PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/12, C07K 14/47, C12N 5/10, A61K 38/17

A2

(11) International Publication Number: **WO 98/36067**

(43) International Publication Date: 20 August 1998 (20.08.98)

(21) International Application Number: PCT/US98/02767

(22) International Filing Date: 13 February 1998 (13.02.98)

(30) Priority Data:

Not furnished	14 February 1997 (14.02.97)	US
Not furnished	12 February 1998 (12.02.98)	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).

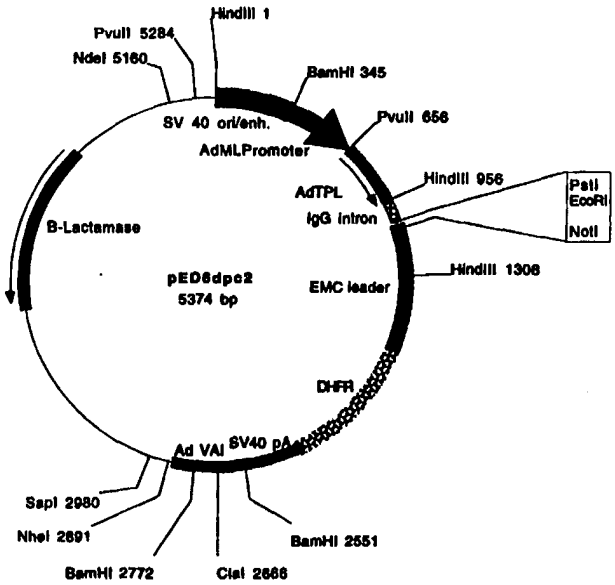
(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, 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

Novel polynucleotides and the proteins encoded thereby are disclosed.



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.

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 Ser. No. 60/XXX,XXX (converted to a provisional application from non-provisional application Ser. No. 08/800,418), filed February 14, 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 128 to nucleotide 1006;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 182 to nucleotide 362;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone BG481_1 deposited under accession number ATCC 98331;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone BG481_1 deposited under accession number ATCC 98331;

(f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone BG481_1 deposited under accession number ATCC 98331;

(g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone BG481_1 deposited under accession number ATCC 98331;

(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;

(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 128 to nucleotide 1006; the nucleotide sequence of SEQ ID NO:1 from nucleotide 182 to nucleotide 362; the nucleotide sequence of the full-length protein coding sequence of clone BG481_1 deposited under accession number ATCC 98331; or the nucleotide sequence of the mature protein coding sequence of clone BG481_1 deposited

under accession number ATCC 98331. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone BG481_1 deposited under accession number ATCC 98331. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:2 from amino acid 1 to amino acid 78.

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) the amino acid sequence of SEQ ID NO:2 from amino acid 1 to amino acid 78;

(c) fragments of the amino acid sequence of SEQ ID NO:2; and

(d) the amino acid sequence encoded by the cDNA insert of clone BG481_1 deposited under accession number ATCC 98331;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:2 or the amino acid sequence of SEQ ID NO:2 from amino acid 1 to amino acid 78.

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 250 to nucleotide 846;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 361 to nucleotide 846;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 1 to nucleotide 564;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone BJ9_1 deposited under accession number ATCC 98331;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone BJ9_1 deposited under accession number ATCC 98331;

(g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone BJ9_1 deposited under accession number ATCC 98331;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone BJ9_1 deposited under accession number ATCC 98331;

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

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:4 having biological activity;

(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:3 from nucleotide 250 to nucleotide 846; the nucleotide sequence of SEQ ID NO:3 from nucleotide 361 to nucleotide 846; the nucleotide sequence of SEQ ID NO:3 from nucleotide 1 to nucleotide 564; the nucleotide sequence of the full-length protein coding sequence of clone BJ9_1 deposited under accession number ATCC 98331; or the nucleotide sequence of the mature protein coding sequence of clone BJ9_1 deposited under accession number ATCC 98331. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone BJ9_1 deposited under accession number ATCC 98331. 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 1 to amino acid 105.

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 1 to amino acid 105;

(c) fragments of the amino acid sequence of SEQ ID NO:4; and

(d) the amino acid sequence encoded by the cDNA insert of clone BJ9_1 deposited under accession number ATCC 98331;

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 1 to amino acid 105.

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 354 to nucleotide 674;

(c) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone BK34_3 deposited under accession number ATCC 98331;

(d) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone BK34_3 deposited under accession number ATCC 98331;

(e) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone BK34_3 deposited under accession number ATCC 98331;

(f) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone BK34_3 deposited under accession number ATCC 98331;

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

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

(i) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(f) above;

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

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

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:5 from nucleotide 354 to nucleotide 674; the nucleotide sequence of the full-length

protein coding sequence of clone BK34_3 deposited under accession number ATCC 98331; or the nucleotide sequence of the mature protein coding sequence of clone BK34_3 deposited under accession number ATCC 98331. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone BK34_3 deposited under accession number ATCC 98331.

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) fragments of the amino acid sequence of SEQ ID NO:6; and

(c) the amino acid sequence encoded by the cDNA insert of clone BK34_3 deposited under accession number ATCC 98331;

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

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 823 to nucleotide 960;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:8 from nucleotide 931 to nucleotide 960;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone BP883_2 deposited under accession number ATCC 98331;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone BP883_2 deposited under accession number ATCC 98331;

(f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone BP883_2 deposited under accession number ATCC 98331;

(g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone BP883_2 deposited under accession number ATCC 98331;

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

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

(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:8 from nucleotide 823 to nucleotide 960; the nucleotide sequence of SEQ ID NO:8 from nucleotide 931 to nucleotide 960; the nucleotide sequence of the full-length protein coding sequence of clone BP883_2 deposited under accession number ATCC 98331; or the nucleotide sequence of the mature protein coding sequence of clone BP883_2 deposited under accession number ATCC 98331. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone BP883_2 deposited under accession number ATCC 98331. 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 1 to amino acid 16.

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 1 to amino acid 16;

(c) fragments of the amino acid sequence of SEQ ID NO:9; and

(d) the amino acid sequence encoded by the cDNA insert of clone BP883_2 deposited under accession number ATCC 98331;

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 1 to amino acid 16.

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 84 to nucleotide 1016;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:10 from nucleotide 786 to nucleotide 1016;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:10 from nucleotide 619 to nucleotide 899;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CI363_1 deposited under accession number ATCC 98331;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CI363_1 deposited under accession number ATCC 98331;

(g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CI363_1 deposited under accession number ATCC 98331;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CI363_1 deposited under accession number ATCC 98331;

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

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:11 having biological activity;

(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:10 from nucleotide 84 to nucleotide 1016; the nucleotide sequence of SEQ ID NO:10 from nucleotide 786 to nucleotide 1016; the nucleotide sequence of SEQ ID NO:10 from nucleotide 619 to nucleotide 899; the nucleotide sequence of the full-length protein coding

sequence of clone CI363_1 deposited under accession number ATCC 98331; or the nucleotide sequence of the mature protein coding sequence of clone CI363_1 deposited under accession number ATCC 98331. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone CI363_1 deposited under accession number ATCC 98331. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:11 from amino acid 180 to amino acid 272.

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) the amino acid sequence of SEQ ID NO:11 from amino acid 180 to amino acid 272;

(c) fragments of the amino acid sequence of SEQ ID NO:11; and

(d) the amino acid sequence encoded by the cDNA insert of clone CI363_1 deposited under accession number ATCC 98331;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:11 or the amino acid sequence of SEQ ID NO:11 from amino acid 180 to amino acid 272.

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 88 to nucleotide 561;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:12 from nucleotide 142 to nucleotide 561;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:12 from nucleotide 1 to nucleotide 554;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CO806_1 deposited under accession number ATCC 98331;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CO806_1 deposited under accession number ATCC 98331;

(g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CO806_1 deposited under accession number ATCC 98331;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CO806_1 deposited under accession number ATCC 98331;

(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;

(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 88 to nucleotide 561; the nucleotide sequence of SEQ ID NO:12 from nucleotide 142 to nucleotide 561; the nucleotide sequence of SEQ ID NO:12 from nucleotide 1 to nucleotide 554; the nucleotide sequence of the full-length protein coding sequence of clone CO806_1 deposited under accession number ATCC 98331; or the nucleotide sequence of the mature protein coding sequence of clone CO806_1 deposited under accession number ATCC 98331. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone CO806_1 deposited under accession number ATCC 98331. 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 112 to amino acid 156.

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 112 to amino acid 156;

(c) fragments of the amino acid sequence of SEQ ID NO:13; and

(d) the amino acid sequence encoded by the cDNA insert of clone CO806_1 deposited under accession number ATCC 98331;

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 112 to amino acid 156.

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 787 to nucleotide 945;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:14 from nucleotide 853 to nucleotide 945;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:14 from nucleotide 324 to nucleotide 945;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CT24_3 deposited under accession number ATCC 98331;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CT24_3 deposited under accession number ATCC 98331;

(g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CT24_3 deposited under accession number ATCC 98331;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CT24_3 deposited under accession number ATCC 98331;

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

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:15 having biological activity;

(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:14 from nucleotide 787 to nucleotide 945; the nucleotide sequence of SEQ ID NO:14 from nucleotide 853 to nucleotide 945; the nucleotide sequence of SEQ ID NO:14 from nucleotide 324 to nucleotide 945; the nucleotide sequence of the full-length protein coding sequence of clone CT24_3 deposited under accession number ATCC 98331; or the nucleotide sequence of the mature protein coding sequence of clone CT24_3 deposited under accession number ATCC 98331. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone CT24_3 deposited under accession number ATCC 98331.

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) fragments of the amino acid sequence of SEQ ID NO:15; and

(c) the amino acid sequence encoded by the cDNA insert of clone CT24_3 deposited under accession number ATCC 98331;

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

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 562 to nucleotide 738;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:16 from nucleotide 1 to nucleotide 729;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone ER366_3 deposited under accession number ATCC 98331;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone ER366_3 deposited under accession number ATCC 98331;

(f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone ER366_3 deposited under accession number ATCC 98331;

(g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone ER366_3 deposited under accession number ATCC 98331;

(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;

(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 562 to nucleotide 738; the nucleotide sequence of SEQ ID NO:16 from nucleotide 1 to nucleotide 729; the nucleotide sequence of the full-length protein coding sequence of clone ER366_3 deposited under accession number ATCC 98331; or the nucleotide sequence of the mature protein coding sequence of clone ER366_3 deposited under accession number ATCC 98331. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone ER366_3 deposited under accession number ATCC 98331. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:17 from amino acid 1 to amino acid 56.

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) the amino acid sequence of SEQ ID NO:17 from amino acid 1 to amino acid 56;

(c) fragments of the amino acid sequence of SEQ ID NO:17; and

(d) the amino acid sequence encoded by the cDNA insert of clone ER366_3 deposited under accession number ATCC 98331;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:17 or the amino acid sequence of SEQ ID NO:17 from amino acid 1 to amino acid 56.

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) 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 "BG481_1"

A polynucleotide of the present invention has been identified as clone "BG481_1". BG481_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. BG481_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "BG481_1 protein").

The nucleotide sequence of BG481_1 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 BG481_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:2.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone BG481_1 should be approximately 2600 bp.

The nucleotide sequence disclosed herein for BG481_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. BG481_1 demonstrated at least some similarity with sequences identified as AB002351 (Human mRNA for KIAA0353 gene, partial cds), R41215 (yf84h04.s1 Homo sapiens cDNA clone 29428 3'), W26369 (26f5 Human retina cDNA randomly primed sublibrary Homo sapiens cDNA), and W27653 (36e7 Human retina cDNA randomly primed sublibrary Homo sapiens cDNA). The predicted amino acid sequence disclosed herein for BG481_1 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted BG481_1 protein demonstrated at least some similarity to sequences identified as AB002351 (KIAA0353 [Homo sapiens]). Based upon sequence similarity, BG481_1 proteins and each similar protein or peptide may share at least some activity.

Clone "BJ9_1"

A polynucleotide of the present invention has been identified as clone "BJ9_1". BJ9_1 was isolated from a human adult ovary (pool of retinoic-acid-treated, activin-treated, and untreated PA-1 teratocarcinoma cells) 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. BJ9_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "BJ9_1 protein").

The nucleotide sequence of BJ9_1 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 BJ9_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:4. Amino acids 25 to 37 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 38.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone BJ9_1 should be approximately 1100 bp.

The nucleotide sequence disclosed herein for BJ9_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. BJ9_1 demonstrated at least some similarity with sequences identified as AA114043 (zm29f09.r1 Stratagene pancreas (#937208) Homo sapiens cDNA clone 527081 5') and H27069 (yl16a06.r1 Homo sapiens cDNA clone 158386 5' similar to gb | M87903 | HUMALNE37 Human carcinoma cell-derived Alu RNA transcript, (rRNA); gb:M87338 ACTIVATOR 1 40 KD SUBUNIT (HUMAN);contains Alu repetitive element). Based upon sequence similarity, BJ9_1 proteins and each similar protein or peptide may share at least some activity. The nucleotide sequence of BJ9_1 indicates that it may contain Alu and CAAA repeat sequences.

Clone "BK34_3"

A polynucleotide of the present invention has been identified as clone "BK34_3". BK34_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. BK34_3 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "BK34_3 protein").

The nucleotide sequence of the 5' portion of BK34_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 BK34_3 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:6. Additional nucleotide sequence from the 3' portion of BK34_3, including the polyA tail, is reported in SEQ ID NO:7.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone BK34_3 should be approximately 1350 bp.

The nucleotide sequence disclosed herein for BK34_3 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. BK34_3 demonstrated at least some similarity with sequences identified as H94111 (yw58h12.s1 Soares placenta 8to9weeks 2NbHP8to9W Homo sapiens cDNA clone 256487 3'). Based upon sequence similarity, BK34_3 proteins and each similar protein or peptide may share at least some activity.

Clone "BP883_2"

A polynucleotide of the present invention has been identified as clone "BP883_2". BP883_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. BP883_2 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "BP883_2 protein").

The nucleotide sequence of BP883_2 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 BP883_2 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:9. Amino acids 24 to 36 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 37, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone BP883_2 should be approximately 1445 bp.

The nucleotide sequence disclosed herein for BP883_2 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. BP883_2 demonstrated at least some similarity with sequences identified as AA164399 (zo96a10.s1 Stratagene ovarian cancer (#937219) Homo sapiens cDNA clone 594714 3'), H04678 (yj10a10.r1 Homo sapiens cDNA clone), H17605 (ym36h02.r1 Homo sapiens cDNA clone 50386 5'), I25658 (Sequence 19 from patent US 5552281), N47764 (yy55e08.r1 Homo sapiens cDNA clone 277478 5'), Q61070 (Human brain Expressed Sequence Tag EST01127), Q72530 (Osteoclast-specific/related expressed gene clone 241B), R37216 (yh96f12.r1 Homo sapiens cDNA clone 137615 5'), and W56113 (zc56g06.r1 Soares parathyroid tumor NbHPA Homo sapiens). Based upon sequence similarity, BP883_2 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts a potential transmembrane domain at the carboxyl terminus of the BP883_2 protein sequence reported as SEQ ID NO:9.

Clone "CI363_1"

A polynucleotide of the present invention has been identified as clone "CI363_1". CI363_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. CI363_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "CI363_1 protein").

The nucleotide sequence of CI363_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 CI363_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:11. Amino acids 222 to 234 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 235, or are a transmembrane domain. Amino acids 133 to 145 are another possible leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 146, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone CI363_1 should be approximately 2300 bp.

The nucleotide sequence disclosed herein for CI363_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. CI363_1 demonstrated at least some similarity with sequences identified as AA024102 (mh98c08.r1 Soares mouse placenta 4NbMP13.5 14.5 Mus musculus cDNA clone 458990 5' similar to PIR:S19586 S19586 N-methyl-D-aspartate receptor glutamate-binding chain - rat), AA413815 (vc67b06.s1 Knowles Solter mouse 2 cell Mus musculus cDNA clone 779603 5'), H06014 (yl76e04.r1 Homo sapiens cDNA clone 43696 5' similar to SP:S19586 S19586 N-METHYL-D-ASPARTATE RECEPTOR GLUTAMATE-BINDING CHAIN), N70951 (za34a02.s1 Homo sapiens cDNA clone 294410 3'), R73412 (yj92f12.r1 Homo sapiens cDNA clone), and S61973 (NMDA receptor glutamate-binding subunit). The predicted amino acid sequence disclosed herein for CI363_1 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted CI363_1 protein demonstrated at least some similarity to sequences identified as S61973 (NMDA receptor glutamate-binding subunit [rats, Peptide, 516 aa] [Rattus sp.]) and V01555 (BWRF1 reading frame 12 [Human herpesvirus]). Based upon sequence similarity, CI363_1 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts seven potential transmembrane domains within the CI363_1 protein sequence, centered around amino acids 110, 130, 180, 200, 230, 250, and 290 of SEQ ID NO:11, respectively.

Clone "CO806_1"

A polynucleotide of the present invention has been identified as clone "CO806_1". CO806_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. CO806_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "CO806_1 protein").

The nucleotide sequence of CO806_1 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 CO806_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:13. Amino acids 6 to 18 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 19, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone CO806_1 should be approximately 1100 bp.

The nucleotide sequence disclosed herein for CO806_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. CO806_1 demonstrated at least some similarity with sequences identified as AA075444 (zm87d06.r1 Stratagene ovarian cancer (#937219) Homo sapiens cDNA clone 544907 5' similar to WP:F22E10.5 CE05695 PHOSPHOTRANSFERASE), F08821 (H. sapiens partial cDNA sequence; clone c-2tc04), H19601 (yn59g02.r1 Homo sapiens cDNA clone 172754 5'), R39687 (yc97c08.s1 Homo sapiens cDNA clone 24193 3'), R44048 (yg22b11.s1 Homo sapiens cDNA clone 33076 3'), T19387 (Human gene signature HUMGS00411), T49354 (ya74g01.r1 Homo sapiens cDNA clone 67440 5'), U12735 (Glycine max aminoalcoholphosphotransferase (AAPT1) mRNA, complete cds), and Y08486 (M.musculus Sycp3 gene). The predicted amino acid sequence disclosed herein for CO806_1 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted CO806_1 protein demonstrated at least some similarity to sequences identified as U12735 (aminoalcoholphosphotransferase [Glycine max]), Y08486 (synaptonemal complex protein [Mus musculus]), Z50797 (F22E10.5 [Caenorhabditis elegans]), and Z67882 (F22E10.5 [Caenorhabditis elegans]). Two amino acid sequence motifs were identified in the predicted CO806_1 protein: a CDP-alcohol phosphatidyltransferase motif and a cytochrome b/b6 motif. Based upon

sequence similarity, CO806_1 proteins and each similar protein or peptide may share at least some activity.

Clone "CT24_3"

A polynucleotide of the present invention has been identified as clone "CT24_3". CT24_3 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. CT24_3 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "CT24_3 protein").

The nucleotide sequence of CT24_3 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 CT24_3 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:15. Amino acids 10 to 22 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 23, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone CT24_3 should be approximately 2000 bp.

The nucleotide sequence disclosed herein for CT24_3 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. CT24_3 demonstrated at least some similarity with sequences identified as AA344698 (EST50612 Gall bladder I Homo sapiens cDNA 5' end), H48143 (yp80h01.s1 Homo sapiens cDNA clone 193777 3'), N72077 (yz97c06.s1 Homo sapiens cDNA clone 290986 3'), T95902 (ye47e08.s1 Homo sapiens cDNA clone 120902 3'), T96003 (ye47e08.r1 Homo sapiens cDNA clone 120902 5'), and Z45277 (H. sapiens partial cDNA sequence; clone c-2ld02). Based upon sequence similarity, CT24_3 proteins and each similar protein or peptide may share at least some activity

Clone "ER366_3"

A polynucleotide of the present invention has been identified as clone "ER366_3". ER366_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. ER366_3 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "ER366_3 protein").

The nucleotide sequence of ER366_3 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 ER366_3 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:17. Amino acids 12 to 24 are a possible leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 25, or are part of a potential transmembrane domain that includes the amino-terminal 20 amino acids of SEQ ID NO:17.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone ER366_3 should be approximately 1000 bp.

The nucleotide sequence disclosed herein for ER366_3 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. No hits were found in the database. The nucleotide and amino acid sequences of ER366_3 indicate that it may contain one or more Alu or MIR repeat sequences.

Deposit of Clones

Clones BG481_1, BJ9_1, BK34_3, BP883_2, CI363_1, CO806_1, CT24_3, and ER366_3 were deposited on February 14, 1997 with the American Type Culture Collection as an original deposit under the Budapest Treaty and were given the accession number ATCC 98331, 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
BG481_1	SEQ ID NO:18
BJ9_1	SEQ ID NO:19
BK34_3	SEQ ID NO:20
BP883_2	SEQ ID NO:21
CI363_1	SEQ ID NO:22
CO806_1	SEQ ID NO:23
CT24_3	SEQ ID NO:24
ER366_3	SEQ ID NO:25

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 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 of the mature form 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 homologs 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, as determined by those of skill in the art. Species homologs 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 homologs are those isolated from mammalian species.

The invention also encompasses allelic variants of the disclosed polynucleotides or proteins; that is, naturally-occurring alternative forms of the isolated polynucleotide which also encode proteins which are identical, homologous, or related to that encoded by the polynucleotides .

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 J.

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

(iii) NUMBER OF SEQUENCES: 25

(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: 2647 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

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

GTCCCCGAAA GGTTTGCAGA CGCCTGTGAA GGATGCTGGT GGTGGGACCG GTAGAGAGGC 60

AGAAGCAAGA GAGCTACGGT TCAGGTTGGG CACCAGTGAT GCCACTGGTT CTCTGCAAGG 120

CGATTCCATG ACAGAAACCG TAGCAGAAAA CATCGTTACC AGTATCCTGA AGCAGTTCAC 180

CCAGTCTCCA GAGACAGAAG CATCTGCTGA TTCTTTTCCA GACACAAAAG TCACTTACGT 240

GGACAGGAAA GAGCTTCCTG GGGAAAGGAA AACAAAGACT GAAATAGTTG TGGAGTCTAA 300

ACTGACTGAG GATGTTGATG TTTCCGATGA AGCTGGCCTG GACTACCTTT TAAGCAAGGA 360

TATTAAGGAA GTGGGGCTGA AAGGCAAGTC AGCCGAGCAG ATGATAGGAG ACATCATCAA 420

CCTCGGCCTG AAAGGGAGGG AGGGGAGAGC AAAGGTCGTC AACGTGGAGA TCGTGGAGGA 480

GCCCGTGAGT TATGTCAGCG GGGAGAAGCC GGAGGAGTTT TCCGTCCCAT TCAAAGTGGA 540

GGAGGTCGAA GATGTGTCGC CAGGCCCCTG GGGGTTGGTT AAGGAGGAGG AAGGTTATGG 600

AGAAAGCGAT GTCACATTCT CAGTTAATCA GCATCGAAGG ACCAAGCAGC CCCAGGAGAA 660

CACGACTCAC GTGGAAGAAG TGACAGAGGC AGGTGATTCA GAGGGCGAGC AGAGTTATTT 720

TGTGTCCACT CCAGATGAAC ACCCCGGGGG GCACGACAGA GATGACGGCT CGGTGTACGG 780

GCAGATCCAC ATCGAGGAGG AATCCACCAT CAGGTACTCT TGGCAGGATG AAATCGTGCA 840

GGGGACTCGA AGGAGGACAC AGAAGGACGG TGCAGTGGGC GAGAAGGTTG TGAAGCCCTT 900

GGATGTCCCA GCGCCCTCTC TGGAGGGGGA CCTGGGTTCC ACTCACTGGA AAGAACAAGC 960

TAGAAGCGGT GAATTTCATG CCGAACCCAC AGTCAGAAAA AGAAATTAAA ATACCCCACG 1020

AATTCCACAC CTCCATGAAG GGCATCTCCT CCAAGGAGCC CCGGCAGCAG CTGGTGGAGG 1080

TCATCGGGCA GCTGGAGGAA ACCCTTCCCG AGCGCATGAG GGAGGAGCTG TCCGCCCTCA 1140

CCAGAGAGGG GCAGGGTGGG CCGGGGAGCG TTTCCGTGGA TGTCAAGAAG GTCCAGGGTG 1200

CTGGTGGCAG TTCCGTGACC CTGGTTGCTG AAGTCAACGT CTCACAAACT GTGGATGCCG 1260

ATCGGTTAGA CCTGGAGGAG CTGAGCAAAG ATGAGGCCAG TGAGATGGAG AAGGCTGTGG 1320

AGTCGGTGGT TCGGGAGAGC CTGAGCAGGC AACGCAGCCC AGCGCCTGGC AGCCCAGATG 1380

AGGAAGGTGG AGCGGAGGCC CCGGCTGCTG GCATTCGCTT CAGGCGTTGG GCCACCCGGG 1440

AGCTGTACAT CCCTTCAGGC GAGAGCGAGG TTGCTGGTGG GGCCTCTCAC AGCTCGGGAC 1500

AGCGCACTCC CCAGGGCCCA GTGTCGGCCA CTGTGGAGGT CAGCAGCCCC ACAGGCTTTG 1560

CCCAGTCACA GGTGCTGGAG GATGTGAGCC AGGCTGCAAG GCACATAAAA CTCGGCCCCT 1620

CTGAAGTCTG GAGGACTGAG CGAATGTCAT ATGAAGGACC CACTGCAGAA GTGGTGGAGA 1680

TGGATGTGAG TAACGTAGAG GCGATCCGCA GCCGGACACA GGAAGCGGGA GCTCTCGGTG 1740

TGTCTGACCG TGGTTCCTGG AGAGACGCGG ACAGTAGGAA TGACCAGGCA GTTGGTGTGA 1800

GCTTTAAGGC CTCTGCTGGG GAAGGAGACC AGGCCCACAG AGAACAGGGC AAGGAGCAGG 1860

CCATGTTTGA TAAGAAGGTG CAGCTCCAGA GAATGGTAGA CCAAAGGTCG GTGATTTCAG 1920

ATGAAAAGAA AGTTGCCCTC CTCTATCTAG ACAATGAGGA GGAGGAGAAT GATGGGCATT 1980

GGTTTTAATA AGCAGAAACA TTTTGTTTTA ATGGCAGCCT GTTGGCGACG TGCCAACATC 2040

CAAAGGCCTT AACTTATTTT AAGAGGCCGA GGGAGTCTAT GAAAATCTCC CCTTTTTTAC 2100

TTTTTTAAAG AGTACTCCCG GCATGGTCAA TTTCCTTTAT AGTTAATCCG TAAAGGTTTC 2160

CAGTTAATTC ATGCCTTAAA AGGCACTGCA ATTTTATTTT TGAGTTGGGA CTTTTACAAA 2220

ACACTTTTTT CCCTGGAGTC TTCTCTCCAC TTCTGGAGAT GAATTTCTAT GTTTTGCACC 2280

TGGTCACAGA CATGGCTTGC ATCTGTTTGA AACTACAATT AATTATAGAT GTCAAAACAT 2340

TAACCAGATT AAAGTAATAT ATTTAAGAGT AAATTTTGCT TGCATGTGCT AATATGAAAT 2400

AACAGACTAA CATTTTAGGG GAAAAATAAA TACAATTTAG ACTCTAAAAA GTCTTTTCAA 2460

AAAGAAATGG GAAATAGGCA GACTGTTTAT GTTAAAAAAA TTCTTGCTAA ATGATTTCAT 2520

CTTTAGGAAA AAATTACTTG CCATATAGAG CTAAATTCAT CTTAAGACTT GAATGAATTG 2580

CTTTCTATGT ACAGAACTTT AAACAATATA GTATTTATGG CGAGGAAAAA AAAAAAAAAA 2640

AAAAAAA 2647

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 293 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

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

```
Met Thr Glu Thr Val Ala Glu Asn Ile Val Thr Ser Ile Leu Lys Gln
1               5                  10                  15

Phe Thr Gln Ser Pro Glu Thr Glu Ala Ser Ala Asp Ser Phe Pro Asp
            20                  25                  30

Thr Lys Val Thr Tyr Val Asp Arg Lys Glu Leu Pro Gly Glu Arg Lys
        35                  40                  45

Thr Lys Thr Glu Ile Val Val Glu Ser Lys Leu Thr Glu Asp Val Asp
    50                  55                  60

Val Ser Asp Glu Ala Gly Leu Asp Tyr Leu Leu Ser Lys Asp Ile Lys
65                  70                  75                  80

Glu Val Gly Leu Lys Gly Lys Ser Ala Glu Gln Met Ile Gly Asp Ile
                85                  90                  95

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

Val Glu Ile Val Glu Glu Pro Val Ser Tyr Val Ser Gly Glu Lys Pro
        115                 120                 125

Glu Glu Phe Ser Val Pro Phe Lys Val Glu Glu Val Glu Asp Val Ser
    130                 135                 140

Pro Gly Pro Trp Gly Leu Val Lys Glu Glu Glu Gly Tyr Gly Glu Ser
145                 150                 155                 160

Asp Val Thr Phe Ser Val Asn Gln His Arg Arg Thr Lys Gln Pro Gln
                165                 170                 175

Glu Asn Thr Thr His Val Glu Glu Val Thr Glu Ala Gly Asp Ser Glu
            180                 185                 190

Gly Glu Gln Ser Tyr Phe Val Ser Thr Pro Asp Glu His Pro Gly Gly
        195                 200                 205

His Asp Arg Asp Asp Gly Ser Val Tyr Gly Gln Ile His Ile Glu Glu
    210                 215                 220

Glu Ser Thr Ile Arg Tyr Ser Trp Gln Asp Glu Ile Val Gln Gly Thr
225                 230                 235                 240

Arg Arg Arg Thr Gln Lys Asp Gly Ala Val Gly Glu Lys Val Val Lys
                245                 250                 255

Pro Leu Asp Val Pro Ala Pro Ser Leu Glu Gly Asp Leu Gly Ser Thr
            260                 265                 270

His Trp Lys Glu Gln Ala Arg Ser Gly Glu Phe His Ala Glu Pro Thr
        275                 280                 285
```

Val Arg Lys Arg Asn
290

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1073 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

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

CACCTGCCTC GGCCTCCCAG AGTGCTGAGA TTACAGGCGT GAGCCTCCGC GCCCGGCCCC 60

CTTGCAGTTC TCTCTGATTT GGTTTGTTCT GTCTCAGGCT TCTGTGGCAG GACTGGCCCA 120

GGGAGGAGGA AGCCAGCAGC ACACCTGGGG AATGGGGTCC CGGCCGGGAG GCTTGGCCTC 180

TGGGCGACCT CGTCCTGTTT TGTTTGTTTG TTTGTTTGTT TTTTTAAAGG TAAACCTCCT 240

GGGCCGCAGA TGGCAAAGGG AGTGCCTGGG CCTGGTGACC CAGGGCTGGA TCCACCCCTG 300

CGGAGCCCTG GGCCAGGCAG GTGTCTGCTG CTCACCTGGC TCTGGAGGGC TGCCCTGCAG 360

CTGGGCCTGG GGACAGGTCG GCTGTGGGGC AGCTCAGTAC CCTCCCTGAG GCTCACGGTG 420

GCTCCGAGCA TGAGCTCTGC CTCCTGGGCG AGACCCAGCA GTGGACAGCA CGGTCCTCAC 480

ACCCAGCTCC CTGCACACCC AGGCCAGCCA CCCCTCCCGC TCGTGCACAG GCACGCAGAT 540

GCGCTCACAC GTACACACAC ACAAATGCAC GCCCACTTGC ACATGCTCAC GCACACGTTC 600

ACACATGCAC ACTCACGCTC ACACATGCTG TCACGCATAC ACACACGCAC ATACTCCTGC 660

ACATGTTCCC ATGCATGTGT GTGCACTCGG ACCGAGCATC TCCCACGCAC CTCTACCCCA 720

CCCCAAGCAC CTCTCTCCCC CCATGCACCT CTCCCCAACA ACACACACAG CCCCCTGCAC 780

CGCCCGCCCC CCGCCCCCAC CAAGGCCCCA GCCTCTGGCC ATCAGTCCTG GTGCCAGAGC 840

TTTGCGTGAA GTTCGGGCCG CAGAGTGGCC CGCTGGGACT CCCATGTGCT GCCGTCTGAT 900

GTGCTCAGAT GGGCTCATCG TTGGTTCGTT TTTACTGTAT ATTTATAGTA ATAAAATCAT 960

GCAGCAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 1020

AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAA 1073

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 199 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

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

```
Met Ala Lys Gly Val Pro Gly Pro Gly Asp Pro Gly Leu Asp Pro Pro
1               5                  10                  15

Leu Arg Ser Pro Gly Pro Gly Arg Cys Leu Leu Leu Thr Trp Leu Trp
            20                  25                  30

Arg Ala Ala Leu Gln Leu Gly Leu Gly Thr Gly Arg Leu Trp Gly Ser
        35                  40                  45

Ser Val Pro Ser Leu Arg Leu Thr Val Ala Pro Ser Met Ser Ser Ala
    50                  55                  60

Ser Trp Ala Arg Pro Ser Ser Gly Gln His Gly Pro His Thr Gln Leu
65                  70                  75                  80

Pro Ala His Pro Gly Gln Pro Pro Leu Pro Leu Val His Arg His Ala
                85                  90                  95

Asp Ala Leu Thr Arg Thr His Thr Gln Met His Ala His Leu His Met
            100                 105                 110

Leu Thr His Thr Phe Thr His Ala His Ser Arg Ser His Met Leu Ser
        115                 120                 125

Arg Ile His Thr Arg Thr Tyr Ser Cys Thr Cys Ser His Ala Cys Val
    130                 135                 140

Cys Thr Arg Thr Glu His Leu Pro Arg Thr Ser Thr Pro Pro Gln Ala
145                 150                 155                 160

Pro Leu Ser Pro His Ala Pro Leu Pro Asn Asn Thr His Ser Pro Leu
                165                 170                 175

His Arg Pro Pro Pro Ala Pro Thr Lys Ala Pro Ala Ser Gly His Gln
            180                 185                 190

Ser Trp Cys Gln Ser Phe Ala
        195
```

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 674 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

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

GGAAATTCCA GCGCTGCTTC ATCTTCCAGT GCTGATAATT ATGGCATTAG CCATCCTGAG 60

TTTCTGCTAT GGTGCTGGAA AATCAGTTCA TGTGCTGAGA CATATAGGCG GTCCTGAGAG 120

CGAACCTCCC CAGGCACTTC GGCCACGGGA TAGAAGACGG CAGGAGGAAA TTGATTATAG 180

ACCTGATGGT GGAGCAGGTG ATGCCGATTT CCATTATAGG GGCCAAATGG GCCCCACTGA 240

GCAAGGCCCT TATGCCAAAA CGTATGAGGG TAGAAGAGAG ATTTTGAGAG AGAGAGATGT 300

TGACTTGAGA TTTCAGACTG GCAACAAGAG CCCTGAAGTG CTCCGGGCAT TTGATGTACC 360

AGACGCAGAG GCACGAGAGC ATCCCACGGT GGTACCCAGT CATAAATCAC CTGTTTTGGA 420

TACAAAGCCC AAGGAGACAG GTGGAATCCT GGGGGAAGGC ACACCGAAAG AAAGCAGTAC 480

TGAAAGCAGC CAGTCGGCCA AGCCTGTCTC TGGCCAAGAC ACATCAGGGA ATACAGAAGG 540

TTCACCCGCA GCGGAAAAGG CCCAGCTCAA GTCTGAAGCC GCAGGCAGCC CAGACCAAGG 600

CAGCACATAC AGCCCCGCAA GAGGTGTGGC TGGACCACGT GGACAGGATC CGGTCAGCAG 660

CCCCTGTGGC TAGA 674

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 107 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

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

Met Tyr Gln Thr Gln Arg His Glu Ser Ile Pro Arg Trp Tyr Pro Val
1 5 10 15

Ile Asn His Leu Phe Trp Ile Gln Ser Pro Arg Arg Gln Val Glu Ser

20 25 30

Trp Gly Lys Ala His Arg Lys Lys Ala Val Leu Lys Ala Ala Ser Arg
35 40 45

Pro Ser Leu Ser Leu Ala Lys Thr His Gln Gly Ile Gln Lys Val His
50 55 60

Pro Gln Arg Lys Arg Pro Ser Ser Ser Leu Lys Pro Gln Ala Ala Gln
65 70 75 80

Thr Lys Ala Ala His Thr Ala Pro Gln Glu Val Trp Leu Asp His Val
85 90 95

Asp Arg Ile Arg Ser Ala Ala Pro Val Ala Arg
100 105

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 169 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

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

AACAATGTTT AAATGGAATA ATGTTGAATG TTTACAGGCT GTAAGAACTA TTGTATACAC 60

AAAATAATAC ACAAAGTTTG TACTTTGTGT ACAAATACAA ATTTGTACTT TGTGTACAAA 120

TAATACAAAA AGTTTGTATA CACAAAAAAA AAAAAAAAAA AAAAAAAAA 169

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1210 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

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

GGTTCTGTGC TCCTGAGATT AGTTCAGATG GTCTAACCAT TGTTCTATAT GTGCATTTTA 60

GTTAATATTG TGTATTAAAG GATAAGTCTT AATGCTCAAA GTATGTTAAA AATAGATGTA 120

GTAAATCAGT CCCTTTGTGA ATGTCCTTTT GTTAGTTTTT AGGAAGGCCT GTCCTCTGGG 180

AGTGACCTTT ATTAGTCCAC CCCTTGGAGC TAGACATCCT GTACTTAGTC ACGGGGATGG 240

TGGAAGAGGG AGAAGAGGAA GGGTGAAGGG AAGGGCTCTT TGCTAGTATC TCCATATCTA 300

GACGATGGTT TTAGATGATA ACCACAGGTC TACAAGAGCG TTTTTAGTAA AGTGCCTGTG 360

TTCATTGTGG ACAAAGTTAT TATTTTGCAA CATCTAAGCT TTACGAATGG GGTGACAACT 420

TATGATAAAA ACTAGAGCTA GTGAATTAGC CTATTTGTAA ATACCTTTGT TATAATTGAT 480

AGGATACATC TTGGACATGG AATTGTTAAG CCACCTCTGA GCAGTGTATG TCAGGACTTG 540

TTCATTAGGT TGGCAGCAGA GGGGCAGAAG GAATTATACA GGTAGAGATG TATGCAGATG 600

TGTCCATATA TGTCCATATT TACATTTTGA TAGCCATTGA TGTATGCATC TCTTGGCTGT 660

ACTATAAGAA CACATTAATT CAATGGAAAT ACACTTTGCT AATATTTTAA TGGTATAGAT 720

CTGCTAATGA ATTCTCTTAA AAACATACTG TATTCTGTTG CTGTGTGTTT CATTTTAAAT 780

TGAGCATTAA GGGAATGCAG CATTTAAATC AGAACTCTGC CAATGCTTTT ATCTAGAGGC 840

GTGTTGCCAT TTTTGTCTTA TATGAAATTT CTGTCCCAAG AAAGGCAGGA TTACATCTTT 900

TTTTTTTTTT TTAGCAGTTT GAGTTGGTGT AGTGTATTCT TGGTTATCAG AATACTCATA 960

TAGCTTTGGG ATTTTGAATT GGTAAATATT CATGATGTGT GAAAAATCAT GATACATACT 1020

GTACAGTCTC AGTCCCATAA AATTGGATGT TGTGCCTACA CACAGGATCT AGAAGAATAT 1080

GTCAAACTAT AAACTGCTTG TGATTGTGAA TGACTTTGTT CTTTGCTTGT GTTTTTCAAT 1140

TTCCTATAAT GCACATACTA ACTTTTAAAA AATAAAGGTT ATTTTAAAAG CCTGTAAAAA 1200

AAAAAAAAAA 1210

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 46 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

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

Met Leu Leu Ser Arg Gly Val Leu Pro Phe Leu Ser Tyr Met Lys Phe
1 5 10 15

Leu Ser Gln Glu Arg Gln Asp Tyr Ile Phe Phe Phe Phe Ser Ser
20 25 30

Leu Ser Trp Cys Ser Val Phe Leu Val Ile Arg Ile Leu Ile
35 40 45

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2309 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

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

GCAACACTCC GAGGCCAGGA ACGCTCCGTC TGGAACGGCG CAGGTCCCAG CAGCTGGGGT 60

TCCCCGTCAG CCCGTGAGCG GCCATGTCCA ACCCCAGCGC CCCACCACCA TATGAAGACC 120

GCAACCCCCT GTACCCAGGC CCTCCGCCCC CTGGGGGCTA TGGGCAGCCA TCTGTCCTGC 180

CAGGAGGGTA TCCTGCCTAC CCTGGCTACC CGCAGCCTGG CTACGGTCAC CCTGCTGGCT 240

ACCCACAGCC CATGCCCCCC ACCCACCCGA TGCCCATGAA CTACGGCCCA GGCCATGGCT 300

ATGATGGGGA GGAGAGAGCG GTGAGTGATA GCTTCGGGCC TGGAGAGTGG GATGACCGGA 360

AAGTGCGACA CACTTTTATC CGAAAGGTTT ACTCCATCAT CTCCGTGCAG CTGCTCATCA 420

CTGTGGCCAT CATTGCTATC TTCACCTTTG TGGAACCTGT CAGCGCCTTT GTGAGGAGAA 480

ATGTGGCTGT CTACTACGTG TCCTATGCTG TCTTCGTTGT CACCTACCTG ATCCTTGCCT 540

GCTGCCAGGG ACCCAGACGC CGTTTCCCAT GGAACATCAT TCTGCTGACC CTTTTTACTT 600

TTGCCATGGG CTTCATGACG GGCACCATTT CCAGTATGTA CCAAACCAAA GCCGTCATCA 660

TTGCAATGAT CATCACTGCG GTGGTATCCA TTTCAGTCAC CATCTTCTGC TTTCAGACCA 720

AGGTGGACTT CACCTCGTGC ACAGGCCTCT TCTGTGTCCT GGGAATTGTG CTCCTGGTGA 780

CTGGGATTGT CACTAGCATT GTGCTCTACT TCCAATACGT TTACTGGCTC CACATGCTCT 840

ATGCTGCTCT GGGGGCCATT TGTTTCACCC TGTTCCTGGC TTACGACACA CAGCTGGTCC 900

TGGGGAACCG GAAGCACACC ATCAGCCCGG AGGACTACAT CACTGGCGCC CTGCAGATTT 960

ACACAGACAT CATCTACATC TTCACCTTTG TGCTGCAGCT GATGGGGGAT CGCAATTAAG 1020

GAGCAAGCCC CCATTTTCAC CCGATCCTGG GCTCTCCCTT CCAAGCTAGA GGGCTGGGCC 1080

CTATGACTGT GGTCTGGGCT TTAGGCCCCT TTCCTTCCCC TTGAGTAACA TGCCCAGTTT 1140

CCTTTCTGTC CTGGAGACAG GTGGCCTCTC TGGCTATGGA TGTGTGGGTA CTTGGTGGGG 1200

ACGGAGGAGC TAGGGACTAA CTGTTGCTCT TGGTGGGCTT GGCAGGGACT AGGCTGAAGA 1260

TGTGTCTTCT CCCCGCCACC TACTGTATGA CACCACATTC TTCCTAACAG CTGGGGTTGT 1320

GAGGAATATG AAAAGAGCCT ATTCGATAGC TAGAAGGGAA TATGAAAGGT AGAAGTGACT 1380

TCAAGGTCAC GAGGTTCCCC TCCCACCTCT GTCACAGGCT TCTTGACTAC GTAGTTGGAG 1440

CTATTTCTTC CCCCAGCAAA GCCAGAGAGC TTTGTCCCCG GCCTCCTGGA CACATAGGCC 1500

ATTATCCTGT ATTCCTTTGG CTTGGCATCT TTTAGCTCAG GAAGGTAGAA GAGATCTGTG 1560

CCCATGGGTC TCCTTGCTTC AATCCCTTCT TGTTTCAGTG ACATATGTAT TGTTTATCTG 1620

GGTTAGGGAT GGGGGACAGA TAATAGAACG AGCAAAGTAA CCTATACAGG CCAGCATGGA 1680

ACAGCATCTC CCCTGGGCTT GCTCCTGGCT TGTGACGCTA TAAGACAGAG CAGGCCACAT 1740

GTGGCCATCT GCTCCCCATT CTTGAAAGCT GCTGGGGCCT CCTTGCAGGC TTCTGGATCT 1800

CTGGTCAGAG TGAACTCTTG CTTCCTGTAT TCAGGCAGCT CAGAGCAGAA AGTAAGGGGC 1860

AGAGTCATAC GTGTGGCCAG GAAGTAGCCA GGGTGAAGAG AGACTCGGTG CGGGCAGGGA 1920

GAATGCCTGG GGGTCCCTCA CCTGGCTAGG GAGATACCGA AGCCTACTGT GGTACTGAAG 1980

ACTTCTGGGT TCTTTCCTTC TGCTAACCCA GGGAGGGTCC TAAGAGGAAG GTGACTTCTC 2040

TCTGTTTGTC TTAAGTTGCA CTGGGGGATT TCTGACTTGA GGCCCATCTC TCCAGCCAGC 2100

CACTGCYTTC TTTGTAATAT TAAGTGCCTT GAGCTGGAAT GGGGAAGGGG GACAAGGGTC 2160

AGTCTGTCGG GTGGGGGCAG AAATCAAATC AGCCCAAGGA TATAGTTAGG ATTAATTACT 2220

TAATAGAGAA ATCCTAACTA TATCACACAA AGGGATACAA CTATAAATGT AATAAAGTTT 2280

ATGTCTAGAA GTTAAAAAAA AAAAAAAAA 2309

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 311 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

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

```
Met Ser Asn Pro Ser Ala Pro Pro Pro Tyr Glu Asp Arg Asn Pro Leu
1               5                  10                  15

Tyr Pro Gly Pro Pro Pro Pro Gly Gly Tyr Gly Gln Pro Ser Val Leu
            20                  25                  30

Pro Gly Gly Tyr Pro Ala Tyr Pro Gly Tyr Pro Gln Pro Gly Tyr Gly
        35                  40                  45

His Pro Ala Gly Tyr Pro Gln Pro Met Pro Pro Thr His Pro Met Pro
    50                  55                  60

Met Asn Tyr Gly Pro Gly His Gly Tyr Asp Gly Glu Glu Arg Ala Val
65                  70                  75                  80

Ser Asp Ser Phe Gly Pro Gly Glu Trp Asp Asp Arg Lys Val Arg His
                85                  90                  95

Thr Phe Ile Arg Lys Val Tyr Ser Ile Ile Ser Val Gln Leu Leu Ile
            100                 105                 110

Thr Val Ala Ile Ile Ala Ile Phe Thr Phe Val Glu Pro Val Ser Ala
        115                 120                 125

Phe Val Arg Arg Asn Val Ala Val Tyr Tyr Val Ser Tyr Ala Val Phe
    130                 135                 140

Val Val Thr Tyr Leu Ile Leu Ala Cys Cys Gln Gly Pro Arg Arg Arg
145                 150                 155                 160

Phe Pro Trp Asn Ile Ile Leu Leu Thr Leu Phe Thr Phe Ala Met Gly
                165                 170                 175

Phe Met Thr Gly Thr Ile Ser Ser Met Tyr Gln Thr Lys Ala Val Ile
            180                 185                 190

Ile Ala Met Ile Ile Thr Ala Val Val Ser Ile Ser Val Thr Ile Phe
        195                 200                 205

Cys Phe Gln Thr Lys Val Asp Phe Thr Ser Cys Thr Gly Leu Phe Cys
    210                 215                 220

Val Leu Gly Ile Val Leu Leu Val Thr Gly Ile Val Thr Ser Ile Val
225                 230                 235                 240

Leu Tyr Phe Gln Tyr Val Tyr Trp Leu His Met Leu Tyr Ala Ala Leu
                245                 250                 255

Gly Ala Ile Cys Phe Thr Leu Phe Leu Ala Tyr Asp Thr Gln Leu Val
            260                 265                 270
```

Leu Gly Asn Arg Lys His Thr Ile Ser Pro Glu Asp Tyr Ile Thr Gly
275 280 285

Ala Leu Gln Ile Tyr Thr Asp Ile Ile Tyr Ile Phe Thr Phe Val Leu
290 295 300

Gln Leu Met Gly Asp Arg Asn
305 310

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1212 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

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

CACCGCTACA GCGCGGCGGG CGTCTCGCTG CTCGAGCCGC CGCTGCAGCT CTACTGGACC 60

TGGCTGCTCC AGTGGATCCC GCTCTGGATG GCCCCCAACT CCATCACCCT GCTGGGGCTC 120

GCCGTCAACG TGGTCACCAC GCTCGTGCTC ATCTCCTACT GTCCCACGGC CACCGAAGAG 180

GCACCATACT GGACATACCT TTTATGTGCA CTGGGACTTT TTATTTACCA GTCACTGGAT 240

GCTATTGATG GGAAACAAGC CAGAAGAACA AACTCTTGTT CCCCTTTAGG GGAGCTCTTT 300

GACCATGGCT GTGACTCTCT TTCCACAGTA TTTATGGCAG TGGGAGCTTC AATTGCCGCT 360

CGCTTAGGAA CTTATCCTGA CTGGTTTTTT TTCTGCTCTT TTATTGGGAT GTTTGTGTTT 420

TATTGCGCTC ATTGGCAGAC TTATGTTTCA GGCATGTTGA GATTTGGAAA AGTGGATGTA 480

ACTGAAATTC AGATAGCTTT AGTGATTGTC TTTGTGTTGT CTGCATTTGG AGGAGCAACA 540

ATGTGGGACT ATACGTTTTC TTGAACAGCA GTTGACCAAC AGATTCCTAT TCTAGAAATA 600

AAATTGAAGA TCCTTCCAGT TCTTGGATTT CTAGGGCACC AGTGTCTTGT CACCTGGACT 660

CCACATAGGA CTAATTATTA TACTGGCAAT AATGATCTAT AAAAAGTCAG CAACTGATGT 720

GTTTGAAAAG CATCCTTGTC TTTATATCCT AATGTTTGGA TGTGTCTTTG CTAAAGTCTC 780

ACAAAAATTA GTGGTAGCTC ACATGACCAA AAGTGAACTA TATCTTCAAG ACACTGTCTT 840

TTTGGGGCCA GGTCTTTTGT TTTTAGACCA GTACTTTAAT AACTTTATAG ACGAATATGT 900

TGTTCTATGG ATGGCAATGG TGATTTCTTC ATTTGATATG GTGATATACT TTAGTGCTTT 960

GTGCCTGCAA ATTTCAAGAC ACCTTCATCT AAATATATTC AAGACTGCAT GTCATCAAGC 1020

ACCTGAACAG GTTCAAGTTC TTTCTTCAAA GAGTCATCAG AATAACATGG ATTGAAGAGA 1080

CTTCCGAACA CTTGCTATCT CTTGCTGCTG CTGTTTCATG GAAGGAGATA TTAAACATTT 1140

GTTTAATTTT TATTTAAGTG TTATACCTAT TTCAGCAAAT AAAATATTTC ATTGCTTAAA 1200

AAAAAAAAAA AA 1212

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 158 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

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

Met Ala Pro Asn Ser Ile Thr Leu Leu Gly Leu Ala Val Asn Val Val
1 5 10 15

Thr Thr Leu Val Leu Ile Ser Tyr Cys Pro Thr Ala Thr Glu Glu Ala
20 25 30

Pro Tyr Trp Thr Tyr Leu Leu Cys Ala Leu Gly Leu Phe Ile Tyr Gln
35 40 45

Ser Leu Asp Ala Ile Asp Gly Lys Gln Ala Arg Arg Thr Asn Ser Cys
50 55 60

Ser Pro Leu Gly Glu Leu Phe Asp His Gly Cys Asp Ser Leu Ser Thr
65 70 75 80

Val Phe Met Ala Val Gly Ala Ser Ile Ala Ala Arg Leu Gly Thr Tyr
85 90 95

Pro Asp Trp Phe Phe Phe Cys Ser Phe Ile Gly Met Phe Val Phe Tyr
100 105 110

Cys Ala His Trp Gln Thr Tyr Val Ser Gly Met Leu Arg Phe Gly Lys
115 120 125

Val Asp Val Thr Glu Ile Gln Ile Ala Leu Val Ile Val Phe Val Leu
130 135 140

Ser Ala Phe Gly Gly Ala Thr Met Trp Asp Tyr Thr Phe Ser
145 150 155

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1831 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

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

AGACAACTAA GCAACATCTG AATACAAATG TAAAGTATGC TGGCTCAGTG TATCTTTTAG 60

GTCACTAAAG AGGACACATC AGTAAATAAA TTGTTTTCAG TGTAATATTT TTACATAAAA 120

ATARATTTCC CCCCCAAAAT TTAARATTTA TTTARATTCT TGAGACATTT TTTTATGAAA 180

AGGAATTAAT AACATACATC ATCAAAACTT TTCTTAGGAG CGTTGAAACA ATATAAGAAG 240

GTTACAGAAG GCCTTCTAAA ATATGAATTC CCATCTGTAT GAGGTACTTC TTTCCAACTC 300

AGTGACTTTT TGCCCAGAGA CTTGGAAGAC TGGGCAATGT TTTCTTTCCC CTTTCCCAAC 360

TCCTGGAGAC AAGGCCCTAG GAGCAATTTC CTCCAGTAGG GCAGGCCGTA GTCCTGCTCT 420

GTCTTACTTC CCACTGCAGC AGGAACATGT CCTCGCCTCT GATTCCTTCT GTACCAAGCC 480

ATTGGTCCAT TGAGGTCAAT TGAGGACTCA CTGTAAATTT TAAATCTGTT AATAAAGCAA 540

GGATATTGGC ATGTTCCTCT TCTCATCAAT ATCCTAAAAG ACATTTATTT TTTACACACT 600

CCTTGGGAAA AATTAACTTT TTTTCACTGA AAATATTTCC TTTTTTGGTT ATCTTGATCT 660

CAGATTATTT TGTGAAAGAA TTTTACTGTA CTTAGTTCAA AAGAGTAGAA AGAATGATTT 720

ACTATTGCAG ACATATGTAG GGTAAAATCA TAATTTATTT AAACTGACTG TACAACACCA 780

TTTAGAGTTG ATATTGACAT AAATGTTATT AGCCTACTAA TTTGGAACTG CATTTCTCAA 840

CAATGCTGGC AAGCATCTTC CGTACTTAGC ATACCAAGTT GTAGGGGAGA GACTGTGTAT 900

ATATTTTTTA AAAGCAATCC AATGGATTTG TTTTTGTTCA TATTTTGAAA ACAACTCGAA 960

GGATTTTTCT TATTTTAGAA GGTAGAAAAT ATGTGGACCT GTGTTTTAAA TTAATATGTA 1020

TTTATAAATT ATAACATTGA ATGATACAAC AGAGCTCTAA TATTGATGTT CTCACTTTCT 1080

GATATTTAAT TTTTATAAAT GTTTTTGTAA GTAACTAATT GTAGTATTGC TTTTAAATAG 1140

TTTTAAAATC TATCTCAAAT AAGTTCTGCC CAGACCTATT TCCTTAGGAC AGTATTCTAA 1200

AGTTCAGTAG TCCAGTGTGA GCTTGAAATA CTCCTAGATC CCGTCAGTGA TTTTTATATA 1260

TACCAATACA CAAGCAAATT ATTAATATAC CAAAAACCTA ATTTCTGCAT TGCATTAATA 1320

AAATACTTGT TATTTTGCCT TTATCCAATT CAGTTGATTA GGTGAAATAG AAAATTAATC 1380

TTTTCAAATG TGCATATGAT CATAATTTTT CAAAATGTTA CTGTGGTCAT TTTTGGTGCA 1440

TGATGCCAAG TTTATCTTTA GTTAGCCATT GCCACCTGAT ATGTAATGAC AAATGTTTTA 1500

CTATCTGATC ATTGGGTTGT AGATTAAACT ATTTTTTTTC TCTGTAGATT CTCACTACAT 1560

TTTCAAATAT GTACTTGGAA AAAGCTGCAA AGTGCTACCT ATAAGCAAAT TAGTAACTTT 1620

GGCCTTATCT ATGCATAGTG TTCGCTCAGT TTGGTTTTGA TAAATTGATT GACTTGCAGT 1680

TCCTTTGTAG TATTTTAACT TATTGTGATG TAATGAATTT TTTGAAATGT TTATTTGATT 1740

AGCTGTTAAA ATATAGGAAA GAATGTAGCT TTGCATGAAA TAAATTACAT TTCTACAAGT 1800

TAAAAAAAAA RAAAAAAAAA AAAAAAAAAA A 1831

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 53 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

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

Val Asp Ile Asp Ile Asn Val Ile Ser Leu Leu Ile Trp Asn Cys Ile
1 5 10 15

Ser Gln Gln Cys Trp Gln Ala Ser Ser Val Leu Ser Ile Pro Ser Cys
20 25 30

Arg Gly Glu Thr Val Tyr Ile Phe Phe Lys Ser Asn Pro Met Asp Leu
35 40 45

Phe Leu Phe Ile Phe
50

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 980 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

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

CTAAAGTTTA CTGGTGAGGA GCTTGAACTG AGTAGAAGCA AAGAATCATA CTCCAGCCTG 60

GGTGACAGAG CAAGACTCTG TCTCAAAAAA AAAAATTATT TAAACAATAG TATCCATCTT 120

CCCACAATTA CTATTAAAGA TTCTTAAAAT TTCTGGAAAT ATTGAAGGAA ACGAATTATG 180

CTTTAAGATT TTAAACCTAC AGAGTAAGTT AATTCTGTCT TTGCCATTTA CTAGCTGTGT 240

AAACTTGGAG AGATCAGCTT GAGACTCTCT GAGCCTCATT TTTTTAATCT GTCAAATGGG 300

TCTCAGAAAA TTGTGAGGAT TAAATGAAAT ATTACCTGTA AGTTGTTTAG AACAGTTCTT 360

GGAACATAGT AAAAACTCAG AAAGTGTTCA GTAAAATGTT GAGTATATTA TTGTGATGTG 420

GGAATTGCTG ATCCAGTTGA CTCAATTCAC ACTGCAAATT ACTGGATTTG AATAAGGAGA 480

TGAAGATTGA GTCACATTGT TCTCTTTGAT CCATTTTCTT TTAGACAACA TCACTCAAAC 540

AACCCTATTA CTTCAGTAGC AATGAACTGT CTCTGGATTT TATTAAGCAT CTCACTAGTT 600

CCTTTTCTTC AGCTGTATGG CACCCTGTCT TCCTGTACAC CAGAGGCTCC TCAGCTGGGT 660

AAGGTGAGCC AACGTTACCA GGAGTATATG CTGAGAGGCC ATTTCAAAGT CTTTCATAGA 720

AGGCTGTGCT TGGGCAAGTA GAACTTTTCA TCATACAGTC CCAGAGATGT GAAGTTATCA 780

AGGTCAGAGA AGAGGAAAAG AGACTCAGAG AAACTGTGTT CTCTAGTCTT TCATTCGGGA 840

ACAAATGGTG TCGTACCAAA TGGCTGAAAA GATCCCAGGC TTACTCTGAG AAACATCTCT 900

CTCTTCTGAA TTAAACTTGC ACAGTGGCAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 960

AAAAAAAAAA AAAAAAAAAA 980

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 59 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

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

Met Asn Cys Leu Trp Ile Leu Leu Ser Ile Ser Leu Val Pro Phe Leu
1 5 10 15

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

Gly Lys Val Ser Gln Arg Tyr Gln Glu Tyr Met Leu Arg Gly His Phe
35 40 45

Lys Val Phe His Arg Arg Leu Cys Leu Gly Lys
50 55

(2) INFORMATION FOR SEQ ID NO:18:

(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:18:

TNTCTGTCATG GAATCGCCTT GCAGAGAA 29

(2) INFORMATION FOR SEQ ID NO:19:

(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:19:

CNGCCACAGAA GCCTGAGACA GAACAAAC 29

(2) INFORMATION FOR SEQ ID NO:20:

(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:20:

TNAAATGCCCG GAGCACTTCA GGGCTCTT 29

(2) INFORMATION FOR SEQ ID NO:21:

(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:21:

TNTTATAGTAC AGCCAAGAGA TGCATACA 29

(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:

ANCAGAAGATG GTGACTGAAA TGGATACC 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:

GNAGACAACAC AAAGACAATC ACTAAAGC 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:

CNTACAACTTG GTATGCTAAG TACGGAAG 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:

GNAGTAATAGG GTTGTTTGAG TGATGTTG 29

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 128 to nucleotide 1006;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 182 to nucleotide 362;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone BG481_1 deposited under accession number ATCC 98331;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone BG481_1 deposited under accession number ATCC 98331;

(f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone BG481_1 deposited under accession number ATCC 98331;

(g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone BG481_1 deposited under accession number ATCC 98331;

(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;

(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) the amino acid sequence of SEQ ID NO:2 from amino acid 1 to amino acid 78;

(c) fragments of the amino acid sequence of SEQ ID NO:2; and

(d) the amino acid sequence encoded by the cDNA insert of clone BG481_1 deposited under accession number ATCC 98331;

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. The protein of claim 8, wherein said protein comprises the amino acid sequence of SEQ ID NO:2 from amino acid 1 to amino acid 78.

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

12. 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 11.

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

14. 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 250 to nucleotide 846;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 361 to nucleotide 846;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 1 to nucleotide 564;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone BJ9_1 deposited under accession number ATCC 98331;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone BJ9_1 deposited under accession number ATCC 98331;

(g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone BJ9_1 deposited under accession number ATCC 98331;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone BJ9_1 deposited under accession number ATCC 98331;

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

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:4 having biological activity;

(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).

15. 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 1 to amino acid 105;

(c) fragments of the amino acid sequence of SEQ ID NO:4; and

(d) the amino acid sequence encoded by the cDNA insert of clone BJ9_1 deposited under accession number ATCC 98331;

the protein being substantially free from other mammalian proteins.

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

17. 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 354 to nucleotide 674;

(c) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone BK34_3 deposited under accession number ATCC 98331;

(d) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone BK34_3 deposited under accession number ATCC 98331;

(e) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone BK34_3 deposited under accession number ATCC 98331;

(f) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone BK34_3 deposited under accession number ATCC 98331;

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

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

(i) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(f) above;

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

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

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

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

(b) fragments of the amino acid sequence of SEQ ID NO:6; and

(c) the amino acid sequence encoded by the cDNA insert of clone BK34_3 deposited under accession number ATCC 98331;

the protein being substantially free from other mammalian proteins.

19. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:5 and SEQ ID NO:7.

20. 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 823 to nucleotide 960;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:8 from nucleotide 931 to nucleotide 960;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone BP883_2 deposited under accession number ATCC 98331;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone BP883_2 deposited under accession number ATCC 98331;

(f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone BP883_2 deposited under accession number ATCC 98331;

(g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone BP883_2 deposited under accession number ATCC 98331;

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

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

(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).

21. 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 1 to amino acid 16;

(c) fragments of the amino acid sequence of SEQ ID NO:9; and

(d) the amino acid sequence encoded by the cDNA insert of clone BP883_2 deposited under accession number ATCC 98331;

the protein being substantially free from other mammalian proteins.

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

23. 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 84 to nucleotide 1016;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:10 from nucleotide 786 to nucleotide 1016;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:10 from nucleotide 619 to nucleotide 899;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CI363_1 deposited under accession number ATCC 98331;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CI363_1 deposited under accession number ATCC 98331;

(g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CI363_1 deposited under accession number ATCC 98331;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CI363_1 deposited under accession number ATCC 98331;

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

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:11 having biological activity;

(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).

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

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

(b) the amino acid sequence of SEQ ID NO:11 from amino acid 180 to amino acid 272;

(c) fragments of the amino acid sequence of SEQ ID NO:11; and

(d) the amino acid sequence encoded by the cDNA insert of clone CI363_1 deposited under accession number ATCC 98331;

the protein being substantially free from other mammalian proteins.

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

26. 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 88 to nucleotide 561;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:12 from nucleotide 142 to nucleotide 561;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:12 from nucleotide 1 to nucleotide 554;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CO806_1 deposited under accession number ATCC 98331;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CO806_1 deposited under accession number ATCC 98331;

(g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CO806_1 deposited under accession number ATCC 98331;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CO806_1 deposited under accession number ATCC 98331;

(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;

(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).

27. 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 112 to amino acid 156;

(c) fragments of the amino acid sequence of SEQ ID NO:13; and

(d) the amino acid sequence encoded by the cDNA insert of clone CO806_1 deposited under accession number ATCC 98331;

the protein being substantially free from other mammalian proteins.

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

29. 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 787 to nucleotide 945;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:14 from nucleotide 853 to nucleotide 945;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:14 from nucleotide 324 to nucleotide 945;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CT24_3 deposited under accession number ATCC 98331;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CT24_3 deposited under accession number ATCC 98331;

(g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CT24_3 deposited under accession number ATCC 98331;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CT24_3 deposited under accession number ATCC 98331;

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

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:15 having biological activity;

(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).

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

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

(b) fragments of the amino acid sequence of SEQ ID NO:15; and

(c) the amino acid sequence encoded by the cDNA insert of clone CT24_3 deposited under accession number ATCC 98331;

the protein being substantially free from other mammalian proteins.

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

32. 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 562 to nucleotide 738;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:16 from nucleotide 1 to nucleotide 729;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone ER366_3 deposited under accession number ATCC 98331;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone ER366_3 deposited under accession number ATCC 98331;

(f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone ER366_3 deposited under accession number ATCC 98331;

(g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone ER366_3 deposited under accession number ATCC 98331;

(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;

(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).

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

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

(b) the amino acid sequence of SEQ ID NO:17 from amino acid 1 to amino acid 56;

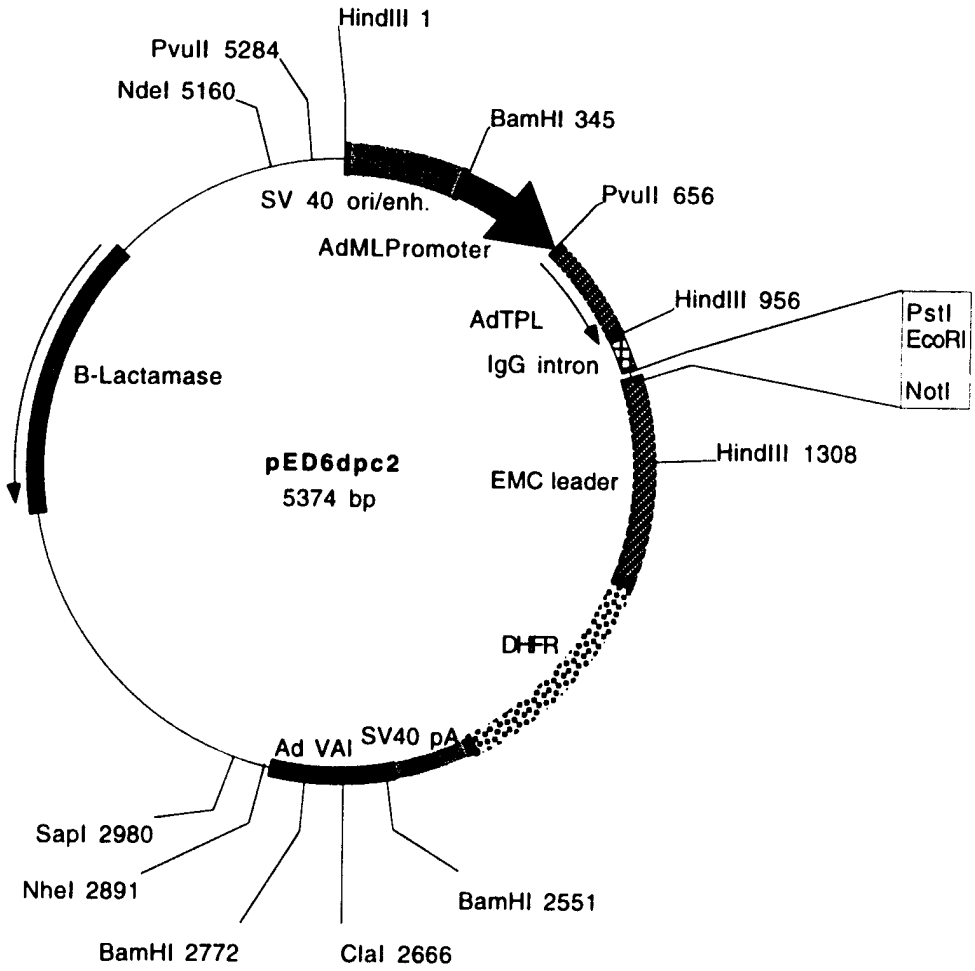
(c) fragments of the amino acid sequence of SEQ ID NO:17; and

(d) the amino acid sequence encoded by the cDNA insert of clone ER366_3 deposited under accession number ATCC 98331;

the protein being substantially free from other mammalian proteins.

34. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:16.

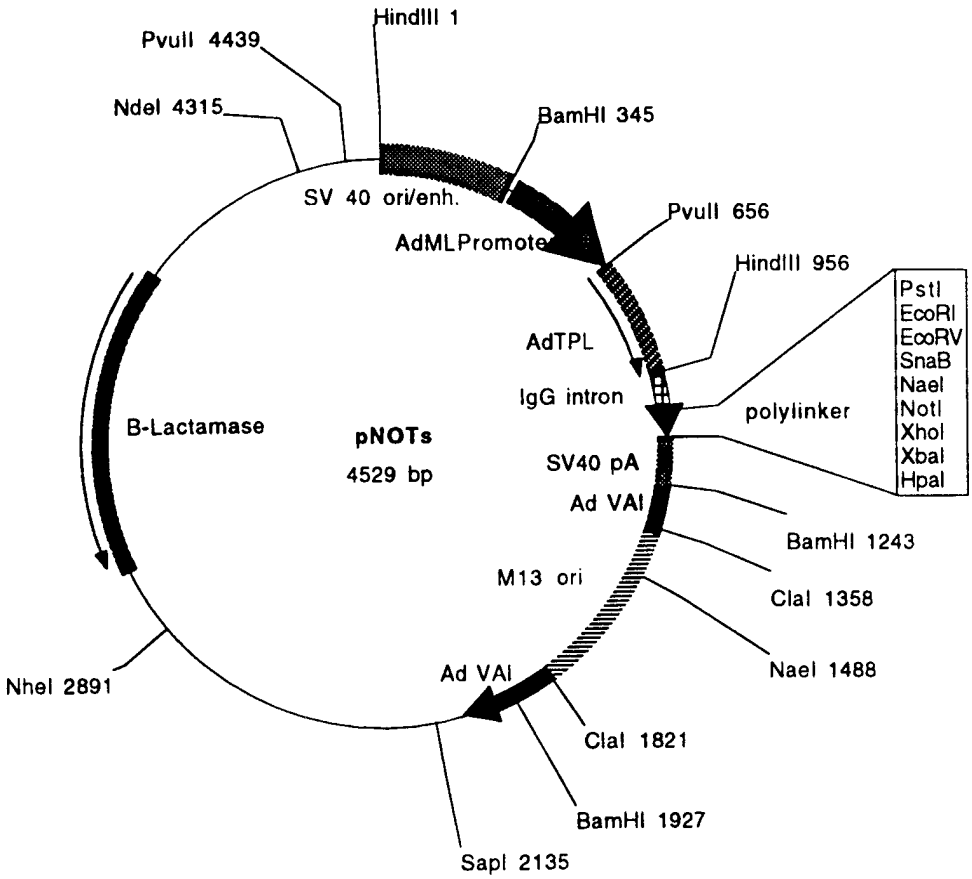
FIGURE 1A



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