by the present invention.

Methods are also provided for preventing, treating or ameliorating a medical condition which comprises administering to a mammalian subject a therapeutically effective amount of a composition comprising a protein of the present invention and a pharmaceutically acceptable carrier.

BRIEF DESCRIPTION OF THE DRAWINGS

Figures 1A and 1B are schematic representations of the pED6 and pNOTs vectors, respectively, used for deposit of clones disclosed herein.

DETAILED DESCRIPTION

ISOLATED PROTEINS AND POLYNUCLEOTIDES

Nucleotide and amino acid sequences, as presently determined, are reported below for each clone and protein disclosed in the present application. The nucleotide sequence of each clone can readily be determined by sequencing of the deposited clone in accordance with known methods. The predicted amino acid sequence (both full-length and mature forms) can then be determined from such nucleotide sequence. The amino acid sequence of the protein encoded by a particular clone can also be determined by expression of the clone in a suitable host cell, collecting the protein and determining its sequence. For each disclosed protein applicants have identified what they have determined to be the reading frame best identifiable with sequence information available at the time of filing.

As used herein a "secreted" protein is one which, when expressed in a suitable host cell, is transported across or through a membrane, including transport as a result of signal sequences in its amino acid sequence. "Secreted" proteins include without limitation proteins secreted wholly (e.g., soluble proteins) or partially (e.g. , receptors) from the cell in which they are expressed. "Secreted" proteins also include without limitation proteins which are transported across the membrane of the endoplasmic reticulum.

Clone "bd164_7"

A polynucleotide of the present invention has been identified as clone "bd164_7". bd164_7 was isolated from a human fetal kidney cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. bd164_7 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "bd164_7 protein").

The nucleotide sequence of bd164_7 as presently determined is reported in SEQ ID NO:1. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the bd164_7 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:2. Another potential bd164_7 reading frame and predicted amino acid sequence is encoded by basepairs 610 to 762 of SEQ ID NO:1 and is reported in SEQ ID NO:32.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone bd164_7 should be approximately 1950 bp.

The nucleotide sequence disclosed herein for bd164_7 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. bd164_7 demonstrated at least some similarity with sequences identified as AF001540 (Human clone alpha1 mRNA, partial sequence), C05823 (similar to none), G22994 (human STS WI-30658), H03651 (yj37e12.s1 Homo sapiens cDNA clone 150958 3'), H26492 (EST51a22 Homo sapiens cDNA clone 51a22), H90721 (yv96f02.r1 Homo sapiens cDNA clone 250587 5'), N58545 (yv73d07.s1 Homo sapiens cDNA clone 248365 3'), R10191 (yf35d07.r1 Homo sapiens cDNA clone 128845 5'), and X17272 (Human heterogenous nuclear RNA W16W). Based upon sequence similarity, bd164_7 proteins and each similar protein or peptide may share at least some activity.

Clone "bi129_2"

A polynucleotide of the present invention has been identified as clone "bi129_2". bi129_2 was isolated from a human fetal kidney cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. bi129_2 is a full-length clone,

including the entire coding sequence of a secreted protein (also referred to herein as "bi129_2 protein").

The nucleotide sequence of bi129_2 as presently determined is reported in SEQ ID NO:3. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the bi129_2 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:4. Amino acids 91 to 103 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 104, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone bi129_2 should be approximately 1100 bp.

The nucleotide sequence disclosed herein for bi129_2 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. bi129_2 demonstrated at least some similarity with sequences identified as H88684 (yw23b01.r1 Homo sapiens cDNA), R59623 (yh02g07.s1 Homo sapiens cDNA clone 42126 3'), T17199 (NIB515 Homo sapiens cDNA 3'end), T24786 (Human gene signature HUMGS06869), T65550 (yc76b12.s1 Homo sapiens cDNA clone 21611 3'), and T65617 (yc76b12.r1 Homo sapiens cDNA clone 21611 5'). The predicted amino acid sequence disclosed herein for bi129_2 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted bi129_2 protein demonstrated at least some similarity to sequences identified as AF016712 (testicular condensing enzyme [Mus musculus]) and U43375 (Similar to sugar transporter (Caenorhabditis elegans cosmid K09C4)). Based upon sequence similarity, bi129_2 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts six potential transmembrane domains within the bi129_2 protein sequence, centered around amino acids 11, 36, 69, 100, 131, and 185 of SEQ ID NO:4, respectively.

Clone "bk95_3"

A polynucleotide of the present invention has been identified as clone "bk95_3". bk95_3 was isolated from a human adult retina cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. bk95_3 is a full-length clone,

including the entire coding sequence of a secreted protein (also referred to herein as "bk95_3 protein").

The nucleotide sequence of the 5' portion of bk95_3 as presently determined is reported in SEQ ID NO:5. What applicants presently believe is the proper reading frame for the coding region is indicated in SEQ ID NO:6. The predicted amino acid sequence of the bk95_3 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:6. Amino acids 87 to 99 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 100, or are a transmembrane domain. Additional nucleotide sequence from the 3' portion of bk95_3, including the polyA tail, is reported in SEQ ID NO:7.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone bk95_3 should be approximately 2400 bp.

The nucleotide sequence disclosed herein for bk95_3 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. bk95_3 demonstrated at least some similarity with sequences identified as AA521036 (aa71b06.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:826355 3' similar to SW:SYB2_XENLA P47193 SYNAPTOBREVIN 2), N29686 (yw78a05.s1 Homo sapiens cDNA clone 258320 3' similar to SP:SW:SYB2_XENLA P47193 SYNAPTOBREVIN 2), T33715 (Cellubrevin-2 coding sequence), U14567 (***ALU WARNING Human Alu-J subfamily consensus sequence), and U60150 (Mus musculus vesicle-associated membrane protein VAMP-2 mRNA, complete cds). The predicted amino acid sequence disclosed herein for bk95_3 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted bk95_3 protein demonstrated at least some similarity to sequences identified as L14270 (synaptobrevin [Drosophila melanogaster]), M36205 (synaptobrevin 2 (SYB2) [Homo sapiens]), U60961 (cellubrevin [Mus musculus]), U64520 (synaptobrevin-3 [Homo sapiens]), W04181 (Cellubrevin-2), and X76199 (synaptobrevin [Bos taurus]). Based upon sequence similarity, bk95_3 proteins and each similar protein or peptide may share at least some activity. The nucleotide sequence of bk95_3 indicates that it may contain an Alu repetitive element.

Clone "cg160_6"

A polynucleotide of the present invention has been identified as clone "cg160_6". cg160_6 was isolated from a human adult testes cDNA library using methods which are

selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. cg160_6 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "cg160_6 protein").

The nucleotide sequence of cg160_6 as presently determined is reported in SEQ ID NO:8. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the cg160_6 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:9. Amino acids 11 to 23 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 24, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone cg160_6 should be approximately 1400 bp.

The nucleotide sequence disclosed herein for cg160_6 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. cg160_6 demonstrated at least some similarity with sequences identified as AA405957 (zu66c07.r1 Soares testis NHT Homo sapiens cDNA clone 742956 5') and T19219 (f02011t Testis 1 Homo sapiens cDNA clone f02011 5' end). Based upon sequence similarity, cg160_6 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts three additional potential transmembrane domains within the cg160_6 protein sequence, centerd around amino acids 148, 195, and 236 of SEQ ID NO:9, respectively.

Clone "cw775_1"

A polynucleotide of the present invention has been identified as clone "cw775_1". cw775_1 was isolated from a human fetal brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. cw775_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "cw775_1 protein").

The nucleotide sequence of cw775_1 as presently determined is reported in SEQ ID NO:10. What applicants presently believe to be the proper reading frame and the

predicted amino acid sequence of the cw775_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:11.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone cw775_1 should be approximately 4200 bp.

The nucleotide sequence disclosed herein for cw775_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. cw775_1 demonstrated at least some similarity with sequences identified as AA104324 (mo50d06.r1 Life Tech mouse embryo 10 5dpc 10665016 Mus musculus cDNA clone 557003 5'), AA373350 (EST85423 HSC172 cells I Homo sapiens cDNA 5' end), H30439 (ym58f10.r1 Homo sapiens cDNA clone 52688 5'), N28734 (yx67c10.r1 Homo sapiens cDNA clone 266802 5'), and N57005 (yy56h03.s1 Homo sapiens cDNA clone 277589 3'). Based upon sequence similarity, cw775_1 proteins and each similar protein or peptide may share at least some activity.

Clone "dn740_3"

A polynucleotide of the present invention has been identified as clone "dn740_3". dn740_3 was isolated from a human fetal brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. dn740_3 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "dn740_3 protein").

The nucleotide sequence of dn740_3 as presently determined is reported in SEQ ID NO:12. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the dn740_3 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:13. Amino acids 38 to 50 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 51, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone dn740_3 should be approximately 1650 bp.

The nucleotide sequence disclosed herein for dn740_3 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. dn740_3 demonstrated at least some similarity with sequences identified as AA053844 (zf53h07.r1 Soares retina N2b4HR Homo sapiens cDNA clone

380701 5'), AA056525 (zl65g08.r1 Stratagene colon (#937204) Homo sapiens cDNA clone 509534 5'), H70470 (yr91c07.s1 Homo sapiens cDNA clone 212652 3'), N53038 (yv53d09.s1 Homo sapiens cDNA clone 246449 3'), R56318 (yg90e03.r1 Homo sapiens cDNA clone 40653 5'), and W73718 (zd50f06.s1 Soares fetal heart NbHH19W Homo sapiens cDNA clone 344099 3'). The predicted amino acid sequence disclosed herein for dn740_3 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted dn740_3 protein demonstrated at least some similarity to sequences identified as M34651 (ORF-3 protein [Suid herpesvirus 1]), U15306 (NFX1 [Homo sapiens]), and Z81103 (M04G12.1 [Caenorhabditis elegans]). Based upon sequence similarity, dn740_3 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts two potential transmembrane domains within the dn740_3 protein sequence, centerd around amino acids 110 and 180 of SEQ ID NO:13, respectively. The nucleotide sequence of dn740_3 indicates that it may contain a simple AT repeat sequence.

Clone "dn904_2"

A polynucleotide of the present invention has been identified as clone "dn904_2". dn904_2 was isolated from a human fetal brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. dn904_2 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "dn904_2 protein").

The nucleotide sequence of dn904_2 as presently determined is reported in SEQ ID NO:14. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the dn904_2 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:15.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone dn904_2 should be approximately 2700 bp.

The nucleotide sequence disclosed herein for dn904_2 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. dn904_2 demonstrated at least some similarity with sequences identified as N66026 (za28g05.s1 Homo sapiens cDNA clone 293912 3' similar to contains Alu repetitive element;contains element MER6 repetitive element) and U67221 (Human

clone HS4.14 Alu-Ya5 sequence). The predicted amino acid sequence disclosed herein for dn904_2 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted dn904_2 protein demonstrated at least some similarity to sequences identified as U79260 (unknown [Homo sapiens]). Based upon sequence similarity, dn904_2 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts a potential transmembrane domain within the dn904_2 protein sequence centered around amino acid 15 of SEQ ID NO:15. The nucleotide sequence of dn904_2 indicates that it may contain an Alu repetitive element.

Clone "do568_11"

A polynucleotide of the present invention has been identified as clone "do568_11". do568_11 was isolated from a human adult testes cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. do568_11 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "do568_11 protein").

The nucleotide sequence of do568_11 as presently determined is reported in SEQ ID NO:16. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the do568_11 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:17.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone do568_11 should be approximately 2300 bp.

The nucleotide sequence disclosed herein for do568_11 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. do568_11 demonstrated at least some similarity with sequences identified as AA399248 (zt57d07.s1 Soares testis NHT Homo sapiens cDNA clone 726445 3'), AA552222 (nk06a07.s1 NCI_CGAP_Co2 Homo sapiens cDNA clone IMAGE:1012692), H41337 (yn91d06.r1 Homo sapiens cDNA clone), H56978 (yr07a01.r1 Homo sapiens cDNA clone 204552 5'), J05096 (Human Na,K-ATPase subunit alpha 2 (ATP1A2) gene, complete cds), N95160 (zb52c09.s1 Soares fetal lung NbHL19W Homo sapiens cDNA clone 307216 3'similar to contains element MER22 repetitive element), R42239 (yf98a10.s1 Homo sapiens cDNA clone 30435 3'), T15786 (IB1892 Infant brain, Bento Soares Homo

sapiens cDNA 3'end), and T20399 (Human gene signature HUMGS01552). Based upon sequence similarity, do568_11 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts two potential transmembrane domains within the do568_11 protein sequence, one at the amino terminus and another centered around amino acid 230 of SEQ ID NO:17.

Clone "ek626_3"

A polynucleotide of the present invention has been identified as clone "ek626_3". ek626_3 was isolated from a human fetal brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. ek626_3 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "ek626_3 protein").

The nucleotide sequence of ek626_3 as presently determined is reported in SEQ ID NO:18. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the ek626_3 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:19.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone ek626_3 should be approximately 1900 bp.

The nucleotide sequence disclosed herein for ek626_3 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. ek626_3 demonstrated at least some similarity with sequences identified as AA112543 (zm28a12.r1 Stratagene pancreas (#937208) Homo sapiens cDNA clone 526942 5'), AA160534 (zo73f06.s1 Stratagene pancreas (#937208) Homo sapiens cDNA clone 592547 3'), AA160629 (zo73f06.r1 Stratagene pancreas (#937208) Homo sapiens cDNA clone 592547 5'), AA168779 (ms37g07.r1 Stratagene mouse heart (#937316) Mus musculus cDNA clone 613788 5'), AA211632 (zn56b09.r1 Stratagene muscle 937209 Homo sapiens cDNA clone 562169 5'), AA224303 (zr15e10.r1 Stratagene NT2 neuronal precursor 937230 Homo sapiens cDNA clone 663498 5'), AA429442 (zw47b06.r1 Soares total fetus Nb2HF8 9w Homo sapiens cDNA clone 773171 5'), H22161 (yl38g02.s1 Homo sapiens cDNA clone), T52832 (Human gene signature HUMGS08061), U21718 (Rattus norvegicus clone C426 intestinal epithelium proliferating cell-associated mRNA sequence), and W26019 (18b9 Human retina cDNA randomly primed sublibrary Homo sapiens

cDNA). The predicted amino acid sequence disclosed herein for dn904_2 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted dn904_2 protein demonstrated at least some similarity to sequences identified as R99052 (Spider dragline variant, DP-1A.9 monomer) and Z97342 (nuclear antigen homolog [Arabidopsis thaliana]). Based upon sequence similarity, ek626_3 proteins and each similar protein or peptide may share at least some activity.

Clone "fe366_1"

A polynucleotide of the present invention has been identified as clone "fe366_1". fe366_1 was isolated from a human adult brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. fe366_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "fe366_1 protein").

The nucleotide sequence of fe366_1 as presently determined is reported in SEQ ID NO:20. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the fe366_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:21. Amino acids 11 to 23 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 24, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone fe366_1 should be approximately 3100 bp.

The nucleotide sequence disclosed herein for fe366_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. fe366_1 demonstrated at least some similarity with sequences identified as AA139623 (mq40b07.r1 Barstead MPLRB1 Mus musculus cDNA clone 581173 5' similar to WP:F43E2.7 CE07243), AA306766 (EST177699 Jurkat T-cells VI Homo sapiens cDNA 5' end), AA663899 (ae74d05.s1 Stratagene schizo brain S11 Homo sapiens cDNA clone 969897 3'), H29956 (yp44b03.r1 Homo sapiens cDNA clone 190253 5'), H93431 (ys76d10.r1 Homo sapiens cDNA clone 220723 5'), and M61937 (R.norvegicus dihydrodiol dehydrogenase mRNA, complete cds). Based upon sequence similarity, fe366_1 proteins and each similar protein or peptide may share at least some activity. The nucleotide

sequence of fe366_1 indicates that it may contain one or more of the following: CAA repeat, Alu repetitive element.

Deposit of Clones

Clones bd164_7, bi129_2, bk95_3, cg160_6, cw775_1, dn740_3, dn904_2, do568_11, ek626_3, and fe366_1 were deposited on March 19, 1997 with the American Type Culture Collection as an original deposit under the Budapest Treaty and were given the accession number ATCC 98364, from which each clone comprising a particular polynucleotide is obtainable. All restrictions on the availability to the public of the deposited material will be irrevocably removed upon the granting of the patent, except for the requirements specified in 37 C.F.R. § 1.808(b).

Each clone has been transfected into separate bacterial cells (*E. coli*) in this composite deposit. Each clone can be removed from the vector in which it was deposited by performing an EcoRI/NotI digestion (5' site, EcoRI; 3' site, NotI) to produce the appropriate fragment for such clone. Each clone was deposited in either the pED6 or pNOTs vector depicted in Fig. 1. The pED6dpc2 vector ("pED6") was derived from pED6dpc1 by insertion of a new polylinker to facilitate cDNA cloning (Kaufman *et al.*, 1991, *Nucleic Acids Res.* **19**: 4485-4490); the pNOTs vector was derived from pMT2 (Kaufman *et al.*, 1989, *Mol. Cell. Biol.* 9: 946-958) by deletion of the DHFR sequences, insertion of a new polylinker, and insertion of the M13 origin of replication in the ClaI site. In some instances, the deposited clone can become "flipped" (i.e., in the reverse orientation) in the deposited isolate. In such instances, the cDNA insert can still be isolated by digestion with EcoRI and NotI. However, NotI will then produce the 5' site and EcoRI will produce the 3' site for placement of the cDNA in proper orientation for expression in a suitable vector. The cDNA may also be expressed from the vectors in which they were deposited.

Bacterial cells containing a particular clone can be obtained from the composite deposit as follows:

An oligonucleotide probe or probes should be designed to the sequence that is known for that particular clone. This sequence can be derived from the sequences provided herein, or from a combination of those sequences. The sequence of the oligonucleotide probe that was used to isolate each full-length clone is identified below, and should be most reliable in isolating the clone of interest.

Clone	Probe Sequence
bd164_7	SEQ ID NO:22
bi129_2	SEQ ID NO:23
bk95_3	SEQ ID NO:24
cg160_6	SEQ ID NO:25
cw775_1	SEQ ID NO:26
dn740_3	SEQ ID NO:27
dn904_2	SEQ ID NO:28
do568_11	SEQ ID NO:29
ek626_3	SEQ ID NO:30
fe366_1	SEQ ID NO:31

In the sequences listed above which include an N at position 2, that position is occupied in preferred probes/primers by a biotinylated phosphoaramidite residue rather than a nucleotide (such as , for example, that produced by use of biotin phosphoramidite (1-dimethoxytrityloxy-2-(N-biotinyl-4-aminobutyl)-propyl-3-O-(2-cyanoethyl)-(N,N-diisopropyl)-phosphoramadite) (Glen Research, cat. no. 10-1953)).

The design of the oligonucleotide probe should preferably follow these parameters:

(a) It should be designed to an area of the sequence which has the fewest ambiguous bases ("N's"), if any;

(b) It should be designed to have a T_m of approx. 80 ° C (assuming 2° for each A or T and 4 degrees for each G or C).

The oligonucleotide should preferably be labeled with g-^{32}P ATP (specific activity 6000 Ci/mmole) and T4 polynucleotide kinase using commonly employed techniques for labeling oligonucleotides. Other labeling techniques can also be used. Unincorporated label should preferably be removed by gel filtration chromatography or other established methods. The amount of radioactivity incorporated into the probe should be quantitated by measurement in a scintillation counter. Preferably, specific activity of the resulting probe should be approximately 4e+6 dpm/pmole.

The bacterial culture containing the pool of full-length clones should preferably be thawed and 100 µl of the stock used to inoculate a sterile culture flask containing 25 ml of sterile L-broth containing ampicillin at 100 µg/ml. The culture should preferably be grown to saturation at 37°C, and the saturated culture should preferably be diluted in

fresh L-broth. Aliquots of these dilutions should preferably be plated to determine the dilution and volume which will yield approximately 5000 distinct and well-separated colonies on solid bacteriological media containing L-broth containing ampicillin at 100 µg/ml and agar at 1.5% in a 150 mm petri dish when grown overnight at 37°C. Other known methods of obtaining distinct, well-separated colonies can also be employed.

Standard colony hybridization procedures should then be used to transfer the colonies to nitrocellulose filters and lyse, denature and bake them.

The filter is then preferably incubated at 65°C for 1 hour with gentle agitation in 6X SSC (20X stock is 175.3 g NaCl/liter, 88.2 g Na citrate/liter, adjusted to pH 7.0 with NaOH) containing 0.5% SDS, 100 µg/ml of yeast RNA, and 10 mM EDTA (approximately 10 mL per 150 mm filter). Preferably, the probe is then added to the hybridization mix at a concentration greater than or equal to 1e+6 dpm/mL. The filter is then preferably incubated at 65°C with gentle agitation overnight. The filter is then preferably washed in 500 mL of 2X SSC/0.5% SDS at room temperature without agitation, preferably followed by 500 mL of 2X SSC/0.1% SDS at room temperature with gentle shaking for 15 minutes. A third wash with 0.1X SSC/0.5% SDS at 65°C for 30 minutes to 1 hour is optional. The filter is then preferably dried and subjected to autoradiography for sufficient time to visualize the positives on the X-ray film. Other known hybridization methods can also be employed.

The positive colonies are picked, grown in culture, and plasmid DNA isolated using standard procedures. The clones can then be verified by restriction analysis, hybridization analysis, or DNA sequencing.

Fragments of the proteins of the present invention which are capable of exhibiting biological activity are also encompassed by the present invention. Fragments of the protein may be in linear form or they may be cyclized using known methods, for example, as described in H.U. Saragovi, *et al.*, Bio/Technology 10, 773-778 (1992) and in R.S. McDowell, *et al.*, J. Amer. Chem. Soc. 114, 9245-9253 (1992), both of which are incorporated herein by reference. Such fragments may be fused to carrier molecules such as immunoglobulins for many purposes, including increasing the valency of protein binding sites. For example, fragments of the protein may be fused through "linker" sequences to the Fc portion of an immunoglobulin. For a bivalent form of the protein, such a fusion could be to the Fc portion of an IgG molecule. Other immunoglobulin isotypes may also be used to generate such fusions. For example, a protein - IgM fusion would generate a decavalent form of the protein of the invention.

The present invention also provides both full-length and mature forms of the disclosed proteins. The full-length form of the such proteins is identified in the sequence listing by translation of the nucleotide sequence of each disclosed clone. The mature form(s) of such protein may be obtained by expression of the disclosed full-length polynucleotide (preferably those deposited with ATCC) in a suitable mammalian cell or other host cell. The sequence(s) of the mature form(s) of the protein may also be determinable from the amino acid sequence of the full-length form.

The present invention also provides genes corresponding to the polynucleotide sequences disclosed herein. "Corresponding genes" are the regions of the genome that are transcribed to produce the mRNAs from which cDNA polynucleotide sequences are derived and may include contiguous regions of the genome necessary for the regulated expression of such genes. Corresponding genes may therefore include but are not limited to coding sequences, 5' and 3' untranslated regions, alternatively spliced exons, introns, promoters, enhancers, and silencer or suppressor elements. The corresponding genes can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include the preparation of probes or primers from the disclosed sequence information for identification and/or amplification of genes in appropriate genomic libraries or other sources of genomic materials. An "isolated gene" is a gene that has been separated from the adjacent coding sequences, if any, present in the genome of the organism from which the gene was isolated.

Organisms that have enhanced, reduced, or modified expression of the gene(s) corresponding to the polynucleotide sequences disclosed herein are provided. The desired change in gene expression can be achieved through the use of antisense polynucleotides or ribozymes that bind and/or cleave the mRNA transcribed from the gene (Albert and Morris, 1994, *Trends Pharmacol. Sci.* **15(7)**: 250-254; Lavarosky *et al.*, 1997, *Biochem. Mol. Med.* **62(1)**: 11-22; and Hampel, 1998, *Prog. Nucleic Acid Res. Mol. Biol.* **58**: 1-39; all of which are incorporated by reference herein). Transgenic animals that have multiple copies of the gene(s) corresponding to the polynucleotide sequences disclosed herein, preferably produced by transformation of cells with genetic constructs that are stably maintained within the transformed cells and their progeny, are provided. Transgenic animals that have modified genetic control regions that increase or reduce gene expression levels, or that change temporal or spatial patterns of gene expression, are also provided (see European Patent No. 0 649 464 B1, incorporated by reference herein). In addition, organisms are provided in which the gene(s) corresponding to the

polynucleotide sequences disclosed herein have been partially or completely inactivated, through insertion of extraneous sequences into the corresponding gene(s) or through deletion of all or part of the corresponding gene(s). Partial or complete gene inactivation can be accomplished through insertion, preferably followed by imprecise excision, of transposable elements (Plasterk, 1992, *Bioessays* **14(9)**: 629-633; Zwaal *et al.*, 1993, *Proc. Natl. Acad. Sci. USA* **90(16)**: 7431-7435; Clark *et al.*, 1994, *Proc. Natl. Acad. Sci. USA* **91(2)**: 719-722; all of which are incorporated by reference herein), or through homologous recombination, preferably detected by positive/negative genetic selection strategies (Mansour *et al.*, 1988, *Nature* **336**: 348-352; U.S. Patent Nos. 5,464,764; 5,487,992; 5,627,059; 5,631,153; 5,614, 396; 5,616,491; and 5,679,523; all of which are incorporated by reference herein). These organisms with altered gene expression are preferably eukaryotes and more preferably are mammals. Such organisms are useful for the development of non-human models for the study of disorders involving the corresponding gene(s), and for the development of assay systems for the identification of molecules that interact with the protein product(s) of the corresponding gene(s).

Where the protein of the present invention is membrane-bound (e.g., is a receptor), the present invention also provides for soluble forms of such protein. In such forms part or all of the intracellular and transmembrane domains of the protein are deleted such that the protein is fully secreted from the cell in which it is expressed. The intracellular and transmembrane domains of proteins of the invention can be identified in accordance with known techniques for determination of such domains from sequence information.

Proteins and protein fragments of the present invention include proteins with amino acid sequence lengths that are at least 25%(more preferably at least 50%, and most preferably at least 75%) of the length of a disclosed protein and have at least 60% sequence identity (more preferably, at least 75% identity; most preferably at least 90% or 95% identity) with that disclosed protein, where sequence identity is determined by comparing the amino acid sequences of the proteins when aligned so as to maximize overlap and identity while minimizing sequence gaps. Also included in the present invention are proteins and protein fragments that contain a segment preferably comprising 8 or more (more preferably 20 or more, most preferably 30 or more) contiguous amino acids that shares at least 75% sequence identity (more preferably, at least 85% identity; most preferably at least 95% identity) with any such segment of any of the disclosed proteins.

Species homologues of the disclosed polynucleotides and proteins are also provided by the present invention. As used herein, a "species homologue" is a protein or

polynucleotide with a different species of origin from that of a given protein or polynucleotide, but with significant sequence similarity to the given protein or polynucleotide. Preferably, polynucleotide species homologues have at least 60% sequence identity (more preferably, at least 75% identity; most preferably at least 90% identity) with the given polynucleotide, and protein species homologues have at least 30% sequence identity (more preferably, at least 45% identity; most preferably at least 60% identity) with the given protein, where sequence identity is determined by comparing the nucleotide sequences of the polynucleotides or the amino acid sequences of the proteins when aligned so as to maximize overlap and identity while minimizing sequence gaps. Species homologues may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source from the desired species. Preferably, species homologues are those isolated from mammalian species. Most preferably, species homologues are those isolated from certain mammalian species such as, for example, *Pan troglodytes, Gorilla gorilla, Pongo pygmaeus, Hylobates concolor, Macaca mulatta, Papio papio, Papio hamadryas, Cercopithecus aethiops, Cebus capucinus, Aotus trivirgatus, Sanguinus oedipus, Microcebus murinus, Mus musculus, Rattus norvegicus, Cricetulus griseus, Felis catus, Mustela vison, Canis familiaris, Oryctolagus cuniculus, Bos taurus, Ovis aries, Sus scrofa,* and *Equus caballus,* for which genetic maps have been created allowing the identification of syntenic relationships between the genomic organization of genes in one species and the genomic organization of the related genes in another species (O'Brien and Seuánez, 1988, *Ann. Rev. Genet.* **22**: 323-351; O'Brien *et al.*, 1993, *Nature Genetics* **3**:103-112; Johansson *et al.*, 1995, *Genomics* **25**: 682-690; Lyons *et al.*, 1997, *Nature Genetics* **15**: 47-56; O'Brien *et al.*, 1997, *Trends in Genetics* **13(10)**: 393-399; Carver and Stubbs, 1997, *Genome Research* **7**:1123-1137; all of which are incorporated by reference herein).

The invention also encompasses allelic variants of the disclosed polynucleotides or proteins; that is, naturally-occurring alternative forms of the isolated polynucleotides which also encode proteins which are identical or have significantly similar sequences to those encoded by the disclosed polynucleotides. Preferably, allelic variants have at least 60% sequence identity (more preferably, at least 75% identity; most preferably at least 90% identity) with the given polynucleotide, where sequence identity is determined by comparing the nucleotide sequences of the polynucleotides when aligned so as to maximize overlap and identity while minimizing sequence gaps. Allelic variants may be isolated and

identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source from individuals of the appropriate species.

The invention also includes polynucleotides with sequences complementary to those of the polynucleotides disclosed herein.

The present invention also includes polynucleotides capable of hybridizing under reduced stringency conditions, more preferably stringent conditions, and most preferably highly stringent conditions, to polynucleotides described herein. Examples of stringency conditions are shown in the table below: highly stringent conditions are those that are at least as stringent as, for example, conditions A-F; stringent conditions are at least as stringent as, for example, conditions G-L; and reduced stringency conditions are at least as stringent as, for example, conditions M-R.

Stringency Condition	Polynucleotide Hybrid	Hybrid Length (bp)‡	Hybridization Temperature and Buffer†	Wash Temperature and Buffer†
A	DNA:DNA	≥ 50	65°C; 1xSSC -or- 42°C; 1xSSC, 50% formamide	65°C; 0.3xSSC
B	DNA:DNA	<50	T_B*; 1xSSC	T_B*; 1xSSC
C	DNA:RNA	≥ 50	67°C; 1xSSC -or- 45°C; 1xSSC, 50% formamide	67°C; 0.3xSSC
D	DNA:RNA	<50	T_D*; 1xSSC	T_D*; 1xSSC
E	RNA:RNA	≥ 50	70°C; 1xSSC -or- 50°C; 1xSSC, 50% formamide	70°C; 0.3xSSC
F	RNA:RNA	<50	T_F*; 1xSSC	T_F*; 1xSSC
G	DNA:DNA	≥ 50	65°C; 4xSSC -or- 42°C; 4xSSC, 50% formamide	65°C; 1xSSC
H	DNA:DNA	<50	T_H*; 4xSSC	T_H*; 4xSSC
I	DNA:RNA	≥ 50	67°C; 4xSSC -or- 45°C; 4xSSC, 50% formamide	67°C; 1xSSC
J	DNA:RNA	<50	T_J*; 4xSSC	T_J*; 4xSSC
K	RNA:RNA	≥ 50	70°C; 4xSSC -or- 50°C; 4xSSC, 50% formamide	67°C; 1xSSC
L	RNA:RNA	<50	T_L*; 2xSSC	T_L*; 2xSSC
M	DNA:DNA	≥ 50	50°C; 4xSSC -or- 40°C; 6xSSC, 50% formamide	50°C; 2xSSC
N	DNA:DNA	<50	T_N*; 6xSSC	T_N*; 6xSSC
O	DNA:RNA	≥ 50	55°C; 4xSSC -or- 42°C; 6xSSC, 50% formamide	55°C; 2xSSC
P	DNA:RNA	<50	T_P*; 6xSSC	T_P*; 6xSSC
Q	RNA:RNA	≥ 50	60°C; 4xSSC -or- 45°C; 6xSSC, 50% formamide	60°C; 2xSSC
R	RNA:RNA	<50	T_R*; 4xSSC	T_R*; 4xSSC

‡: The hybrid length is that anticipated for the hybridized region(s) of the hybridizing polynucleotides. When hybridizing a polynucleotide to a target polynucleotide of unknown sequence, the hybrid length is assumed to be that of the hybridizing polynucleotide. When polynucleotides of known sequence are hybridized, the hybrid length can be determined by aligning the sequences of the polynucleotides and identifying the region or regions of optimal sequence complementarity.

†: SSPE (1xSSPE is 0.15M NaCl, 10mM NaH_2PO_4, and 1.25mM EDTA, pH 7.4) can be substituted for SSC (1xSSC is 0.15M NaCl and 15mM sodium citrate) in the hybridization and wash buffers; washes are performed for 15 minutes after hybridization is complete.

*T_B - T_R: The hybridization temperature for hybrids anticipated to be less than 50 base pairs in length should be 5-10°C less than the melting temperature (T_m) of the hybrid, where T_m is determined according to the following equations. For hybrids less than 18 base pairs in length, T_m(°C) = 2(# of A + T bases) + 4(# of G + C bases). For hybrids between 18 and 49 base pairs in length, T_m(°C) = 81.5 + 16.6($\log_{10}$[Na^+]) + 0.41(%G+C) - (600/N), where N is the number of bases in the hybrid, and [Na^+] is the concentration of sodium ions in the hybridization buffer ([Na^+] for 1xSSC = 0.165 M).

Additional examples of stringency conditions for polynucleotide hybridization are provided in Sambrook, J., E.F. Fritsch, and T. Maniatis, 1989, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, chapters 9 and 11, and *Current Protocols in Molecular Biology*, 1995, F.M. Ausubel et al., eds., John Wiley & Sons, Inc., sections 2.10 and 6.3-6.4, incorporated herein by reference.

Preferably, each such hybridizing polynucleotide has a length that is at least 25%(more preferably at least 50%, and most preferably at least 75%) of the length of the polynucleotide of the present invention to which it hybridizes, and has at least 60% sequence identity (more preferably, at least 75% identity; most preferably at least 90% or 95% identity) with the polynucleotide of the present invention to which it hybridizes, where sequence identity is determined by comparing the sequences of the hybridizing polynucleotides when aligned so as to maximize overlap and identity while minimizing sequence gaps.

The isolated polynucleotide of the invention may be operably linked to an expression control sequence such as the pMT2 or pED expression vectors disclosed in Kaufman *et al.*, Nucleic Acids Res. 19, 4485-4490 (1991), in order to produce the protein recombinantly. Many suitable expression control sequences are known in the art. General methods of expressing recombinant proteins are also known and are exemplified in R. Kaufman, Methods in Enzymology 185, 537-566 (1990). As defined herein "operably linked" means that the isolated polynucleotide of the invention and an expression control sequence are situated within a vector or cell in such a way that the protein is expressed by a host cell which has been transformed (transfected) with the ligated polynucleotide/expression control sequence.

A number of types of cells may act as suitable host cells for expression of the protein. Mammalian host cells include, for example, monkey COS cells, Chinese Hamster Ovary (CHO) cells, human kidney 293 cells, human epidermal A431 cells, human Colo205 cells, 3T3 cells, CV-1 cells, other transformed primate cell lines, normal diploid cells, cell strains derived from in vitro culture of primary tissue, primary explants, HeLa cells, mouse L cells, BHK, HL-60, U937, HaK or Jurkat cells.

Alternatively, it may be possible to produce the protein in lower eukaryotes such as yeast or in prokaryotes such as bacteria. Potentially suitable yeast strains include *Saccharomyces cerevisiae, Schizosaccharomyces pombe, Kluyveromyces* strains, *Candida*, or any yeast strain capable of expressing heterologous proteins. Potentially suitable bacterial strains include *Escherichia coli, Bacillus subtilis, Salmonella typhimurium*, or any bacterial

strain capable of expressing heterologous proteins. If the protein is made in yeast or bacteria, it may be necessary to modify the protein produced therein, for example by phosphorylation or glycosylation of the appropriate sites, in order to obtain the functional protein. Such covalent attachments may be accomplished using known chemical or enzymatic methods.

The protein may also be produced by operably linking the isolated polynucleotide of the invention to suitable control sequences in one or more insect expression vectors, and employing an insect expression system. Materials and methods for baculovirus/insect cell expression systems are commercially available in kit form from, *e.g.*, Invitrogen, San Diego, California, U.S.A. (the MaxBac® kit), and such methods are well known in the art, as described in Summers and Smith, Texas Agricultural Experiment Station Bulletin No. 1555 (1987), incorporated herein by reference. As used herein, an insect cell capable of expressing a polynucleotide of the present invention is "transformed."

The protein of the invention may be prepared by culturing transformed host cells under culture conditions suitable to express the recombinant protein. The resulting expressed protein may then be purified from such culture (i.e., from culture medium or cell extracts) using known purification processes, such as gel filtration and ion exchange chromatography. The purification of the protein may also include an affinity column containing agents which will bind to the protein; one or more column steps over such affinity resins as concanavalin A-agarose, heparin-toyopearl® or Cibacrom blue 3GA Sepharose®; one or more steps involving hydrophobic interaction chromatography using such resins as phenyl ether, butyl ether, or propyl ether; or immunoaffinity chromatography.

Alternatively, the protein of the invention may also be expressed in a form which will facilitate purification. For example, it may be expressed as a fusion protein, such as those of maltose binding protein (MBP), glutathione-S-transferase (GST) or thioredoxin (TRX). Kits for expression and purification of such fusion proteins are commercially available from New England BioLab (Beverly, MA), Pharmacia (Piscataway, NJ) and InVitrogen, respectively. The protein can also be tagged with an epitope and subsequently purified by using a specific antibody directed to such epitope. One such epitope ("Flag") is commercially available from Kodak (New Haven, CT).

Finally, one or more reverse-phase high performance liquid chromatography (RP-HPLC) steps employing hydrophobic RP-HPLC media, e.g., silica gel having pendant

methyl or other aliphatic groups, can be employed to further purify the protein. Some or all of the foregoing purification steps, in various combinations, can also be employed to provide a substantially homogeneous isolated recombinant protein. The protein thus purified is substantially free of other mammalian proteins and is defined in accordance with the present invention as an "isolated protein."

The protein of the invention may also be expressed as a product of transgenic animals, e.g., as a component of the milk of transgenic cows, goats, pigs, or sheep which are characterized by somatic or germ cells containing a nucleotide sequence encoding the protein.

The protein may also be produced by known conventional chemical synthesis. Methods for constructing the proteins of the present invention by synthetic means are known to those skilled in the art. The synthetically-constructed protein sequences, by virtue of sharing primary, secondary or tertiary structural and/or conformational characteristics with proteins may possess biological properties in common therewith, including protein activity. Thus, they may be employed as biologically active or immunological substitutes for natural, purified proteins in screening of therapeutic compounds and in immunological processes for the development of antibodies.

The proteins provided herein also include proteins characterized by amino acid sequences similar to those of purified proteins but into which modification are naturally provided or deliberately engineered. For example, modifications in the peptide or DNA sequences can be made by those skilled in the art using known techniques. Modifications of interest in the protein sequences may include the alteration, substitution, replacement, insertion or deletion of a selected amino acid residue in the coding sequence. For example, one or more of the cysteine residues may be deleted or replaced with another amino acid to alter the conformation of the molecule. Techniques for such alteration, substitution, replacement, insertion or deletion are well known to those skilled in the art (see, e.g., U.S. Patent No. 4,518,584). Preferably, such alteration, substitution, replacement, insertion or deletion retains the desired activity of the protein.

Other fragments and derivatives of the sequences of proteins which would be expected to retain protein activity in whole or in part and may thus be useful for screening or other immunological methodologies may also be easily made by those skilled in the art given the disclosures herein. Such modifications are believed to be encompassed by the present invention.

USES AND BIOLOGICAL ACTIVITY

The polynucleotides and proteins of the present invention are expected to exhibit one or more of the uses or biological activities (including those associated with assays cited herein) identified below. Uses or activities described for proteins of the present invention may be provided by administration or use of such proteins or by administration or use of polynucleotides encoding such proteins (such as, for example, in gene therapies or vectors suitable for introduction of DNA).

Research Uses and Utilities

The polynucleotides provided by the present invention can be used by the research community for various purposes. The polynucleotides can be used to express recombinant protein for analysis, characterization or therapeutic use; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in disease states); as molecular weight markers on Southern gels; as chromosome markers or tags (when labeled) to identify chromosomes or to map related gene positions; to compare with endogenous DNA sequences in patients to identify potential genetic disorders; as probes to hybridize and thus discover novel, related DNA sequences; as a source of information to derive PCR primers for genetic fingerprinting; as a probe to "subtract-out" known sequences in the process of discovering other novel polynucleotides; for selecting and making oligomers for attachment to a "gene chip" or other support, including for examination of expression patterns; to raise anti-protein antibodies using DNA immunization techniques; and as an antigen to raise anti-DNA antibodies or elicit another immune response. Where the polynucleotide encodes a protein which binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the polynucleotide can also be used in interaction trap assays (such as, for example, that described in Gyuris et al., Cell 75:791-803 (1993)) to identify polynucleotides encoding the other protein with which binding occurs or to identify inhibitors of the binding interaction.

The proteins provided by the present invention can similarly be used in assay to determine biological activity, including in a panel of multiple proteins for high-throughput screening; to raise antibodies or to elicit another immune response; as a reagent (including the labeled reagent) in assays designed to quantitatively determine levels of the protein (or its receptor) in biological fluids; as markers for tissues in which

the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in a disease state); and, of course, to isolate correlative receptors or ligands. Where the protein binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the protein can be used to identify the other protein with which binding occurs or to identify inhibitors of the binding interaction. Proteins involved in these binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction.

Any or all of these research utilities are capable of being developed into reagent grade or kit format for commercialization as research products.

Methods for performing the uses listed above are well known to those skilled in the art. References disclosing such methods include without limitation "Molecular Cloning: A Laboratory Manual", 2d ed., Cold Spring Harbor Laboratory Press, Sambrook, J., E.F. Fritsch and T. Maniatis eds., 1989, and "Methods in Enzymology: Guide to Molecular Cloning Techniques", Academic Press, Berger, S.L. and A.R. Kimmel eds., 1987.

Nutritional Uses

Polynucleotides and proteins of the present invention can also be used as nutritional sources or supplements. Such uses include without limitation use as a protein or amino acid supplement, use as a carbon source, use as a nitrogen source and use as a source of carbohydrate. In such cases the protein or polynucleotide of the invention can be added to the feed of a particular organism or can be administered as a separate solid or liquid preparation, such as in the form of powder, pills, solutions, suspensions or capsules. In the case of microorganisms, the protein or polynucleotide of the invention can be added to the medium in or on which the microorganism is cultured.

Cytokine and Cell Proliferation/Differentiation Activity

A protein of the present invention may exhibit cytokine, cell proliferation (either inducing or inhibiting) or cell differentiation (either inducing or inhibiting) activity or may induce production of other cytokines in certain cell populations. Many protein factors discovered to date, including all known cytokines, have exhibited activity in one or more factor dependent cell proliferation assays, and hence the assays serve as a convenient confirmation of cytokine activity. The activity of a protein of the present invention is evidenced by any one of a number of routine factor dependent cell proliferation assays

for cell lines including, without limitation, 32D, DA2, DA1G, T10, B9, B9/11, BaF3, MC9/G, M+ (preB M+), 2E8, RB5, DA1, 123, T1165, HT2, CTLL2, TF-1, Mo7e and CMK.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for T-cell or thymocyte proliferation include without limitation those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Bertagnolli et al., J. Immunol. 145:1706-1712, 1990; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Bertagnolli, et al., J. Immunol. 149:3778-3783, 1992; Bowman et al., J. Immunol. 152: 1756-1761, 1994.

Assays for cytokine production and/or proliferation of spleen cells, lymph node cells or thymocytes include, without limitation, those described in: Polyclonal T cell stimulation, Kruisbeek, A.M. and Shevach, E.M. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 3.12.1-3.12.14, John Wiley and Sons, Toronto. 1994; and Measurement of mouse and human Interferon γ, Schreiber, R.D. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.8.1-6.8.8, John Wiley and Sons, Toronto. 1994.

Assays for proliferation and differentiation of hematopoietic and lymphopoietic cells include, without limitation, those described in: Measurement of Human and Murine Interleukin 2 and Interleukin 4, Bottomly, K., Davis, L.S. and Lipsky, P.E. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.3.1-6.3.12, John Wiley and Sons, Toronto. 1991; deVries et al., J. Exp. Med. 173:1205-1211, 1991; Moreau et al., Nature 336:690-692, 1988; Greenberger et al., Proc. Natl. Acad. Sci. U.S.A. 80:2931-2938, 1983; Measurement of mouse and human interleukin 6 - Nordan, R. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.6.1-6.6.5, John Wiley and Sons, Toronto. 1991; Smith et al., Proc. Natl. Acad. Sci. U.S.A. 83:1857-1861, 1986; Measurement of human Interleukin 11 - Bennett, F., Giannotti, J., Clark, S.C. and Turner, K. J. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.15.1 John Wiley and Sons, Toronto. 1991; Measurement of mouse and human Interleukin 9 - Ciarletta, A., Giannotti, J., Clark, S.C. and Turner, K.J. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.13.1, John Wiley and Sons, Toronto. 1991.

Assays for T-cell clone responses to antigens (which will identify, among others, proteins that affect APC-T cell interactions as well as direct T-cell effects by measuring proliferation and cytokine production) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function; Chapter 6, Cytokines and their cellular receptors; Chapter 7, Immunologic studies in Humans); Weinberger et al., Proc. Natl. Acad. Sci. USA 77:6091-6095, 1980; Weinberger et al., Eur. J. Immun. 11:405-411, 1981; Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988.

Immune Stimulating or Suppressing Activity

A protein of the present invention may also exhibit immune stimulating or immune suppressing activity, including without limitation the activities for which assays are described herein. A protein may be useful in the treatment of various immune deficiencies and disorders (including severe combined immunodeficiency (SCID)), e.g., in regulating (up or down) growth and proliferation of T and/or B lymphocytes, as well as effecting the cytolytic activity of NK cells and other cell populations. These immune deficiencies may be genetic or be caused by viral (e.g., HIV) as well as bacterial or fungal infections, or may result from autoimmune disorders. More specifically, infectious diseases causes by viral, bacterial, fungal or other infection may be treatable using a protein of the present invention, including infections by HIV, hepatitis viruses, herpesviruses, mycobacteria, Leishmania spp., malaria spp. and various fungal infections such as candidiasis. Of course, in this regard, a protein of the present invention may also be useful where a boost to the immune system generally may be desirable, *i.e.*, in the treatment of cancer.

Autoimmune disorders which may be treated using a protein of the present invention include, for example, connective tissue disease, multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent diabetes mellitis, myasthenia gravis, graft-versus-host disease and autoimmune inflammatory eye disease. Such a protein of the present invention may also to be useful in the treatment of allergic reactions and conditions, such as asthma (particularly allergic asthma) or other respiratory problems. Other conditions, in which immune suppression is desired (including, for

example, organ transplantation), may also be treatable using a protein of the present invention.

Using the proteins of the invention it may also be possible to immune responses, in a number of ways. Down regulation may be in the form of inhibiting or blocking an immune response already in progress or may involve preventing the induction of an immune response. The functions of activated T cells may be inhibited by suppressing T cell responses or by inducing specific tolerance in T cells, or both. Immunosuppression of T cell responses is generally an active, non-antigen-specific, process which requires continuous exposure of the T cells to the suppressive agent. Tolerance, which involves inducing non-responsiveness or anergy in T cells, is distinguishable from immunosuppression in that it is generally antigen-specific and persists after exposure to the tolerizing agent has ceased. Operationally, tolerance can be demonstrated by the lack of a T cell response upon reexposure to specific antigen in the absence of the tolerizing agent.

Down regulating or preventing one or more antigen functions (including without limitation B lymphocyte antigen functions (such as , for example, B7)), *e.g.*, preventing high level lymphokine synthesis by activated T cells, will be useful in situations of tissue, skin and organ transplantation and in graft-versus-host disease (GVHD). For example, blockage of T cell function should result in reduced tissue destruction in tissue transplantation. Typically, in tissue transplants, rejection of the transplant is initiated through its recognition as foreign by T cells, followed by an immune reaction that destroys the transplant. The administration of a molecule which inhibits or blocks interaction of a B7 lymphocyte antigen with its natural ligand(s) on immune cells (such as a soluble, monomeric form of a peptide having B7-2 activity alone or in conjunction with a monomeric form of a peptide having an activity of another B lymphocyte antigen (*e.g.*, B7-1, B7-3) or blocking antibody), prior to transplantation can lead to the binding of the molecule to the natural ligand(s) on the immune cells without transmitting the corresponding costimulatory signal. Blocking B lymphocyte antigen function in this matter prevents cytokine synthesis by immune cells, such as T cells, and thus acts as an immunosuppressant. Moreover, the lack of costimulation may also be sufficient to anergize the T cells, thereby inducing tolerance in a subject. Induction of long-term tolerance by B lymphocyte antigen-blocking reagents may avoid the necessity of repeated administration of these blocking reagents. To achieve sufficient immunosuppression or

tolerance in a subject, it may also be necessary to block the function of a combination of B lymphocyte antigens.

The efficacy of particular blocking reagents in preventing organ transplant rejection or GVHD can be assessed using animal models that are predictive of efficacy in humans. Examples of appropriate systems which can be used include allogeneic cardiac grafts in rats and xenogeneic pancreatic islet cell grafts in mice, both of which have been used to examine the immunosuppressive effects of CTLA4Ig fusion proteins *in vivo* as described in Lenschow *et al.*, Science 257:789-792 (1992) and Turka *et al.*, Proc. Natl. Acad. Sci USA, *89*:11102-11105 (1992). In addition, murine models of GVHD (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 846-847) can be used to determine the effect of blocking B lymphocyte antigen function *in vivo* on the development of that disease.

Blocking antigen function may also be therapeutically useful for treating autoimmune diseases. Many autoimmune disorders are the result of inappropriate activation of T cells that are reactive against self tissue and which promote the production of cytokines and autoantibodies involved in the pathology of the diseases. Preventing the activation of autoreactive T cells may reduce or eliminate disease symptoms. Administration of reagents which block costimulation of T cells by disrupting receptor:ligand interactions of B lymphocyte antigens can be used to inhibit T cell activation and prevent production of autoantibodies or T cell-derived cytokines which may be involved in the disease process. Additionally, blocking reagents may induce antigen-specific tolerance of autoreactive T cells which could lead to long-term relief from the disease. The efficacy of blocking reagents in preventing or alleviating autoimmune disorders can be determined using a number of well-characterized animal models of human autoimmune diseases. Examples include murine experimental autoimmune encephalitis, systemic lupus erythmatosis in MRL/*lpr/lpr* mice or NZB hybrid mice, murine autoimmune collagen arthritis, diabetes mellitus in NOD mice and BB rats, and murine experimental myasthenia gravis (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 840-856).

Upregulation of an antigen function (preferably a B lymphocyte antigen function), as a means of up regulating immune responses, may also be useful in therapy. Upregulation of immune responses may be in the form of enhancing an existing immune response or eliciting an initial immune response. For example, enhancing an immune response through stimulating B lymphocyte antigen function may be useful in cases of

viral infection. In addition, systemic viral diseases such as influenza, the common cold, and encephalitis might be alleviated by the administration of stimulatory forms of B lymphocyte antigens systemically.

Alternatively, anti-viral immune responses may be enhanced in an infected patient by removing T cells from the patient, costimulating the T cells *in vitro* with viral antigen-pulsed APCs either expressing a peptide of the present invention or together with a stimulatory form of a soluble peptide of the present invention and reintroducing the *in vitro* activated T cells into the patient. Another method of enhancing anti-viral immune responses would be to isolate infected cells from a patient, transfect them with a nucleic acid encoding a protein of the present invention as described herein such that the cells express all or a portion of the protein on their surface, and reintroduce the transfected cells into the patient. The infected cells would now be capable of delivering a costimulatory signal to, and thereby activate, T cells *in vivo*.

In another application, up regulation or enhancement of antigen function (preferably B lymphocyte antigen function) may be useful in the induction of tumor immunity. Tumor cells (*e.g.*, sarcoma, melanoma, lymphoma, leukemia, neuroblastoma, carcinoma) transfected with a nucleic acid encoding at least one peptide of the present invention can be administered to a subject to overcome tumor-specific tolerance in the subject. If desired, the tumor cell can be transfected to express a combination of peptides. For example, tumor cells obtained from a patient can be transfected *ex vivo* with an expression vector directing the expression of a peptide having B7-2-like activity alone, or in conjunction with a peptide having B7-1-like activity and/or B7-3-like activity. The transfected tumor cells are returned to the patient to result in expression of the peptides on the surface of the transfected cell. Alternatively, gene therapy techniques can be used to target a tumor cell for transfection *in vivo*.

The presence of the peptide of the present invention having the activity of a B lymphocyte antigen(s) on the surface of the tumor cell provides the necessary costimulation signal to T cells to induce a T cell mediated immune response against the transfected tumor cells. In addition, tumor cells which lack MHC class I or MHC class II molecules, or which fail to reexpress sufficient amounts of MHC class I or MHC class II molecules, can be transfected with nucleic acid encoding all or a portion of (*e.g.*, a cytoplasmic-domain truncated portion) of an MHC class I α chain protein and β_2 microglobulin protein or an MHC class II α chain protein and an MHC class II β chain protein to thereby express MHC class I or MHC class II proteins on the cell surface.

Expression of the appropriate class I or class II MHC in conjunction with a peptide having the activity of a B lymphocyte antigen (*e.g.*, B7-1, B7-2, B7-3) induces a T cell mediated immune response against the transfected tumor cell. Optionally, a gene encoding an antisense construct which blocks expression of an MHC class II associated protein, such as the invariant chain, can also be cotransfected with a DNA encoding a peptide having the activity of a B lymphocyte antigen to promote presentation of tumor associated antigens and induce tumor specific immunity. Thus, the induction of a T cell mediated immune response in a human subject may be sufficient to overcome tumor-specific tolerance in the subject.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for thymocyte or splenocyte cytotoxicity include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Herrmann et al., Proc. Natl. Acad. Sci. USA 78:2488-2492, 1981; Herrmann et al., J. Immunol. 128:1968-1974, 1982; Handa et al., J. Immunol. 135:1564-1572, 1985; Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Herrmann et al., Proc. Natl. Acad. Sci. USA 78:2488-2492, 1981; Herrmann et al., J. Immunol. 128:1968-1974, 1982; Handa et al., J. Immunol. 135:1564-1572, 1985; Takai et al., J. Immunol. 137:3494-3500, 1986; Bowmanet al., J. Virology 61:1992-1998; Takai et al., J. Immunol. 140:508-512, 1988; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Brown et al., J. Immunol. 153:3079-3092, 1994.

Assays for T-cell-dependent immunoglobulin responses and isotype switching (which will identify, among others, proteins that modulate T-cell dependent antibody responses and that affect Th1/Th2 profiles) include, without limitation, those described in: Maliszewski, J. Immunol. 144:3028-3033, 1990; and Assays for B cell function: *In vitro* antibody production, Mond, J.J. and Brunswick, M. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 3.8.1-3.8.16, John Wiley and Sons, Toronto. 1994.

Mixed lymphocyte reaction (MLR) assays (which will identify, among others, proteins that generate predominantly Th1 and CTL responses) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter

7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Bertagnolli et al., J. Immunol. 149:3778-3783, 1992.

Dendritic cell-dependent assays (which will identify, among others, proteins expressed by dendritic cells that activate naive T-cells) include, without limitation, those described in: Guery et al., J. Immunol. 134:536-544, 1995; Inaba et al., Journal of Experimental Medicine 173:549-559, 1991; Macatonia et al., Journal of Immunology 154:5071-5079, 1995; Porgador et al., Journal of Experimental Medicine 182:255-260, 1995; Nair et al., Journal of Virology 67:4062-4069, 1993; Huang et al., Science 264:961-965, 1994; Macatonia et al., Journal of Experimental Medicine 169:1255-1264, 1989; Bhardwaj et al., Journal of Clinical Investigation 94:797-807, 1994; and Inaba et al., Journal of Experimental Medicine 172:631-640, 1990.

Assays for lymphocyte survival/apoptosis (which will identify, among others, proteins that prevent apoptosis after superantigen induction and proteins that regulate lymphocyte homeostasis) include, without limitation, those described in: Darzynkiewicz et al., Cytometry 13:795-808, 1992; Gorczyca et al., Leukemia 7:659-670, 1993; Gorczyca et al., Cancer Research 53:1945-1951, 1993; Itoh et al., Cell 66:233-243, 1991; Zacharchuk, Journal of Immunology 145:4037-4045, 1990; Zamai et al., Cytometry 14:891-897, 1993; Gorczyca et al., International Journal of Oncology 1:639-648, 1992.

Assays for proteins that influence early steps of T-cell commitment and development include, without limitation, those described in: Antica et al., Blood 84:111-117, 1994; Fine et al., Cellular Immunology 155:111-122, 1994; Galy et al., Blood 85:2770-2778, 1995; Toki et al., Proc. Nat. Acad Sci. USA 88:7548-7551, 1991.

Hematopoiesis Regulating Activity

A protein of the present invention may be useful in regulation of hematopoiesis and, consequently, in the treatment of myeloid or lymphoid cell deficiencies. Even marginal biological activity in support of colony forming cells or of factor-dependent cell lines indicates involvement in regulating hematopoiesis, e.g. in supporting the growth and proliferation of erythroid progenitor cells alone or in combination with other cytokines, thereby indicating utility, for example, in treating various anemias or for use in conjunction with irradiation/chemotherapy to stimulate the production of erythroid precursors and/or erythroid cells; in supporting the growth and proliferation of myeloid cells such as granulocytes and monocytes/macrophages (i.e., traditional CSF activity) useful, for example, in conjunction with chemotherapy to prevent or treat consequent

myelo-suppression; in supporting the growth and proliferation of megakaryocytes and consequently of platelets thereby allowing prevention or treatment of various platelet disorders such as thrombocytopenia, and generally for use in place of or complimentary to platelet transfusions; and/or in supporting the growth and proliferation of hematopoietic stem cells which are capable of maturing to any and all of the above-mentioned hematopoietic cells and therefore find therapeutic utility in various stem cell disorders (such as those usually treated with transplantation, including, without limitation, aplastic anemia and paroxysmal nocturnal hemoglobinuria), as well as in repopulating the stem cell compartment post irradiation/chemotherapy, either *in-vivo* or *ex-vivo* (i.e., in conjunction with bone marrow transplantation or with peripheral progenitor cell transplantation (homologous or heterologous)) as normal cells or genetically manipulated for gene therapy.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for proliferation and differentiation of various hematopoietic lines are cited above.

Assays for embryonic stem cell differentiation (which will identify, among others, proteins that influence embryonic differentiation hematopoiesis) include, without limitation, those described in: Johansson et al. Cellular Biology 15:141-151, 1995; Keller et al., Molecular and Cellular Biology 13:473-486, 1993; McClanahan et al., Blood 81:2903-2915, 1993.

Assays for stem cell survival and differentiation (which will identify, among others, proteins that regulate lympho-hematopoiesis) include, without limitation, those described in: Methylcellulose colony forming assays, Freshney, M.G. In *Culture of Hematopoietic Cells*. R.I. Freshney, *et al.* eds. Vol pp. 265-268, Wiley-Liss, Inc., New York, NY. 1994; Hirayama et al., Proc. Natl. Acad. Sci. USA 89:5907-5911, 1992; Primitive hematopoietic colony forming cells with high proliferative potential, McNiece, I.K. and Briddell, R.A. In *Culture of Hematopoietic Cells*. R.I. Freshney, *et al.* eds. Vol pp. 23-39, Wiley-Liss, Inc., New York, NY. 1994; Neben et al., Experimental Hematology 22:353-359, 1994; Cobblestone area forming cell assay, Ploemacher, R.E. In *Culture of Hematopoietic Cells*. R.I. Freshney, *et al.* eds. Vol pp. 1-21, Wiley-Liss, Inc.., New York, NY. 1994; Long term bone marrow cultures in the presence of stromal cells, Spooncer, E., Dexter, M. and Allen, T. In *Culture of Hematopoietic Cells*. R.I. Freshney, *et al.* eds. Vol pp. 163-179, Wiley-Liss, Inc., New York, NY. 1994; Long term culture initiating cell assay, Sutherland,

H.J. In *Culture of Hematopoietic Cells*. R.I. Freshney, *et al.* eds. Vol pp. 139-162, Wiley-Liss, Inc., New York, NY. 1994.

Tissue Growth Activity

A protein of the present invention also may have utility in compositions used for bone, cartilage, tendon, ligament and/or nerve tissue growth or regeneration, as well as for wound healing and tissue repair and replacement, and in the treatment of burns, incisions and ulcers.

A protein of the present invention, which induces cartilage and/or bone growth in circumstances where bone is not normally formed, has application in the healing of bone fractures and cartilage damage or defects in humans and other animals. Such a preparation employing a protein of the invention may have prophylactic use in closed as well as open fracture reduction and also in the improved fixation of artificial joints. *De novo* bone formation induced by an osteogenic agent contributes to the repair of congenital, trauma induced, or oncologic resection induced craniofacial defects, and also is useful in cosmetic plastic surgery.

A protein of this invention may also be used in the treatment of periodontal disease, and in other tooth repair processes. Such agents may provide an environment to attract bone-forming cells, stimulate growth of bone-forming cells or induce differentiation of progenitors of bone-forming cells. A protein of the invention may also be useful in the treatment of osteoporosis or osteoarthritis, such as through stimulation of bone and/or cartilage repair or by blocking inflammation or processes of tissue destruction (collagenase activity, osteoclast activity, etc.) mediated by inflammatory processes.

Another category of tissue regeneration activity that may be attributable to the protein of the present invention is tendon/ligament formation. A protein of the present invention, which induces tendon/ligament-like tissue or other tissue formation in circumstances where such tissue is not normally formed, has application in the healing of tendon or ligament tears, deformities and other tendon or ligament defects in humans and other animals. Such a preparation employing a tendon/ligament-like tissue inducing protein may have prophylactic use in preventing damage to tendon or ligament tissue, as well as use in the improved fixation of tendon or ligament to bone or other tissues, and in repairing defects to tendon or ligament tissue. De novo tendon/ligament-like tissue formation induced by a composition of the present invention contributes to the repair of

congenital, trauma induced, or other tendon or ligament defects of other origin, and is also useful in cosmetic plastic surgery for attachment or repair of tendons or ligaments. The compositions of the present invention may provide an environment to attract tendon- or ligament-forming cells, stimulate growth of tendon- or ligament-forming cells, induce differentiation of progenitors of tendon- or ligament-forming cells, or induce growth of tendon/ligament cells or progenitors *ex vivo* for return *in vivo* to effect tissue repair. The compositions of the invention may also be useful in the treatment of tendinitis, carpal tunnel syndrome and other tendon or ligament defects. The compositions may also include an appropriate matrix and/or sequestering agent as a carrier as is well known in the art.

The protein of the present invention may also be useful for proliferation of neural cells and for regeneration of nerve and brain tissue, *i.e.* for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders, which involve degeneration, death or trauma to neural cells or nerve tissue. More specifically, a protein may be used in the treatment of diseases of the peripheral nervous system, such as peripheral nerve injuries, peripheral neuropathy and localized neuropathies, and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome. Further conditions which may be treated in accordance with the present invention include mechanical and traumatic disorders, such as spinal cord disorders, head trauma and cerebrovascular diseases such as stroke. Peripheral neuropathies resulting from chemotherapy or other medical therapies may also be treatable using a protein of the invention.

Proteins of the invention may also be useful to promote better or faster closure of non-healing wounds, including without limitation pressure ulcers, ulcers associated with vascular insufficiency, surgical and traumatic wounds, and the like.

It is expected that a protein of the present invention may also exhibit activity for generation or regeneration of other tissues, such as organs (including, for example, pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac) and vascular (including vascular endothelium) tissue, or for promoting the growth of cells comprising such tissues. Part of the desired effects may be by inhibition or modulation of fibrotic scarring to allow normal tissue to regenerate. A protein of the invention may also exhibit angiogenic activity.

A protein of the present invention may also be useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, and conditions resulting from systemic cytokine damage.

A protein of the present invention may also be useful for promoting or inhibiting differentiation of tissues described above from precursor tissues or cells; or for inhibiting the growth of tissues described above.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for tissue generation activity include, without limitation, those described in: International Patent Publication No. WO95/16035 (bone, cartilage, tendon); International Patent Publication No. WO95/05846 (nerve, neuronal); International Patent Publication No. WO91/07491 (skin, endothelium).

Assays for wound healing activity include, without limitation, those described in: Winter, Epidermal Wound Healing, pps. 71-112 (Maibach, HI and Rovee, DT, eds.), Year Book Medical Publishers, Inc., Chicago, as modified by Eaglstein and Mertz, J. Invest. Dermatol 71:382-84 (1978).

Activin/Inhibin Activity

A protein of the present invention may also exhibit activin- or inhibin-related activities. Inhibins are characterized by their ability to inhibit the release of follicle stimulating hormone (FSH), while activins and are characterized by their ability to stimulate the release of follicle stimulating hormone (FSH). Thus, a protein of the present invention, alone or in heterodimers with a member of the inhibin α family, may be useful as a contraceptive based on the ability of inhibins to decrease fertility in female mammals and decrease spermatogenesis in male mammals. Administration of sufficient amounts of other inhibins can induce infertility in these mammals. Alternatively, the protein of the invention, as a homodimer or as a heterodimer with other protein subunits of the inhibin-β group, may be useful as a fertility inducing therapeutic, based upon the ability of activin molecules in stimulating FSH release from cells of the anterior pituitary. See, for example, United States Patent 4,798,885. A protein of the invention may also be useful for advancement of the onset of fertility in sexually immature mammals, so as to increase the lifetime reproductive performance of domestic animals such as cows, sheep and pigs.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for activin/inhibin activity include, without limitation, those described in: Vale et al., Endocrinology 91:562-572, 1972; Ling et al., Nature 321:779-782, 1986; Vale et al., Nature 321:776-779, 1986; Mason et al., Nature 318:659-663, 1985; Forage et al., Proc. Natl. Acad. Sci. USA 83:3091-3095, 1986.

Chemotactic/Chemokinetic Activity

A protein of the present invention may have chemotactic or chemokinetic activity (e.g., act as a chemokine) for mammalian cells, including, for example, monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells. Chemotactic and chemokinetic proteins can be used to mobilize or attract a desired cell population to a desired site of action. Chemotactic or chemokinetic proteins provide particular advantages in treatment of wounds and other trauma to tissues, as well as in treatment of localized infections. For example, attraction of lymphocytes, monocytes or neutrophils to tumors or sites of infection may result in improved immune responses against the tumor or infecting agent.

A protein or peptide has chemotactic activity for a particular cell population if it can stimulate, directly or indirectly, the directed orientation or movement of such cell population. Preferably, the protein or peptide has the ability to directly stimulate directed movement of cells. Whether a particular protein has chemotactic activity for a population of cells can be readily determined by employing such protein or peptide in any known assay for cell chemotaxis.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for chemotactic activity (which will identify proteins that induce or prevent chemotaxis) consist of assays that measure the ability of a protein to induce the migration of cells across a membrane as well as the ability of a protein to induce the adhesion of one cell population to another cell population. Suitable assays for movement and adhesion include, without limitation, those described in: Current Protocols in Immunology, Ed by J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W.Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 6.12, Measurement of alpha and beta Chemokines 6.12.1-6.12.28; Taub et al. J. Clin. Invest. 95:1370-1376, 1995; Lind et al. APMIS 103:140-146, 1995; Muller et al Eur. J. Immunol. 25: 1744-1748; Gruber et al. J. of Immunol. 152:5860-5867, 1994; Johnston et al. J. of Immunol. 153: 1762-1768, 1994.

Hemostatic and Thrombolytic Activity

A protein of the invention may also exhibit hemostatic or thrombolytic activity. As a result, such a protein is expected to be useful in treatment of various coagulation disorders (including hereditary disorders, such as hemophilias) or to enhance coagulation and other hemostatic events in treating wounds resulting from trauma, surgery or other causes. A protein of the invention may also be useful for dissolving or inhibiting formation of thromboses and for treatment and prevention of conditions resulting therefrom (such as, for example, infarction of cardiac and central nervous system vessels (e.g., stroke).

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assay for hemostatic and thrombolytic activity include, without limitation, those described in: Linet et al., J. Clin. Pharmacol. 26:131-140, 1986; Burdick et al., Thrombosis Res. 45:413-419, 1987; Humphrey et al., Fibrinolysis 5:71-79 (1991); Schaub, Prostaglandins 35:467-474, 1988.

Receptor/Ligand Activity

A protein of the present invention may also demonstrate activity as receptors, receptor ligands or inhibitors or agonists of receptor/ligand interactions. Examples of such receptors and ligands include, without limitation, cytokine receptors and their ligands, receptor kinases and their ligands, receptor phosphatases and their ligands, receptors involved in cell-cell interactions and their ligands (including without limitation, cellular adhesion molecules (such as selectins, integrins and their ligands) and receptor/ligand pairs involved in antigen presentation, antigen recognition and development of cellular and humoral immune responses). Receptors and ligands are also useful for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction. A protein of the present invention (including, without limitation, fragments of receptors and ligands) may themselves be useful as inhibitors of receptor/ligand interactions.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for receptor-ligand activity include without limitation those described in:Current Protocols in Immunology, Ed by J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W.Strober, Pub. Greene Publishing Associates and

Wiley-Interscience (Chapter 7.28, Measurement of Cellular Adhesion under static conditions 7.28.1-7.28.22), Takai et al., Proc. Natl. Acad. Sci. USA 84:6864-6868, 1987; Bierer et al., J. Exp. Med. 168:1145-1156, 1988; Rosenstein et al., J. Exp. Med. 169:149-160 1989; Stoltenborg et al., J. Immunol. Methods 175:59-68, 1994; Stitt et al., Cell 80:661-670, 1995.

Anti-Inflammatory Activity

Proteins of the present invention may also exhibit anti-inflammatory activity. The anti-inflammatory activity may be achieved by providing a stimulus to cells involved in the inflammatory response, by inhibiting or promoting cell-cell interactions (such as, for example, cell adhesion), by inhibiting or promoting chemotaxis of cells involved in the inflammatory process, inhibiting or promoting cell extravasation, or by stimulating or suppressing production of other factors which more directly inhibit or promote an inflammatory response. Proteins exhibiting such activities can be used to treat inflammatory conditions including chronic or acute conditions), including without limitation inflammation associated with infection (such as septic shock, sepsis or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine-induced lung injury, inflammatory bowel disease, Crohn's disease or resulting from over production of cytokines such as TNF or IL-1. Proteins of the invention may also be useful to treat anaphylaxis and hypersensitivity to an antigenic substance or material.

Cadherin/Tumor Invasion Suppressor Activity

Cadherins are calcium-dependent adhesion molecules that appear to play major roles during development, particularly in defining specific cell types. Loss or alteration of normal cadherin expression can lead to changes in cell adhesion properties linked to tumor growth and metastasis. Cadherin malfunction is also implicated in other human diseases, such as pemphigus vulgaris and pemphigus foliaceus (auto-immune blistering skin diseases), Crohn's disease, and some developmental abnormalities.

The cadherin superfamily includes well over forty members, each with a distinct pattern of expression. All members of the superfamily have in common conserved extracellular repeats (cadherin domains), but structural differences are found in other parts of the molecule. The cadherin domains bind calcium to form their tertiary structure and thus calcium is required to mediate their adhesion. Only a few amino acids in the

first cadherin domain provide the basis for homophilic adhesion; modification of this recognition site can change the specificity of a cadherin so that instead of recognizing only itself, the mutant molecule can now also bind to a different cadherin. In addition, some cadherins engage in heterophilic adhesion with other cadherins.

E-cadherin, one member of the cadherin superfamily, is expressed in epithelial cell types. Pathologically, if E-cadherin expression is lost in a tumor, the malignant cells become invasive and the cancer metastasizes. Transfection of cancer cell lines with polynucleotides expressing E-cadherin has reversed cancer-associated changes by returning altered cell shapes to normal, restoring cells' adhesiveness to each other and to their substrate, decreasing the cell growth rate, and drastically reducing anchorage-independent cell growth. Thus, reintroducing E-cadherin expression reverts carcinomas to a less advanced stage. It is likely that other cadherins have the same invasion suppressor role in carcinomas derived from other tissue types. Therefore, proteins of the present invention with cadherin activity, and polynucleotides of the present invention encoding such proteins, can be used to treat cancer. Introducing such proteins or polynucleotides into cancer cells can reduce or eliminate the cancerous changes observed in these cells by providing normal cadherin expression.

Cancer cells have also been shown to express cadherins of a different tissue type than their origin, thus allowing these cells to invade and metastasize in a different tissue in the body. Proteins of the present invention with cadherin activity, and polynucleotides of the present invention encoding such proteins, can be substituted in these cells for the inappropriately expressed cadherins, restoring normal cell adhesive properties and reducing or eliminating the tendency of the cells to metastasize.

Additionally, proteins of the present invention with cadherin activity, and polynucleotides of the present invention encoding such proteins, can used to generate antibodies recognizing and binding to cadherins. Such antibodies can be used to block the adhesion of inappropriately expressed tumor-cell cadherins, preventing the cells from forming a tumor elsewhere. Such an anti-cadherin antibody can also be used as a marker for the grade, pathological type, and prognosis of a cancer, i.e. the more progressed the cancer, the less cadherin expression there will be, and this decrease in cadherin expression can be detected by the use of a cadherin-binding antibody.

Fragments of proteins of the present invention with cadherin activity, preferably a polypeptide comprising a decapeptide of the cadherin recognition site, and polynucleotides of the present invention encoding such protein fragments, can also be used

to block cadherin function by binding to cadherins and preventing them from binding in ways that produce undesirable effects. Additionally, fragments of proteins of the present invention with cadherin activity, preferably truncated soluble cadherin fragments which have been found to be stable in the circulation of cancer patients, and polynucleotides encoding such protein fragments, can be used to disturb proper cell-cell adhesion.

Assays for cadherin adhesive and invasive suppressor activity include, without limitation, those described in: Hortsch et al. J Biol Chem 270 (32): 18809-18817, 1995; Miyaki et al. Oncogene 11: 2547-2552, 1995; Ozawa et al. Cell 63: 1033-1038, 1990.

Tumor Inhibition Activity

In addition to the activities described above for immunological treatment or prevention of tumors, a protein of the invention may exhibit other anti-tumor activities. A protein may inhibit tumor growth directly or indirectly (such as, for example, via ADCC). A protein may exhibit its tumor inhibitory activity by acting on tumor tissue or tumor precursor tissue, by inhibiting formation of tissues necessary to support tumor growth (such as, for example, by inhibiting angiogenesis), by causing production of other factors, agents or cell types which inhibit tumor growth, or by suppressing, eliminating or inhibiting factors, agents or cell types which promote tumor growth.

Other Activities

A protein of the invention may also exhibit one or more of the following additional activities or effects: inhibiting the growth, infection or function of, or killing, infectious agents, including, without limitation, bacteria, viruses, fungi and other parasites; effecting (suppressing or enhancing) bodily characteristics, including, without limitation, height, weight, hair color, eye color, skin, fat to lean ratio or other tissue pigmentation, or organ or body part size or shape (such as, for example, breast augmentation or diminution, change in bone form or shape); effecting biorhythms or caricadic cycles or rhythms; effecting the fertility of male or female subjects; effecting the metabolism, catabolism, anabolism, processing, utilization, storage or elimination of dietary fat, lipid, protein, carbohydrate, vitamins, minerals, cofactors or other nutritional factors or component(s); effecting behavioral characteristics, including, without limitation, appetite, libido, stress, cognition (including cognitive disorders), depression (including depressive disorders) and violent behaviors; providing analgesic effects or other pain reducing effects; promoting differentiation and growth of embryonic stem cells in lineages other than hematopoietic

lineages; hormonal or endocrine activity; in the case of enzymes, correcting deficiencies of the enzyme and treating deficiency-related diseases; treatment of hyperproliferative disorders (such as, for example, psoriasis); immunoglobulin-like activity (such as, for example, the ability to bind antigens or complement); and the ability to act as an antigen in a vaccine composition to raise an immune response against such protein or another material or entity which is cross-reactive with such protein.

ADMINISTRATION AND DOSING

A protein of the present invention (from whatever source derived, including without limitation from recombinant and non-recombinant sources) may be used in a pharmaceutical composition when combined with a pharmaceutically acceptable carrier. Such a composition may also contain (in addition to protein and a carrier) diluents, fillers, salts, buffers, stabilizers, solubilizers, and other materials well known in the art. The term "pharmaceutically acceptable" means a non-toxic material that does not interfere with the effectiveness of the biological activity of the active ingredient(s). The characteristics of the carrier will depend on the route of administration. The pharmaceutical composition of the invention may also contain cytokines, lymphokines, or other hematopoietic factors such as M-CSF, GM-CSF, TNF, IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IFN, TNF0, TNF1, TNF2, G-CSF, Meg-CSF, thrombopoietin, stem cell factor, and erythropoietin. The pharmaceutical composition may further contain other agents which either enhance the activity of the protein or compliment its activity or use in treatment. Such additional factors and/or agents may be included in the pharmaceutical composition to produce a synergistic effect with protein of the invention, or to minimize side effects. Conversely, protein of the present invention may be included in formulations of the particular cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent to minimize side effects of the cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent.

A protein of the present invention may be active in multimers (e.g., heterodimers or homodimers) or complexes with itself or other proteins. As a result, pharmaceutical compositions of the invention may comprise a protein of the invention in such multimeric or complexed form.

The pharmaceutical composition of the invention may be in the form of a complex of the protein(s) of present invention along with protein or peptide antigens. The protein and/or peptide antigen will deliver a stimulatory signal to both B and T lymphocytes. B lymphocytes will respond to antigen through their surface immunoglobulin receptor. T lymphocytes will respond to antigen through the T cell receptor (TCR) following presentation of the antigen by MHC proteins. MHC and structurally related proteins including those encoded by class I and class II MHC genes on host cells will serve to present the peptide antigen(s) to T lymphocytes. The antigen components could also be supplied as purified MHC-peptide complexes alone or with co-stimulatory molecules that can directly signal T cells. Alternatively antibodies able to bind surface immunolgobulin and other molecules on B cells as well as antibodies able to bind the TCR and other molecules on T cells can be combined with the pharmaceutical composition of the invention.

The pharmaceutical composition of the invention may be in the form of a liposome in which protein of the present invention is combined, in addition to other pharmaceutically acceptable carriers, with amphipathic agents such as lipids which exist in aggregated form as micelles, insoluble monolayers, liquid crystals, or lamellar layers in aqueous solution. Suitable lipids for liposomal formulation include, without limitation, monoglycerides, diglycerides, sulfatides, lysolecithin, phospholipids, saponin, bile acids, and the like. Preparation of such liposomal formulations is within the level of skill in the art, as disclosed, for example, in U.S. Patent No. 4,235,871; U.S. Patent No. 4,501,728; U.S. Patent No. 4,837,028; and U.S. Patent No. 4,737,323, all of which are incorporated herein by reference.

As used herein, the term "therapeutically effective amount" means the total amount of each active component of the pharmaceutical composition or method that is sufficient to show a meaningful patient benefit, i.e., treatment, healing, prevention or amelioration of the relevant medical condition, or an increase in rate of treatment, healing, prevention or amelioration of such conditions. When applied to an individual active ingredient, administered alone, the term refers to that ingredient alone. When applied to a combination, the term refers to combined amounts of the active ingredients that result in the therapeutic effect, whether administered in combination, serially or simultaneously.

In practicing the method of treatment or use of the present invention, a therapeutically effective amount of protein of the present invention is administered to a mammal having a condition to be treated. Protein of the present invention may be

administered in accordance with the method of the invention either alone or in combination with other therapies such as treatments employing cytokines, lymphokines or other hematopoietic factors. When co-administered with one or more cytokines, lymphokines or other hematopoietic factors, protein of the present invention may be administered either simultaneously with the cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors, or sequentially. If administered sequentially, the attending physician will decide on the appropriate sequence of administering protein of the present invention in combination with cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors.

Administration of protein of the present invention used in the pharmaceutical composition or to practice the method of the present invention can be carried out in a variety of conventional ways, such as oral ingestion, inhalation, topical application or cutaneous, subcutaneous, intraperitoneal, parenteral or intravenous injection. Intravenous administration to the patient is preferred.

When a therapeutically effective amount of protein of the present invention is administered orally, protein of the present invention will be in the form of a tablet, capsule, powder, solution or elixir. When administered in tablet form, the pharmaceutical composition of the invention may additionally contain a solid carrier such as a gelatin or an adjuvant. The tablet, capsule, and powder contain from about 5 to 95% protein of the present invention, and preferably from about 25 to 90% protein of the present invention. When administered in liquid form, a liquid carrier such as water, petroleum, oils of animal or plant origin such as peanut oil, mineral oil, soybean oil, or sesame oil, or synthetic oils may be added. The liquid form of the pharmaceutical composition may further contain physiological saline solution, dextrose or other saccharide solution, or glycols such as ethylene glycol, propylene glycol or polyethylene glycol. When administered in liquid form, the pharmaceutical composition contains from about 0.5 to 90% by weight of protein of the present invention, and preferably from about 1 to 50% protein of the present invention.

When a therapeutically effective amount of protein of the present invention is administered by intravenous, cutaneous or subcutaneous injection, protein of the present invention will be in the form of a pyrogen-free, parenterally acceptable aqueous solution. The preparation of such parenterally acceptable protein solutions, having due regard to pH, isotonicity, stability, and the like, is within the skill in the art. A preferred

pharmaceutical composition for intravenous, cutaneous, or subcutaneous injection should contain, in addition to protein of the present invention, an isotonic vehicle such as Sodium Chloride Injection, Ringer's Injection, Dextrose Injection, Dextrose and Sodium Chloride Injection, Lactated Ringer's Injection, or other vehicle as known in the art. The pharmaceutical composition of the present invention may also contain stabilizers, preservatives, buffers, antioxidants, or other additives known to those of skill in the art.

The amount of protein of the present invention in the pharmaceutical composition of the present invention will depend upon the nature and severity of the condition being treated, and on the nature of prior treatments which the patient has undergone. Ultimately, the attending physician will decide the amount of protein of the present invention with which to treat each individual patient. Initially, the attending physician will administer low doses of protein of the present invention and observe the patient's response. Larger doses of protein of the present invention may be administered until the optimal therapeutic effect is obtained for the patient, and at that point the dosage is not increased further. It is contemplated that the various pharmaceutical compositions used to practice the method of the present invention should contain about 0.01 µg to about 100 mg (preferably about 0.1ng to about 10 mg, more preferably about 0.1 µg to about 1 mg) of protein of the present invention per kg body weight.

The duration of intravenous therapy using the pharmaceutical composition of the present invention will vary, depending on the severity of the disease being treated and the condition and potential idiosyncratic response of each individual patient. It is contemplated that the duration of each application of the protein of the present invention will be in the range of 12 to 24 hours of continuous intravenous administration. Ultimately the attending physician will decide on the appropriate duration of intravenous therapy using the pharmaceutical composition of the present invention.

Protein of the invention may also be used to immunize animals to obtain polyclonal and monoclonal antibodies which specifically react with the protein. Such antibodies may be obtained using either the entire protein or fragments thereof as an immunogen. The peptide immunogens additionally may contain a cysteine residue at the carboxyl terminus, and are conjugated to a hapten such as keyhole limpet hemocyanin (KLH). Methods for synthesizing such peptides are known in the art, for example, as in R.P. Merrifield, J. Amer.Chem.Soc. 85, 2149-2154 (1963); J.L. Krstenansky, *et al.*, FEBS Lett. 211, 10 (1987). Monoclonal antibodies binding to the protein of the invention may be useful diagnostic agents for the immunodetection of the protein. Neutralizing monoclonal

antibodies binding to the protein may also be useful therapeutics for both conditions associated with the protein and also in the treatment of some forms of cancer where abnormal expression of the protein is involved. In the case of cancerous cells or leukemic cells, neutralizing monoclonal antibodies against the protein may be useful in detecting and preventing the metastatic spread of the cancerous cells, which may be mediated by the protein.

For compositions of the present invention which are useful for bone, cartilage, tendon or ligament regeneration, the therapeutic method includes administering the composition topically, systematically, or locally as an implant or device. When administered, the therapeutic composition for use in this invention is, of course, in a pyrogen-free, physiologically acceptable form. Further, the composition may desirably be encapsulated or injected in a viscous form for delivery to the site of bone, cartilage or tissue damage. Topical administration may be suitable for wound healing and tissue repair. Therapeutically useful agents other than a protein of the invention which may also optionally be included in the composition as described above, may alternatively or additionally, be administered simultaneously or sequentially with the composition in the methods of the invention. Preferably for bone and/or cartilage formation, the composition would include a matrix capable of delivering the protein-containing composition to the site of bone and/or cartilage damage, providing a structure for the developing bone and cartilage and optimally capable of being resorbed into the body. Such matrices may be formed of materials presently in use for other implanted medical applications.

The choice of matrix material is based on biocompatibility, biodegradability, mechanical properties, cosmetic appearance and interface properties. The particular application of the compositions will define the appropriate formulation. Potential matrices for the compositions may be biodegradable and chemically defined calcium sulfate, tricalciumphosphate, hydroxyapatite, polylactic acid, polyglycolic acid and polyanhydrides. Other potential materials are biodegradable and biologically well-defined, such as bone or dermal collagen. Further matrices are comprised of pure proteins or extracellular matrix components. Other potential matrices are nonbiodegradable and chemically defined, such as sintered hydroxapatite, bioglass, aluminates, or other ceramics. Matrices may be comprised of combinations of any of the above mentioned types of material, such as polylactic acid and hydroxyapatite or collagen and tricalciumphosphate. The bioceramics may be altered in composition, such as in calcium-

aluminate-phosphate and processing to alter pore size, particle size, particle shape, and biodegradability.

Presently preferred is a 50:50 (mole weight) copolymer of lactic acid and glycolic acid in the form of porous particles having diameters ranging from 150 to 800 microns. In some applications, it will be useful to utilize a sequestering agent, such as carboxymethyl cellulose or autologous blood clot, to prevent the protein compositions from disassociating from the matrix.

A preferred family of sequestering agents is cellulosic materials such as alkylcelluloses (including hydroxyalkylcelluloses), including methylcellulose, ethylcellulose, hydroxyethylcellulose, hydroxypropylcellulose, hydroxypropyl-methylcellulose, and carboxymethylcellulose, the most preferred being cationic salts of carboxymethylcellulose (CMC). Other preferred sequestering agents include hyaluronic acid, sodium alginate, poly(ethylene glycol), polyoxyethylene oxide, carboxyvinyl polymer and poly(vinyl alcohol). The amount of sequestering agent useful herein is 0.5-20 wt%, preferably 1-10 wt% based on total formulation weight, which represents the amount necessary to prevent desorbtion of the protein from the polymer matrix and to provide appropriate handling of the composition, yet not so much that the progenitor cells are prevented from infiltrating the matrix, thereby providing the protein the opportunity to assist the osteogenic activity of the progenitor cells.

In further compositions, proteins of the invention may be combined with other agents beneficial to the treatment of the bone and/or cartilage defect, wound, or tissue in question. These agents include various growth factors such as epidermal growth factor (EGF), platelet derived growth factor (PDGF), transforming growth factors (TGF-α and TGF-β), and insulin-like growth factor (IGF).

The therapeutic compositions are also presently valuable for veterinary applications. Particularly domestic animals and thoroughbred horses, in addition to humans, are desired patients for such treatment with proteins of the present invention.

The dosage regimen of a protein-containing pharmaceutical composition to be used in tissue regeneration will be determined by the attending physician considering various factors which modify the action of the proteins, e.g., amount of tissue weight desired to be formed, the site of damage, the condition of the damaged tissue, the size of a wound, type of damaged tissue (e.g., bone), the patient's age, sex, and diet, the severity of any infection, time of administration and other clinical factors. The dosage may vary with the type of matrix used in the reconstitution and with inclusion of other proteins in

the pharmaceutical composition. For example, the addition of other known growth factors, such as IGF I (insulin like growth factor I), to the final composition, may also effect the dosage. Progress can be monitored by periodic assessment of tissue/bone growth and/or repair, for example, X-rays, histomorphometric determinations and tetracycline labeling.

Polynucleotides of the present invention can also be used for gene therapy. Such polynucleotides can be introduced either *in vivo* or *ex vivo* into cells for expression in a mammalian subject. Polynucleotides of the invention may also be administered by other known methods for introduction of nucleic acid into a cell or organism (including, without limitation, in the form of viral vectors or naked DNA).

Cells may also be cultured *ex vivo* in the presence of proteins of the present invention in order to proliferate or to produce a desired effect on or activity in such cells. Treated cells can then be introduced *in vivo* for therapeutic purposes.

Patent and literature references cited herein are incorporated by reference as if fully set forth.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Jacobs, Kenneth
McCoy, John M.
LaVallie, Edward R.
Racie, Lisa A.
Merberg, David
Treacy, Maurice
Spaulding, Vikki
Agostino, Michael

(ii) TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

(iii) NUMBER OF SEQUENCES: 32

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Genetics Institute, Inc.
(B) STREET: 87 CambridgePark Drive
(C) CITY: Cambridge
(D) STATE: MA
(E) COUNTRY: U.S.A.
(F) ZIP: 02140

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Sprunger, Suzanne A.
(B) REGISTRATION NUMBER: 41,323

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (617) 498-8284
(B) TELEFAX: (617) 876-5851

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1800 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

TTTTTTTTTT TACAGACTTC ACAGAGAATG CAGTTGTCTT GACTTCAGGT CTGTCTGTTC 60

TGTTGGCAAG TAAATGCAGT ACTGTTCTGA TCCCGCTGCT ATTAGAATGC ATTGTGAAAC 120

GACTGGAGTA TGATTAAAAG TTGTGTTCCC CAATGCTTGG AGTAGTGATT GTTGAAGGAA 180

AAAATCCAGC TGAGTGATAA AGGCTGAGTG TTGAGGAAAT TTCTGCAGTT TTAAGCAGTC 240

GTATTTGTGA TTGAAGCTGA GTACATTTTG CTGGTGTATT TTTAGGTAAA ATGCTTTTTG 300

TTCATTTCTG GTGGTGGGAG GGGACTGAAG CCTTTAGTCT TTTCCAGATG CAACCTTAAA 360

ATCAGTGACA AGAAACATTC CAAACAAGCA ACAGTCTTCA AGAAATTAAA CTGGCAAGTG 420

GAAATGTTTA AACAGTTCAG TGATCTTTAG TGCATTGTTT ATGTGTGGGT TTCTCTCTCC 480

CCTCCCTTGG TCTTAATTCT TACATGCAGG AACACTCAGC AGACACACGT ATGCGAAGGG 540

CCAGAGAAGC CAGACCCAGT AAGAAAAAAT AGCCTATTTA CTTTAAATAA ACCAAACATT 600

CCATTTTAAA TGTGGGGATT GGGAACCACT AGTTCTTTCA GATGGTATTC TTCAGACTAT 660

AGAAGGAGCT TCCAGTTGAA TTCACCAGTG GACAAAATGA GGAAAACAGG TGAACAAGCT 720

TTTTCTGTAT TTACATACAA AGTCAGATCA GTTATGGGAC AATAGTATTG AATAGATTTC 780

AGCTTTATGC TGGAGTAACT GGCATGTGAG CAAACTGTGT TGGCGTGGGG GTGGAGGGGT 840

GAGGTGGGCG CTAAGCTTTT TTTAAGATTT TTCAGGTACC CTTCACTAAA GGCACCGAAG 900

GCTTAAAGTA GGACAACCAT GGAGCTTCCT GTGGCAGGAG AGACAACAAA GCGCTATTAT 960

CCTAAGGTCA AGAGAAGTGT CAGCCTCACC TGATTTTTAT TAGTAATGAG GACTTGCCTC 1020

AACTCCCTCT TTCTGGAGTG AAGCATCCGA AGGAATGCTT GAAGTACCCC TGGGCTTCTC 1080

TTAACATTTA AGCAAGCTGT TTTTATAGCA GCTCTTAATA ATAAAGCCCA AATCTCAAGC 1140

GGTGCTTGAA GGGGAGGGAA AGGGGGAAAG CGGGCAACCA CTTTTCCCTA GCTTTTCCAG 1200

AAGCCTGTTA AAAGCAAGGT CTCCCCACAA GCAACTTCTC TGCCACATCG CCACCCCGTG 1260

CCTTTTGATC TAGCACAGAC CCTTCACCCC TCACCTCGAT GCAGCCAGTA GCTTGGATCC 1320

TTGTGGGCAT GATCCATAAT CGGTTTCAAG GTAACGATGG TGTCGAGKTC TTTGGTGGGT 1380

TGAACTATGT TAGAAAAGGC CATTAATTTG CCTGCAAATT GTTAACAGAA GGGTATTAAA 1440

ACCACAGCTA AGTAGCTCTA TTATAATACT TATCCAGTGA CTAAAACCAA CTTAAACCAG 1500

TAAGTGGAGA AATAACATGT TCAAGAACTG TAATGCTGGG TGGGAACATG TAACTTGTAG 1560

ACTGGAGAAG ATAGGCATTT GAGTGGCTGA GAGGGCTTTT GGGTGGGAAT GCAAAAATTC 1620

TCTGCTAAGA CTTTTTCAGG TGAACATAAC AGACTTGGCC AAGCTAGCAT CTTAGCGGAA 1680

GCTGATCTCC AATGCTCTTC AGTAGGGTCA TGAAGGTTTT TCTTTTCCTG AGAAAACAAC 1740

ACGTATTGTT TTCTCAGGTT TTGCTTTTTG GCCTTTTTCT AGCTTAAAAA AAAAAAAAAA 1800

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 48 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Val Trp Val Ser Leu Ser Pro Pro Leu Val Leu Ile Leu Thr Cys Arg
1 5 10 15

Asn Thr Gln Gln Thr His Val Cys Glu Gly Pro Glu Lys Pro Asp Pro
20 25 30

Val Arg Lys Asn Ser Leu Phe Thr Leu Asn Lys Pro Asn Ile Pro Phe
35 40 45

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1063 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

AAAGTTCCAT CTCTAGAACT GATTTTTATC CGTTCTGTTT TTCAGGTCTT ATCTGTGTTA 60

GTTGTGTGTT ACTATCAGGA GGCCCCCTTT GGACCCAGTG GATACAGATT ACGACTCTTC 120

TTTTATGGTG TATGCAATGT CATTTCTATC ACTTGTGCTT ATACATCATT TTCAATAGTT 180

CCTCCCAGCA ATGGGACCAC TATGTGGAGA GCCACAACTA CAGTCTTCAG TGCCATTTTG 240

GCTTTTTTAC TCGTAGATGA GAAAATGGCT TATGTTGACA TGGCTACAGT TGTTTGCAGC 300

ATCTTAGGTG TTTGTCTTGT CATGATCCCA AACATTGTTG ATGAAGACAA TTCTTTGTTA 360

AATGCCTGGA AAGAAGCCTT TGGGTACACC ATGACTGTGA TGGCTGGACT GACCACTGCT 420

CTCTCAATGA TAGTATACAG ATCCATCAAG GAGAAGATCA GCATGTGGAC TGCACTGTTT 480

ACTTTTGGTT GGACTGGGAC AATTTGGGGA ATATCTACTA TGTTTATTCT TCAAGAACCC 540

ATCATCCCAT TAGATGGAGA AACCTGGAGT TATCTCATTG CTATATGTGT CTGTTCTACT 600

GCAGCATTCT TAGGAGTTTA TTATGCCTTG GACAAATTCC ATCCAGCTTT GGTTAGCACA 660

GTACAACATT TGGAGATTGT GGTAGCTATG GTCTTGCAGC TTCTCGTGCT GCACATATTT 720

CCTAGCATCT ATGATGTTTT TGGAGGGGTA ATCATTATGA TTAGTGTTTT TGTCCTTGCT 780

GGCTATAAAC TTTACTGGAG GAATTTAAGA AGGCAGGACT ACCAGGAAAT ATTAGACTCT 840

CCCATTAAAT GAATACCTGA TTATTATTGT CTCATTAATG TTCAGTTATT AATATGTATA 900

CTGCCATTTT AATGTTTACC TATGAATGTC TTTTGTGTTA TATAACTGAC AGAGTGCTAT 960

AAAATATATA ATATATACAA ATGCAGAAAA TTTATTCTAG TCTAATATAT TCAAATACAA 1020

ATATTAAATA TATGAAATAC GTTAAAAAAA AAAAAAAAAA AAA 1063

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 216 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Trp Arg Ala Thr Thr Thr Val Phe Ser Ala Ile Leu Ala Phe Leu
1 5 10 15

Leu Val Asp Glu Lys Met Ala Tyr Val Asp Met Ala Thr Val Val Cys
20 25 30

```
Ser Ile Leu Gly Val Cys Leu Val Met Ile Pro Asn Ile Val Asp Glu
        35                  40                  45

Asp Asn Ser Leu Leu Asn Ala Trp Lys Glu Ala Phe Gly Tyr Thr Met
    50                  55                  60

Thr Val Met Ala Gly Leu Thr Thr Ala Leu Ser Met Ile Val Tyr Arg
65                  70                  75                  80

Ser Ile Lys Glu Lys Ile Ser Met Trp Thr Ala Leu Phe Thr Phe Gly
                85                  90                  95

Trp Thr Gly Thr Ile Trp Gly Ile Ser Thr Met Phe Ile Leu Gln Glu
            100                 105                 110

Pro Ile Ile Pro Leu Asp Gly Glu Thr Trp Ser Tyr Leu Ile Ala Ile
        115                 120                 125

Cys Val Cys Ser Thr Ala Ala Phe Leu Gly Val Tyr Tyr Ala Leu Asp
    130                 135                 140

Lys Phe His Pro Ala Leu Val Ser Thr Val Gln His Leu Glu Ile Val
145                 150                 155                 160

Val Ala Met Val Leu Gln Leu Leu Val Leu His Ile Phe Pro Ser Ile
                165                 170                 175

Tyr Asp Val Phe Gly Gly Val Ile Ile Met Ile Ser Val Phe Val Leu
            180                 185                 190

Ala Gly Tyr Lys Leu Tyr Trp Arg Asn Leu Arg Arg Gln Asp Tyr Gln
        195                 200                 205

Glu Ile Leu Asp Ser Pro Ile Lys
    210                 215
```

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 356 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

```
TGGCCAAAGA GGCCTAGCCG GGAGCGGGCG AGGCGGCGGC GGCAGCAGCG ATGGCAGGAA     60

TAGAGTTGGA GCGGTGCCAG CAGCAGGCGA ACGAGGTGAC GGAAATTATG CGTAACAACT     120
```

TCGGCAAGGT CCTGGAGCGT GGTGTGAAGC TGGCCGAACT GCAGCAGCGT TCAGACCAAC 180

TCCTGGATAT GAGCTCAACC TTCAACAAGA CTACACAGAA CCTGGCCCAG AAGAAGTGCT 240

GGGAGAACAT CCGTTACCGG ATCTGCGTGG GGCTGGTGGT GGTTGGTGTC CTGCTCATCA 300

TCCTGATTGT GCTGCTGGTC GTCTTTCTCC CTCAGAGCAG TGACAGCAGT AGTGCC 356

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 102 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met Ala Gly Ile Glu Leu Glu Arg Cys Gln Gln Gln Ala Asn Glu Val
1 5 10 15

Thr Glu Ile Met Arg Asn Asn Phe Gly Lys Val Leu Glu Arg Gly Val
20 25 30

Lys Leu Ala Glu Leu Gln Gln Arg Ser Asp Gln Leu Leu Asp Met Ser
35 40 45

Ser Thr Phe Asn Lys Thr Thr Gln Asn Leu Ala Gln Lys Lys Cys Trp
50 55 60

Glu Asn Ile Arg Tyr Arg Ile Cys Val Gly Leu Val Val Val Gly Val
65 70 75 80

Leu Leu Ile Ile Leu Ile Val Leu Leu Val Val Phe Leu Pro Gln Ser
85 90 95

Ser Asp Ser Ser Ser Ala
100

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 92 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 60

AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AA 92

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1131 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GGGCCTCAAC TTTGGCGTCG TGAGATTCTT GTGAGGCGTC TGCCTGGAAG CCGGCAGCAA 60

TTTTGCTTCT TTAAAGAGAA AAAGAAGGCT AGGGACTCAG ATTCCTGGAT TCTGAGATCC 120

AGACCAGCTC CTCCCAGACC TCTCCAGAAG AAGCCATGGG AACCCCTCGT ATCCAGCATT 180

TGCTGATCCT CCTGGTCCTA GGAGCCTCCC TCCTGACCTC GGGCCTAGAG CTGTATTGTC 240

AAAAGGGTCT GTCCATGACT GTGGAAGCAG ATCCAGCCAA TATGTTTAAC TGGACCACAG 300

AGGAAGTGGA GACTTGTGAC AAAGGGGCAC TTTGCCAGGA AACCATACTA ATAATTAAAG 360

CAGGGACTGA GACAGCCATT TTGGCCACGA AGGGCTGCAT CCCGGAAGGG GAGGAGGCCA 420

TAACAATTGT CCAGCACTCT TCACCTCCCG GCCTGATCGT GACCTCCTAC AGTAACTACT 480

GTGAGGATTC CTTCTGTAAT GACAAAGACA GCCTGTCTCA GTTTTGGGAG TTCAGTGAGA 540

CCACAGCTTC CACTGTGTCA ACAACCCTCC ATTGTCCAAC CTGTGTGGCT TTGGGGACCT 600

GTTTCAGTGC TCCTTCTCTT CCCTGTCCCA ATGGTACAAC TCGATGCTAT CAAGGAAAAC 660

TTGAGATCAC TGGAGGTGGC ATTGAGTCGT CTGTGGAGGT CAAAGGCTGT ACAGCCATGA 720

TTGGCTGCAG GCTGATGTCT GGAATCTTAG CAGTAGGACC CATGTTTGTG AGGGAAGCGT 780

GCCCACATCA GCTGCTCACT CAACCTCGAA AGACTGAAAA TGGGGCCACC TGTCTTCCCA 840

TTCCTGTTTG GGGGTTACAG CTACTGCTGC CATTGCTGCT GCCATCATTT ATTCACTTTT 900

CCTAAGAAGG CACTTCTGGG CCTGGGTCTG AGGACATCTT TTTTGACTGG GAGCCTTCTT 960

ACTGTTGAGG TTCAACAAGC TGAGGAGTAG ATGGGAATTT GAGGGAGAAT ACAGAGATAC 1020

TATGAACGTA TTTGACATTT TTAATACAAT TTCTGCTATA ATTTTTGTAT GCAGTAGGCG 1080

TTACTAATAA ACATTTCTGC TGTGAAAAAA AAAAAAAAAA AAAAAAAAAA A 1131

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 249 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

```
Met Gly Thr Pro Arg Ile Gln His Leu Leu Ile Leu Leu Val Leu Gly
1               5                  10                  15

Ala Ser Leu Leu Thr Ser Gly Leu Glu Leu Tyr Cys Gln Lys Gly Leu
            20                  25                  30

Ser Met Thr Val Glu Ala Asp Pro Ala Asn Met Phe Asn Trp Thr Thr
        35                  40                  45

Glu Glu Val Glu Thr Cys Asp Lys Gly Ala Leu Cys Gln Glu Thr Ile
    50                  55                  60

Leu Ile Ile Lys Ala Gly Thr Glu Thr Ala Ile Leu Ala Thr Lys Gly
65                  70                  75                  80

Cys Ile Pro Glu Gly Glu Glu Ala Ile Thr Ile Val Gln His Ser Ser
                85                  90                  95

Pro Pro Gly Leu Ile Val Thr Ser Tyr Ser Asn Tyr Cys Glu Asp Ser
            100                 105                 110

Phe Cys Asn Asp Lys Asp Ser Leu Ser Gln Phe Trp Glu Phe Ser Glu
        115                 120                 125

Thr Thr Ala Ser Thr Val Ser Thr Thr Leu His Cys Pro Thr Cys Val
    130                 135                 140

Ala Leu Gly Thr Cys Phe Ser Ala Pro Ser Leu Pro Cys Pro Asn Gly
145                 150                 155                 160

Thr Thr Arg Cys Tyr Gln Gly Lys Leu Glu Ile Thr Gly Gly Gly Ile
                165                 170                 175

Glu Ser Ser Val Glu Val Lys Gly Cys Thr Ala Met Ile Gly Cys Arg
            180                 185                 190
```

Leu Met Ser Gly Ile Leu Ala Val Gly Pro Met Phe Val Arg Glu Ala
195 200 205

Cys Pro His Gln Leu Leu Thr Gln Pro Arg Lys Thr Glu Asn Gly Ala
210 215 220

Thr Cys Leu Pro Ile Pro Val Trp Gly Leu Gln Leu Leu Leu Pro Leu
225 230 235 240

Leu Leu Pro Ser Phe Ile His Phe Ser
245

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 3527 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GCTCCGGGCC GGCTGCGGAG CGACTCCCCG CCGCCAAGTG GGCGGCGTGG CTGTCGGGAA 60

AGAAGGGCTG GGGCCTGCCG TTCTTCCTCC CGAGTATCCC CTCCAGCTGG ACGACCCCAC 120

GCTGCAGCAC GGGCTTCCGG CTTCTCTCCT CAGTGGCCAA TTCGAGGGCA CAGCGGGCTC 180

CGGAGGCGCG GCGGCAAGCC TATCCCGCCT CCCAACCACA GCCTCCAGCA CCCGAGAGAA 240

CGGCCGCCCA CAGCACACGT TCTCCGGACA GGAGGGCGAA GGCCCAAGAC CTGGAGAGAT 300

GGTCAGCTCT CAAAAAAGGC ACAAACAATT GAAGGATGGA TACCATGGCA TATGTTAAAA 360

GCGTGTTGAA AGGAAAATAA GAAAGCCAGG AATCTCAGGA TGAATCAGTC TAGATCGAGA 420

TCAGATGGTG GCAGTGAAGA AACCTTACCT CAAGACCATA ATCATCATGA AAATGAGAGA 480

AGATGGCAGC AAGAGCGTCT CCACAGAGAA GAGGCCTATT ATCAGTTTAT TAATGAACTC 540

AATGATGAAG ATTATCGGCT TATGAGAGAC CATAATCTTT TAGGCACCCC TGGAGAAATA 600

ACATCAGAAG AACTGCAACA GCGGTTAGAT GGCGTCAAGG AACAACTAGC ATCTCAGCCT 660

GACTTGAGAG ATGGAACGAA TTACAGAGAC TCAGAAGTCC CTAGAGAAAG TTCACATGAA 720

GATTCTCTTC TAGAATGGTT GAACACCTTT CGGCGCACAG GAAATGCAAC TCGAAGTGGA 780

CAAAATGGGA ACCAAACTTG GAGAGCTGTG AGTCGAACAA ACCCGAACAA TGGAGAGTTT 840

CGGTTTAGTT TGGAAATCCA CGTAAATCAT GAAAATAGAG GATTTGAAAT TCATGGAGAA 900

GATTATACAG ACATTCCACT TTCAGATAGT AACAGAGATC ATACTGCAAA TAGGCAACAA 960

AGGTCAACTA GTCCTGTGGC TAGGCGAACA AGAAGCCAAA CCTCAGTGAA TTTCAATGGT 1020

AGTAGTTCCA ACATTCCAAG GACTAGGCTT GCTTCAAGGG GGCAAAATCC AGCTGAAGGA 1080

TCTTTCTCAA CATTGGGAAG GTTAAGAAAT GGAATTGGGG GAGCAGCTGG CATTCCTCGA 1140

GCTAACGCTT CACGCACTAA TTTCAGTAGT CACACAAACC AATCAGGTGG TAGTGAACTC 1200

AGGCAAAGGG AGGGGCAACG GTTTGGAGCA GCACATGTTT GGGAAAATGG GGCTAGAAGT 1260

AATGTTACAG TGAGGAATAC AAACCAAAGA TTAGAGCCAA TAAGATTACG ATCTACTTCC 1320

AATAGTCGAA GCCGTTCACC AATTCAGAGA CAGAGTGGCA CTGTTTATCA TAATTCCCAA 1380

AGGGAAAGTA GACCAGTACA GCAAACCACT AGAAGATCTG TTAGGAGGAG AGGTAGAACT 1440

CGAGTCTTTT TAGAGCAAGA TAGAGAACGA GAACGCAGAG GTACTGCATA TACCCCATTC 1500

TCTAATTCAA GGCTTGTGTC AAGAATAACA GTAGAAGAAG GAGAAGAATC CAGCAGATCC 1560

TCAACTGCTG TACGACGACA TCCAACAATC ACACTGGACC TTCAAGTGAG AAGGATCCGT 1620

CCTGGAGAAA ATAGAGATCG GGATAGTATT GCAAATAGAA CTCGATCCAG AGTAGGGCTA 1680

GCAGAAAATA CAGTCACTAT TGAAAGCAAT AGTGGGGGCT TTCGCCGAAC CATTTCTCGT 1740

TTAGAGCGGT CAGGTATTCG AACCTATGTT AGTACCATAA CAGTTCCCCT TCGTAGGATT 1800

TCTGAGAATG AGCTTGTTGA GCCATCATCA GTGGCTCTTC GGTCAATTTT AAGGCAGATC 1860

ATGACTGGGT TTGGAGAACT GAGTTCTCTA ATGGAGGCCG ATTCTGAGTC AGAACTTCAA 1920

AGAAATGGCC AGCATTTACC AGACATGCAC TCAGAACTGA GTAACTTAGG TACAGATAAC 1980

AACAGGAGCC AGCACAGGGA AGGTTCCTCT CAAGACAGGC AGGCCCAAGG AGACAGCACT 2040

GAAATGCATG GTGAAAACGA GACCACCCAG CCTCATACTC GAAACAGTGA CAGTAGGGGT 2100

GGCAGGCAGT TGCGAAATCC AAACAATTTA GTTGAAACTG GAACACTACC CATTCTTCGC 2160

CTTGCTCACT TTTTTTTACT AAATGAAAGT GATGATGATG ATCGAATACG TGGTTTAACC 2220

AAAGAGCAGA TTGACAATCT TTCCACCAGG CACTATGAGC ATAACAGTAT TGATAGTGAA 2280

CTAGGTAAAA TCTGTAGTGT TTGTATTAGT GACTATGTAA CTGGAAACAA GCTCAGGCAA 2340

TTACCTTGCA TGCATGAATT TCACATTCAT TGTATTGACC GATGGCTCTC AGAGAATTGC 2400

ACTTGTCCGA TCTGTCGGCA GCCTGTTTTA GGGTCTAACA TAGCAAACAA TGGGTAAGGT 2460

GATGGGATCT ACTCAAATAC TGTTTTTTAG TAGAACTGAA TGTTCAAGCA TTGTTTTGCT 2520

GAGTTATTTG TGATTAGCTA ACCAGGATGA AAAATAACAG ATTATATATA GTTTGAACTA 2580

TTTTTCGTGT GCTTTTTTAA ACTTGTTAAA AAGAAATTTA TATAAAATTT AAAATACAAA 2640

TGTTAAATTA TCCAGAAATA CAGAATAGTT AATATTGCTA GAACCAAATA ACCTCTAAAA 2700

TGTTTTTATT TTGGTAATTT TGTCATGCTA AGCACTTTTG TATCTGCACA ATTCAGTAGG 2760

TTAAGAATCA ATCTTCTTTT TCTTAATAGT ACAGCAGACT TTAGCTTCAA GTTTCATAGG 2820

CTTAGTACTT ATATCTAGAC ATTTGTGTCT AAATAAGCTT TTCATTAACT TTTTATTTTA 2880

AGGACAGTAT CTTTTCATGA AAGAGTATTT GGCTGAATGT TTGCTATATA TATGTTACTT 2940

GAAATGTTAA ATTTAATATG CAGCATACCA TAGGTGTATA TATAGGTATA TAATTTTAAG 3000

GTTAAAATAT TCAGTCTAAC AAGTTTGGTT CTTATTTAAG CTTTTGGGCT AATACTGCAT 3060

ATGGCACAAT GTTTAATATT GGCAAGTTCA TCTCAGAGAA AGGGGATTCA GATATAATTT 3120

TAAAGTAGAG ATAATTTACT GAAGCGTCTC TGACAATCTA ACTTATTAGA CAGCAAGCAA 3180

TATATAATAC TGAAAAAGTA TTCAGAAATG GAAAATTTAC ATCATATAGG TTATTTAACT 3240

TGTGTTCAGC CTTTTTGTAA CTTTTTTGAA AGTGCAAACA ATTCTTTGGA TTATTAAATA 3300

AGGTATACAG TATGCATGGT TTCTCAAATT TAGCTTTAAA ATCTAAAAGT CTATAAAGAA 3360

TCAGATGCAT AGGCAATATG TTAAGTTCAC TTGGAGGCTA AAAATCTCCA GTGAAAACAA 3420

AACGAAAACC TTTAAGAGAA TGTAGAGTTT ATATAAACAC AAAGTATGCA TTGAAGATCT 3480

GTTTCTACCA ATAAACATTA AAACAAAAAA AAAAAAAAAA AAAAAAA 3527

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 685 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Met Asn Gln Ser Arg Ser Arg Ser Asp Gly Gly Ser Glu Glu Thr Leu
1 5 10 15

Pro Gln Asp His Asn His His Glu Asn Glu Arg Arg Trp Gln Gln Glu
20 25 30

```
Arg Leu His Arg Glu Glu Ala Tyr Tyr Gln Phe Ile Asn Glu Leu Asn
        35                  40                  45

Asp Glu Asp Tyr Arg Leu Met Arg Asp His Asn Leu Leu Gly Thr Pro
    50                  55                  60

Gly Glu Ile Thr Ser Glu Glu Leu Gln Gln Arg Leu Asp Gly Val Lys
65                  70                  75                  80

Glu Gln Leu Ala Ser Gln Pro Asp Leu Arg Asp Gly Thr Asn Tyr Arg
                85                  90                  95

Asp Ser Glu Val Pro Arg Glu Ser Ser His Glu Asp Ser Leu Leu Glu
            100                 105                 110

Trp Leu Asn Thr Phe Arg Arg Thr Gly Asn Ala Thr Arg Ser Gly Gln
        115                 120                 125

Asn Gly Asn Gln Thr Trp Arg Ala Val Ser Arg Thr Asn Pro Asn Asn
    130                 135                 140

Gly Glu Phe Arg Phe Ser Leu Glu Ile His Val Asn His Glu Asn Arg
145                 150                 155                 160

Gly Phe Glu Ile His Gly Glu Asp Tyr Thr Asp Ile Pro Leu Ser Asp
                165                 170                 175

Ser Asn Arg Asp His Thr Ala Asn Arg Gln Gln Arg Ser Thr Ser Pro
            180                 185                 190

Val Ala Arg Arg Thr Arg Ser Gln Thr Ser Val Asn Phe Asn Gly Ser
        195                 200                 205

Ser Ser Asn Ile Pro Arg Thr Arg Leu Ala Ser Arg Gly Gln Asn Pro
    210                 215                 220

Ala Glu Gly Ser Phe Ser Thr Leu Gly Arg Leu Arg Asn Gly Ile Gly
225                 230                 235                 240

Gly Ala Ala Gly Ile Pro Arg Ala Asn Ala Ser Arg Thr Asn Phe Ser
                245                 250                 255

Ser His Thr Asn Gln Ser Gly Gly Ser Glu Leu Arg Gln Arg Glu Gly
            260                 265                 270

Gln Arg Phe Gly Ala Ala His Val Trp Glu Asn Gly Ala Arg Ser Asn
        275                 280                 285

Val Thr Val Arg Asn Thr Asn Gln Arg Leu Glu Pro Ile Arg Leu Arg
    290                 295                 300

Ser Thr Ser Asn Ser Arg Ser Arg Ser Pro Ile Gln Arg Gln Ser Gly
305                 310                 315                 320

Thr Val Tyr His Asn Ser Gln Arg Glu Ser Arg Pro Val Gln Gln Thr
```

325 330 335

Thr Arg Arg Ser Val Arg Arg Arg Gly Arg Thr Arg Val Phe Leu Glu
340 345 350

Gln Asp Arg Glu Arg Glu Arg Arg Gly Thr Ala Tyr Thr Pro Phe Ser
355 360 365

Asn Ser Arg Leu Val Ser Arg Ile Thr Val Glu Glu Gly Glu Glu Ser
370 375 380

Ser Arg Ser Ser Thr Ala Val Arg Arg His Pro Thr Ile Thr Leu Asp
385 390 395 400

Leu Gln Val Arg Arg Ile Arg Pro Gly Glu Asn Arg Asp Arg Asp Ser
405 410 415

Ile Ala Asn Arg Thr Arg Ser Arg Val Gly Leu Ala Glu Asn Thr Val
420 425 430

Thr Ile Glu Ser Asn Ser Gly Gly Phe Arg Arg Thr Ile Ser Arg Leu
435 440 445

Glu Arg Ser Gly Ile Arg Thr Tyr Val Ser Thr Ile Thr Val Pro Leu
450 455 460

Arg Arg Ile Ser Glu Asn Glu Leu Val Glu Pro Ser Ser Val Ala Leu
465 470 475 480

Arg Ser Ile Leu Arg Gln Ile Met Thr Gly Phe Gly Glu Leu Ser Ser
485 490 495

Leu Met Glu Ala Asp Ser Glu Ser Glu Leu Gln Arg Asn Gly Gln His
500 505 510

Leu Pro Asp Met His Ser Glu Leu Ser Asn Leu Gly Thr Asp Asn Asn
515 520 525

Arg Ser Gln His Arg Glu Gly Ser Ser Gln Asp Arg Gln Ala Gln Gly
530 535 540

Asp Ser Thr Glu Met His Gly Glu Asn Glu Thr Thr Gln Pro His Thr
545 550 555 560

Arg Asn Ser Asp Ser Arg Gly Gly Arg Gln Leu Arg Asn Pro Asn Asn
565 570 575

Leu Val Glu Thr Gly Thr Leu Pro Ile Leu Arg Leu Ala His Phe Phe
580 585 590

Leu Leu Asn Glu Ser Asp Asp Asp Asp Arg Ile Arg Gly Leu Thr Lys
595 600 605

Glu Gln Ile Asp Asn Leu Ser Thr Arg His Tyr Glu His Asn Ser Ile
610 615 620

```
Asp Ser Glu Leu Gly Lys Ile Cys Ser Val Cys Ile Ser Asp Tyr Val
625                 630                 635                 640

Thr Gly Asn Lys Leu Arg Gln Leu Pro Cys Met His Glu Phe His Ile
                645                 650                 655

His Cys Ile Asp Arg Trp Leu Ser Glu Asn Cys Thr Cys Pro Ile Cys
            660                 665                 670

Arg Gln Pro Val Leu Gly Ser Asn Ile Ala Asn Asn Gly
        675                 680                 685
```

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1463 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

```
CAGCCTGGGC TCCGCGCAGC CCACCGATCT GGGCGCCCAC AAGCGGCCGG CATCCGTGTC     60

GAGCAGCGCT GCCGTGGAGC ACGAGCAGCG TGAGGCGGCA GCCAAGGAGA AACAACCGCC    120

GCCGCCTGCG CACCGGGGCC CGGCCGACAG CCTGTCCACC GCGGCCGGGG CCGCCGAGCT    180

GAGCGCGGAA GGTGCGGGCA AGAGCCGCGG GTCTGGAGAG CAGGACTGGG TCAACAGGCC    240

CAAGACCGTG CGCGACACGC TGCTGGCGCT GCACCAGCAC GGCCACTCGG GGCCCTTCGA    300

GAGCAAGTTT AAGAAGGAGC CGGCCCTGAC TGCAGGCAGG TTGTTGGGTT TCGAGGCCAA    360

CGGGGCCAAC GGGTCTAAAG CAGTTGCAAG AACAGCAAGG AAAAGGAAGC CCTCTCCAGA    420

ACCAGAAGGT GAAGTCGGGC CCCCTAAGAT CAACGGAGAG GCCCAGCCGT GGCTGTCCAC    480

ATCCACAGAG GGGCTCAAGA TCCCCATGAC TCCTACATCC TCTTTTGTGT CTCCGCCACC    540

ACCCACTGCC TCACCTCATT CCAACCGGAC CACACCGCCT GAAGCGGCCC AGAATGGCCA    600

GTCCCCCATG GCAGCCCTGA TCTTAGTAGC AGACAATGCA GGGGGCAGTC ATGCCTCAAA    660

AGATGCCAAC CAGGTTCACT CCACTACCAG GAGGAATAGC AACAGTCCGC CCTCTCCGTC    720

CTCTATGAAC CAAAGAAGGC TGGGCCCCAG AGAGGTGGGG GGCCAGGGAG CAGGCAACAC    780

AGGAGGACTG GAGCCAGTGC ACCCTGCCAG CCTCCCGGAC TCCTCTCTGG CAACCAGTGC    840
```

CCCGCTGTGC TGCACCCTCT GCCACGAGCG GCTGGAGGAC ACCCATTTTG TGCAGTGCCC 900

GTCCGTCCCT TCGCACAAGT TCTGCTTCCC TTGCTCCAGA CAAAGCATCA AACAGCAGGG 960

AGCTAGTGGA GAGGTCTATT GTCCCAGTGG GGAAAAATGC CCTCTTGTGG GCTCCAATGT 1020

CCCCTGGGCC TTTATGCAAG GGGAAATTGC AACCATCCTT GCTGGAGATG TGAAAGTGAA 1080

AAAAGAGAGA GACTCGTGAC TTTTCCGGTT TCAGAAAAAC CCAATGATTA CCCTTAATTA 1140

AAACTGCTTG AATTGTATAT ATATCTCCAT ATATATATAT ATCCAAGACA AGGGAAATGT 1200

AGACTTCATA AACATGGCTG TATAATTTTG ATTTTTTTTG AATACATTGT GTTTCTATAT 1260

TTTTTTTGAC GACAAAAGGT ATGTACTTAT AAAGACATTT TTTTCTTTTG TTAACGTTAT 1320

TAGCATATCT TTGTGCTTTA TTATCCTGGT GACAGTTACC GTTCTATGTA GGCTGTGACT 1380

TGCGCTGCTT TTTTAGAGCA CTTGGCAAAT CAGAAATGCT TCTAGCTGTA TTTGTATGCA 1440

CTTATTTTAA AAAAAAAAAA AAA 1463

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 197 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Met Thr Pro Thr Ser Ser Phe Val Ser Pro Pro Pro Pro Thr Ala Ser
1 5 10 15

Pro His Ser Asn Arg Thr Thr Pro Pro Glu Ala Ala Gln Asn Gly Gln
20 25 30

Ser Pro Met Ala Ala Leu Ile Leu Val Ala Asp Asn Ala Gly Gly Ser
35 40 45

His Ala Ser Lys Asp Ala Asn Gln Val His Ser Thr Thr Arg Arg Asn
50 55 60

Ser Asn Ser Pro Pro Ser Pro Ser Ser Met Asn Gln Arg Arg Leu Gly
65 70 75 80

Pro Arg Glu Val Gly Gly Gln Gly Ala Gly Asn Thr Gly Gly Leu Glu
85 90 95

```
Pro Val His Pro Ala Ser Leu Pro Asp Ser Ser Leu Ala Thr Ser Ala
            100                 105                 110

Pro Leu Cys Cys Thr Leu Cys His Glu Arg Leu Glu Asp Thr His Phe
        115                 120                 125

Val Gln Cys Pro Ser Val Pro Ser His Lys Phe Cys Phe Pro Cys Ser
    130                 135                 140

Arg Gln Ser Ile Lys Gln Gln Gly Ala Ser Gly Glu Val Tyr Cys Pro
145                 150                 155                 160

Ser Gly Glu Lys Cys Pro Leu Val Gly Ser Asn Val Pro Trp Ala Phe
                165                 170                 175

Met Gln Gly Glu Ile Ala Thr Ile Leu Ala Gly Asp Val Lys Val Lys
            180                 185                 190

Lys Glu Arg Asp Ser
        195
```

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2547 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

```
CATTTTTCTG GTCCTTCTTA AAAGTAATCA CTCTTAAATT TTGTGCTTAT TCTGTTGTTT     60

TAAAAAATAG TTTAAACAAA TATGTGTGTA CTCATAAACA TAGGTTACTT TTGCTTCTTT    120

TTGAGATATA TTTAAATTTT ATTGTGGTCT ACATATTCTT CAGCAGTTTG TTTTTTTACC    180

CAATATTATG TTTCATCTGT ATTACTGCAT TTACTATCCC TAGTTGATTC ACTTCCCTGA    240

AGTACAATAT TCAGTTGTGT GGCTATACCA TAATTTAGTT ATTCATTTTG TTGTCAGTAA    300

AATTTGGGTG ATTATCAGAT TTTTTTCTAG CATGAAAAAT GCTACTARGA ACATTCSTGT    360

ATGTGTCTAA TGGTATACAC TTTCAAGTGT TTTTTTATAT ATGTGAGAGT AGATTACTTG    420

GACCTTGAAG ATGAACATGC TATCTTTTCC AGATACTGCC AATTATTTCA GCAAGATATG    480

AGTTCCCATC ATTTTATATT TGTCAGCATT TGATATTTCC AGGCCTAGTG ATTTCCAGTC    540

ATTTACTGGA TATAATATGA TTATCTCTGT AGGGAGTTGA TTTCCATCTC CTCAATTACT    600
```

AATAAAGTTA AAAATCTTTT CATATGTTTT ATTGCCATTT TTATTTCTTC TGTAAAGTAC 660

CTACTCATGG CTTTTTCTCA TTTTTTGTTT GTCATCATTG AATTATAGGA GTTTTGAGAG 720

AGTGAGCAAG CTAGTCTGTG TGTGTGTGTG TGTGCGTGTG TGTGTATCTC CTTAATGTGT 780

TATATGTGAT TGGAACTTCT TCTCCCACCT TGATGCTTCC TTTCTTCCCC ACTTGTTTTA 840

GGTATCTTCT GATGAAGTGG AGTTATTTAT GGTATGTTCT CAGGAGCTAC AATTTTTAAT 900

TTCAATATAA TCAGTGTTTT TAATTATCTT ATGTTTAGCT CTTTTGGGTC ATGCTTAGGA 960

AATTCTTCTT AAATTTCATT GATAACAGTC TTCCATACTT TCTTCTAAAG TCTTATATTT 1020

TGGCCTTTCA TATTTATTCC TTTAATCCAM CTGGAGTAGA TTTTTTTTTT CCCTCTGTAG 1080

AGTTTGGAGT AGAGATTTTA TTTCCTTTTT TTTTTTTTTT TTTTTTTCTT TTTTTTTGAG 1140

ACAGAGTCTT GCTCTGTCGC CCAGGCTGGA GTGCAGTGGC ACTATCTCAG CTCACTGCAA 1200

CCTCCACCTC CTGGGTTCAA GCGATTCTCC TGCCTCCGCC TCCCGAGTAG CTGGGACTAC 1260

AGGCATGTGC CACCACGCCC AGCTAATTTT TTGTATTTTT TTTAGTAGAG ATGGGGTTCC 1320

ACCATGTTAG CTAGGATGAT TTCGATTTCC TGACCTTGTG ATCCGCCCGC CTCGGCCTCC 1380

CAAAATGCTG GGATTATAGG TGTGAGCCAC CACGTGGCCT CATTTCATTC TTTCATGTGG 1440

ATAGGCAGTT GTTCCAGAAG TATATAGTGA GGAGCTTCTT CTTTCTCTAA TGATCTGCAA 1500

TGTCACCTTC ATCATTTATG AAGGTTGCAC ATATACATGG GAATTTTTTA GTCTGGCATT 1560

AAATGTTCTT CAAAAGAGTT CCTGCAAACG TTTTTGTTTT TATTTCCTAC TGTTCCCTTC 1620

ACGTACTCTC TACTGAACTA AACTCTGTAA TGTGTCTCGA AACTGTCCCA CAATTTTCCT 1680

TGTCTTAAGA GTTTAATGCT TTCATACACC TCTCACATTC AGCCTTGTGC TATTGTCTTA 1740

GGTATATTTA TTTCTCTTTT GCTCCCAATT ATGTTGTAAA CTTTTGGAAG CAGGAAGGAT 1800

ATATTGTTCA TCTTTGGTAG CATTAAACAA TGAATACAGT GTTTTTTACT TAATAGATAT 1860

TTGGTAAATC ATTGAACTAA ATTGGGGTTT GGAATTGAAG GTCTTAGAAA TTACCTGACC 1920

ACTCCCATTA TATTTGCCCA TCCATGATCA CTGAGATTTA TAGAGATTAG ATGCAATGCC 1980

CAGTTTCACA TATGTTTTTG CATCACTGTC TCTTTTTTTC TTGAGCTTAT TCCAGAGTGT 2040

CTTTTAATAT CCATTCCATG ATCAAATGGC TGAACTATTA AAATGCTGTC CAGAAGTGTA 2100

AAGCAATATG AAGATGCTAG AAAAGTTGAA GAGACACATA TATGGTAGGT CCAAGACCAT 2160

TACACTTACT GAGTCCATTA CTAAAAATGA TGTTCACTTA ACATCAAAAC ACTCAGGATT 2220

ACCCAAGCAC AATATACTGA TTTGCACCTC TGCCTTTGTT CATGCCCCTT GTTCAGGAGA 2280

ACTGCTTTCA TGTGCTACTG TCCATAGATC TTCTCTATCC TTACAGATTA ATTTCTTCCT 2340

TTTGAATGCT ATGTTTCCAT ACTTTGACAT TCCTTCTGCA CCATTCAGAC CATATTTTAG 2400

TTCTTTTTTA TGGTATCTCT CACTTTTGAT TGTCACCCCT TAAGTCAAAG ACAATTTTTT 2460

CATCTGTGTC TTCTCAACAC CCAGCACAGG GCTATGTTTG GTAAAAATTA GGTATCCAAG 2520

ATGTACTAAA TGAAAAAAAA AAAAAAA 2547

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 41 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Met Phe Phe Lys Arg Val Pro Ala Asn Val Phe Val Phe Ile Ser Tyr
1 5 10 15

Cys Ser Leu His Val Leu Ser Thr Glu Leu Asn Ser Val Met Cys Leu
20 25 30

Glu Thr Val Pro Gln Phe Ser Leu Ser
35 40

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2245 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

GCTCAACGGC CTCTTCTGGT TGCTGTCTTC CTCGTCCCTC CGGCCCTTCT TCCTACTCAG 60

CGTCTCACTT TTGGCCTATT TTCTGCTGGA TCTCTGGCAG CCTCGCTTTC TCCCTGACGT 120

TTCAGCATCA TCCCCAGAGG AGCCACACTC TGACAGTGAG GGTGCGGGGT CAGGCGCCCG 180

GCCGCACCTG CTGAGTGTGC CCGAGTTGTG CAGATACCTG GCTGAGAGCT GGCTCACCTT 240

CCAGATTCAC CTGCAGGAGC TGCTGCAGTA CAAGAGGCAG AATCCAGCTC AGTTCTGCGT 300

TCGARTCTGC TCTGGCTGTG CTGTGTTGGC TGTGTTGGGA CACTATGTTC CAGGGATTAT 360

GATTTCCTAC ATTGTCTTGT TGAGTATCCT GCTGTGGCCC CTGGTGGTTT ATCATGARCT 420

GATCCAGAGG ATGTWCACTC GCCTGGAGCC CCTGCTCATG CAGCTGGACT ACAGCATGAA 480

GGCAGAAKCC AATGCCCTGC ATCACAAACA CGACAAGAGG AAGCGTCAGG GGAAGAATGC 540

ACCCCCAGGA GGTGATGAGC CACTGGCAGA GACAGAGAGT GAAAGCGAGG CAGAGCTGGC 600

TGGCTTCTCC CCAGTGGTGG ATGTGAAGAA AACAGCATTG GCCTTGGCCA TTACAGACTC 660

AGAGCTGTCA GATGAGGAGG CTTCTATCTT GGAGAGTGGT GGCTTCTCCG TATCCCGGGC 720

CACAACTCCG CAGCTGACTG ATGTCTCCGA GGATTTGGAC CAGCAGAGCC TGCCAAGTGA 780

ACCAGAGGAG ACCCTAAGCC GGGACCTAGG GGAGGGAGAG GAGGGAGAGC TGGCCCCTCC 840

CGAAGACCTA CTAGGCCGTC CTCAAGCTCT GTCAAGGCAA GCCCTGGACT TGGAGGAAGA 900

GGAAGAGGAT GTGGCAGCTA AGGAAACCTT GTTGCGGCTC TCATCCCCCC TCCACTTTGT 960

GAACACGCAC TTCAATGGGG CAGGGTCCCC CCCAGATGGA GTGAAATGCT CCCCTGGAGG 1020

ACCAGTGGAG ACACTGAGCC CCGAGACAGT GAGTGGTGGC CTCACTGCTC TGCCCGGCAC 1080

CCTGTCACCT CCACTTTGCC TTGTTGGAAG TGACCCAGCC CCCTCCCCTT CCATTCTCCC 1140

ACCTGTTCCC CAGGACTCAC CCCAGCCCCT GCCTGCCCCT GAGGAAGAAG AGGCACTCAC 1200

CACTGAGGAC TTTGAGTTGC TGGATCAGGG GGAGCTGGAG CAGCTGAATG CAGAGCTGGG 1260

CTTGGAGCCA GAGACACCGC CAAAACCCCC TGATGCTCCA CCCCTGGGGC CCGACATCCA 1320

TTCTCTGGTA CAGTCAGACC AAGAAGCTCA GGCCGTGGCA GAGCCATGAG CCAGCCGTTG 1380

AGGAAGGAGC TGCAGGCACA GTAGGGCTTC CTGGCTAGGA GTGTTGCTGT TTCCTCCTTT 1440

GCCTACCACT CTGGGGTGGG GCAGTGTGTG GGGAAGCTGG CTGTCGGATG GTAGCTATTC 1500

CACCYTCTGC CTGCCTGCCT GCCTGCTGTC CTGGGCATGG TGCAGTACCT GTGCCTAGGA 1560

TTGGTTTTAA ATTTGTAAAT AATTTTCCAT TTGGGTTAGT GGATGTGAAC AGGGCTAGGG 1620

AAGTCCTTCC CACAGCCTGC GCTTGCCTCC CTGCCTCATC TCTATTCTCA TTCCACTATG 1680

CCCCAAGCCC TGGTGGTCTG GCCCTTTCTT TTTCCTCCTA TCCTCAGGGA CCTGTGCTGC 1740

TCTGCCCTCA TGTCCCACTT GGTTGTTTAG TTGAGGCACT TTATAATTTT TCTCTTGTCT 1800

TGTGTTCCTT TCTGCTTTAT TTCCCTGCTG TGTCCTGTCC TTAGCAGCTC AACCCCATCC 1860

TTTGCCAGCT CCTCCTATCC CGTGGGCACT GGCCAAGCTT TAGGGAGGCT CCTGGTCTGG 1920

GAAGTAAAGA GTAAACCTGG GGCAGTGGGT CAGGCCAGTA GTTACACTCT TAGGTCACTG 1980

TAGTCTGTGT AACCTTCACT GCATCCTTGC CCCATTCAGC CCGGCCTTTC ATGATGCAGG 2040

AGAGCAGGGA TCCCGCAGTA CATGGCGCCA GCACTGGAGT TGGTGAGCAT GTGCTCTYTY 2100

TTGAGATTAG GAGCTTCCTT ACTGCTCCTC TGGGTGATCC AAGTGTAGTG GGACCCCCTA 2160

CTAGGGTCAG GAAGTGGACA CTAACATCTG TGCAGGTGTT GACTTGAAAA ATAAAGTGTT 2220

GATTGGCTAG AAAAAAAAAA AAAAA 2245

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 336 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

Met Ile Ser Tyr Ile Val Leu Leu Ser Ile Leu Leu Trp Pro Leu Val
1 5 10 15

Val Tyr His Glu Leu Ile Gln Arg Met Xaa Thr Arg Leu Glu Pro Leu
20 25 30

Leu Met Gln Leu Asp Tyr Ser Met Lys Ala Glu Xaa Asn Ala Leu His
35 40 45

His Lys His Asp Lys Arg Lys Arg Gln Gly Lys Asn Ala Pro Pro Gly
50 55 60

Gly Asp Glu Pro Leu Ala Glu Thr Glu Ser Glu Ser Glu Ala Glu Leu
65 70 75 80

Ala Gly Phe Ser Pro Val Val Asp Val Lys Lys Thr Ala Leu Ala Leu
85 90 95

Ala Ile Thr Asp Ser Glu Leu Ser Asp Glu Glu Ala Ser Ile Leu Glu
100 105 110

Ser Gly Gly Phe Ser Val Ser Arg Ala Thr Thr Pro Gln Leu Thr Asp
115 120 125

Val Ser Glu Asp Leu Asp Gln Gln Ser Leu Pro Ser Glu Pro Glu Glu
130 135 140

```
Thr Leu Ser Arg Asp Leu Gly Glu Gly Glu Glu Gly Glu Leu Ala Pro
145                 150                 155                 160

Pro Glu Asp Leu Leu Gly Arg Pro Gln Ala Leu Ser Arg Gln Ala Leu
                165                 170                 175

Asp Leu Glu Glu Glu Glu Glu Asp Val Ala Ala Lys Glu Thr Leu Leu
            180                 185                 190

Arg Leu Ser Ser Pro Leu His Phe Val Asn Thr His Phe Asn Gly Ala
        195                 200                 205

Gly Ser Pro Pro Asp Gly Val Lys Cys Ser Pro Gly Gly Pro Val Glu
    210                 215                 220

Thr Leu Ser Pro Glu Thr Val Ser Gly Gly Leu Thr Ala Leu Pro Gly
225                 230                 235                 240

Thr Leu Ser Pro Pro Leu Cys Leu Val Gly Ser Asp Pro Ala Pro Ser
                245                 250                 255

Pro Ser Ile Leu Pro Pro Val Pro Gln Asp Ser Pro Gln Pro Leu Pro
            260                 265                 270

Ala Pro Glu Glu Glu Glu Ala Leu Thr Thr Glu Asp Phe Glu Leu Leu
        275                 280                 285

Asp Gln Gly Glu Leu Glu Gln Leu Asn Ala Glu Leu Gly Leu Glu Pro
    290                 295                 300

Glu Thr Pro Pro Lys Pro Pro Asp Ala Pro Pro Leu Gly Pro Asp Ile
305                 310                 315                 320

His Ser Leu Val Gln Ser Asp Gln Glu Ala Gln Ala Val Ala Glu Pro
                325                 330                 335
```

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1406 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

```
CTTGTGGGAA GAGCTGAAGC AGGCGCTCTT GGCTCGGCGC GGCCCGCTGC AATCCGTGGA     60

GGAACGCGCC GCCGAGCCAC CATCATGCCT GGGCACTTAC AGGAAGGCTT CGGCTGCGTG     120
```

GTCACCAACC GATTCGACCA GTTATTTGAC GACGAATCGG ACCCCTTCGA GGTGCTGAAG 180

GCAGCAGAGA ACAAGAAAAA AGAAGCCGGC GGGGGCGGCG TTGGGGGCCC TGGGGCCAAG 240

AGCGCAGCTC AGGCCGCGGC CCAGACCAAC TCCAACGCGG CAGGCAAACA GCTGCGCAAG 300

GAGTCCCAGA AAGACCGCAA GAACCCGCTG CCCCCCAGCG TTGGCGTGGT TGACAAGAAA 360

GAGGAGACGC AGCCGCCCGT GGCGCTTAAG AAAGAAGGAA TAAGACGAGT TGGAAGAAGA 420

CCTGATCAAC AACTTCAGGG TGAAGGGAAA ATAATTGATA GAAGACCAGA AAGGCGACCA 480

CCTCGTGAAC GAAGATTCGA AAAGCCACTT GAAGAAAAGG GTGAAGGAGG CGAATTTTCA 540

GTTGATAGAC CGATTATTGA CCGACCTATT CGAGGTCGTG GTGGTCTTGG AAGAGGTCGA 600

GGGGGCCGTG GACGTGGAAT GGGCCGAGGA GATGGATTTG ATTCTCGTGG CAAACGTGAA 660

TTTGATAGGC ATAGTGGAAG TGATAGATCT TCTTTTTCAC ATTACAGTGG CCTGAAGCAC 720

GAGGACAAAC GTGGAGGTAG CGGATCTCAC AACTGGGGAA CTGTCAAAGA CGAATTAACT 780

GACTTGGATC AATCAAATGT GACTGAGGAA ACACCTGAAG GTGAAGAACA TCATCCAGTG 840

GCAGACACTG AAAATAAGGA GAATGAAGTT GAAGAGGTAA AAGAGGAGGG TCCAAAAGAG 900

ATGACTTTGG ATGAGTGGAA GGCTATTCAA AATAAGGACC GGGCAAAAGT AGAATTTAAT 960

ATCCGAAAAC CAAATGAAGG TGCTGATGGG CAGTGGAAGA AGGGATTTGT TCTTCATAAA 1020

TCAAAGAGTG AAGAGGCTCA TGCTGAAGAT TCGGTTATGG ACCATCATTT CCGGAAGCCA 1080

GCAAATGATA TAACGTTTCA GCTGGAGATC AATTTTGGAG ACCTTGGCCG CCCAGGACGT 1140

GGCGGCAGGG GAGGACGAGG TGGACGTGGG CGTGGTGGGC GCCCAAACCG TGGCAGCAGG 1200

ACCGACAAGT CAAGTGCTTT TGCTCCTGAT GTGGATGACC CAGAGGCATT CCCAGTTTTG 1260

GCTTAAMTGG ATGCCATAAG ACAACCCTGG TTCCTTTGTG AACCCTTTTG TTCAAAGCTT 1320

TTGCATGCTT AAGGATTCCA AACGACTAAG AAATTAAAAA AAAAAAAAAA AAAAAAAAAA 1380

AAAAAAAAAA AAAAAAAAAA AAAAAA 1406

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 393 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

```
Met Pro Gly His Leu Gln Glu Gly Phe Gly Cys Val Val Thr Asn Arg
1               5                  10                  15

Phe Asp Gln Leu Phe Asp Asp Glu Ser Asp Pro Phe Glu Val Leu Lys
            20                  25                  30

Ala Ala Glu Asn Lys Lys Lys Glu Ala Gly Gly Gly Gly Val Gly Gly
        35                  40                  45

Pro Gly Ala Lys Ser Ala Ala Gln Ala Ala Ala Gln Thr Asn Ser Asn
    50                  55                  60

Ala Ala Gly Lys Gln Leu Arg Lys Glu Ser Gln Lys Asp Arg Lys Asn
65                  70                  75                  80

Pro Leu Pro Pro Ser Val Gly Val Val Asp Lys Lys Glu Glu Thr Gln
                85                  90                  95

Pro Pro Val Ala Leu Lys Lys Glu Gly Ile Arg Arg Val Gly Arg Arg
            100                 105                 110

Pro Asp Gln Gln Leu Gln Gly Glu Gly Lys Ile Ile Asp Arg Arg Pro
        115                 120                 125

Glu Arg Arg Pro Pro Arg Glu Arg Arg Phe Glu Lys Pro Leu Glu Glu
    130                 135                 140

Lys Gly Glu Gly Gly Glu Phe Ser Val Asp Arg Pro Ile Ile Asp Arg
145                 150                 155                 160

Pro Ile Arg Gly Arg Gly Gly Leu Gly Arg Gly Arg Gly Gly Arg Gly
                165                 170                 175

Arg Gly Met Gly Arg Gly Asp Gly Phe Asp Ser Arg Gly Lys Arg Glu
            180                 185                 190

Phe Asp Arg His Ser Gly Ser Asp Arg Ser Ser Phe Ser His Tyr Ser
        195                 200                 205

Gly Leu Lys His Glu Asp Lys Arg Gly Gly Ser Gly Ser His Asn Trp
    210                 215                 220

Gly Thr Val Lys Asp Glu Leu Thr Asp Leu Asp Gln Ser Asn Val Thr
225                 230                 235                 240

Glu Glu Thr Pro Glu Gly Glu Glu His His Pro Val Ala Asp Thr Glu
                245                 250                 255

Asn Lys Glu Asn Glu Val Glu Glu Val Lys Glu Glu Gly Pro Lys Glu
            260                 265                 270

Met Thr Leu Asp Glu Trp Lys Ala Ile Gln Asn Lys Asp Arg Ala Lys
```

```
         275                 280                 285

Val Glu Phe Asn Ile Arg Lys Pro Asn Glu Gly Ala Asp Gly Gln Trp
    290                 295                 300

Lys Lys Gly Phe Val Leu His Lys Ser Lys Ser Glu Glu Ala His Ala
305                 310                 315                 320

Glu Asp Ser Val Met Asp His His Phe Arg Lys Pro Ala Asn Asp Ile
                325                 330                 335

Thr Phe Gln Leu Glu Ile Asn Phe Gly Asp Leu Gly Arg Pro Gly Arg
            340                 345                 350

Gly Gly Arg Gly Gly Arg Gly Gly Arg Gly Arg Gly Gly Arg Pro Asn
        355                 360                 365

Arg Gly Ser Arg Thr Asp Lys Ser Ser Ala Phe Ala Pro Asp Val Asp
    370                 375                 380

Asp Pro Glu Ala Phe Pro Val Leu Ala
385                 390
```

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 4237 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

```
GCGGACGCGG CCAGTCAGGT GCTCCTGGGC TCCGGTCTCA CCATCCTGTC CCAGCCGCTC     60

ATGTACGTGA AAGTGCTCAT CCAGGTGGGA TATGAGCCTC TTCCTCCAAC AATAGGACGA    120

AATATTTTTG GGCGGCAAGT GTGTCAGCTT CCTGGTCTCT TTAGTTATGC TCAGCACATT    180

GCCAGTATCG ATGGGAGGCG CGGGTTGTTC ACAGGCTTAA CTCCAAGACT GTGTTCGGGA    240

GTCCTTGGAA CTGTGGTCCA TGGTAAAGTT TTACAGCATT ACCAGGAGAG TGACAAGGGT    300

GAGGAGTTAG GAMCTGGAAA TGTACARAAA GAAGTCTCAT CTTCCTTTGA MCACGTTATC    360

AAGGAGACAA CTCGAGAGAT GATCGCTCGT TCTGCTGCTA CCCTCATCAC ACATCCCTTC    420

CATGTTGATC ACTCTGAGAT CTATGGTACA RTTCATTGGC AGAGAATCCA AGTACTGTGG    480

ACTTTGTGAT TCCATAATAA CCATCTATCG GGAAGAGGGC ATTCTAGGAT TTTTCGCGGG    540
```

TCTTGTTCCT CGCCTTCTAG GTGACATCCT TTCTTTGTGG CTGTGTAACT CACTGGCCTA 600

CCTCGTCAAT ACCTATGCAC TGGACAGTGG GGTTTCTACC ATGAATGAAA TGAAGAGTTA 660

TTCTCAAGCT GTCACAGGAT TTTTTGCGAG TATGTTGACC TATCCCTTTG TGCTTGTCTC 720

CAATCTTATG GCTGTCAACA ACTGTGGTCT TGCTGGTGGA TGCCCTCCTT ACTCCCCAAT 780

ATATACGTCT TGGATAGACT GTTGGTGCAT GCTACAAAAA GAGGGGAATA TGAGCCGAGG 840

AAATAGCTTA TTTTTCCGGA AGGTCCCCTT TGGGAAGACT TATTGTTGTG ACCTGAAAAT 900

GTTAATTTGA AGATGTGGGG CAGGGACAGT GACATTTCTG TAGTCCCAGA TGCACAGAAT 960

TATGGGAGAG AATGTTGATT TCTATACAGT GTGGCGCGCT TTTTTAATAA TCATTTAATC 1020

TTGGGAAAAT TCAGGTGTTT GGTGTCTGCC TTTTTTGTTC TTTTTTCCAG CACAACATAA 1080

CTTACCACTG ATACTCCCCC TTTAGTTATT CTGAATTAGG ATATTTTTGC TCCAAATTCT 1140

TATTTTACTT AACCAGAAGG GAAAAAAAGT TGTATTTTCC TGAAGCTACA GGCACTTTGT 1200

CATGTGATTT TTGAGTCTCA ATTTAAGGCT TTGTAAAATG AAGAGTAGAA TTCCAAGAAA 1260

AATGAGAAAT AATTTTGTAA AACTTAACAA AATCACTAAA TTAAACTATA TGGGAGGTTA 1320

TGAATTACTT TTTCTTGGGT AGACCCTAAA ATGTCAGTAG CATGCACCAG AATCTGACTC 1380

CCATTATGCT TCTAAGCACA TTTCATTGAC CTTGTCTCTC ATACTTCAAG AAAAGGACAG 1440

TACATTGCTA CATTACCCTA GAAAGTCTGT GTGAGGATCT GCCCCTTCAG TCTGTTATTG 1500

CAAAGTAATA AAATGTCACC TACAGGGAGC CTCTGAGCCT ACTCTAGTTC AAGAGGCTAC 1560

CTGAAAAAAA ATAAATAAGA TAAAGGGTCA GCAACAACAA AGAAAAAGAC AATTACAGAA 1620

AATAAGCAAG ATTTGGAAAG GAAGTATAAT GGCACTTTTT TCCTCAAAGG AAGTTCTTGT 1680

TTTCACATAA AATATGAAAA GCAGATCCTG CAGGAGTAAC CCCCTTCTTT AAGAGCCAAG 1740

TATTTGCCAG TGCTTAAATT ACACCATACC GTTCTAATTA TATATAATCT TTTGTTCTTC 1800

AGTTTTTTGT TTTGTTTCCT TTTTGTTATT GTTGCCGAAG GTGAGTAGTT TTGCATTTCT 1860

GATGACAGCC TTGGAAAGTA TATTTGTAAC TCCATGTCTG GTAATGCCAA CCCAAGTCGA 1920

CATGGGTCTT AGGACACTGA CCACCTCACA TGCCATACCC TCAGTTAAGC ATGTTAACAT 1980

TTATAGGAGG AAAAAAATCA CTTTGGGAGA AAATAAAATT CAACTCAAGC ATAAAGCTTC 2040

TGTTTACTCA GGCCTTCTAA AAAGCAGGTT AAAATGCTCT AAAATGAGAA AGCCTGTGGT 2100

TTCACTTATT TATATAACTC ACTGGGACAT TGCCAAATGA GTAAGCACTT AATTCGCTGC 2160

TTCTGAGACT TCTCTGTCAA AACAGCCCCA CTGATAATAT TAGACAGAAC GAGAATGCAG 2220

GGGTCTCTTC CCTCCCCTGG GGTTTAGGAA GCTCATGAGG AGCTCGGCTT AAAATGTCTT 2280

TGATGTCTCT TCCTTTGTCT CAAAAAGTAA TGTCAATTTT ATATACTATT TCAATATTAC 2340

TATCTGCATT TGTTTTAATA TAAAAATGTT TGCTGCCTAC CTTTTTCTCC CAAAAAATCT 2400

TTAAGTAAAG ATGATCTGGG AAAATGAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 2460

AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 2520

AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 2580

AAAAAAAAAA AAAAAAAAGC GGCCGCAGGT CTAGAATTCA ATCGGAAGGT ATATAGCTTA 2640

TTTGTTGCTT TTCATTGTAA TTTAACATGG TTAATGGTTA ATTACTATTT AACACACATT 2700

TCAAATGAAT ATTATTTGGG GGATTAGATT GAGTGAAATT AACCTGCTAT TAAATAGTAA 2760

ACTTTTCCTC TGGAGTCACT TTTTTCCCCC TTCAAAGTAT GTTACTGAGG AAGTAAACTT 2820

TTTTTTTTTT TTTTGGTTTT TGTTTTTTGA GACACAGTCT CGCTCTGTTG CCCAGGCTGC 2880

TGGAGTGCCG TGGCGCAATC TCGGCTCACT GCAACCTCCG CCTCCTGGAT TCAAACAATT 2940

CTCCTGCTTC AGCCTCCTGA GTAGCTGGGA TTACAGGCAC ATGCCACCAC GCCCGGCTAA 3000

TTTTTGTATT TTTAGTAGAG ACTGGGTTTC ACCATGTTGG TCAGGCTGGT CTCAAACTCC 3060

TGACCTCGTG ATCCACCCGC CTCGGCCTCC CAAAATCCTG GGATTACAGG CGCGAGCCAC 3120

CACACCCGGC TGGAAGTAAA CATTTTTAAA GCTACTTTTA CTCATTCTAG CCTTGTAGAA 3180

TGACCATTGC AGCTTGAGGG ACCTAGTTCT TACCTTTTCT TGCAACCAAC ACACTTGCAA 3240

TTGTGTCTGG TATGCTTGTT CCTGCTGCTA ATAAAGTAAG GCCCATTACT GTATCGGGAA 3300

TTTCTAGTGT TTCCCCTGTA ATAAACAGAT ATTTCAAGTT ACAAATCTTA AAGATTCACT 3360

AACCATCCTT TGCAGTTATT TTGGATATTT CCTTCGTGAA CAAAAATAAA ATAGGCACAT 3420

TTAGAATTCA GAGCCAATAT GTGCTTGCTT ATTAGTTTTT TAGCTAGCAA CATATTTGAA 3480

TCAGGCTGGT AATTCGGGTA ACCCAGGTAG CACAGATTTT TAATGACATA TYTAAAGATA 3540

CGTAACAGCT AAAATTTCTG CCAGTGAGAA ATTTTCCTGT TTGATATTTC TTACAAAAGA 3600

TGTTTATGTC CACCATTATY TTCATTCAGG GGCTGTGCTG AATATTTGAT AATGAGACTG 3660

ATCATTCCGC TTTTTCTTTC TTAAAAATAT TAGGCAGAGT TAAGCAAATT AATTATAGCT 3720

ATCTTTAAGC TATAAATGTG TTAACATGTA TATATACCAT TTATTATGTT CTACTTTAGT 3780

GATATACCTT AATTTAGTGG GCTTTGGCAG GGCGGGGGAG GGGGAACGTT CATTAATCTC 3840

TGAGGAAAAC AAAACCTGTT TTCTACTTGA GTCTAACATA TGGTCCCAAT TTATTAATAC 3900

TTCTGTTAAA TTTGATGTCA GGTCAACATT TTTCAGAAAT GTATTTATTC TCAGAAACAG 3960

AACCAGAGAG AAGTTAAACA AAAGGTTATG TAACTGTTCC TTTAATGTTG TAATTGAAAA 4020

CTTGGTTTAG CGTCTTTTTT TTCTTTCTCT TTTTTTTTCT TAAAATGCCA ACTAAAATAA 4080

TTAGAAAGTA GCTTATTTAT TGCATGCTTA TACATTGATA TTGGAATTGG AATTGGTTGT 4140

TAATTTCTGT TACTGGCTTT GCTAGAATTC ATATGTGCAT AAATAACACT AATATTTATC 4200

ATCTTGGAAA AAAAAAAAAA AAAAAAAAAA AAAAAAA 4237

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 94 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Met Tyr Ile Tyr His Leu Leu Cys Ser Thr Leu Val Ile Tyr Leu Asn
1 5 10 15

Leu Val Gly Phe Gly Arg Ala Gly Glu Gly Glu Arg Ser Leu Ile Ser
20 25 30

Glu Glu Asn Lys Thr Cys Phe Leu Leu Glu Ser Asn Ile Trp Ser Gln
35 40 45

Phe Ile Asn Thr Ser Val Lys Phe Asp Val Arg Ser Thr Phe Phe Arg
50 55 60

Asn Val Phe Ile Leu Arg Asn Arg Thr Arg Glu Lys Leu Asn Lys Arg
65 70 75 80

Leu Cys Asn Cys Ser Phe Asn Val Val Ile Glu Asn Leu Val
85 90

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

TNTTGAAGACT GTTGCTTGTT TGGAATGT 29

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

CNCCATCTAAT GGGATGATGG GTTCTTGA 29

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

ANTTTCCGTCA CCTCGTTCGC CTGCTGCT 29

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

GNATACGAGGG GTTCCCATGG CTTCTTCT 29

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

TNTACGACGAC ATCCAACAAT CACACTGG 29

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

TNGTCCGGTTG GAATGAGGTG AGGCAGTG 29

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "olgionucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

GNTCCTCACTA TATACTTCTG GAACAACT 29

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

GNCCTAAGAGT GTAACTACTG GCCTGACC 29

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "olgionucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

TNTCCTCGTGC TTCAGGCCAC TGTAATGT 29

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

ANGCCCACTAA ATTAAGGTAT ATCACTAA 29

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 51 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

Met Trp Gly Leu Gly Thr Thr Ser Ser Phe Arg Trp Tyr Ser Ser Asp
1 5 10 15

Tyr Arg Arg Ser Phe Gln Leu Asn Ser Pro Val Asp Lys Met Arg Lys
20 25 30

Thr Gly Glu Gln Ala Phe Ser Val Phe Thr Tyr Lys Val Arg Ser Val
35 40 45

Met Gly Gln
50

What is claimed is:

1. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 463 to nucleotide 606;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 1 to nucleotide 501;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone bd164_7 deposited under accession number ATCC 98364;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone bd164_7 deposited under accession number ATCC 98364;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone bd164_7 deposited under accession number ATCC 98364;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone bd164_7 deposited under accession number ATCC 98364;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:2;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:2 having biological activity, the fragment comprising the amino acid sequence from amino acid 19 to amino acid 28 of SEQ ID NO:2;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

2. The polynucleotide of claim 1 wherein said polynucleotide is operably linked to at least one expression control sequence.

3. A host cell transformed with the polynucleotide of claim 2.

4. The host cell of claim 3, wherein said cell is a mammalian cell.

5. A process for producing a protein encoded by the polynucleotide of claim 2, which process comprises:

(a) growing a culture of the host cell of claim 3 in a suitable culture medium; and

(b) purifying said protein from the culture.

6. A protein produced according to the process of claim 5.

7. The protein of claim 6 comprising a mature protein.

8. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:2;

(b) fragments of the amino acid sequence of SEQ ID NO:2 comprising the amino acid sequence from amino acid 19 to amino acid 28 of SEQ ID NO:2; and

(c) the amino acid sequence encoded by the cDNA insert of clone bd164_7 deposited under accession number ATCC 98364;

the protein being substantially free from other mammalian proteins.

9. The protein of claim 8, wherein said protein comprises the amino acid sequence of SEQ ID NO:2.

10. A composition comprising the protein of claim 8 and a pharmaceutically acceptable carrier.

11. A method for preventing, treating or ameliorating a medical condition which comprises administering to a mammalian subject a therapeutically effective amount of a composition of claim 10.

12. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:1.

13. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 202 to nucleotide 849;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 511 to nucleotide 849;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone bi129_2 deposited under accession number ATCC 98364;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone bi129_2 deposited under accession number ATCC 98364;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone bi129_2 deposited under accession number ATCC 98364;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone bi129_2 deposited under accession number ATCC 98364;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:4;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:4 having biological activity, the fragment comprising the amino acid sequence from amino acid 103 to amino acid 112 of SEQ ID NO:4;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

14. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:4;

(b) the amino acid sequence of SEQ ID NO:4 from amino acid 88 to amino acid 209;

(c) fragments of the amino acid sequence of SEQ ID NO:4 comprising the amino acid sequence from amino acid 103 to amino acid 112 of SEQ ID NO:4; and

(d) the amino acid sequence encoded by the cDNA insert of clone bi129_2 deposited under accession number ATCC 98364;

the protein being substantially free from other mammalian proteins.

15. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:3.

16. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:8;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:8 from nucleotide 156 to nucleotide 902;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:8 from nucleotide 225 to nucleotide 902;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:8 from nucleotide 237 to nucleotide 654;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone cg160_6 deposited under accession number ATCC 98364;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone cg160_6 deposited under accession number ATCC 98364;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone cg160_6 deposited under accession number ATCC 98364;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone cg160_6 deposited under accession number ATCC 98364;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:9;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:9 having biological activity, the fragment

comprising the amino acid sequence from amino acid 119 to amino acid 128 of SEQ ID NO:9;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

17. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:9;

(b) the amino acid sequence of SEQ ID NO:9 from amino acid 28 to amino acid 166;

(c) fragments of the amino acid sequence of SEQ ID NO:9 comprising the amino acid sequence from amino acid 119 to amino acid 128 of SEQ ID NO:9; and

(d) the amino acid sequence encoded by the cDNA insert of clone cg160_6 deposited under accession number ATCC 98364;

the protein being substantially free from other mammalian proteins.

18. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:8.

19. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:10;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:10 from nucleotide 400 to nucleotide 2454;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:10 from nucleotide 1454 to nucleotide 1787;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone cw775_1 deposited under accession number ATCC 98364;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone cw775_1 deposited under accession number ATCC 98364;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone cw775_1 deposited under accession number ATCC 98364;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone cw775_1 deposited under accession number ATCC 98364;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:11;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:11 having biological activity, the fragment comprising the amino acid sequence from amino acid 337 to amino acid 346 of SEQ ID NO:11;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

20. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:11;

(b) fragments of the amino acid sequence of SEQ ID NO:11 comprising the amino acid sequence from amino acid 337 to amino acid 346 of SEQ ID NO:11; and

(c) the amino acid sequence encoded by the cDNA insert of clone cw775_1 deposited under accession number ATCC 98364;

the protein being substantially free from other mammalian proteins.

21. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:10.

22. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:12;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:12 from nucleotide 506 to nucleotide 1096;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:12 from nucleotide 656 to nucleotide 1096;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:12 from nucleotide 2 to nucleotide 1078;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone dn740_3 deposited under accession number ATCC 98364;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone dn740_3 deposited under accession number ATCC 98364;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone dn740_3 deposited under accession number ATCC 98364;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone dn740_3 deposited under accession number ATCC 98364;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:13;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:13 having biological activity, the fragment comprising the amino acid sequence from amino acid 93 to amino acid 102 of SEQ ID NO:13;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

23. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:13;

(b) the amino acid sequence of SEQ ID NO:13 from amino acid 1 to amino acid 191;

(c) fragments of the amino acid sequence of SEQ ID NO:13 comprising the amino acid sequence from amino acid 93 to amino acid 102 of SEQ ID NO:13; and

(d) the amino acid sequence encoded by the cDNA insert of clone dn740_3 deposited under accession number ATCC 98364;

the protein being substantially free from other mammalian proteins.

24. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:12.

25. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:14;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:14 from nucleotide 1563 to nucleotide 1685;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:14 from nucleotide 1100 to nucleotide 1646;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone dn904_2 deposited under accession number ATCC 98364;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone dn904_2 deposited under accession number ATCC 98364;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone dn904_2 deposited under accession number ATCC 98364;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone dn904_2 deposited under accession number ATCC 98364;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:15;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:15 having biological activity, the fragment comprising the amino acid sequence from amino acid 15 to amino acid 24 of SEQ ID NO:15;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

26. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:15;

(b) the amino acid sequence of SEQ ID NO:15 from amino acid 1 to amino acid 28;

(c) fragments of the amino acid sequence of SEQ ID NO:15 comprising the amino acid sequence from amino acid 15 to amino acid 24 of SEQ ID NO:15; and

(d) the amino acid sequence encoded by the cDNA insert of clone dn904_2 deposited under accession number ATCC 98364;

the protein being substantially free from other mammalian proteins.

27. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:14.

28. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:16;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:16 from nucleotide 359 to nucleotide 1369;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:16 from nucleotide 1547 to nucleotide 1868;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone do568_11 deposited under accession number ATCC 98364;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone do568_11 deposited under accession number ATCC 98364;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone do568_11 deposited under accession number ATCC 98364;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone do568_11 deposited under accession number ATCC 98364;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:17;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:17 having biological activity, the fragment comprising the amino acid sequence from amino acid 163 to amino acid 172 of SEQ ID NO:17;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

29. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:17;

(b) fragments of the amino acid sequence of SEQ ID NO:17 comprising the amino acid sequence from amino acid 163 to amino acid 172 of SEQ ID NO:17; and

(c) the amino acid sequence encoded by the cDNA insert of clone do568_11 deposited under accession number ATCC 98364;

the protein being substantially free from other mammalian proteins.

30. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:16.

31. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:18;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:18 from nucleotide 85 to nucleotide 1263;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:18 from nucleotide 265 to nucleotide 608;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone ek626_3 deposited under accession number ATCC 98364;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone ek626_3 deposited under accession number ATCC 98364;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone ek626_3 deposited under accession number ATCC 98364;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone ek626_3 deposited under accession number ATCC 98364;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:19;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:19 having biological activity, the fragment comprising the amino acid sequence from amino acid 191 to amino acid 200 of SEQ ID NO:19;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

32. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:19;

(b) the amino acid sequence of SEQ ID NO:19 from amino acid 61 to amino acid 175;

(c) fragments of the amino acid sequence of SEQ ID NO:19 comprising the amino acid sequence from amino acid 191 to amino acid 200 of SEQ ID NO:19; and

(d) the amino acid sequence encoded by the cDNA insert of clone ek626_3 deposited under accession number ATCC 98364;

the protein being substantially free from other mammalian proteins.

33. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:18.

34. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:20;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:20 from nucleotide 3746 to nucleotide 4027;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:20 from nucleotide 3815 to nucleotide 4027;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:20 from nucleotide 3640 to nucleotide 3940;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone fe366_1 deposited under accession number ATCC 98364;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone fe366_1 deposited under accession number ATCC 98364;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone fe366_1 deposited under accession number ATCC 98364;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone fe366_1 deposited under accession number ATCC 98364;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:21;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:21 having biological activity, the fragment comprising the amino acid sequence from amino acid 42 to amino acid 51 of SEQ ID NO:21;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

35. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:21;

(b) the amino acid sequence of SEQ ID NO:21 from amino acid 1 to amino acid 65;

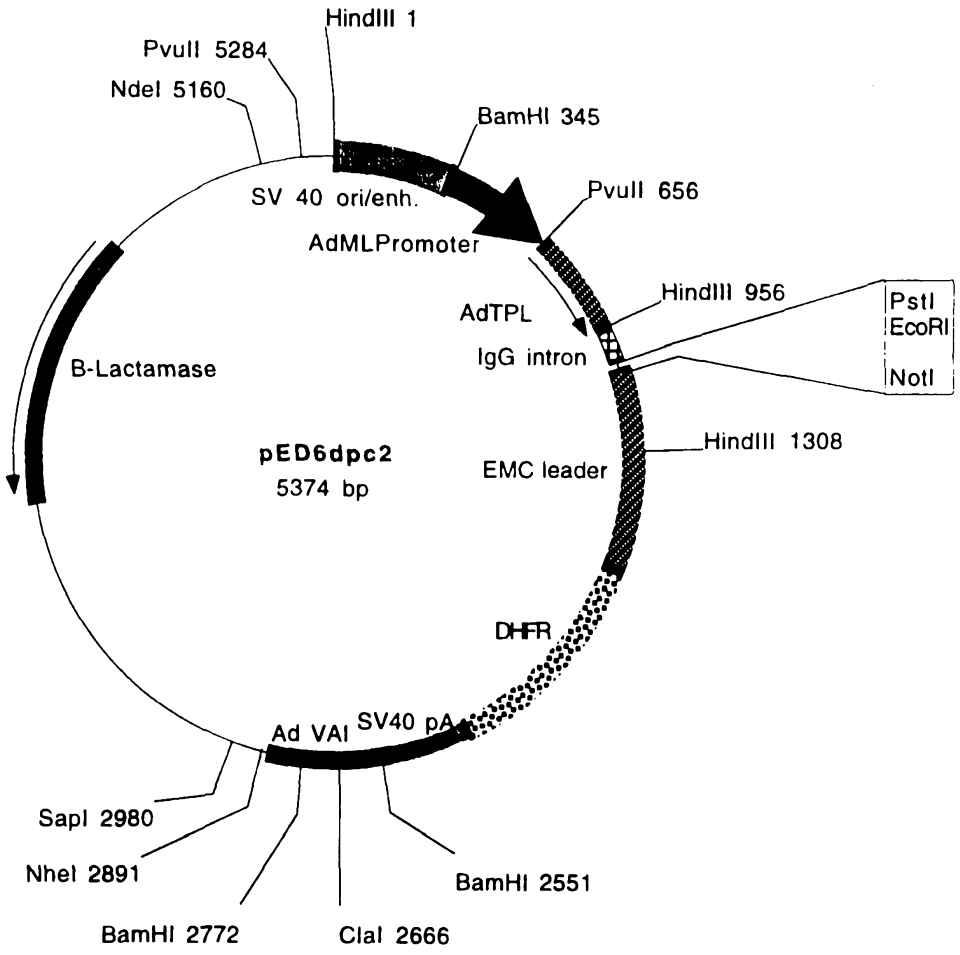
(c) fragments of the amino acid sequence of SEQ ID NO:21 comprising the amino acid sequence from amino acid 42 to amino acid 51 of SEQ ID NO:21; and

(d) the amino acid sequence encoded by the cDNA insert of clone fe366_1 deposited under accession number ATCC 98364;

the protein being substantially free from other mammalian proteins.

36. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:20.

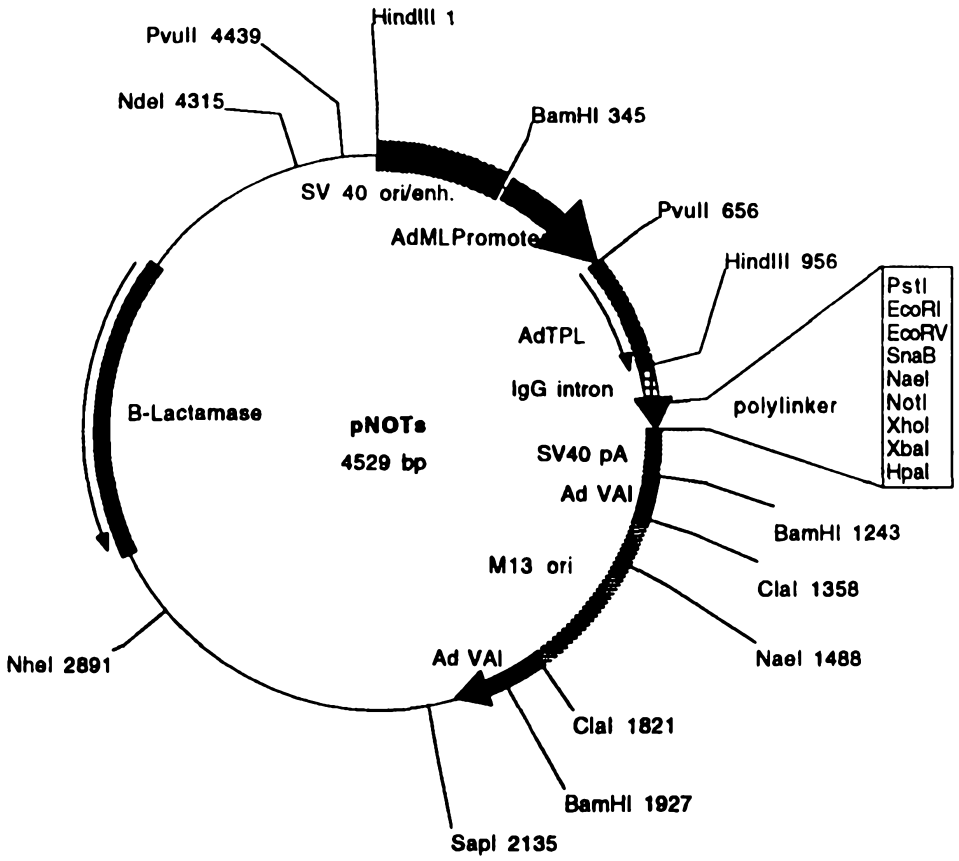
FIGURE 1A



Plasmid name: pED6dpc2
Plasmid size: 5374 bp

Comments/References: pED6dpc2 is derived from pED6dpc1 by insertion of a new polylinker to facilitate cDNA cloning. SST cDNAs are cloned between EcoRI and NotI. pED vectors are described in Kaufman et al.(1991), NAR 19: 4485-4490.

FIGURE 1B



Plasmid name: pNOTs
Plasmid size: 4529 bp

Comments/References: pNOTs is a derivative of pMT2 (Kaufman et al,1989. Mol.Cell.Biol.9:1741-1750). DHFR was deleted and a new polylinker was inserted between EcoRI and HpaI. M13 origin of replication was inserted in the ClaI site. SST cDNAs are cloned between EcoRI and NotI