PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/12, 5/10, C07K 14/47, C12Q 1/68, A61K 38/17

A2

(11) International Publication Number: **WO 98/24905**

(43) International Publication Date: 11 June 1998 (11.06.98)

(21) International Application Number: PCT/US97/22211

(22) International Filing Date: 5 December 1997 (05.12.97)

(30) Priority Data:

08/762,216	6 December 1996 (06.12.96)	US
08/984,516	3 December 1997 (03.12.97)	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, HU, 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, 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).

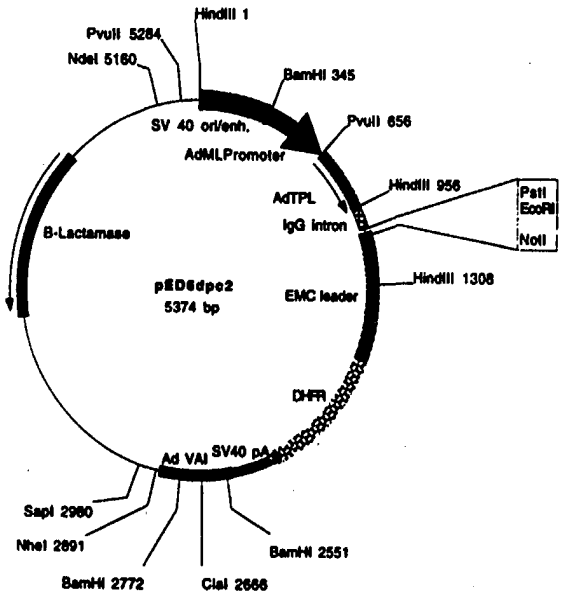
Published

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. 08/762,216, filed December 6, 1996, 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 202 to nucleotide 759;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 391 to nucleotide 759;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone AM795_4 deposited under accession number ATCC 98271;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone AM795_4 deposited under accession number ATCC 98271;

(f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone AM795_4 deposited under accession number ATCC 98271;

(g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone AM795_4 deposited under accession number ATCC 98271;

(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 202 to nucleotide 759; the nucleotide sequence of SEQ ID NO:1 from nucleotide 391 to nucleotide 759; the nucleotide sequence of the full-length protein coding sequence of clone AM795_4 deposited under accession number ATCC 98271; or the nucleotide sequence of the mature protein coding sequence of clone AM795_4 deposited

under accession number ATCC 98271. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone AM795_4 deposited under accession number ATCC 98271. 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 53 to amino acid 186.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:1 or 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:2;

(b) the amino acid sequence of SEQ ID NO:2 from amino acid 53 to amino acid 186;

(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 AM795_4 deposited under accession number ATCC 98271;

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 53 to amino acid 186.

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 19 to nucleotide 262;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 91 to nucleotide 262;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone AT340_1 deposited under accession number ATCC 98271;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone AT340_1 deposited under accession number ATCC 98271;

(f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone AT340_1 deposited under accession number ATCC 98271;

(g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone AT340_1 deposited under accession number ATCC 98271;

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

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

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

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

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

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:5 from nucleotide 19 to nucleotide 262; the nucleotide sequence of SEQ ID NO:5 from nucleotide 91 to nucleotide 262; the nucleotide sequence of the full-length protein coding sequence of clone AT340_1 deposited under accession number ATCC 98271; or the nucleotide sequence of the mature protein coding sequence of clone AT340_1 deposited under accession number ATCC 98271. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone AT340_1 deposited under accession number ATCC 98271. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:6 from amino acid 1 to amino acid 66.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:5 or SEQ ID NO:4.

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

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

(b) the amino acid sequence of SEQ ID NO:6 from amino acid 1 to amino acid 66;

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

(d) the amino acid sequence encoded by the cDNA insert of clone AT340_1 deposited under accession number ATCC 98271;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:6 or the amino acid sequence of SEQ ID NO:6 from amino acid 1 to amino acid 66.

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:7;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 2 to nucleotide 601;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 401 to nucleotide 601;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone BG132_1 deposited under accession number ATCC 98271;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone BG132_1 deposited under accession number ATCC 98271;

(f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone BG132_1 deposited under accession number ATCC 98271;

(g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone BG132_1 deposited under accession number ATCC 98271;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:8;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:8 having biological activity;

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

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

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:7 from nucleotide 2 to nucleotide 601; the nucleotide sequence of SEQ ID NO:7 from nucleotide 401 to nucleotide 601; the nucleotide sequence of the full-length protein coding

sequence of clone BG132_1 deposited under accession number ATCC 98271; or the nucleotide sequence of the mature protein coding sequence of clone BG132_1 deposited under accession number ATCC 98271. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone BG132_1 deposited under accession number ATCC 98271. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:8 from amino acid 119 to amino acid 200.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:7 or SEQ ID NO:9.

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:8;

(b) the amino acid sequence of SEQ ID NO:8 from amino acid 119 to amino acid 200;

(c) fragments of the amino acid sequence of SEQ ID NO:8; and

(d) the amino acid sequence encoded by the cDNA insert of clone BG132_1 deposited under accession number ATCC 98271;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:8 or the amino acid sequence of SEQ ID NO:8 from amino acid 119 to amino acid 200.

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 225 to nucleotide 701;

(c) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone BG219_2 deposited under accession number ATCC 98271;

(d) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone BG219_2 deposited under accession number ATCC 98271;

(e) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone BG219_2 deposited under accession number ATCC 98271;

(f) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone BG219_2 deposited under accession number ATCC 98271;

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

(h) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:11 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:10 from nucleotide 225 to nucleotide 701; the nucleotide sequence of the full-length protein coding sequence of clone BG219_2 deposited under accession number ATCC 98271; or the nucleotide sequence of the mature protein coding sequence of clone BG219_2 deposited under accession number ATCC 98271. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone BG219_2 deposited under accession number ATCC 98271. 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 1 to amino acid 97.

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 1 to amino acid 97;

(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 BG219_2 deposited under accession number ATCC 98271;

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 1 to amino acid 97.

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 2115 to nucleotide 2510;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:12 from nucleotide 1 to nucleotide 324;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone BG366_2 deposited under accession number ATCC 98271;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone BG366_2 deposited under accession number ATCC 98271;

(f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone BG366_2 deposited under accession number ATCC 98271;

(g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone BG366_2 deposited under accession number ATCC 98271;

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

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

(j) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:34;

(k) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:34 having biological activity;

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

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

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

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:12 from nucleotide 2115 to nucleotide 2510; the nucleotide sequence of SEQ ID NO:12 from nucleotide 1 to nucleotide 324; the nucleotide sequence of the full-length protein coding sequence of clone BG366_2 deposited under accession number ATCC 98271; or the nucleotide sequence of the mature protein coding sequence of clone BG366_2 deposited under accession number ATCC 98271. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone BG366_2 deposited under accession number ATCC 98271.

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

(c) the amino acid sequence encoded by the cDNA insert of clone BG366_2 deposited under accession number ATCC 98271;

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

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:34; and

(b) fragments of the amino acid sequence of SEQ ID NO:34;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:34.

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 27 to nucleotide 215;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:14 from nucleotide 27 to nucleotide 181;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone BV172_2 deposited under accession number ATCC 98271;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone BV172_2 deposited under accession number ATCC 98271;

(f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone BV172_2 deposited under accession number ATCC 98271;

(g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone BV172_2 deposited under accession number ATCC 98271;

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

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

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

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

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

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:14 from nucleotide 27 to nucleotide 215; the nucleotide sequence of SEQ ID NO:14 from nucleotide 27 to nucleotide 181; the nucleotide sequence of the full-length protein coding sequence of clone BV172_2 deposited under accession number ATCC 98271; or the nucleotide sequence of the mature protein coding sequence of clone BV172_2 deposited under accession number ATCC 98271. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone BV172_2 deposited under accession number ATCC 98271. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:15 from amino acid 1 to amino acid 51.

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

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

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

(b) the amino acid sequence of SEQ ID NO:15 from amino acid 1 to amino acid 51;

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

(d) the amino acid sequence encoded by the cDNA insert of clone BV172_2 deposited under accession number ATCC 98271;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:15 or the amino acid sequence of SEQ ID NO:15 from amino acid 1 to amino acid 51.

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 338 to nucleotide 409;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:16 from nucleotide 362 to nucleotide 409;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:16 from nucleotide 270 to nucleotide 419;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CC247_10 deposited under accession number ATCC 98271;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CC247_10 deposited under accession number ATCC 98271;

(g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CC247_10 deposited under accession number ATCC 98271;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CC247_10 deposited under accession number ATCC 98271;

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

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:17 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:16 from nucleotide 338 to nucleotide 409; the nucleotide sequence of SEQ ID NO:16 from nucleotide 362 to nucleotide 409; the nucleotide sequence of SEQ ID NO:16 from nucleotide 270 to nucleotide 419; the nucleotide sequence of the full-length protein coding sequence of clone CC247_10 deposited under accession number ATCC 98271; or the nucleotide sequence of the mature protein coding sequence of clone CC247_10 deposited under accession number ATCC 98271. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone CC247_10 deposited under accession number ATCC 98271.

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

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

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

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

(c) the amino acid sequence encoded by the cDNA insert of clone CC247_10 deposited under accession number ATCC 98271;

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

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

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

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:18 from nucleotide 128 to nucleotide 1600;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:18 from nucleotide 281 to nucleotide 1600;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:18 from nucleotide 62 to nucleotide 373;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CI480_9 deposited under accession number ATCC 98271;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CI480_9 deposited under accession number ATCC 98271;

(g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CI480_9 deposited under accession number ATCC 98271;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CI480_9 deposited under accession number ATCC 98271;

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

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:19 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:18 from nucleotide 128 to nucleotide 1600; the nucleotide sequence of SEQ ID NO:18 from nucleotide 281 to nucleotide 1600; the nucleotide sequence of SEQ ID NO:18 from nucleotide 62 to nucleotide 373; the nucleotide sequence of the full-length protein coding sequence of clone CI480_9 deposited under accession number ATCC 98271; or the nucleotide sequence of the mature protein coding sequence of clone CI480_9 deposited under accession number ATCC 98271. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone CI480_9

deposited under accession number ATCC 98271. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:19 from amino acid 1 to amino acid 82.

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

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

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

(b) the amino acid sequence of SEQ ID NO:19 from amino acid 1 to amino acid 82;

(c) fragments of the amino acid sequence of SEQ ID NO:19; and

(d) the amino acid sequence encoded by the cDNA insert of clone CI480_9 deposited under accession number ATCC 98271;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:19 or the amino acid sequence of SEQ ID NO:19 from amino acid 1 to amino acid 82.

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

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

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:20 from nucleotide 383 to nucleotide 3958;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:20 from nucleotide 470 to nucleotide 3958;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:20 from nucleotide 271 to nucleotide 488;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CO722_1 deposited under accession number ATCC 98271;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CO722_1 deposited under accession number ATCC 98271;

(g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CO722_1 deposited under accession number ATCC 98271;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CO722_1 deposited under accession number ATCC 98271;

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

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

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

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

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

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:20 from nucleotide 383 to nucleotide 3958; the nucleotide sequence of SEQ ID NO:20 from nucleotide 470 to nucleotide 3958; the nucleotide sequence of SEQ ID NO:20 from nucleotide 271 to nucleotide 488; the nucleotide sequence of the full-length protein coding sequence of clone CO722_1 deposited under accession number ATCC 98271; or the nucleotide sequence of the mature protein coding sequence of clone CO722_1 deposited under accession number ATCC 98271. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone CO722_1 deposited under accession number ATCC 98271. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:21 from amino acid 1 to amino acid 34.

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

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

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

(b) the amino acid sequence of SEQ ID NO:21 from amino acid 1 to amino acid 34;

(c) fragments of the amino acid sequence of SEQ ID NO:21; and

(d) the amino acid sequence encoded by the cDNA insert of clone CO722_1 deposited under accession number ATCC 98271;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:21 or the amino acid sequence of SEQ ID NO:21 from amino acid 1 to amino acid 34.

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:22;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:22 from nucleotide 914 to nucleotide 2353;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:22 from nucleotide 1793 to nucleotide 2353;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:22 from nucleotide 1037 to nucleotide 1260;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CT748_2 deposited under accession number ATCC 98271;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CT748_2 deposited under accession number ATCC 98271;

(g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CT748_2 deposited under accession number ATCC 98271;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CT748_2 deposited under accession number ATCC 98271;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:23;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:23 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:22 from nucleotide 914 to nucleotide 2353; the nucleotide sequence of SEQ ID NO:22 from nucleotide 1793 to nucleotide 2353; the nucleotide sequence of SEQ ID NO:22 from nucleotide 1037 to nucleotide 1260; the nucleotide sequence of the full-length protein coding sequence of clone CT748_2 deposited under accession number ATCC 98271; or the nucleotide sequence of the mature protein coding sequence of clone CT748_2 deposited under accession number ATCC 98271. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone CT748_2 deposited under accession number ATCC 98271. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:23 from amino acid 22 to amino acid 116.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:22.

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:23;

(b) the amino acid sequence of SEQ ID NO:23 from amino acid 22 to amino acid 116;

(c) fragments of the amino acid sequence of SEQ ID NO:23; and

(d) the amino acid sequence encoded by the cDNA insert of clone CT748_2 deposited under accession number ATCC 98271;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:23 or the amino acid sequence of SEQ ID NO:23 from amino acid 22 to amino acid 116.

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.

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 "AM795_4"

A polynucleotide of the present invention has been identified as clone "AM795_4". AM795_4 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. AM795_4 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "AM795_4 protein").

The nucleotide sequence of the 5' portion of AM795_4 as presently determined is reported in SEQ ID NO:1. What applicants presently believe is the proper reading frame for the coding region is indicated in SEQ ID NO:2. The predicted amino acid sequence of the AM795_4 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:2. Amino acids 51 to 63 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 64, or are a transmembrane domain. Additional nucleotide sequence from the 3' portion of AM795_4, including the polyA tail, is reported in SEQ ID NO:3.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone AM795_4 should be approximately 1900 bp.

The nucleotide sequence disclosed herein for AM795_4 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. AM795_4 demonstrated at least some similarity with sequences identified as H05619 (yl70a10.s1 Homo sapiens cDNA clone 43207 3'), U46493 (Cloning vector pFlp recombinase gene, complete cds), U59486 (Rattus norvegicus GDNF receptor alpha mRNA, complete cds), and W73633 (zd55h01.s1 Soares fetal heart NbHH19W Homo sapiens cDNA). The predicted amino acid sequence disclosed herein for AM795_4 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted AM795_4 protein demonstrated at least some similarity to sequences identified as U59486 (GDNF receptor alpha [Rattus norvegicus]). Based upon sequence similarity, AM795_4 proteins and each similar protein or peptide may share at least some activity.

Clone "AT340_1"

A polynucleotide of the present invention has been identified as clone "AT340_1". AT340_1 was isolated from a human adult blood (lymphocytes and dendritic cells treated

with mixed lymphocyte reaction) 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. AT340_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "AT340_1 protein").

The partial nucleotide sequence of AT340_1, including its 3' end and any identified polyA tail, 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 AT340_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:6. Amino acids 12 to 24 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 25, or are a transmembrane domain. Additional nucleotide sequence from the 5' portion of AT340_1 is reported in SEQ ID NO:4.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone AT340_1 should be approximately 1100 bp.

The nucleotide sequence disclosed herein for AT340_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. AT340_1 demonstrated at least some similarity with sequences identified as AA039343 (zk39g04.s1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone 485238 3'), R68951 (yi43g06.r1 Homo sapiens cDNA clone 142042 5' similar to SP:C35D10.1 CE01190), R77532 (yi76c01.r1 Homo sapiens cDNA), R92619 (yq04a04.r1 Homo sapiens cDNA clone 195918 5' similar to SP:C35D10.1 CE01190), and W60997 (zc99f09.s1 Pancreatic Islet Homo sapiens cDNA clone 339305 3'). The predicted amino acid sequence disclosed herein for AT340_1 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted AT340_1 protein demonstrated at least some similarity to sequences identified as U21324 (similar to S. cerevisiae hypothetical protein YKL166 [Caenorhabditis elegans]). Based upon sequence similarity, AT340_1 proteins and each similar protein or peptide may share at least some activity.

Clone "BG132_1"

A polynucleotide of the present invention has been identified as clone "BG132_1". BG132_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. BG132_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "BG132_1 protein").

The nucleotide sequence of the 5' portion of BG132_1 as presently determined is reported in SEQ ID NO:7. What applicants presently believe is the proper reading frame for the coding region is indicated in SEQ ID NO:8. The predicted amino acid sequence of the BG132_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:8. Amino acids 119 to 133 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 134, or are a transmembrane domain. Additional nucleotide sequence from the 3' portion of BG132_1, including the polyA tail, is reported in SEQ ID NO:9.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone BG132_1 should be approximately 2000 bp.

The nucleotide sequence disclosed herein for BG132_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. BG132_1 demonstrated at least some similarity with sequences identified as AA078587 (7P05H12 Chromosome 7 Placental cDNA Library Homo sapiens cDNA clone 7P05H12), H14301 (ym63c04.r1 Homo sapiens cDNA clone 163590 5' similar to gb:U03642_cds1 PROBABLE G PROTEIN-COUPLED RECEPTOR APJ (HUMAN)), L09249 (putative G-protein coupled receptor, rhodopsin family), S79811 (adrenomedullin receptor [rats, lung, mRNA]), T36034 (rchd523 gene differentially expressed in cardiovascular disease), U58828 (Human IL8-related receptor (DRY12) mRNA, complete cds), and Y08162 (H.sapiens mRNA for heptahelix receptor). The predicted amino acid sequence disclosed herein for BG132_1 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted BG132_1 protein demonstrated at least some similarity to sequences identified as L06109 (G protein-coupled receptor [Gallus gallus]), L34339 (galanin receptor [Homo sapiens]), U30290 (galanin receptor GALR1 [Rattus norvegicus]), U58828 (IL8-related receptor [Homo sapiens]), W03739 (rchd523 gene product (G protein-coupled receptor)), X98510 (G protein-coupled receptor [Homo sapiens]), and Y08162 (heptahelix receptor [Homo sapiens]). Based upon sequence similarity, BG132_1 proteins and each similar protein or peptide may share at least some activity.

Clone "BG219_2"

A polynucleotide of the present invention has been identified as clone "BG219_2". BG219_2 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. BG219_2 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "BG219_2 protein").

The nucleotide sequence of BG219_2 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 BG219_2 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:11.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone BG219_2 should be approximately 700 bp.

The nucleotide sequence disclosed herein for BG219_2 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. BG219_2 demonstrated at least some similarity with sequences identified as AA210695 (zr88b05.s1 Soares NbHTGBC Homo sapiens cDNA clone 682737 3'), C01459 (HUMGS0008450, Human Gene Signature, 3'-directed cDNA sequence), N22628 (EST49p115 Homo sapiens cDNA clone 49p115), and T26211 (Human gene signature HUMGS08450). Based upon sequence similarity, BG219_2 proteins and each similar protein or peptide may share at least some activity.

Clone "BG366_2"

A polynucleotide of the present invention has been identified as clone "BG366_2". BG366_2 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. BG366_2 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "BG366_2 protein").

The nucleotide sequence of BG366_2 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 BG366_2 protein corresponding to the foregoing nucleotide

sequence is reported in SEQ ID NO:13. The amino acid sequence of another protein that could be encoded by BG366_2 is reported in SEQ ID NO:34.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone BG366_2 should be approximately 3000 bp.

The nucleotide sequence disclosed herein for BG366_2 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. BG366_2 demonstrated at least some similarity with sequences identified as N39453 (yy49h03.s1 Homo sapiens cDNA clone 276917 3'). Based upon sequence similarity, BG366_2 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts a potential transmembrane domain within the BG366_2 protein sequence centered around amino acid 92 of SEQ ID NO:13.

Clone "BV172_2"

A polynucleotide of the present invention has been identified as clone "BV172_2". BV172_2 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. BV172_2 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "BV172_2 protein").

The nucleotide sequence of BV172_2 as presently determined is reported in SEQ ID NO:14. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the BV172_2 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:15.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone BV172_2 should be approximately 1700 bp.

The nucleotide sequence disclosed herein for BV172_2 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 TopPredII computer program predicts a potential transmembrane domain within the BV172_2 protein sequence centered around amino acid 19 of SEQ ID NO:15. The nucleotide sequence of BV172_2 indicates that it may contain one or more of the following types of repetitive elements: an element similar to chicken CR1, human L1, Mer33.

Clone "CC247_10"

A polynucleotide of the present invention has been identified as clone "CC247_10". CC247_10 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. CC247_10 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "CC247_10 protein").

The nucleotide sequence of CC247_10 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 CC247_10 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:17. Amino acids 1 to 8 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 9, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone CC247_10 should be approximately 550 bp.

The nucleotide sequence disclosed herein for CC247_10 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. CC247_10 demonstrated at least some similarity with sequences identified as AA291226 (zs47d03.r1 NCI_CGAP_GCB1 Homo sapiens cDNA clone 700613 5'), T05738 (EST03627 Homo sapiens cDNA clone HFBDF64), W51195 (ma14b04.r1 Life Tech mouse brain Mus musculus cDNA clone 304495 5'), and W93640 (zd95d09.s1 Soares fetal heart NbHH19W Homo sapiens cDNA clone 357233 3'). The predicted amino acid sequence disclosed herein for CC247_10 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted CC247_10 protein demonstrated at least some similarity to sequences identified as M62424 (thrombin receptor [Homo sapiens]). The predicted CC247_10 protein is highly hydrophobic. Based upon sequence similarity, CC247_10 proteins and each similar protein or peptide may share at least some activity.

Clone "CI480_9"

A polynucleotide of the present invention has been identified as clone "CI480_9". CI480_9 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. CI480_9 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "CI480_9 protein").

The nucleotide sequence of CI480_9 as presently determined is reported in SEQ ID NO:18. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the CI480_9 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:19. Amino acids 39 to 51 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 52, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone CI480_9 should be approximately 1940 bp.

The nucleotide sequence disclosed herein for CI480_9 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. CI480_9 demonstrated at least some similarity with sequences identified as N99342 (IMAGE:20093 Homo sapiens cDNA clone 20093), R89725 (ym99d09.r1 Homo sapiens cDNA clone 167057 5'), and U60644 (Human HU-K4 mRNA, complete cds). The predicted amino acid sequence disclosed herein for CI480_9 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted CI480_9 protein demonstrated at least some similarity to sequences identified as U60644 (HU-K4 [Homo sapiens]). Based upon sequence similarity, CI480_9 proteins and each similar protein or peptide may share at least some activity.

Clone "CO722_1"

A polynucleotide of the present invention has been identified as clone "CO722_1". CO722_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. CO722_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "CO722_1 protein").

The nucleotide sequence of CO722_1 as presently determined is reported in SEQ ID NO:20. What applicants presently believe to be the proper reading frame and the

predicted amino acid sequence of the CO722_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:21. Amino acids 17 to 29 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 30, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone CO722_1 should be approximately 6800 bp.

The nucleotide sequence disclosed herein for CO722_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. CO722_1 demonstrated at least some similarity with sequences identified as AA186616 (zp71a08.s1 Stratagene endothelial cell 937223 Homo sapiens cDNA clone 625622 3' similar to contains Alu repetitive element), H10376 (ym08a03.s1 Homo sapiens cDNA clone 47067 3'), N86013 (J5997F Fetal heart, Lambda ZAP Express Homo sapiens cDNA), U55258 (Human hBRAVO/Nr-CAM precursor (hBRAVO/Nr-CAM) gene, complete cds), W19770 (zb39d01.r1 Soares parathyroid tumor NbHPA Homo sapiens), W31608 (zb91d09.r1 Soares parathyroid tumor NbHPA Homo sapiens cDNA clone), and X58482 (Chicken mRNA for neuronal transmembrane protein Nr-CAM, ng-CAM related). The predicted amino acid sequence disclosed herein for CO722_1 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted CO722_1 protein demonstrated at least some similarity to sequences identified as AB002341 (KIAA0343 [Homo sapiens]) and X58482 (Nr-CAM protein [Gallus gallus]). Based upon sequence similarity, CO722_1 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts two potential transmembrane domains within the CO722_1 protein sequence, centered around amino acids 610 and 1070 of SEQ ID NO:21.

Clone "CT748_2"

A polynucleotide of the present invention has been identified as clone "CT748_2". CT748_2 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. CT748_2 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "CT748_2 protein").

The nucleotide sequence of CT748_2 as presently determined is reported in SEQ ID NO:22. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the CT748_2 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:23. Amino acids 281 to 293 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 294, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone CT748_2 should be approximately 5500 bp.

The nucleotide sequence disclosed herein for CT748_2 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. CT748_2 demonstrated at least some similarity with sequences identified as T48063 (yb24f03.s1 Homo sapiens cDNA clone 72125 3') and X54175 (Human specific Alu element (HS C4N2) DNA). The predicted amino acid sequence disclosed herein for CT748_2 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted CT748_2 protein demonstrated at least some similarity to sequences identified as Z36714 (cyclin F [Homo sapiens]). Based upon sequence similarity, CT748_2 proteins and each similar protein or peptide may share at least some activity. The nucleotide sequence of CT748_2 indicates that it may contain an Alu repetitive element.

Deposit of Clones

Clones AM795_4, AT340_1, BG132_1, BG219_2, BG366_2, BV172_2, CC247_10, CI480_9, CO722_1 and CT748_2 were deposited on December 5, 1996 with the American Type Culture Collection as an original deposit under the Budapest Treaty and were given the accession number ATCC 98271, 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
AM795_4	SEQ ID NO:24
AT340_1	SEQ ID NO:25
BG132_1	SEQ ID NO:26
BG219_2	SEQ ID NO:27
BG366_2	SEQ ID NO:28
BV172_2	SEQ ID NO:29
CC247_10	SEQ ID NO:30
CI480_9	SEQ ID NO:31
CO722_1	SEQ ID NO:32
CT748_2	SEQ ID NO:33

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 cDNA sequences disclosed herein. "Corresponding genes" are the regions of the genome that are transcribed to produce the mRNAs from which the cDNA 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.

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.

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, $T_m(°C) = 2(\#\text{ of A + T bases}) + 4(\#\text{ of G + C bases})$. 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: 34

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

(ii) MOLECULE TYPE: cDNA

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

GGGAGTAATT CCTTTTCAAA AGCGGGCATG ACTTNTGCGC TAAGATTGTC AGTTTCCAAA 60

AANGAGGAGG ATTTGATATT CACCTGGCCC GCGGTGATGC CTTTGAGGGT GGCCGCGTCC 120

ATCTGGTCAG AAAAGACAAT CTTTTTGTTG TCAAGCTTGA GGTGTGGCAG GCTTGAGATC 180

TGGCCANACA CTTGAGTGAC AATGACATCC ACTTTGCNTT TNTCTCCACA GGTGTCCACT 240

CCCAGGTCCA ACTGCAGANT TNGAATTCGG CCAAAGAGGC NNACATCGGG TGGACTARCT 300

GGGATCTCCG CATTGGATTT GGGGCTGATT ACCACTGCTT GCCTATTATT ATTGTTGTTT 360

TTACTACTAT TATTTTTTTT TACCCAAGGG AGAAAGACAA AAAAACGGTG GGATTTATTT 420

AACATGATCT TGGCAAACGT CTTCTGCCTC TTCTTCTTTC TAGACGAGAC CCTCCGCTCT 480

TTGGCCAGCC CTTCCTCCCT GCAGGGCCCC GAGCTCCACG GCTGGCGCCC CCCAGTGGAC 540

TGTGTCCGGG CCAATGAGCT GTGTGCCGCC GAATCCAACT GCAGCTCTCG CTACCGCACT 600

CTGCGGCAGT GCCTGGCAGG CCGCGACCGC AACACCATGC TGGCCAACAA GGAGTGCCAG 660

GCGGCCTTGG AGGTCTTGCA GGAGAGCCCG CTGTACGACT GCCGCTGCAA GCGGGGCATG 720

AAGAAGGAGC TGCAGTGTCT GCAGATCTAC TGGAGCATC 759

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 186 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

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

Met Thr Ser Thr Leu Xaa Xaa Ser Pro Gln Val Ser Thr Pro Arg Ser
1 5 10 15

Asn Cys Arg Xaa Xaa Ile Arg Pro Lys Arg Xaa Thr Ser Gly Gly Leu
20 25 30

Xaa Gly Ile Ser Ala Leu Asp Leu Gly Leu Ile Thr Thr Ala Cys Leu
35 40 45

Leu Leu Leu Leu Phe Leu Leu Leu Leu Phe Phe Phe Thr Gln Gly Arg
50 55 60

Lys Thr Lys Lys Arg Trp Asp Leu Phe Asn Met Ile Leu Ala Asn Val
65 70 75 80

Phe Cys Leu Phe Phe Phe Leu Asp Glu Thr Leu Arg Ser Leu Ala Ser
85 90 95

Pro Ser Ser Leu Gln Gly Pro Glu Leu His Gly Trp Arg Pro Pro Val
100 105 110

Asp Cys Val Arg Ala Asn Glu Leu Cys Ala Ala Glu Ser Asn Cys Ser
115 120 125

Ser Arg Tyr Arg Thr Leu Arg Gln Cys Leu Ala Gly Arg Asp Arg Asn
130 135 140

Thr Met Leu Ala Asn Lys Glu Cys Gln Ala Ala Leu Glu Val Leu Gln
145 150 155 160

Glu Ser Pro Leu Tyr Asp Cys Arg Cys Lys Arg Gly Met Lys Lys Glu
165 170 175

Leu Gln Cys Leu Gln Ile Tyr Trp Ser Ile
180 185

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 300 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

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

ATTTCTGTCC AGGAGCAGGG GCTGAAGCCC AACAANTCCA AAGAGTTAAG CATGTGNTTC 60

ACAGAGCTCA NGACAAATAT CATCCCAGGG AGTAACAAGG TGATCAAACN TAACTCAGGC 120

CCCAGCAGAG CCAGACNGTC GGCTGCTTTG ACCGTGNTGT CTGTCCTGAT GCTGAAACAG 180

CCTTTGTAGG CTGTGGGAAC CGAGTCAGAA GATTTTTGAA AGNTACGCAG ACAAGAACAG 240

CCGCCTGANG AAATGGAAAC ACACACAGAC ACACACACAC CTTGCAAAAA AAAAAAAAAA 300

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 355 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

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

CTCTCATTTT TCCTTTCGAA GTGATATCCA CTCGAACGAG ATCAAATCTG TAAGCTGGAG 60

ATACAACTTC CACTACATAA GATCCAGAAG GTATATCATG AACCACAAAA CTCCCATCTG 120

TCTTAAGGAA ACNGACGTGC TCTTCTCCGT NTACCAGCAC TCGGGCCGCC GAGATCCAGT 180

CCTGAGGCTT CACCCNTGGA ACAACTGCAC GCCCCTCAAT CTTGAAGNGA TCTCCTATGC 240

CGACCCCACT CCCTCCCGAT CCCTCAGCAG CAGCCCCGGG CACCTCCGAG TTCTGGACAT 300

CCCCCGATAG CAGCAGCAGC AGCAGGACGG GAAAGAAGCC CCACAGAGCG GCCGC 355

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 587 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

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

GGCTGGACAG ACTTTCTAAT GAACCCAATG GTTATGATGA TGGTTCTTCC TTTATTGATA 60

TTTGTGCTTC TGCCTAAAGT GGTCAACACA AGTGATCCTG ACATGAGACG GGAAATGGAG 120

CAGTCAATGA ATATGCTGAA TTCCAACCAT GAGTTGCCTG ATGTTTCTGA GTTCATGACA 180

AGACTCTTCT CTTCAAAATC ATCTGGCAAA TCTAGCAGCG GCAGCAGTAA AACAGGCAAA 240

AGTGGGGCTG GCAAAAGGAG GTAGTCAGGC CGTCCARAGC TGGCATTTGC ACAAACACGG 300

CAACACTGGG TGGCATCCAA GTCTTGGAAA ACCGTGTGAA GCAACTACTA TAAACTTGAG 360

TCATCCCGAC GTTGATCTCT TACAACTGTG TATGTTAACT TTTTAGCACA TGTTTTGTAC 420

TTGGTACACG AGAAAACCCA GCTTTCATCT TTTGTCTGTA TGAGGTCAAT ATTGATGTCA 480

CTGAATTAAT TACAGTGTCC TATAGAAAAT GCCATTAATA AATTATATGA ACTACTATAC 540

ATTATGTATA TTAATTAAAA CATCTTAATC CAGAAAAAAA AAAAAAA 587

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 81 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

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

```
Met Asn Pro Met Val Met Met Met Val Leu Pro Leu Leu Ile Phe Val
1               5                  10                  15

Leu Leu Pro Lys Val Val Asn Thr Ser Asp Pro Asp Met Arg Arg Glu
            20                  25                  30

Met Glu Gln Ser Met Asn Met Leu Asn Ser Asn His Glu Leu Pro Asp
        35                  40                  45

Val Ser Glu Phe Met Thr Arg Leu Phe Ser Ser Lys Ser Ser Gly Lys
    50                  55                  60

Ser Ser Ser Gly Ser Ser Lys Thr Gly Lys Ser Gly Ala Gly Lys Arg
65                  70                  75                  80

Arg
```

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 601 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

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

```
GATGTACCTA GGCACCGCGC AGCCTGCGGC CCCCAACACC ACCTCCCCCG AGCTCAACCT      60

GTCCCACCCG CTCCTGGGCA CCGCCCTGGC CAATGGGACA GGTGAGCTCT CGGAGCACCA      120

GCAGTACGTG ATCGGCCTGT TCCTCTCGTG CCTCTACACC ATCTTCCTCT TCCCCATCGG      180

CTTTGTGGGC AACATCCTGA TCCTGGTGGT GAACATCAGC TTCCGCGAGA AGATGACCAT      240

CCCCGACCTG TACTTCATCA ACCTGGCGGT GGCGGACCTC ATCCTGGTGG CCGACTCCCT      300

CATTGAGGTG TTCAACCTGC ACGAGCGGTA CTACGACATC GCCGTCCTGT GCACCTTCAT      360
```

GTCGCTCTTC CTGCAGGTCA ACATGTACAG CAGCGTCTTC TTCCTCACCT GGATGAGCTT 420

CGACCGCTAC ATCGCCCTGG CCAGGGCCAT GCGCTGCAGC CTGTTCCGCA CCAAGCACCA 480

CGCCCGGCTG AGCTGTGGCC TCATCTGGAT GGCATCCGTG TCAGCCACGC TGGTGCCCTT 540

CACCGCCGTG CACCTGCAGC ACACCGACGA GGCCTGCTTC TGTTTCGCGG ATGTCCGGGA 600

G 601

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 200 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

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

Met Tyr Leu Gly Thr Ala Gln Pro Ala Ala Pro Asn Thr Thr Ser Pro
1 5 10 15

Glu Leu Asn Leu Ser His Pro Leu Leu Gly Thr Ala Leu Ala Asn Gly
20 25 30

Thr Gly Glu Leu Ser Glu His Gln Gln Tyr Val Ile Gly Leu Phe Leu
35 40 45

Ser Cys Leu Tyr Thr Ile Phe Leu Phe Pro Ile Gly Phe Val Gly Asn
50 55 60

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

Pro Asp Leu Tyr Phe Ile Asn Leu Ala Val Ala Asp Leu Ile Leu Val
85 90 95

Ala Asp Ser Leu Ile Glu Val Phe Asn Leu His Glu Arg Tyr Tyr Asp
100 105 110

Ile Ala Val Leu Cys Thr Phe Met Ser Leu Phe Leu Gln Val Asn Met
115 120 125

Tyr Ser Ser Val Phe Phe Leu Thr Trp Met Ser Phe Asp Arg Tyr Ile
130 135 140

Ala Leu Ala Arg Ala Met Arg Cys Ser Leu Phe Arg Thr Lys His His
145 150 155 160

Ala Arg Leu Ser Cys Gly Leu Ile Trp Met Ala Ser Val Ser Ala Thr
165 170 175

Leu Val Pro Phe Thr Ala Val His Leu Gln His Thr Asp Glu Ala Cys
180 185 190

Phe Cys Phe Ala Asp Val Arg Glu
195 200

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 419 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

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

ACGCCTGAGC GTCCTCCATC TTCCAGGATG GCAGCAATGG CGCTGTGCGG CCTCACCAGG 60

CCCACGAGGA GCAGCAGCGN TCGGCCCGGA GCAGCAGGAA GGCCCTTTTG TGGAGCGCCC 120

GCCGTCTGCT CCGGGGTGGT TCAGTCACTG CTTGTTGACA TCAACATGGC AATTGCANTC 180

ATGTGGACTG GGACCGTGCG AGCTGCCGTG TGGGTTAGTC GGGTGCCAGG ACAATGAAAT 240

ACTCCAGCAC GTGTGGCTGA CGAATTTGTT TTTACAGAAA TAACAGCTGG GGACAACTGC 300

GGTGATGATG TAAAAACCTT CCCATAAAAT GTAAGAAAAG CTGATGAGGC TGGTGACGTT 360

CAGCCTTTGT CAATAAACCT GTCATGTGCG GAAAAAAAAA AAAAAAAAAA AAAAAAAAA 419

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 714 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

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

AGCCATTGGG ACAGGAAATG CCAAACAACA CCCAGATAAG GTTGCTGAAG CCATAATTGA 60

TGCCATTGAA GACTTTGTCC AGAAAGGATC AGCCCAGTCT GTGAAAAAAG TTAAAGTTGT 120

TATCTTTCTG CCTCAAGTAC TGGATGTGTT TTATGCCAAC ATGAAGAAAA GAGAAGGGAC 180

TCAGCTTTCT TCCCAACAGY CTARTSWTSY YTWWMYTTKY AKCATTTTTG GGCTTTTCAA 240

AGCAATCTCC CCAAAAAAAG AATCATTTGG TTTTGGAAAA GAAAACAGAA TCAGCAACTT 300

TTCGGGTGTG TGGTGAAAAT GTCACGTGTG TGGAATACGC TATCTCCTGG CTACAAGACC 360

TGATTGAAAA AGAACAGTGT CCTTACACCA GTGAAGATGA GTGCATCAAA GACTTTGATG 420

AAAAGGAGTA TCAGGAGTTG AATGAGCTGC AGAAGAAGTT AAATATTAAC ATTTCCCTGG 480

ACCATAAGAG ACCTTTGATT AAGGTTTTGG GAATTAGCAG AGATGTGATG CAGGCTAGAG 540

ATGAAATTGA GGCGATGATC AAGAGAGTTC GATTGGCCAA AGAACAGGAA TCCCGGGCAG 600

ATTGTATCAG TGAGTTTATA GAATGGCAGT ATAATGACAA TAACANTTNT CATTGTTTTA 660

ACAAAATGAC CAATCTGAAA TTAGAGGATG CAAGGAGAGA AAAAAAAAAA AAAA 714

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 159 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

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

Phe Leu Gly Phe Ser Lys Gln Ser Pro Gln Lys Lys Asn His Leu Val
1 5 10 15

Leu Glu Lys Lys Thr Glu Ser Ala Thr Phe Arg Val Cys Gly Glu Asn
20 25 30

Val Thr Cys Val Glu Tyr Ala Ile Ser Trp Leu Gln Asp Leu Ile Glu
35 40 45

Lys Glu Gln Cys Pro Tyr Thr Ser Glu Asp Glu Cys Ile Lys Asp Phe
50 55 60

Asp Glu Lys Glu Tyr Gln Glu Leu Asn Glu Leu Gln Lys Lys Leu Asn
65 70 75 80

Ile Asn Ile Ser Leu Asp His Lys Arg Pro Leu Ile Lys Val Leu Gly
85 90 95

Ile Ser Arg Asp Val Met Gln Ala Arg Asp Glu Ile Glu Ala Met Ile
100 105 110

Lys Arg Val Arg Leu Ala Lys Glu Gln Glu Ser Arg Ala Asp Cys Ile
115 120 125

Ser Glu Phe Ile Glu Trp Gln Tyr Asn Asp Asn Asn Xaa Xaa His Cys
130 135 140

Phe Asn Lys Met Thr Asn Leu Lys Leu Glu Asp Ala Arg Arg Glu
145 150 155

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2681 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

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

CATAACTTCC AATCTTCTTG GCCTGGCACT AAGGACCTAC TACAGACTGG CACTGAAATA 60

TCTTCCCTAC CTTATCATTC CTTGTGCTTC TCCACAAGCC CACTCTTCCC CTTCATCATA 120

CATGTGCCGA CCTTTCCTGT CTCTTTTACT TTGCAGCACC AAATGCTTTC TACTTTGTGG 180

TCTAGGAGGA ACACATGTCA CTTTTGTAAG CTGCTCGAAA GCAGGGGCCA CACCTTCATC 240

CTTGTTTTCC ACACAACACC AAGCACTTAG TAGACACCCA ATAAATCATT GCTGAATGAA 300

TGTATTCAGC CTGGAATTGC ACTAGGATTT TTTGGCCAAC ACATTGTATT CTTAACTGAT 360

ACCAGACTTC CAATCAAATA AAATCCTTAA GCCTTTTTCA TAGTCTTTAA TTAAACTACT 420

TCTCTTCCAT TATTTCCCTT TGCCTACTTT TGAACTGATA TTCAGAACTT TTCTGTTAAT 480

GTTTAATTTT CATCCATTAT TCTTGTCTGT ACAGATCTTT TTGATTTTTG ACTCTCTTAT 540

CTAGTTTTTT TTTTTTTTTT TTGGCTTCCC GTTGTGTTAT CCACAGGCAC AATGGGTATA 600

TTGACTATCG TTCATTGAAG TTGTTGATAC AAATGTTGAA CAGGAGAAAA ACCAGTAGCT 660

TGCCAACTTG GCACCTCATT CTCTAGCTTG ACAGTAATCT CTTATCTAGT AGTTTTTAGA 720

TATGGTTATT TAATCAAATA TCACTAGCTT CTAATTTTGT TATAATTCAT TCATGTATAC 780

AGACAATGAA GAATCATCTT TCTTCATTCA ACACATTTAT TGGGAACCAA TCATTGTCCA 840

AGGCAATATG CTAGGCATTG TGTAATAGAA AGCGATTAAG CTTTTACCCC TGTCCTCTTT 900

GGCCTGCAAA GAGCAGGTAA GATTCATGCA CAGACAAATG TAGTACCAGG TAAATAGTGA 960

GGCCAAACAC AAGAGGCAGG TCAAAGGGCT TTGGATTCAG AGGATGGAGA AGCTCTTTCT 1020

AGGCTGGTGA AATATTTGCA GAATGAGTAG GCTCATAAGT GGAGGTGAGG TACAGGCTAA 1080

GCTGCTGTAA GAAAGAGTCC AAAATGCAAA GGATTTACCA CAATGCAGTT AATTTTTCTC 1140

ACAACGGTCG AGGTAGGCAG GTGGTCCAAG TCCAGTTAAT CAGCCCTCCT CAACACAAGG 1200

CTTCCCCTTG TAAGCTCAGG ATGACCGCTC CAGTTCTCAT CATCTCCCAG GCAGAAAAAC 1260

GGGAGAAAGA GAACTAGACG TACCACCCAG TCATTTTAAG GACATAGCCA GGCCAGGTGC 1320

AGTGACTCAC ATCTGTAATC CAAGCACTTT GGGAGGTCGG AGGCTGGAGG ATAGCTTGAA 1380

GCCAGGAGTT TGAGACCAGC TTGGATAAGA AAGCAAGACC TTGTCTCTAC AAAAAATTTA 1440

AAAATTAACC ATGTGTGGTG GTGTGCACCT ATAGTCCCAG CTACTCAGGA GGCTGAGGTG 1500

GAAGGATCAC TTGAGCACAG GAGTTGGAGG ATACAGTGAG CCATGATCAC ACCACCTCAC 1560

TCCAGCCTGG GCAACATAGT GAGACCCCAT CTCTTTAAAA AACAAAACAA AAAAACATCA 1620

CCAGGAAGTT GCTTACATGT CTTCTATTCC TATAACATCG GCCTGAGCTT AGTCACATGG 1680

CCACACCCAG CAGAAAGAAC TCTAGGAAAT ATGGTCTTAT GCTAGGTAAC CCCAAACCCA 1740

GCTAAAACTG TTGCTTTTGA AGAAGGGTGA AACAGACAAT GTGGAGGAGA ATTACCAGTC 1800

TGCCACAAAG AGAAAGAATT CTATGTGAGG AAAACGCTCT AGAAGAGGGA GCTGATTAAT 1860

TAGTTATATC TCAGCCGAGA GGATGTTATG AGAAACACAG ATTTGGGTTT AATAGAAAAA 1920

CTTAGAACCC TTCAAGAAGG AATGGGCGGA CTTGAGAGTG TCCTCCGCTT CTGAAAATGA 1980

TCAAGAAAAA TGGTAAGAAT GTACACCTCA AGGGTGGATG GTAAAGGTCA GCCTTTAAGC 2040

CTCAATGGCG GTGTGGTTGT AGGTGGGGTG GGGGTGGAGG CCCCGAAGAC AGGCAGACTA 2100

CCACAGTAGC CCCGATGGAA GCAGTGAGAA TGAACTGGAA GGAGCGGCTG TGGGAGCGAC 2160

AACGTGATGA GAATAAACCC GGCTTGGCTC TGCCCTGTGC ACACACAGGT GAGCTGTGTG 2220

CTCCCGGATG TGTCAGCTGG TATATGCGTC TGTCAGAAGG CAGCTGGGGA GCACTGTTAG 2280

CTCAGAGACT AAGAGGCAGA CCCAGGAAGC CCTTCTTTGC ATTGGTCAGG GTTTGCTGCA 2340

TTTTCCCAAG CCCAGGAAAT GGAACCCAGT TCTTCTTCTT TCTGTGTAAA ATAATAAGCA 2400

TCACAATAGG CTGTGCTCAT GAGAACGCAT TTTGTTTCCA CAGAAATGTT TTTTCTCACT 2460

CTGTCCTGAT TTTGATTTCT GTTAAACTCA GTAAACACAT TACCAAATTT TAAAATAAGG 2520

TGACTTGTTT TCCCCAACTC ACAGTTCACC AAAGGTATTT CATCTGTTTG TTCTGAAAAT 2580

GCAGCTGCTG TCTAGATTTA TGTGTGCTCT GACAAGAAAT GTTTTGTGTA ACAATAAAAA 2640

TCATTTCCTT TGATGNAAAA AAAAAAAAAA AAAAAAAAAA A 2681

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 132 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

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

Met Glu Ala Val Arg Met Asn Trp Lys Glu Arg Leu Trp Glu Arg Gln
1 5 10 15

Arg Asp Glu Asn Lys Pro Gly Leu Ala Leu Pro Cys Ala His Thr Gly
20 25 30

Glu Leu Cys Ala Pro Gly Cys Val Ser Trp Tyr Met Arg Leu Ser Glu
35 40 45

Gly Ser Trp Gly Ala Leu Leu Ala Gln Arg Leu Arg Gly Arg Pro Arg
50 55 60

Lys Pro Phe Phe Ala Leu Val Arg Val Cys Cys Ile Phe Pro Ser Pro
65 70 75 80

Gly Asn Gly Thr Gln Phe Phe Phe Phe Leu Cys Lys Ile Ile Ser Ile
85 90 95

Thr Ile Gly Cys Ala His Glu Asn Ala Phe Cys Phe His Arg Asn Val
100 105 110

Phe Ser His Ser Val Leu Ile Leu Ile Ser Val Lys Leu Ser Lys His
115 120 125

Ile Thr Lys Phe
130

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1585 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

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

CTAAGCTATT TGATTCTAGG TCTAGAATGT TATCTCTTAT TAGAGGATAT GTTAATTTTC 60

CTGCATTTTA TTCATTTATT AACTTAACAT CTCTGATTGC CTACCATGTG TCAGGCTCTG 120

TACTAAGGAT TGAGGACCCA AAGATGAACA AAACATGGGG CCTAATTCAA AGATTTCACA 180

ATCTGGAGAG AAAGTCAGCC ACATACAAAA AATTATAAGG TAGAATGTGC TATAAAAAAT 240

GATTTAGGTA CAGTGGAAGT TTAGAAAGGC TTCACTAAAG ATGTCGTATT TGAATTGGGT 300

CTGAGCATGA ATTAGTCTTC AGGTGAGGGA GGGGGTTAAA GAAAACTCTA ATGAAGGAAC 360

TCTAGTATGT GTGAAGGCAA GGAGGCTGGT GTGTTTAGGT TGCAGCAAAC CAAGTACAAC 420

CAGGCTTGGA CTTGAGTGAC TGGAAAGACA GGAAGATGCC ATGCTGAGAA AAACTGCCCA 480

TGCCAAGCTG AGAAATGTTC AGCAGAAACA TAAGGTGAGC TGCATATGTC ATTTAAAATG 540

TTCTAGGAGC CACATTTTTA AAAAATCAAA ATTAACAAGT CAAAAAATAA AAAGCAATGG 600

GGGGAGATTA AATGCATATT ACTAAGTGAA AGAAGCCAAT CCAAAAAGGC TACATACCTG 660

TATGATTCCA ACTATATGAC ATTCTGAAAA AGGCCAAACT ATAGAGACAG TAAATGATCA 720

GTGGTTGCCA GGAGTTAGGG AGGAGGGAGG GATGAACAGG CAGAGCACAG AAGATTTTTA 780

GAGCAGTAGA ACTATTTTGT ATGATATTAT AATATTGTAG ATACATATCA TTATAAATTT 840

GTCCAAACCC ATAGAATGTA CACCAAGAGT GAACTCTAAT GTAAACTGTG GACTTTGGGT 900

GATAATGATG TGTCAGTGTA GATTGATCAG TTGTAACAAA TGTACCATTC TGGTGGGGAA 960

TGTTGATAAT GGAGTAGGCT ATGCATGTGT GGGGCAGTGG GTATATGGGA AATATCTATA 1020

CCTTCTGGTC AGTTTTGCTG TGAACTTGAT CTAAAAAATA GCCTACTAAG AAACACAAGT 1080

CAAATTAATT TTAATAATAC ATTTTATTTA ACCCAATTTA TCAGAAATAC TAATATTTTA 1140

ACATGTAATT GATATAAAAG TTATTAACTA GATATTTTAC TTTTTTTGGT ACTGAGTCTT 1200

TGAAATCTGG TCTGTATTTT ACATTTACAG TACATCTCAA TTCACATTAG CCACATTTCA 1260

GATACTCAGT ACATACATGA GTACCTATGG CTAGTGGCTG CTGTGTTGGA CAGGGCAGGT 1320

CTTGAAACCT GGACTTGCCT GACTCAGAAG CCTCAATTCT CAGCCACAGT GATATCCTGC 1380

TCCCTAAGTA CTATAATGAT AAACACAAGA GGAGAGGAGC TTTCAGATGA TCATCTAATC 1440

CCATGACGTT AGCTGTTGCT CTCCACACTG CCCGGTGGCT CCAGTCTGAA GCATCTAGGC 1500

AGTGCTGTCC AACAGAAATA CAATGAGAGC CAATACGTGA TAAGTGTCCT ATGGGCCACA 1560

TTGAAACAGT AAAAAAAAAA AAAAA 1585

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 63 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

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

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

Phe Ile Asn Leu Thr Ser Leu Ile Ala Tyr His Val Ser Gly Ser Val
20 25 30

Leu Arg Ile Glu Asp Pro Lys Met Asn Lys Thr Trp Gly Leu Ile Gln
35 40 45

Arg Phe His Asn Leu Glu Arg Lys Ser Ala Thr Tyr Lys Lys Leu
50 55 60

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 625 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

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

GATCCCACCA GTTCTGCCTG GCTTCCTCCA TCCCCAGAGG CACTAAAAGC AGTATTTTAA 60

GGTTGGTGTC TTACTCCCTG GAAGCCTGAA ATGGGTGGAA TAGCGGTAAG GCTTGAGTAA 120

AACTAGGGGA CAGAGGTTCT TATTTGTCGA TTTTATTTTA TAATTTGACC ACAGCATCTG 180

AACTCCCTCT CTCCCTGGAA TAAGTATTTT TCCCACATTT TTGGATATAT GTATGGTAGA 240

CAATTTTTTT TTAAGACACA GAGATAAATG TTTTCCTGCT TTGGTTACCT TTCCTTTCCC 300

CTTTAAAAGG AATTAGCTAT AGAACTGCTT TGTAAAGATG CTTCTTGATA TTTTACTTTT 360

GTTCCTTTTC CCTAATCATT CCCTTTTCTC CCCACTCCTC CAGAAGGCAT AACCCTTCTC 420

TCCACACCCC CTACCCCCAC CCCCGTCCTA GGCTCCCATC CTTTCCATCA AGACCTTCAT 480

TAGCTTATGA TATTTGCTGC CGAGATGTTA TAACAAGGAC TCGTTCATGT ATATAAGCTA 540

TTTCTTGATC CATTTAAAAG GAATTGTACA TTGTGTAGGA AAAAAAAAAA AAAAAAAAAA 600

AAAAAAAAAA AAAAAAAAAA AAAAA 625

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 amino acids
(B) TYPE: amino acid

(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

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

Met Leu Leu Asp Ile Leu Leu Leu Phe Leu Phe Pro Asn His Ser Leu
1 5 10 15

Phe Ser Pro Leu Leu Gln Lys Ala
20

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1946 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

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

GAAGCTGGGG CCCAGCGGAG GTAGCAGCAG ACGCCTGAGA GCGAGGCCGA GGCCCCTCAG 60

GGTTTGGAGA CCCTGACACA CCCACCTTCT CACCTGGGCT CTGCGTATCC CCCAGCCTTG 120

AGGGAAGATG AAGCCTAAAC TGATGTACCA GGAGCTGAAG GTGCCTGCAG AGGAGCCCGC 180

CAATGAGCTG CCCATGAATG AGATTGAGGC GTGGAAGGCT GCGGAAAAGA AAGCCCGCTG 240

GGTCCTGCTG GTCCTCATTC TGGCGGTTGT GGGCTTCGGA GCCCTGATGA CTCAGCTGTT 300

TCTATGGGAA TACGGCGACT TGCATCTCTT TGGGCCCAAC CAGCGCCCAG CCCCCTGCTA 360

TGACCCTTGC GAARCAGTGC TGGTGGAAAG CATTCCTGAR GGCCTGGACT TCCCCAATGC 420

CTCCACGGGG AACCCTTCCA CCAGCCAGGC CTGGCTGGGC CTGCTCGCCG GTGCGCACAG 480

CAGCCTGGAC ATCGCCTCCT TCTACTGGAC CCTCACCAAC AATGACACCC ACACGCAGGA 540

GCCCTCTGCC CAGCAGGGTG AGGAGGTCCT CCGGCAGCTG CAGACCCTGG CACCAAAGGG 600

CGTGAACGTC CGCATCGCTG TGAGCAAGCC CAGCGGGCCC CAGCCACAGG CGGACCTGCA 660

GGCTCTGCTG CAGAGCGGTG CCCAGGTCCG CATGGTGGAC ATGCAGAAGC TGACCCATGG 720

CGTCCTGCAT ACCAAGTTCT GGGTGGTGGA CCAGACCCAC TTCTACCTGG GCAGTGCCAA 780

CATGGACTGG CGTTCACTGA CCCAGGTCAA GGAGCTGGGC GTGGTCATGT ACAACTGCAG 840

CTGCCTGGCT CGAGACCTGA CCAAGATCTT TGAGGCCTAC TGGTTCCTGG GCCAGGCAGG 900

CAGCTCCATC CCATCAACTT GGCCCCGGTT CTATGACACC CGCTACAACC AAGAGACACC 960

AATGGAGATC TGCCTCAATG GAACCCCTGC TCTGGCCTAC CTGGCGAGTG CGCCCCCACC 1020

CCTGTGTCCA AGTGGCCGCA CTCCAGACCT GAAGGCTCTA CTCAACGTGG TGGACAATGC 1080

CCGGAGTTTC ATCTACGTCG CTGTCATGAA CTACCTGCCC ACTCTGGAGT TCTCCCACCC 1140

TCACAGGTTC TGGCCTGCCA TTGACGATGG GCTGCGGCGG GCCACCTACG AGCGTGGCGT 1200

CAAGGTGCGC CTGCTCATCA GCTGCTGGGG ACACTCGGAG CCATCCATGC GGGCCTTCCT 1260

GCTCTCTCTG GCTGCCCTGC GTGACAACCA TACCCACTCT GACATCCAGG TGAAACTCTT 1320

TGTGGTCCCC GCGGATGAGG CCCAGGCTCG AATCCCATAT GCCCGTGTCA ACCACAACAA 1380

GTACATGGTG ACTGAACGCG CCACCTACAT CGGAACCTCC AACTGGTCTG GCAACTACTT 1440

CACGGAGACG GCGGGCACCT CGCTGCTGGT GACGCAGAAT GGGAGGGGCG GCCTGCGGAG 1500

CCAGCTGGAG GCCATTTTCC TGAGGGACTG GGACTCCCCT TACAGCCATG ACCTTGACAC 1560

YTCAGYTGAC AGSGTGGGCA ACGCCTGCCG CYTGYTCTGA GGCCCGATCC AGTGGGCAGG 1620

CCAAGGCCTG CTGGGCCCCC GCGGACCCAG GTGYTCTGGG TCACGGTCCC TGTCCCCGCA 1680

CCCCCGYTTY TGTYTGCCCC ATTGTGGCTC CTCAGGYTYT YTCCCCTGYT CTCCCACCTY 1740

TACCTCCACC CCCACCGGCC TGACGCTGTG GCCCCGGGAC CCAGCAGAGC TGGGGGAGGG 1800

ATCAGCCCCC AAAGAAATGG GGGTGCATGC TGGGCCTGGC CCCCTGGCCC ACCCCCAYTT 1860

TTCAGGGCAA AAAGGGCCCA GGGTTATAAT AAGTAAATAA CTTGTCTGTA AAAAAAAAAA 1920

AAAAAAAAAA AAAAAAAAAA AAAAAA 1946

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 490 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

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

Met Lys Pro Lys Leu Met Tyr Gln Glu Leu Lys Val Pro Ala Glu Glu
1 5 10 15

```
Pro Ala Asn Glu Leu Pro Met Asn Glu Ile Glu Ala Trp Lys Ala Ala
            20                  25                  30

Glu Lys Lys Ala Arg Trp Val Leu Leu Val Leu Ile Leu Ala Val Val
        35                  40                  45

Gly Phe Gly Ala Leu Met Thr Gln Leu Phe Leu Trp Glu Tyr Gly Asp
    50                  55                  60

Leu His Leu Phe Gly Pro Asn Gln Arg Pro Ala Pro Cys Tyr Asp Pro
65                  70                  75                  80

Cys Glu Xaa Val Leu Val Glu Ser Ile Pro Glu Gly Leu Asp Phe Pro
                85                  90                  95

Asn Ala Ser Thr Gly Asn Pro Ser Thr Ser Gln Ala Trp Leu Gly Leu
            100                 105                 110

Leu Ala Gly Ala His Ser Ser Leu Asp Ile Ala Ser Phe Tyr Trp Thr
        115                 120                 125

Leu Thr Asn Asn Asp Thr His Thr Gln Glu Pro Ser Ala Gln Gln Gly
    130                 135                 140

Glu Glu Val Leu Arg Gln Leu Gln Thr Leu Ala Pro Lys Gly Val Asn
145                 150                 155                 160

Val Arg Ile Ala Val Ser Lys Pro Ser Gly Pro Gln Pro Gln Ala Asp
                165                 170                 175

Leu Gln Ala Leu Leu Gln Ser Gly Ala Gln Val Arg Met Val Asp Met
            180                 185                 190

Gln Lys Leu Thr His Gly Val Leu His Thr Lys Phe Trp Val Val Asp
        195                 200                 205

Gln Thr His Phe Tyr Leu Gly Ser Ala Asn Met Asp Trp Arg Ser Leu
    210                 215                 220

Thr Gln Val Lys Glu Leu Gly Val Val Met Tyr Asn Cys Ser Cys Leu
225                 230                 235                 240

Ala Arg Asp Leu Thr Lys Ile Phe Glu Ala Tyr Trp Phe Leu Gly Gln
                245                 250                 255

Ala Gly Ser Ser Ile Pro Ser Thr Trp Pro Arg Phe Tyr Asp Thr Arg
            260                 265                 270

Tyr Asn Gln Glu Thr Pro Met Glu Ile Cys Leu Asn Gly Thr Pro Ala
        275                 280                 285

Leu Ala Tyr Leu Ala Ser Ala Pro Pro Pro Leu Cys Pro Ser Gly Arg
    290                 295                 300

Thr Pro Asp Leu Lys Ala Leu Leu Asn Val Val Asp Asn Ala Arg Ser
305                 310                 315                 320
```

Phe Ile Tyr Val Ala Val Met Asn Tyr Leu Pro Thr Leu Glu Phe Ser
325 330 335

His Pro His Arg Phe Trp Pro Ala Ile Asp Asp Gly Leu Arg Arg Ala
340 345 350

Thr Tyr Glu Arg Gly Val Lys Val Arg Leu Leu Ile Ser Cys Trp Gly
355 360 365

His Ser Glu Pro Ser Met Arg Ala Phe Leu Leu Ser Leu Ala Ala Leu
370 375 380

Arg Asp Asn His Thr His Ser Asp Ile Gln Val Lys Leu Phe Val Val
385 390 395 400

Pro Ala Asp Glu Ala Gln Ala Arg Ile Pro Tyr Ala Arg Val Asn His
405 410 415

Asn Lys Tyr Met Val Thr Glu Arg Ala Thr Tyr Ile Gly Thr Ser Asn
420 425 430

Trp Ser Gly Asn Tyr Phe Thr Glu Thr Ala Gly Thr Ser Leu Leu Val
435 440 445

Thr Gln Asn Gly Arg Gly Gly Leu Arg Ser Gln Leu Glu Ala Ile Phe
450 455 460

Leu Arg Asp Trp Asp Ser Pro Tyr Ser His Asp Leu Asp Thr Ser Xaa
465 470 475 480

Asp Xaa Val Gly Asn Ala Cys Arg Leu Xaa
485 490

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 6254 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

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

AGGGTCCTGG GCAGTCGCGC CAGAGCTGAG CGGAGGGCGC GGCGCGAGAA CGAATCTTTG 60

TGACATTCTC TCTCAGCATT CTTTATCCCC TGTTTGCTGA AGACTTCGAC CAAAGCTGGT 120

CTTAGCTGTT GGCATTCTCC TGAGAAAAGG ATAGCTTCAG AAATCAGAAA AACATTTGGG 180

AGGTGTCTAG CCCAGTGGAC CTTCTGAAGA GCAATGCTAA GAAGACGTTT GGTTTAAAGA 240

ATTAAAAGGA AGAACAACTT AAGAGCTTCT TCAAAGTTCC CCGCATGAAA ATTACTTAAA 300

CGTTGCACAC AACGTTTCAC AAAATCTTTT GTGAAAGAAG AAAAGGAAAT TCAGTGTGTG 360

AGTCTCAGCA GGAGTTAAGC TAATGCAGCT TAACATCATG CCGACAAAGA AGCGCTTATC 420

TGCGGGCAGA GTGCCCCTGA TTCTCTTCCT GTGCCAGATG ATTAGTGCAC TGGAAGTACC 480

TCTTGATCCA AAACTTCTTG AAGACTTGGT ACAGCCTCCA ACCATCACCC AACAGTCTCC 540

AAAAGATTAC ATTATTGACC CTCGGGAGAA TATTGTAATC CAGTGTGAAG CCAAAGGGAA 600

ACCGCCCCCA AGCTTTTCCT GGACCCGTAA TGGGACTCAT TTTGACATCG ATAAAGACCC 660

TCTGGTCACC ATGAAGCCTG GCACAGGAAC GCTCATAATT AACATCATGA GCGAAGGGAA 720

AGCTGAGACC TATGAAGGAG TCTATCAGTG TACAGCAAGG AACGAACGCG GAGCTGCAGT 780

TTCTAATAAC ATTGTTGTCC GCCCATCCAG ATCACCATTG TGGACCAAAG AAAAACTTGA 840

ACCAATCACA CTTCAAAGTG GTCAGTCTTT AGTACTTCCC TGCAGACCCC CAATTGGATT 900

ACCACCACCT ATAATATTTT GGATGGATAA TTCCTTTCAA AGACTTCCAC AAAGTGAGAG 960

AGTTTCTCAA GGTTTGAATG GGGACCTTTA TTTTTCCAAT GTCCTCCCAG AGGACACCCG 1020

CGAAGACTAT ATCTGTTATG CTAGATTTAA TCATACTCAA ACCATACAGC AGAAGCAACC 1080

TATTTCTGTG AAGGTGATTT CAGCTAAATC AAGTAGAGAG AGGCCACCAA CATTTTTAAC 1140

TCCAGAAGGC AATGCAAGTA ACAAAGAGGA ATTAAGAGGA AATGTGCTTT CACTGGAGTG 1200

CATTGCAGAA GGACTGCCTA CCCCAATTAT TTACTGGGCA AAGGAAGATG GAATGCTACC 1260

CAAAAACAGG ACAGTTTATA AGAACTTTGA GAAAACCTTG CAGATCATTC ATGTTTCAGA 1320

AGCAGACTCT GGAAATTACC AATGTATAGC AAAAAATGCA TTAGGAGCCA TCCACCATAC 1380

CATTTCTGTT AGAGTTAAAG CGGCTCCATA CTGGATCACA GCCCCTCAAA ATCTTGTGCT 1440

GTCCCCAGGA GAGGATGGGA CCTTGATCTG CAGAGCTAAT GGCAACCCCA AACCCAGAAT 1500

TAGCTGGTTA ACAAATGGAG TCCCAATAGA AATTGCCCCT GATGACCCCA GCAGAAAAAT 1560

AGATGGCGAT ACCATTATTT TTTCAAATGT TCAAGAAAGA TCAAGTGCAG TATATCAGTG 1620

CAATGCCTCT AATGAATATG GATATTTACT GGCAAACGCA TTTGTAAATG TGCTGGCTGA 1680

GCCACCACGA ATCCTCACAC CTGCAAACAC ACTCTACCAG GTCATTGCAA ACAGGCCTGC 1740

TTTACTAGAC TGTGCCTTCT TTGGGTCTCC TCTCCCAACC ATCGAGTGGT TTAAAGGAGC 1800

TAAAGGAAGT GCTCTTCATG AAGATATTTA TGTTTTACAT GAAAATGGAA CTTTGGAAAT 1860

TCCTGTGGCC CAAAAGGACA GTACAGGAAC TTATACGTGT GTTGCAAGGA ATAAATTAGG 1920

GATGGCAAAG AATGAAGTTC ACTTAGAAAT CAAAGATGCT ACATGGATCG TTAAACAGCC 1980

CGAATATGCA GTTGTGCAAA GAGGGAGCAT GGTGTCCTTT GAATGCAAAG TGAAACATGA 2040

TCACACCTTA TCCCTCACTG TCCTGTGGCT GAAGGACAAC AGGGAACTGC CCAGTGATGA 2100

AAGGTTCACT GTTGACAAGG ATCATCTAGT GGTAGCTGAT GTCAGTGACG ATGACAGCGG 2160

GACCTACACG TGTGTGGCCA ACACCACTCT GGACAGCGTC TCCGCCAGCG CTGTGCTTAG 2220

CGTTGTTGCT CCTACTCCAA CTCCAGCTCC CGTTTACGAT GTCCCAAATC CTCCCTTTGA 2280

CTTAGAACTG ACAGATCAAC TTGACAAAAG TGTTCAGCTG TCATGGACCC CAGGCGATGA 2340

CAACAATAGC CCCATTACAA AATTCATCAT CGAATATGAA GATGCAATGC ACAAGCCAGG 2400

GCTGTGGCAC CACCAAACTG AAGTTTCTGG AACACAGACC ACAGCCCAGC TGAAGCTGTC 2460

TCCTTACGTG AACTACTCCT TCCGCGTGAT GGCAGTGAAC AGCATTGGGA AGAGCTTGCC 2520

CAGCGAGGCC TCTGAGCAGT ATTTGACGAA AGCCTCAGAA CCAGATAAAA ACCCCACAGC 2580

TGTGGAAGGA CTGGGATCAG AGCCTGATAA TTTGGTGATT ACGTGGAAGC CCTTGAATGG 2640

TTTCGAATCT AATGGGCCAG GCCTTCAGTA CAAAGTTAGC TGGCGCCAGA AAGATGGTGA 2700

TGATGAATGG ACATCTGTGG TTGTGGCAAA TGTATCCAAA TATATTGTCT CAGGCACGCC 2760

AACCTTTGTT CCATACCTGA TCAAAGTTCA GGCCCTGAAT GACATGGGGT TTGCCCCCGA 2820

GCCAGCTGTA GTCATGGGAC ATTCTGGAGA AGACCTCCCA ATGGTGGCTC CTGGGAACGT 2880

GCGTGTGAAT GTGGTGAACA GTACCTTAGC CGAGGTGCAC TGGGACCCAG TACCTCTGAA 2940

AAGCATCCGA GGACACCTAC AAGGCTATCG GATTTACTAT TGGAAGACCC AGAGTTCATC 3000

TAAAAGAAAC AGACGTCACA TTGAGAAAAA GATCCTCACC TTCCAAGGCA GCAAGACTCA 3060

TGGCATGTTG CCGGGGCTAG AGCCCTTTAG CCACTACACA CTGAATGTCC GAGTGGTCAA 3120

TGGGAAAGGG GAGGGCCCAG CCAGCCCTGA CAGAGTCTTT AATACTCCAG AAGGAGTCCC 3180

CAGTGCTCCC TCGTCTTTGA AGATTGTGAA TCCAACACTG GACTCTCTCA CTTTGGAATG 3240

GGATCCACCG AGCCACCCGA ATGGCATTTT GACAGAGTAC ACCTTAAAGT ATCAGCCAAT 3300

TAACAGCACA CATGAATTAG GCCCTCTGGT AGATTTGAAA ATTCCTGCCA ACAAGACACG 3360

GTGGACTTTA AAAAATTTAA ATTTCAGCAC TCGATATAAG TTTTATTTCT ATGCACAAAC 3420

ATCAGCAGGA TCAGGAAGTC AAATTACAGA GGAAGCAGTA ACAACTGTGG ATGAAGCTGG 3480

TATTCTTCCA CCTGATGTAG GTGCAGGCAA AGCGATGGCA AGCCGGCAGG TGGATATTGC 3540

AACTCAGGGC TGGTTCATTG GTCTGATGTG TGCTGTTGCT CTCCTTATCT TAATTTTGCT 3600

GATTGTTTGC TTCATCAGAA GAAACAAGGG TGGTAAATAT CCAGTTAAAG AAAAGGAAGA 3660

TGCCCATGCT GACCCTGAAA TCCAGCCTAT GAAGGAAGAT GATGGGACAT TTGGAGAATA 3720

CAGTGATGCA GAAGACCACA AGCCTTTGAA AAAAGGAAGT CGAACTCCTT CAGACAGGAC 3780

TGTGAAAAAA GAAGATAGTG ACGACAGCCT AGTTGACTAT GGAGAAGGGG TTAATGGCCA 3840

GTTCAATGAG GATGGCTCCT TTATTGGACA ATACAGTGGT AAGAAAGAGA AAGAGCCGGC 3900

TGAAGGAAAC GAAAGCTCAG AGGCACCTTC TCCTGTCAAC GCCATGAATT CCTTTGTTTA 3960

ATTTTTAAGC TCTTTGCCAA TATTCCATTT CTCTAGAATG TTTATCCTAA GCACTTGTTT 4020

GTCAGCCCTC TCATACTATG AACATATGGG TAGAGAGTAT ATTTTCTGCT GTATGTTAGT 4080

ATTATGAGAA TAGTTACAGC AAAAACATAA CTCAGTCAAA TGATATGTTA ATATGAACTG 4140

GAATGCAAAG TGCATACTTT TTCATTCAAA ATGGGTATTC TTGATTTCCT CAGAACTGAT 4200

AAAAAATAAT GCAACATCAC CAACAGATCC TGTTATTTCC TCTGCAGGAT ACAGTTCAAT 4260

ATGATGCATG AAAAATGCTC CACATTTAAA GGACATACCC GTGTATGTTA TGAAAACATG 4320

GTTTGATACT TTGTTTATAC TACCCTCAGC TGAACCCCTA TATATGAATT CCGTTTTCAT 4380

TGTCAAGAAT GTTACTGTAG TATTCTCTAG AACTTCAATG TCTTTGTGGA CATTGTTGTG 4440

AAATTGGTGA CTATGTATAG CTGTCGTTAG TCTTTTTGGG AGACTGTTAG GAACAGTTTG 4500

TACAGTATAT ACTTGCTAAA TGAGTTCATT ATGACAGTCA CATTGCTGAT GCTTACTGAG 4560

AACTATTACC TACTCTTGGC TCCTGTTACT CCGTAGGCTT CTTAATCTTC CAGGCATTAC 4620

AGCAGCACAG TGTTCTACTT TTTACATCAT TTCTATGTTC GGTTGTTTTT AGGCATAAAC 4680

AATGTGTATT GCAGTGCATT TCGGCATTTG TGCCATACTG AAAGAATCAA AAACAAATCA 4740

TCCAAATTAA ATTTCAAACA TTATTTCAGA GAACACAGGG CAAGACACAT ACAGTGCCTT 4800

CAGATATTAA GCATTCCACA ACATCGTGCA TTCTGTATCA GCTGGTCCAG TCCATTCTGG 4860

GTCCTAGATT ACTGTCATTG TCTAAAAGTA ACTTTTAAAA AGCAGAGTTC ATGAAAACTG 4920

CAATGCTGGG AAAAGAAGGA AACATGAAAA TAAAAATAAG ACAGTTTATT AGAAATAGCA 4980

TTTCCTCATA AGCATAAAAA GAAATCTTTG TTGCCAACTG AAGCACATGG GATTTTGTGG 5040

TCCTTTATGG TTTCTATAAC ATTCAGTAAG AAAGATGTCA ACATGCTAGA AAATTAATTT 5100

TAAAACTAAG TTATTCCAAC ACTAAAAGCA TACAACAGCA TGCCAACAGT AATATATTAT 5160

TCTCCAAGAC TTTACCTATG TAAGTGTTCA AAACTCTGCA GCATTAAACA ACGTGTATGC 5220

AAATTGTTAT GGATACATTT CAGAATCTAA GAAATCAGGC AAGTGCTTAA AAGGCCAACG 5280

GTCCAAGGGA TTACATCTGC AGTTTAAAAA GTAAATATAT ATTCTATCGT ATTCATAAAC 5340

AATATCTATC AAATGGGTTA CCTCCAAATA TGAAAATCTA TAACAACCTA TGGTTGAAGG 5400

AATGCTCAGT TTCATTTGCC AATAAATTGG TTTCTCATAA CTTGCATCAA GTTTAATTTT 5460

AAGTAAAGCT TTTTATATGT AGATATTTTG TTGAATTTGT AAATACACTT AAAATGTAGA 5520

TGCTATATGC TTATAGGTGT TACATACAAA TAAACATGCA ATGTTTATGT TGTACTGTAT 5580

AAGAGGTAAG CTAATTAATG CAGTGAATGG GATTGGAAAG CATCTACTTA AATATCTATT 5640

GGGTTCCCCC CTCCCCCCAC CTTTTTTGCT GTGAAACTGA AATAGTGAAC TTTTCTACGT 5700

ATTGACAGCA GATTTTTCGA TGAAATCTTC AGAGCTTTGC CTATGGGGCA CAGTAGGCCT 5760

AGTAACCTGG CATGTTTGAT ATATGTAGGT AAAGCATAAT TTAAAGTAAT CCCAGGTAAA 5820

GATGGCCCTA AATACTTTCA TGTCTCTATA TTCATTTTTC ACAGATCCAC CTGTCTCTTG 5880

AAAATATAAA AAGACAAAAC AGGTTTGCCT TGGCATCAGA GAGCACAAAG ATTAAAAGTT 5940

ACTTTAAATT TGCCAATATT TTGGGAGAAC AATAAAACTA CATTTTTTCY TYTTCCATAC 6000

TGGTAGATGC GAAATTTATC TGTGCATGAA AGGGTCACTT CTGTAATAGT GCACAGATTT 6060

GGTATTAAAA ATTAAATGTG GTTTTAAAAG TTCCTYTYTC TTTTGTAATT TATGTTCCCA 6120

ATTGAGTGTG AATGTCCAAG TAATGGTGTA TGTAATGGTA CAGGCAAATG TGACTGGATT 6180

TCCCTCAAAA AAGTAACTAT TAAACAGTYT TGATCTCAAA AAAAAAAAAA AAAAAAAAAA 6240

AAAAAAAAAA AAAA 6254

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1192 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

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

Met Gln Leu Asn Ile Met Pro Thr Lys Lys Arg Leu Ser Ala Gly Arg
1 5 10 15

Val Pro Leu Ile Leu Phe Leu Cys Gln Met Ile Ser Ala Leu Glu Val
20 25 30

Pro Leu Asp Pro Lys Leu Leu Glu Asp Leu Val Gln Pro Pro Thr Ile
35 40 45

Thr Gln Gln Ser Pro Lys Asp Tyr Ile Ile Asp Pro Arg Glu Asn Ile
50 55 60

Val Ile Gln Cys Glu Ala Lys Gly Lys Pro Pro Pro Ser Phe Ser Trp
65 70 75 80

Thr Arg Asn Gly Thr His Phe Asp Ile Asp Lys Asp Pro Leu Val Thr

85 90 95

Met Lys Pro Gly Thr Gly Thr Leu Ile Ile Asn Ile Met Ser Glu Gly
100 105 110

Lys Ala Glu Thr Tyr Glu Gly Val Tyr Gln Cys Thr Ala Arg Asn Glu
115 120 125

Arg Gly Ala Ala Val Ser Asn Asn Ile Val Val Arg Pro Ser Arg Ser
130 135 140

Pro Leu Trp Thr Lys Glu Lys Leu Glu Pro Ile Thr Leu Gln Ser Gly
145 150 155 160

Gln Ser Leu Val Leu Pro Cys Arg Pro Pro Ile Gly Leu Pro Pro Pro
165 170 175

Ile Ile Phe Trp Met Asp Asn Ser Phe Gln Arg Leu Pro Gln Ser Glu
180 185 190

Arg Val Ser Gln Gly Leu Asn Gly Asp Leu Tyr Phe Ser Asn Val Leu
195 200 205

Pro Glu Asp Thr Arg Glu Asp Tyr Ile Cys Tyr Ala Arg Phe Asn His
210 215 220

Thr Gln Thr Ile Gln Gln Lys Gln Pro Ile Ser Val Lys Val Ile Ser
225 230 235 240

Ala Lys Ser Ser Arg Glu Arg Pro Pro Thr Phe Leu Thr Pro Glu Gly
245 250 255

Asn Ala Ser Asn Lys Glu Glu Leu Arg Gly Asn Val Leu Ser Leu Glu
260 265 270

Cys Ile Ala Glu Gly Leu Pro Thr Pro Ile Ile Tyr Trp Ala Lys Glu
275 280 285

Asp Gly Met Leu Pro Lys Asn Arg Thr Val Tyr Lys Asn Phe Glu Lys
290 295 300

Thr Leu Gln Ile Ile His Val Ser Glu Ala Asp Ser Gly Asn Tyr Gln
305 310 315 320

Cys Ile Ala Lys Asn Ala Leu Gly Ala Ile His His Thr Ile Ser Val
325 330 335

Arg Val Lys Ala Ala Pro Tyr Trp Ile Thr Ala Pro Gln Asn Leu Val
340 345 350

Leu Ser Pro Gly Glu Asp Gly Thr Leu Ile Cys Arg Ala Asn Gly Asn
355 360 365

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

Ala Pro Asp Asp Pro Ser Arg Lys Ile Asp Gly Asp Thr Ile Ile Phe

385 390 395 400

Ser Asn Val Gln Glu Arg Ser Ser Ala Val Tyr Gln Cys Asn Ala Ser
405 410 415

Asn Glu Tyr Gly Tyr Leu Leu Ala Asn Ala Phe Val Asn Val Leu Ala
420 425 430

Glu Pro Pro Arg Ile Leu Thr Pro Ala Asn Thr Leu Tyr Gln Val Ile
435 440 445

Ala Asn Arg Pro Ala Leu Leu Asp Cys Ala Phe Phe Gly Ser Pro Leu
450 455 460

Pro Thr Ile Glu Trp Phe Lys Gly Ala Lys Gly Ser Ala Leu His Glu
465 470 475 480

Asp Ile Tyr Val Leu His Glu Asn Gly Thr Leu Glu Ile Pro Val Ala
485 490 495

Gln Lys Asp Ser Thr Gly Thr Tyr Thr Cys Val Ala Arg Asn Lys Leu
500 505 510

Gly Met Ala Lys Asn Glu Val His Leu Glu Ile Lys Asp Ala Thr Trp
515 520 525

Ile Val Lys Gln Pro Glu Tyr Ala Val Val Gln Arg Gly Ser Met Val
530 535 540

Ser Phe Glu Cys Lys Val Lys His Asp His Thr Leu Ser Leu Thr Val
545 550 555 560

Leu Trp Leu Lys Asp Asn Arg Glu Leu Pro Ser Asp Glu Arg Phe Thr
565 570 575

Val Asp Lys Asp His Leu Val Val Ala Asp Val Ser Asp Asp Asp Ser
580 585 590

Gly Thr Tyr Thr Cys Val Ala Asn Thr Thr Leu Asp Ser Val Ser Ala
595 600 605

Ser Ala Val Leu Ser Val Val Ala Pro Thr Pro Thr Pro Ala Pro Val
610 615 620

Tyr Asp Val Pro Asn Pro Pro Phe Asp Leu Glu Leu Thr Asp Gln Leu
625 630 635 640

Asp Lys Ser Val Gln Leu Ser Trp Thr Pro Gly Asp Asp Asn Asn Ser
645 650 655

Pro Ile Thr Lys Phe Ile Ile Glu Tyr Glu Asp Ala Met His Lys Pro
660 665 670

Gly Leu Trp His His Gln Thr Glu Val Ser Gly Thr Gln Thr Thr Ala
675 680 685

Gln Leu Lys Leu Ser Pro Tyr Val Asn Tyr Ser Phe Arg Val Met Ala

690 695 700

Val Asn Ser Ile Gly Lys Ser Leu Pro Ser Glu Ala Ser Glu Gln Tyr
705 710 715 720

Leu Thr Lys Ala Ser Glu Pro Asp Lys Asn Pro Thr Ala Val Glu Gly
725 730 735

Leu Gly Ser Glu Pro Asp Asn Leu Val Ile Thr Trp Lys Pro Leu Asn
740 745 750

Gly Phe Glu Ser Asn Gly Pro Gly Leu Gln Tyr Lys Val Ser Trp Arg
755 760 765

Gln Lys Asp Gly Asp Asp Glu Trp Thr Ser Val Val Val Ala Asn Val
770 775 780

Ser Lys Tyr Ile Val Ser Gly Thr Pro Thr Phe Val Pro Tyr Leu Ile
785 790 795 800

Lys Val Gln Ala Leu Asn Asp Met Gly Phe Ala Pro Glu Pro Ala Val
805 810 815

Val Met Gly His Ser Gly Glu Asp Leu Pro Met Val Ala Pro Gly Asn
820 825 830

Val Arg Val Asn Val Val Asn Ser Thr Leu Ala Glu Val His Trp Asp
835 840 845

Pro Val Pro Leu Lys Ser Ile Arg Gly His Leu Gln Gly Tyr Arg Ile
850 855 860

Tyr Tyr Trp Lys Thr Gln Ser Ser Ser Lys Arg Asn Arg Arg His Ile
865 870 875 880

Glu Lys Lys Ile Leu Thr Phe Gln Gly Ser Lys Thr His Gly Met Leu
885 890 895

Pro Gly Leu Glu Pro Phe Ser His Tyr Thr Leu Asn Val Arg Val Val
900 905 910

Asn Gly Lys Gly Glu Gly Pro Ala Ser Pro Asp Arg Val Phe Asn Thr
915 920 925

Pro Glu Gly Val Pro Ser Ala Pro Ser Ser Leu Lys Ile Val Asn Pro
930 935 940

Thr Leu Asp Ser Leu Thr Leu Glu Trp Asp Pro Pro Ser His Pro Asn
945 950 955 960

Gly Ile Leu Thr Glu Tyr Thr Leu Lys Tyr Gln Pro Ile Asn Ser Thr
965 970 975

His Glu Leu Gly Pro Leu Val Asp Leu Lys Ile Pro Ala Asn Lys Thr
980 985 990

Arg Trp Thr Leu Lys Asn Leu Asn Phe Ser Thr Arg Tyr Lys Phe Tyr

995 1000 1005

Phe Tyr Ala Gln Thr Ser Ala Gly Ser Gly Ser Gln Ile Thr Glu Glu
1010 1015 1020

Ala Val Thr Thr Val Asp Glu Ala Gly Ile Leu Pro Pro Asp Val Gly
1025 1030 1035 1040

Ala Gly Lys Ala Met Ala Ser Arg Gln Val Asp Ile Ala Thr Gln Gly
1045 1050 1055

Trp Phe Ile Gly Leu Met Cys Ala Val Ala Leu Leu Ile Leu Ile Leu
1060 1065 1070

Leu Ile Val Cys Phe Ile Arg Arg Asn Lys Gly Gly Lys Tyr Pro Val
1075 1080 1085

Lys Glu Lys Glu Asp Ala His Ala Asp Pro Glu Ile Gln Pro Met Lys
1090 1095 1100

Glu Asp Asp Gly Thr Phe Gly Glu Tyr Ser Asp Ala Glu Asp His Lys
1105 1110 1115 1120

Pro Leu Lys Lys Gly Ser Arg Thr Pro Ser Asp Arg Thr Val Lys Lys
1125 1130 1135

Glu Asp Ser Asp Asp Ser Leu Val Asp Tyr Gly Glu Gly Val Asn Gly
1140 1145 1150

Gln Phe Asn Glu Asp Gly Ser Phe Ile Gly Gln Tyr Ser Gly Lys Lys
1155 1160 1165

Glu Lys Glu Pro Ala Glu Gly Asn Glu Ser Ser Glu Ala Pro Ser Pro
1170 1175 1180

Val Asn Ala Met Asn Ser Phe Val
1185 1190

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 4253 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

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

GCGATCAGAA GACTGAGGCA GTGCTGCAGG TGAACATGGC TGTAGCCTAC CAATCCTTCA 60

TTTTGGCTGA CTTCATAATT GAACCTGCTC ACTGCCACCC AGTCGTTCAC ATGGCAAACC 120

CCAGTGTTCC TTAAGTGGGC TGAAAAGAAT GGAATCCTAG ACCTGTCTTG TTAGAGTARA 180

AGCTCATTGA GCTGCCTGCC CTTCCATATT GCTTGTGGAC CTATGAATGG TGCTCTTTAA 240

AATGTCAGTG TAGAGGGACT GGGAATCCTG TATTCCTGCC CTATCATTCT GTAACTTTGA 300

GAAGCATGTG GGCTCAGTGC ARACTACTTA TTTTTAGTTC TGTTAGAACA GGATTTTTAA 360

GAAACCATCT GAACAGAAGA AATAAAGCGG AAGGTAAGCT CAGCTCATGT GCCCATCAAG 420

CTGCATAAAT CTCCATCTGC TCTGGGAAAC ATTCATGGGA ACAGAAGGAT TCTTCTTCCT 480

GAGAATGTTG TTTTCATTAT ATAATCATAG ATACAGCATG TGGGTGATTT CTGTGTACAA 540

ATACTTTACC CATGAGAAAT GTATAGCTGT TAGATATGTG TGGGCTCACA CATATATGCA 600

ATAGATGCTA TACATAGCGA TGTGCATGTG TATTTGTATA TATTTATGTT CATATTTTAA 660

CATATGAAAC ATGAAATAGA AGTATAATAT TATTTTTGTA TCCAAATAGT TAACAGAGCA 720

GATGCCCTTA GATTCTAATA TGTTGTTCAG TCCTGAAGAT AGCCAAGTTA AATATAAAAA 780

CAGGAGAAGA GGAAAAATAA AAKTCCATTG AGAAGGAATA GAGAGAAAGA ACRGGTCTTG 840

TGAATGAGAA KTTCATTTGG ATCAACACAG CATTTCCTAR AATGAGTCTG CACTGTGCTT 900

TTTGCARARA CRAATGGGCT CCGTAGGCTC TCCACGGCTT TTGATGGCAR ATGAKTTGTC 960

TCTGTTTCAR ATGAKTCAAG CTCCAGTCCT GGAAGGCAGA TGTCCTCCTC CGATGGTGGG 1020

CCACCGGGCC AGTCARACAC AGACAGCTCC CGTGGAGGAG AGCGACTTCG ACACCATGCC 1080

AGACATTGAG AGTGATAAAA ACATCATCCG GACCAAGATG TTCCTTTACC TGTCAGATTT 1140

GTCCAGGAAG GACCGGAGAA TTGTCAGCAA AAAATATAAA ATTTATTTTT GGAACATCAT 1200

CACCATTGCT GTGTTTTACG CGCTGCCCGT GATCCAGCTG GTCATTACCT ATCAGACAGT 1260

GGTAAATGTC ACTGGCAACC AGGACATCTG TTACTACAAC TTCCTCTGTG CTCACCCCTT 1320

GGGCGTCCTG AGTGCCTTCA ACAACATTCT CAGCAATCTG GGCCACGTGC TTCTGGGCTT 1380

CCTCTTCCTG CTGATAGTCT TGCGCCGCGA CATCCTCCAT CGGAGAGCCC TGGAAGCCAA 1440

GGACATCTTT GCTGTGGAGT ACGGGATTCC CAAACACTTT GGTCTCTTCT ACGCTATGGG 1500

CATTGCATTG ATGATGGAAG GGGTGCTCAG TGCTTGCTAC CATGTCTGCC CTAATTATTC 1560

CAACTTCCAA TTCGACACCT CCTTCATGTA CATGATCGCT GGCCTGTGCA TGCTGAAGCT 1620

CTATCAGACC CGCCACCCAG ACATCAATGC CAGCGCCTAC TCTGCCTATG CCTCCTTTGC 1680

TGTGGTCATC ATGGTCACCG TCCTTGGAGT GGTGTTTGGA AAAAATGACG TATGGTTCTG 1740

GGTCATCTTC TCTGCAATCC ACGTTCTGGC CTCGCTAGCC CTCAGCACCC AAATATATTA 1800

TATGGGTCGT TTCAAGATAG ATGTGTCTGA CACAGATTTG GGAATTTTCC GGCGGGCTGC 1860

CATGGTGTTC TACACAGACT GTATCCAGCA GTGTAGCCGA CCTCTATATA TGGATAGAAT 1920

GGTGTTGCTG GTTGTGGGGA ATCTGGTTAA CTGGTCCTTC GCCCTCTTTG GATTGATATA 1980

CCGCCCCAGG GACTTTGCTT CCTACATGCT GGGCATCTTC ATCTGTAACC TTTTGCTGTA 2040

CCTGGCCTTT TACATCATCA TGAAGCTCCG CAGCTCTGAA AAGGTCCTCC CAGTCCCGCT 2100

CTTCTGCATC GTGGCCACCG CTGTGATGTG GGCTGCCGCC CTATATTTTT TCTTCCAGAA 2160

TCTCAGCAGC TGGGAGGGAA CTCCGGCCGA ATCCCGGGAG AAGAACCGCG AGTGCATTCT 2220

GCTGGATTTC TTCGATGACC ATGACATCTG GCACTTCCTC TCTGCTACTG CTCTGTTTTT 2280

CTCATTCTTG GTTTTGTTAA CTTTGGATGA TGACCTTGAT GTGGTTCGGA GAGACCAGAT 2340

CCCTGTCTTC TGAACCTCCA ACATTAAGAG AGGGGAGGGA GCGATCAATC TTGGTGCTGT 2400

TTCACAAAAA TTACAGTGAC CACAGCAAAG TAACCACTGC CAGATGCTCC ACTCACCCTC 2460

TGTAGAGCCA ACTCTGCATT CACACAGGAA GGAGAGGGGC TGCGGGAGAT TTAAACCTGC 2520

AAGAAAGGAG GCAGAAGGGG AGCCATGTTT TGAGGACAGA CGCAAACCTG AGGAGCTGAG 2580

AAACACTTGC TCCTTCCATC TGCAGCTTTG GGAGTGCAAC AGGGATAGGC ACTGCATCCA 2640

AGTCAACTCA CCATCTTGGG GTCCCTCCCA CCCTCACGGA GACTTGCCAG CAATGGCAGA 2700

ATGCTGCTGC ACACTTCCCT CCAGTTGTCA CCCTGCCCAG AAAGGCCAGC AGCTTGGACT 2760

TCCTGCCCAG AAACTGTGTT GGCCCCCTTC ACACCTCTGC AACACCTGCT GCTCCAGCAA 2820

GAGGATGTGA TTCTTTAGAA TATGGCGGGG AGGTGACCCC AGGCCCTGCC CTACTGGGAT 2880

AGATGTTTTA ATGGCACCAG CTAGTCACCT CCCAGAAGAA ACTCTGTATA TTTCCCCCAG 2940

GTTTCTGATG CCATCAGAAG GGCTCAGGAG TGGGGTTTGT CACACATTCC TCTTAACAAG 3000

TAACTGTCAC TGGGACCGAG TCCTGGGTGC TTACATATTC CTTCGTGTCT TCATCTCACT 3060

GACCTGTGTG GACCTCATCA CTCTGACTCT GCCTTCTTGG AAAGGCCCTG TCACTCCACA 3120

GATGTCTGGC CAGCTTCAAG GCAGAAGGAA AAACAGGAAA AGCTCTTTTA ACAGCAGCAG 3180

GAACAAGAGA AATGACTAAC CATACTAAAA GACTGGTAAC AGCAGCAGCA GCCAGACAGG 3240

CCTCACCTTA AGGACTTGGG CTGCCAGAGC AAATTCAGCA GAGCTTATTT GGCCTCCCAT 3300

TCACACAGCT CAGTTCTGTG CCCACATCAC CTTTGGGGAA GAAATCAGCA TTCTAATCAG 3360

GGACACTACT TCAGGAGTCC TCCACAGCGA GTCCGTCATC TGTCACTTTA TGTAGATCAG 3420

GGTTCTAGAC TTCTTCCCTG AGGTTCTCAG AAGCAGCTCT CAGGATGAAC GTATTGTCCT 3480

CTTCCCCTCT TCTTGCAAAG TGCACAGCTA ATCTAATGTT GTCTCTCGGT TGCACCTGAC 3540

ATTCTCTCCC CAGTAAGGTG TTGGCAAGCT CAGCATCTGG GTTCCACTCT CACACTGTCT 3600

GGCAGCTCTG TGTCTGAGAA GTTCTACATT GACCAGGCCC CCTTGTTGCC TGGAGTATGA 3660

CGTAATCAGA AAATAGACGT ATAAATGTGC ACATGCGTAT GTATTTGCTT GTGAAATTAA 3720

AGTCACCTCT TGCCTCTGCT TTCCTGATCA TTCGTTAGAG AAATGGATCA GGCATTTTTT 3780

TAAATTATTA TTCTTTCTCT AAACTATTTG CATTGTGTTC AAAAACCCAT TTTAGAAGTT 3840

TGAACAGCAA GCTTTTCCTG ATTTTAAAAA CACAAAGTTG CTTTCAATGA AATATTTTGT 3900

GATTTTTTTA AAGTCCCCAA ATGTGTACTT AGCCTTCTGT TATTCCTTAT TCTTTAAGCA 3960

GTGTTGGCTT CCATTGACCA TATGAAGGCC ACCAATTAAA TGGTTGTGTT AATCCAACAT 4020

GTAAAAAACT TTTTGGCAGG GCACAGTGGC TCACGCCTGT AATCCCAACA CTTTGGGAGG 4080

CTGAGGCAGG AGGATCAMTT GAGCCCAGGA GATTGACGCC GCAGTGAACT ATGATTGTGC 4140

CCCTGCACTC CAGCCTGGAT GACAGAGTGA GACCCCATYT YTTAAAAAAT AAAAAAAAAT 4200

AAAAATTTGA ACCTGTTTCA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAA 4253

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 479 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

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

Met Gly Ser Val Gly Ser Pro Arg Leu Leu Met Ala Xaa Xaa Leu Ser
1 5 10 15

Leu Phe Gln Met Xaa Gln Ala Pro Val Leu Glu Gly Arg Cys Pro Pro
20 25 30

Pro Met Val Gly His Arg Ala Ser Gln Thr Gln Thr Ala Pro Val Glu
35 40 45

Glu Ser Asp Phe Asp Thr Met Pro Asp Ile Glu Ser Asp Lys Asn Ile
50 55 60

Ile Arg Thr Lys Met Phe Leu Tyr Leu Ser Asp Leu Ser Arg Lys Asp
65 70 75 80

Arg Arg Ile Val Ser Lys Lys Tyr Lys Ile Tyr Phe Trp Asn Ile Ile
85 90 95

Thr Ile Ala Val Phe Tyr Ala Leu Pro Val Ile Gln Leu Val Ile Thr
100 105 110

Tyr Gln Thr Val Val Asn Val Thr Gly Asn Gln Asp Ile Cys Tyr Tyr
115 120 125

Asn Phe Leu Cys Ala His Pro Leu Gly Val Leu Ser Ala Phe Asn Asn
130 135 140

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

Ile Val Leu Arg Arg Asp Ile Leu His Arg Arg Ala Leu Glu Ala Lys
165 170 175

Asp Ile Phe Ala Val Glu Tyr Gly Ile Pro Lys His Phe Gly Leu Phe
180 185 190

Tyr Ala Met Gly Ile Ala Leu Met Met Glu Gly Val Leu Ser Ala Cys
195 200 205

Tyr His Val Cys Pro Asn Tyr Ser Asn Phe Gln Phe Asp Thr Ser Phe
210 215 220

Met Tyr Met Ile Ala Gly Leu Cys Met Leu Lys Leu Tyr Gln Thr Arg
225 230 235 240

His Pro Asp Ile Asn Ala Ser Ala Tyr Ser Ala Tyr Ala Ser Phe Ala
245 250 255

Val Val Ile Met Val Thr Val Leu Gly Val Val Phe Gly Lys Asn Asp
260 265 270

Val Trp Phe Trp Val Ile Phe Ser Ala Ile His Val Leu Ala Ser Leu
275 280 285

Ala Leu Ser Thr Gln Ile Tyr Tyr Met Gly Arg Phe Lys Ile Asp Val
290 295 300

Ser Asp Thr Asp Leu Gly Ile Phe Arg Arg Ala Ala Met Val Phe Tyr
305 310 315 320

Thr Asp Cys Ile Gln Gln Cys Ser Arg Pro Leu Tyr Met Asp Arg Met
325 330 335

Val Leu Leu Val Val Gly Asn Leu Val Asn Trp Ser Phe Ala Leu Phe
340 345 350

Gly Leu Ile Tyr Arg Pro Arg Asp Phe Ala Ser Tyr Met Leu Gly Ile
355 360 365

Phe Ile Cys Asn Leu Leu Leu Tyr Leu Ala Phe Tyr Ile Ile Met Lys
370 375 380

Leu Arg Ser Ser Glu Lys Val Leu Pro Val Pro Leu Phe Cys Ile Val
385 390 395 400

Ala Thr Ala Val Met Trp Ala Ala Ala Leu Tyr Phe Phe Phe Gln Asn
405 410 415

Leu Ser Ser Trp Glu Gly Thr Pro Ala Glu Ser Arg Glu Lys Asn Arg
420 425 430

Glu Cys Ile Leu Leu Asp Phe Phe Asp Asp His Asp Ile Trp His Phe
435 440 445

Leu Ser Ala Thr Ala Leu Phe Phe Ser Phe Leu Val Leu Leu Thr Leu
450 455 460

Asp Asp Asp Leu Asp Val Val Arg Arg Asp Gln Ile Pro Val Phe
465 470 475

(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:

CGGTGGGATT TATTTAACAT GATCTTGGC 29

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 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:

CCAATGGTTA TGATGATGGT TCTTCCT 27

(2) INFORMATION FOR SEQ ID NO:26:

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

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "oligonucleotide"

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

ANGGATGCCAT CCAGATGAGG CCACAGCT 29

(2) INFORMATION FOR SEQ ID NO:27:

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

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

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

ANCTTCTTCTG CAGCTCATTC AACTCCTG 29

(2) INFORMATION FOR SEQ ID NO:28:

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

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

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

GNTCCTCCTAG ACCACAAAGT AGAAAGCA 29

(2) INFORMATION FOR SEQ ID NO:29:

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

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

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

GNCTGACTTTC TCTCCAGTTT GTGAAATC 29

(2) INFORMATION FOR SEQ ID NO:30:

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

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

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

GNGTGGGGTGA AAAGGGAATG ATTAGGGA 29

(2) INFORMATION FOR SEQ ID NO:31:

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

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

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

TNATTCATGGG CAGCTCATTG GCGGGCTC 29

(2) INFORMATION FOR SEQ ID NO:32:

(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:32:

CNGCTGAGACT CACACACTGA ATTTCCTT 29

(2) INFORMATION FOR SEQ ID NO:33:

(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:33:

CNTCTTGGTCC GGATGATGTT TTTATCAC 29

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 57 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

Met Cys Arg Pro Phe Leu Ser Leu Leu Leu Cys Ser Thr Lys Cys Phe
1 5 10 15

Leu Leu Cys Gly Leu Gly Gly Thr His Val Thr Phe Val Ser Cys Ser
20 25 30

Lys Ala Gly Ala Thr Pro Ser Ser Leu Phe Ser Thr Gln His Gln Ala
35 40 45

Leu Ser Arg His Pro Ile Asn His Cys
50 55

What is claimed is:

1. 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 202 to nucleotide 759;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 391 to nucleotide 759;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone AM795_4 deposited under accession number ATCC 98271;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone AM795_4 deposited under accession number ATCC 98271;

(f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone AM795_4 deposited under accession number ATCC 98271;

(g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone AM795_4 deposited under accession number ATCC 98271;

(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. A composition of claim 1 wherein said polynucleotide is operably linked to at least one expression control sequence.

3. A host cell transformed with a composition 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 a composition 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 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 53 to amino acid 186;

(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 AM795_4 deposited under accession number ATCC 98271;

the protein being substantially free from other mammalian proteins.

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

10. The composition of claim 8, wherein said protein comprises the amino acid sequence of SEQ ID NO:2 from amino acid 53 to amino acid 186.

11. The composition of claim 8, further comprising 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 or SEQ ID NO:3.

14. 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 19 to nucleotide 262;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 91 to nucleotide 262;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone AT340_1 deposited under accession number ATCC 98271;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone AT340_1 deposited under accession number ATCC 98271;

(f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone AT340_1 deposited under accession number ATCC 98271;

(g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone AT340_1 deposited under accession number ATCC 98271;

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

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

(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).

15. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

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

(b) the amino acid sequence of SEQ ID NO:6 from amino acid 1 to amino acid 66;

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

(d) the amino acid sequence encoded by the cDNA insert of clone AT340_1 deposited under accession number ATCC 98271;

the protein being substantially free from other mammalian proteins.

16. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:5 or SEQ ID NO:4.

17. A composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 2 to nucleotide 601;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 401 to nucleotide 601;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone BG132_1 deposited under accession number ATCC 98271;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone BG132_1 deposited under accession number ATCC 98271;

(f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone BG132_1 deposited under accession number ATCC 98271;

(g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone BG132_1 deposited under accession number ATCC 98271;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:8;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:8 having biological activity;

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

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

18. 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:8;

(b) the amino acid sequence of SEQ ID NO:8 from amino acid 119 to amino acid 200;

(c) fragments of the amino acid sequence of SEQ ID NO:8; and

(d) the amino acid sequence encoded by the cDNA insert of clone BG132_1 deposited under accession number ATCC 98271;

the protein being substantially free from other mammalian proteins.

19. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:7 or SEQ ID NO:9.

20. 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 225 to nucleotide 701;

(c) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone BG219_2 deposited under accession number ATCC 98271;

(d) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone BG219_2 deposited under accession number ATCC 98271;

(e) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone BG219_2 deposited under accession number ATCC 98271;

(f) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone BG219_2 deposited under accession number ATCC 98271;

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

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

21. 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 1 to amino acid 97;

(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 BG219_2 deposited under accession number ATCC 98271;

the protein being substantially free from other mammalian proteins.

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

23. 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 2115 to nucleotide 2510;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:12 from nucleotide 1 to nucleotide 324;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone BG366_2 deposited under accession number ATCC 98271;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone BG366_2 deposited under accession number ATCC 98271;

(f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone BG366_2 deposited under accession number ATCC 98271;

(g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone BG366_2 deposited under accession number ATCC 98271;

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

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

24. 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) fragments of the amino acid sequence of SEQ ID NO:13; and

(c) the amino acid sequence encoded by the cDNA insert of clone BG366_2 deposited under accession number ATCC 98271;

the protein being substantially free from other mammalian proteins.

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

26. 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 27 to nucleotide 215;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:14 from nucleotide 27 to nucleotide 181;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone BV172_2 deposited under accession number ATCC 98271;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone BV172_2 deposited under accession number ATCC 98271;

(f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone BV172_2 deposited under accession number ATCC 98271;

(g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone BV172_2 deposited under accession number ATCC 98271;

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

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

(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).

27. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

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

(b) the amino acid sequence of SEQ ID NO:15 from amino acid 1 to amino acid 51;

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

(d) the amino acid sequence encoded by the cDNA insert of clone BV172_2 deposited under accession number ATCC 98271;

the protein being substantially free from other mammalian proteins.

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

29. 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 338 to nucleotide 409;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:16 from nucleotide 362 to nucleotide 409;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:16 from nucleotide 270 to nucleotide 419;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CC247_10 deposited under accession number ATCC 98271;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CC247_10 deposited under accession number ATCC 98271;

(g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CC247_10 deposited under accession number ATCC 98271;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CC247_10 deposited under accession number ATCC 98271;

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

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:17 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 composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

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

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

(c) the amino acid sequence encoded by the cDNA insert of clone CC247_10 deposited under accession number ATCC 98271;

the protein being substantially free from other mammalian proteins.

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

32. A composition comprising an isolated polynucleotide selected from the group consisting of:

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

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:18 from nucleotide 128 to nucleotide 1600;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:18 from nucleotide 281 to nucleotide 1600;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:18 from nucleotide 62 to nucleotide 373;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CI480_9 deposited under accession number ATCC 98271;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CI480_9 deposited under accession number ATCC 98271;

(g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CI480_9 deposited under accession number ATCC 98271;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CI480_9 deposited under accession number ATCC 98271;

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

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:19 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).

33. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

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

(b) the amino acid sequence of SEQ ID NO:19 from amino acid 1 to amino acid 82;

(c) fragments of the amino acid sequence of SEQ ID NO:19; and

(d) the amino acid sequence encoded by the cDNA insert of clone CI480_9 deposited under accession number ATCC 98271;

the protein being substantially free from other mammalian proteins.

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

35. A composition comprising an isolated polynucleotide selected from the group consisting of:

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

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:20 from nucleotide 383 to nucleotide 3958;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:20 from nucleotide 470 to nucleotide 3958;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:20 from nucleotide 271 to nucleotide 488;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CO722_1 deposited under accession number ATCC 98271;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CO722_1 deposited under accession number ATCC 98271;

(g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CO722_1 deposited under accession number ATCC 98271;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CO722_1 deposited under accession number ATCC 98271;

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

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

(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).

36. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

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

(b) the amino acid sequence of SEQ ID NO:21 from amino acid 1 to amino acid 34;

(c) fragments of the amino acid sequence of SEQ ID NO:21; and

(d) the amino acid sequence encoded by the cDNA insert of clone CO722_1 deposited under accession number ATCC 98271;

the protein being substantially free from other mammalian proteins.

37. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:20.

38. A composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:22;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:22 from nucleotide 914 to nucleotide 2353;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:22 from nucleotide 1793 to nucleotide 2353;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:22 from nucleotide 1037 to nucleotide 1260;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CT748_2 deposited under accession number ATCC 98271;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CT748_2 deposited under accession number ATCC 98271;

(g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CT748_2 deposited under accession number ATCC 98271;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CT748_2 deposited under accession number ATCC 98271;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:23;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:23 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).

39. 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:23;

(b) the amino acid sequence of SEQ ID NO:23 from amino acid 22 to amino acid 116;

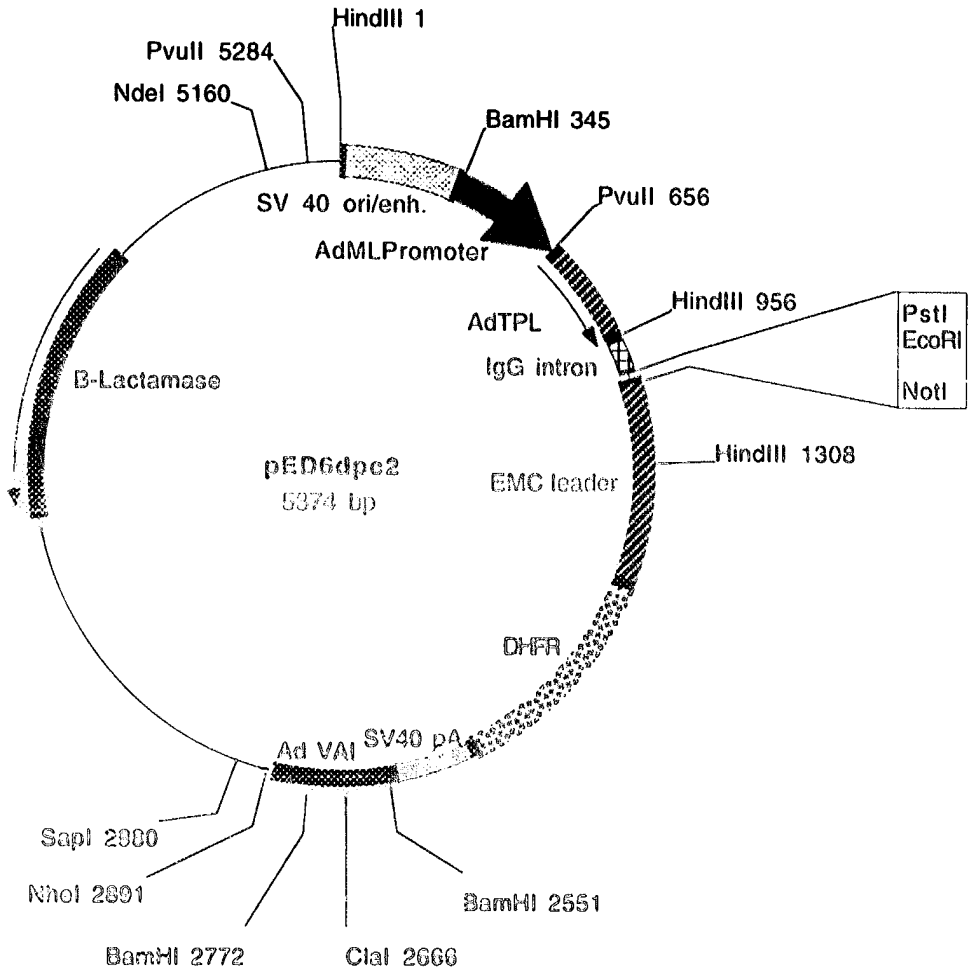
(c) fragments of the amino acid sequence of SEQ ID NO:23; and

(d) the amino acid sequence encoded by the cDNA insert of clone CT748_2 deposited under accession number ATCC 98271;

the protein being substantially free from other mammalian proteins.

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

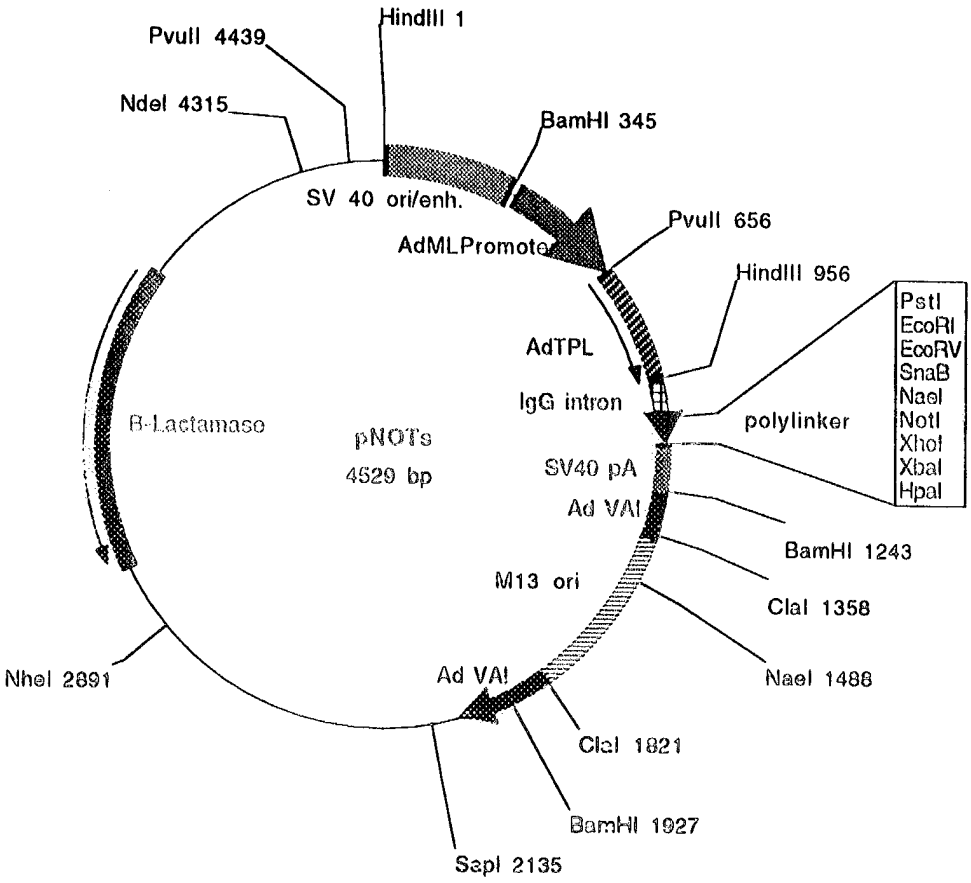
FIGURE 1A



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