PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :		(11) International Publication Number:	WO 98/44113
C12N 15/12, C07K 14/47, A61K 38/17	A1	(43) International Publication Date:	8 October 1998 (08.10.98)

(21) International Application Number: PCT/US98/06176

(22) International Filing Date: 27 March 1998 (27.03.98)

(30) Priority Data:

08/823,330	28 March 1997 (28.03.97)	US
Not furnished	25 March 1998 (25.03.98)	US

(71) Applicant: GENETICS INSTITUTE, INC. [US/US]; 87 Cambridge Park 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.; 113 Ann Lee Road, Harvard, MA 01451 (US). RACIE, Lisa, A.; 124 School Street, Acton, MA 01720 (US). MERBERG, David; 2 Orchard Drive, Acton, MA 01720 (US). TREACY, Maurice; 93 Walcott Road, Chestnut Hill, MA 02167 (US). SPAULDING, Vikki; 11 Meadowbank Road, Billerica, MA 01821 (US). AGOSTINO, Michael, J.; 26 Wolcott Avenue, Andover, MA 01810 (US).

(74) Agent: SPRUNGER, Suzanne, A.; Genetics Institute, Inc., 87 CambridgePark Drive, Cambridge, MA 02140 (US).

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

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

(57) Abstract

Polynucleotides and the proteins encoded thereby are disclosed.

FOR THE PURPOSES OF INFORMATION ONLY

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

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

SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

This application is a continuation-in-part of application Ser. No. 60/XXX,XXX (converted to a provisional application from non-provisional application Ser. No. 08/823,330), filed March 28, 1997, which is incorporated by reference herein.

FIELD OF THE INVENTION

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

BACKGROUND OF THE INVENTION

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

SUMMARY OF THE INVENTION

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

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

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 170 to nucleotide 322;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 218 to nucleotide 322;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 1814 to nucleotide 2355;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone bl209_10 deposited under accession number ATCC 98379;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone bl209_10 deposited under accession number ATCC 98379;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone bl209_10 deposited under accession number ATCC 98379;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone bl209_10 deposited under accession number ATCC 98379;

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

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:2 having biological activity, the fragment comprising the amino acid sequence from amino acid 20 to amino acid 29 of SEQ ID NO:2;

(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:1 from nucleotide 170 to nucleotide 322; the nucleotide sequence of SEQ ID NO:1 from nucleotide 218 to nucleotide 322; the nucleotide sequence of SEQ ID NO:1 from nucleotide 1814 to nucleotide 2355; the nucleotide sequence of the full-length protein coding sequence of clone bl209_10 deposited under accession number ATCC 98379; or the nucleotide sequence of a mature protein coding sequence of clone bl209_10 deposited under accession number ATCC 98379. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone bl209_10 deposited under accession number ATCC 98379.

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

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

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

(b) fragments of the amino acid sequence of SEQ ID NO:2 comprising the amino acid sequence from amino acid 20 to amino acid 29 of SEQ ID NO:2; and

(c) the amino acid sequence encoded by the cDNA insert of clone bl209_10 deposited under accession number ATCC 98379;

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

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

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

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 102 to nucleotide 1295;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 162 to nucleotide 1295;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 804 to nucleotide 1184;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone cr1162_25 deposited under accession number ATCC 98379;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone cr1162_25 deposited under accession number ATCC 98379;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone cr1162_25 deposited under accession number ATCC 98379;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone cr1162_25 deposited under accession number ATCC 98379;

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

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:4 having biological activity, the fragment comprising the amino acid sequence from amino acid 194 to amino acid 203 of SEQ ID NO:4;

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

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

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

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:3 from nucleotide 102 to nucleotide 1295; the nucleotide sequence of SEQ ID NO:3 from nucleotide 162 to nucleotide 1295; the nucleotide sequence of SEQ ID NO:3 from nucleotide 804 to nucleotide 1184; the nucleotide sequence of the full-length protein coding sequence of clone cr1162_25 deposited under accession number ATCC 98379; or the nucleotide sequence of a mature protein coding sequence of clone cr1162_25 deposited under accession number ATCC 98379. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone cr1162_25 deposited under accession number ATCC 98379. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:4 from amino acid 236 to amino acid 361.

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

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

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

(b) the amino acid sequence of SEQ ID NO:4 from amino acid 236 to amino acid 361;

(c) fragments of the amino acid sequence of SEQ ID NO:4 comprising the amino acid sequence from amino acid 194 to amino acid 203 of SEQ ID NO:4; and

(d) the amino acid sequence encoded by the cDNA insert of clone cr1162_25 deposited under accession number ATCC 98379;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:4 or the amino acid sequence of SEQ ID NO:4 from amino acid 236 to amino acid 361.

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 351 to nucleotide 842;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 687 to nucleotide 842;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 1 to nucleotide 689;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone dh40_3 deposited under accession number ATCC 98379;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone dh40_3 deposited under accession number ATCC 98379;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone dh40_3 deposited under accession number ATCC 98379;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone dh40_3 deposited under accession number ATCC 98379;

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

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:6 having biological activity, the fragment comprising the amino acid sequence from amino acid 77 to amino acid 86 of SEQ ID NO:6;

(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:5 from nucleotide 351 to nucleotide 842; the nucleotide sequence of SEQ ID NO:5 from nucleotide 687 to nucleotide 842; the nucleotide sequence of SEQ ID NO:5 from nucleotide 1 to nucleotide 689; the nucleotide sequence of the full-length protein coding sequence of clone dh40_3 deposited under accession number ATCC 98379; or the nucleotide sequence of a mature protein coding sequence of clone dh40_3 deposited under accession number ATCC 98379. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone dh40_3 deposited under accession number ATCC 98379. 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 113.

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

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 113;

(c) fragments of the amino acid sequence of SEQ ID NO:6 comprising the amino acid sequence from amino acid 77 to amino acid 86 of SEQ ID NO:6; and

(d) the amino acid sequence encoded by the cDNA insert of clone dh40_3 deposited under accession number ATCC 98379;

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 113.

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 2205 to nucleotide 2882;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 2262 to nucleotide 2882;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 2494 to nucleotide 3120;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone di39_9 deposited under accession number ATCC 98379;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone di39_9 deposited under accession number ATCC 98379;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone di39_9 deposited under accession number ATCC 98379;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone di39_9 deposited under accession number ATCC 98379;

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

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:8 having biological activity, the fragment comprising the amino acid sequence from amino acid 108 to amino acid 117 of SEQ ID NO:8;

(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:7 from nucleotide 2205 to nucleotide 2882; the nucleotide sequence of SEQ ID NO:7 from nucleotide 2262 to nucleotide 2882; the nucleotide sequence of SEQ ID NO:7 from nucleotide 2494 to nucleotide 3120; the nucleotide sequence of the full-length protein coding sequence of clone di39_9 deposited under accession number ATCC 98379; or the nucleotide sequence of a mature protein coding sequence of clone di39_9 deposited under accession number ATCC 98379. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone di39_9 deposited under accession number ATCC 98379.

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

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

(a) the amino acid sequence of SEQ ID NO:8;

(b) fragments of the amino acid sequence of SEQ ID NO:8 comprising the amino acid sequence from amino acid 108 to amino acid 117 of SEQ ID NO:8; and

(c) the amino acid sequence encoded by the cDNA insert of clone di39_9 deposited under accession number ATCC 98379;

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

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:9;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9 from nucleotide 40 to nucleotide 1503;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9 from nucleotide 863 to nucleotide 1377;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone dt674_2 deposited under accession number ATCC 98379;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone dt674_2 deposited under accession number ATCC 98379;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone dt674_2 deposited under accession number ATCC 98379;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone dt674_2 deposited under accession number ATCC 98379;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:10;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:10 having biological activity, the fragment comprising the amino acid sequence from amino acid 238 to amino acid 247 of SEQ ID NO:10;

(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:9 from nucleotide 40 to nucleotide 1503; the nucleotide sequence of SEQ ID NO:9 from nucleotide 863 to nucleotide 1377; the nucleotide sequence of the full-length protein coding sequence of clone dt674_2 deposited under accession number ATCC 98379; or the nucleotide sequence of a mature protein coding sequence of clone dt674_2 deposited under accession number ATCC 98379. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone dt674_2 deposited under accession number ATCC 98379. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:10 from amino acid 277 to amino acid 446.

Other embodiments provide the gene corresponding to the cDNA sequence of 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:10;

(b) the amino acid sequence of SEQ ID NO:10 from amino acid 277 to amino acid 446;

(c) fragments of the amino acid sequence of SEQ ID NO:10 comprising the amino acid sequence from amino acid 238 to amino acid 247 of SEQ ID NO:10; and

(d) the amino acid sequence encoded by the cDNA insert of clone dt674_2 deposited under accession number ATCC 98379;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:10 or the amino acid sequence of SEQ ID NO:10 from amino acid 277 to amino acid 446.

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:11;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11 from nucleotide 85 to nucleotide 450;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11 from nucleotide 217 to nucleotide 450;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone eh61_1 deposited under accession number ATCC 98379;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone eh61_1 deposited under accession number ATCC 98379;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone eh61_1 deposited under accession number ATCC 98379;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone eh61_1 deposited under accession number ATCC 98379;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:12;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:12 having biological activity, the fragment comprising the amino acid sequence from amino acid 55 to amino acid 64 of SEQ ID NO:12;

(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:11 from nucleotide 85 to nucleotide 450; the nucleotide sequence of SEQ ID NO:11 from nucleotide 217 to nucleotide 450; the nucleotide sequence of the full-length protein coding sequence of clone eh61_1 deposited under accession number ATCC 98379; or the nucleotide sequence of a mature protein coding sequence of clone eh61_1 deposited under accession number ATCC 98379. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone eh61_1 deposited under accession number ATCC 98379. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:12 from amino acid 9 to amino acid 94.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:11 or 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:12;

(b) the amino acid sequence of SEQ ID NO:12 from amino acid 9 to amino acid 94;

(c) fragments of the amino acid sequence of SEQ ID NO:12 comprising the amino acid sequence from amino acid 55 to amino acid 64 of SEQ ID NO:12; and

(d) the amino acid sequence encoded by the cDNA insert of clone eh61_1 deposited under accession number ATCC 98379;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:12 or the amino acid sequence of SEQ ID NO:12 from amino acid 9 to amino acid 94.

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 900 to nucleotide 1073;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:14 from nucleotide 544 to nucleotide 1022;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone fg265_1 deposited under accession number ATCC 98379;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone fg265_1 deposited under accession number ATCC 98379;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone fg265_1 deposited under accession number ATCC 98379;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone fg265_1 deposited under accession number ATCC 98379;

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

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:15 having biological activity, the fragment comprising the amino acid sequence from amino acid 24 to amino acid 33 of SEQ ID NO:15;

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

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

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

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:14 from nucleotide 900 to nucleotide 1073; the nucleotide sequence of SEQ ID NO:14 from nucleotide 544 to nucleotide 1022; the nucleotide sequence of the full-length protein coding sequence of clone fg265_1 deposited under accession number ATCC 98379; or the nucleotide sequence of a mature protein coding sequence of clone fg265_1 deposited under accession number ATCC 98379. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone fg265_1 deposited under accession number ATCC 98379. 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 41.

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 41;

(c) fragments of the amino acid sequence of SEQ ID NO:15 comprising the amino acid sequence from amino acid 24 to amino acid 33 of SEQ ID NO:15; and

(d) the amino acid sequence encoded by the cDNA insert of clone fg265_1 deposited under accession number ATCC 98379;

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 41.

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 119 to nucleotide 2440;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:16 from nucleotide 200 to nucleotide 2440;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:16 from nucleotide 460 to nucleotide 1153;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone fp273_10 deposited under accession number ATCC 98379;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone fp273_10 deposited under accession number ATCC 98379;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone fp273_10 deposited under accession number ATCC 98379;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone fp273_10 deposited under accession number ATCC 98379;

(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, the fragment comprising the amino acid sequence from amino acid 382 to amino acid 391 of SEQ ID NO:17;

(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 119 to nucleotide 2440; the nucleotide sequence of SEQ ID NO:16 from nucleotide 200 to nucleotide 2440; the nucleotide sequence of SEQ ID NO:16 from nucleotide 460 to nucleotide 1153; the nucleotide sequence of the full-length protein coding sequence of clone fp273_10 deposited under accession number ATCC 98379; or the nucleotide sequence of a mature protein coding sequence of clone fp273_10 deposited

under accession number ATCC 98379. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone fp273_10 deposited under accession number ATCC 98379. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:17 from amino acid 115 to amino acid 345.

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

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

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

(b) the amino acid sequence of SEQ ID NO:17 from amino acid 115 to amino acid 345;

(c) fragments of the amino acid sequence of SEQ ID NO:17 comprising the amino acid sequence from amino acid 382 to amino acid 391 of SEQ ID NO:17; and

(d) the amino acid sequence encoded by the cDNA insert of clone fp273_10 deposited under accession number ATCC 98379;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:17 or the amino acid sequence of SEQ ID NO:17 from amino acid 115 to amino acid 345.

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 1187 to nucleotide 1804;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:18 from nucleotide 674 to nucleotide 1014;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone fy243_8 deposited under accession number ATCC 98379;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone fy243_8 deposited under accession number ATCC 98379;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone fy243_8 deposited under accession number ATCC 98379;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone fy243_8 deposited under accession number ATCC 98379;

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

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:19 having biological activity, the fragment comprising the amino acid sequence from amino acid 98 to amino acid 107 of SEQ ID NO:19;

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

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

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

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:18 from nucleotide 1187 to nucleotide 1804; the nucleotide sequence of SEQ ID NO:18 from nucleotide 674 to nucleotide 1014; the nucleotide sequence of the full-length protein coding sequence of clone fy243_8 deposited under accession number ATCC 98379; or the nucleotide sequence of a mature protein coding sequence of clone fy243_8 deposited under accession number ATCC 98379. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone fy243_8 deposited under accession number ATCC 98379. 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 21 to amino acid 69.

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 21 to amino acid 69;

(c) fragments of the amino acid sequence of SEQ ID NO:19 comprising the amino acid sequence from amino acid 98 to amino acid 107 of SEQ ID NO:19; and

(d) the amino acid sequence encoded by the cDNA insert of clone fy243_8 deposited under accession number ATCC 98379;

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 21 to amino acid 69.

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 99 to nucleotide 536;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:20 from nucleotide 1 to nucleotide 370;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone ga205_4 deposited under accession number ATCC 98379;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone ga205_4 deposited under accession number ATCC 98379;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone ga205_4 deposited under accession number ATCC 98379;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone ga205_4 deposited under accession number ATCC 98379;

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

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:21 having biological activity, the fragment comprising the amino acid sequence from amino acid 68 to amino acid 77 of SEQ ID NO:21;

(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:20 from nucleotide 99 to nucleotide 536; the nucleotide sequence of SEQ ID NO:20 from nucleotide 1 to nucleotide 370; the nucleotide sequence of the full-length protein coding sequence of clone ga205_4 deposited under accession number ATCC 98379; or the nucleotide sequence of a mature protein coding sequence of clone ga205_4 deposited under accession number ATCC 98379. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone ga205_4 deposited under accession number ATCC 98379. 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 90.

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 90;

(c) fragments of the amino acid sequence of SEQ ID NO:21 comprising the amino acid sequence from amino acid 68 to amino acid 77 of SEQ ID NO:21; and

(d) the amino acid sequence encoded by the cDNA insert of clone ga205_4 deposited under accession number ATCC 98379;

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 90.

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

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

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

(b) purifying the protein from the culture.

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

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

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

BRIEF DESCRIPTION OF THE DRAWINGS

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

DETAILED DESCRIPTION

ISOLATED PROTEINS AND POLYNUCLEOTIDES

Nucleotide and amino acid sequences, as presently determined, are reported below for each clone and protein disclosed in the present application. The nucleotide sequence of each clone can readily be determined by sequencing of the deposited clone in accordance with known methods. The predicted amino acid sequence (both full-length

and mature forms) can then be determined from such nucleotide sequence. The amino acid sequence of the protein encoded by a particular clone can also be determined by expression of the clone in a suitable host cell, collecting the protein and determining its sequence. For each disclosed protein applicants have identified what they have determined to be the reading frame best identifiable with sequence information available at the time of filing.

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

Clone "bl209_10"

A polynucleotide of the present invention has been identified as clone "bl209_10". bl209_10 was isolated from a human adult testes cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. bl209_10 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "bl209_10 protein").

The nucleotide sequence of bl209_10 as presently determined is reported in SEQ ID NO:1. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the bl209_10 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:2. Amino acids 4 to 16 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 17, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone bl209_10 should be approximately 2400 bp.

The nucleotide sequence disclosed herein for bl209_10 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. bl209_10 demonstrated at least some similarity with sequences identified as AA522436 (ng30g05.s1 NCI_CGAP_Co3 Homo sapiens cDNA clone IMAGE 936344), L06147 (Human (clone SY11) golgin-95 mRNA, complete cds), N29620

(yw67d06.s1 Homo sapiens cDNA clone 257291 3'), N41622 (yw67d06.r1 Homo sapiens cDNA clone 257291 5'), N80172 (za65g07.s1 Homo sapiens cDNA clone 297468 3'), and U35022 (Rattus norvegicus cis-Golgi matrix protein GM130 mRNA, complete cds). The predicted amino acid sequence disclosed herein for bl209_10 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted bl209_10 protein demonstrated at least some similarity to sequences identified as M34651 (immediate-early protein [Suid herpesvirus]). Based upon sequence similarity, bl209_10 proteins and each similar protein or peptide may share at least some activity. [The TopPredII computer program predicts N potential transmembrane domains within the bl209_10 protein sequence, one around amino acid X and another around amino acid Y of SEQ ID NO:2.] [The nucleotide/amino acid sequence of bl209_10 indicates that it may contain an Alu repetitive element.]

Clone "cr1162_25"

A polynucleotide of the present invention has been identified as clone "cr1162_25". Secreted cDNA clones were first isolated from a human adult testes cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or were identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. These cDNA clones were then used to isolate cr1162_25, a full-length human cDNA clone which includes the entire coding sequence of a secreted protein (also referred to herein as "cr1162_25 protein"), from a human fetal brain cDNA library.

The nucleotide sequence of cr1162_25 as presently determined is reported in SEQ ID NO:3. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the cr1162_25 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:4. Amino acids 8 to 20 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 21, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone cr1162_25 should be approximately 3700 bp.

The nucleotide sequence disclosed herein for cr1162_25 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. cr1162_25 demonstrated at least some similarity with sequences identified as H14720 (ym24b05.r1 Homo sapiens cDNA clone 48883 5'), H15268

(ym30d11.r1 Homo sapiens cDNA clone 49904 5'), and N45514 (yy59g07.r1 Homo sapiens cDNA clone 277884 5'). The predicted amino acid sequence disclosed herein for cr1162_25 was searched against the GenPept, GeneSeq, and SwissProt amino acid sequence databases using the BLASTX search protocol. The predicted cr1162_25 protein demonstrated at least some similarity to sequences identified as D12612 (poliovirus receptor gene [Cercopithecus aethiops]), D26156 (hSNF2b; transcriptional activator [Homo sapiens], L12589 (B-lymphocyte activation antigen 7 [Mus musculus]), P51532 (POSSIBLE GLOBAL TRANSCRIPTION ACTIVATOR SNF2L3 (OR SNF2-BETA OR BRG-1) [Homo sapiens]), R07130 (H20B receptor), U29175 (transcriptional activator (BRG1)) [Homo sapiens]), X57516 (poliovirus receptor alpha [Homo sapiens]), X60958 (B lymphocyte activation antigen [Mus musculus]), X64116 (poliovirus receptor alpha [Homo sapiens]), and X68274 (TAG-1/axonin-1 [Homo sapiens]). Based upon sequence similarity, cr1162_25 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts an additional potential transmembrane domain at the carboxy terminus of the cr1162_25 protein sequence, centered around amino acid 342 of SEQ ID NO:4.

Clone "dh40_3"

A polynucleotide of the present invention has been identified as clone "dh40_3". dh40_3 was isolated from a human fetal brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. dh40_3 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "dh40_3 protein").

The nucleotide sequence of dh40_3 as presently determined is reported in SEQ ID NO:5. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the dh40_3 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:6. Amino acids 100 to 112 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 113, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone dh40_3 should be approximately 3000 bp.

The nucleotide sequence disclosed herein for dh40_3 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. dh40_3 demonstrated at least some similarity with sequences identified as AG005063 (Homo sapiens genomic DNA, 21q region, clone T1957SpN11), Z67586 (H.sapiens DNA segment containing (CA) repeat), and Z74023 (Human DNA sequence from cosmid LUCA3 on chromosome 3p21.3. contains ESTs). Based upon sequence similarity, dh40_3 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts an additional potential transmembrane domain within the dh40_3 protein sequence at the extreme carboxy terminus of SEQ ID NO:6.

Clone "di39_9"

A polynucleotide of the present invention has been identified as clone "di39_9". di39_9 was isolated from a human adult testes cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. di39_9 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "di39_9 protein").

The nucleotide sequence of di39_9 as presently determined is reported in SEQ ID NO:7. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the di39_9 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:8. Amino acids 7 to 19 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 20, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone di39_9 should be approximately 3000 bp.

The nucleotide sequence disclosed herein for di39_9 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. di39_9 demonstrated at least some similarity with sequences identified as AA249116 (hfe0042.seq.F Human fetal heart, Lambda ZAP Express Homo sapiens cDNA 5'), AA598667 (ae40a05.s1 Gessler Wilms tumor Homo sapiens cDNA clone 898256 3'), N53166 (yv56e11.s1 Homo sapiens cDNA clone 246764 3'), N80292 (za96h08.s1 Homo sapiens cDNA clone 300447 3'), T86182 (JTV1 coding sequence), U24169 (Human

JTV-1 (JTV-1) mRNA, complete cds), U38964 (Human PMS2 related (hPMSR2) gene, complete cds), and W24630 (zb62g08.r1 Soares fetal lung NbHL19W Homo sapiens cDNA clone 308222 5'). The predicted amino acid sequence disclosed herein for di39_9 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted di39_9 protein demonstrated at least some similarity to sequences identified as U24169 (JTV-1 [Homo sapiens]), U38964 (hPMSR2 [Homo sapiens]), and W25776 (JTV1 protein). The positioning of the regions of similarity to hPMSR2 and JTV-1 relative to each other in the di39_9 sequence is quite similar to that of the JTV-1 and PMS2 sequences in the human genome. Based upon sequence similarity, di39_9 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts two additional potential transmembrane domains within the di39_9 protein sequence, one centered around amino acid 160 and another around amino acid 200 of SEQ ID NO:8.

Clone "dt674_2"

A polynucleotide of the present invention has been identified as clone "dt674_2". dt674_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. dt674_2 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "dt674_2 protein").

The nucleotide sequence of dt674_2 as presently determined is reported in SEQ ID NO:9. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the dt674_2 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:10.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone dt674_2 should be approximately 3500 bp.

The nucleotide sequence disclosed herein for dt674_2 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. dt674_2 demonstrated at least some similarity with sequences identified as T06736 (EST04625 Homo sapiens cDNA clone HFBDX78). The predicted amino acid sequence disclosed herein for dt674_2 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The

predicted dt674_2 protein demonstrated at least some similarity to sequences identified as Z72807 (ORF YGR023w [Saccharomyces cerevisiae]). Based upon sequence similarity, dt674_2 proteins and each similar protein or peptide may share at least some activity. The nucleotide sequence of dt674_2 indicates that it may contain at least one copy of one or more repetitive elements.

Clone "eh61_1"

A polynucleotide of the present invention has been identified as clone "eh61_1". eh61_1 was isolated from a human adult blood (peripheral blood mononuclear cells treated with granulocyte-colony stimulating factor *in vivo*) 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. eh61_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "eh61_1 protein").

The nucleotide sequence of the 5' portion of eh61_1 as presently determined is reported in SEQ ID NO:11. What applicants presently believe is the proper reading frame for the coding region is indicated in SEQ ID NO:12. The predicted amino acid sequence of the eh61_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:12. Amino acids 32 to 44 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 45, or are a transmembrane domain. Additional nucleotide sequence from the 3' portion of eh61_1, including the polyA tail, is reported in SEQ ID NO:13.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone eh61_1 should be approximately 2200 bp.

The nucleotide sequence disclosed herein for eh61_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. eh61_1 demonstrated at least some similarity with sequences identified as AA114131 (zn75g05.s1 Stratagene NT2 neuronal precursor 937230 Homo sapiens cDNA clone 564056 3' similar to contains Alu repetitive element;contains element TAR1 repetitive element), H53674 (yu38e03.r1 Homo sapiens cDNA clone 236092 5'), L24093 (Gorilla gorilla ADP-ribosyltransferase (NAD+) pseudogene, repeat region), N38129 (19356 Arabidopsis thaliana cDNA clone 219I8T7), T04321 (368 Arabidopsis thaliana cDNA clone), U45981 (Schizosaccharomyces pombe Ste20-related protein kinase

(shk2) gene, complete cds), and X97774 (A.thaliana mRNA for light represssible receptor protein kinase). The predicted amino acid sequence disclosed herein for eh61_1 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted eh61_1 protein demonstrated at least some similarity to sequences identified as D10152 (protein tyrosine-serine-threonine kinase [Arabidopsis thaliana]), L24521 (transformation-related protein [Homo sapiens]), and L76191 (interleukin-1 receptor-associated kinase [Homo sapiens]). Based upon sequence similarity, eh61_1 proteins and each similar protein or peptide may share at least some activity. The nucleotide sequence of eh61_1 indicates that it may contain an Alu repetitive element.

Clone "fg265_1"

A polynucleotide of the present invention has been identified as clone "fg265_1". fg265_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. fg265_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "fg265_1 protein").

The nucleotide sequence of fg265_1 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 fg265_1 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 fg265_1 should be approximately 3100 bp.

The nucleotide sequence disclosed herein for fg265_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. fg265_1 demonstrated at least some similarity with sequences identified as AA076592 (zm91h10.r1 Stratagene ovarian cancer (#937219) Homo sapiens cDNA clone 545347 5'), AA482600 (zt34a12.s1 Soares ovary tumor NbHOT Homo sapiens cDNA), N23393 (yx83d12.s1 Homo sapiens cDNA clone 268343 3'), R10011 (yf34g05.r1 Homo sapiens cDNA clone 128792 5'), R41186 (yf84c08.s1 Homo sapiens cDNA clone 29313 3'), and W87844 (zh68a05.r1 Soares fetal liver spleen 1NFLS S1 Homo sapiens cDNA

clone 417200 5'). Based upon sequence similarity, fg265_1 proteins and each similar protein or peptide may share at least some activity.

Clone "fp273_10"

A polynucleotide of the present invention has been identified as clone "fp273_10". fp273_10 was isolated from a human adult placenta 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. fp273_10 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "fp273_10 protein").

The nucleotide sequence of fp273_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 fp273_10 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:17. Amino acids 15 to 27 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 28, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone fp273_10 should be approximately 3800 bp.

The nucleotide sequence disclosed herein for fp273_10 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. fp273_10 demonstrated at least some similarity with sequences identified as R16387 (yf91g01.r1 Homo sapiens cDNA clone 29825 5'), R17806 (yg09b06.r1 Homo sapiens cDNA clone 31763 5'), and T65784 (yc11f10.s1 Homo sapiens cDNA clone 80395 3' similar to contains L1 repetitive element). Based upon sequence similarity, fp273_10 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts four additional potential transmembrane domains within the fp273_10 protein sequence, centered around amino acids 140, 530, 560, and 720 of SEQ ID NO:17, respectively. At amino acid 449 of SEQ ID NO:17, the fp273_10 protein has a C-5 cytosine-specific DNA methylase motif.

Clone "fy243_8"

A polynucleotide of the present invention has been identified as clone "fy243_8". fy243_8 was isolated from a human adult placenta 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. fy243_8 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "fy243_8 protein").

The nucleotide sequence of fy243_8 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 fy243_8 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:19. Additional open reading frames for fy243_8 are predicted at basepairs 297 to 635, at basepairs 826 to 1014, and at basepairs 1102 to 1248 of SEQ ID NO:18; the predicted amino acid sequences corresponding to the foregoing nucleotide sequences are reported in SEQ ID NO:32, SEQ ID NO:33, and SEQ ID NO:34, respectively. The open reading frame for SEQ ID NO:19 could be joined to those for SEQ ID NO:32, SEQ ID NO:33, and SEQ ID NO:34 if the intervening nucleotide sequences of SEQ ID NO:18 were removed.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone fy243_8 should be approximately 2200 bp.

The nucleotide sequence disclosed herein for fy243_8 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. fy243_8 demonstrated at least some similarity with sequences identified as AA121177 (zl88h03.s1 Stratagene colon (#937204) Homo sapiens cDNA clone 511733 3'), AA121218 (zl88h03.r1 Stratagene colon (#937204) Homo sapiens cDNA clone 511733 5' similar to WP F44B9.5 CE00552), AA126582 (zn86g12.s1 Stratagene lung carcinoma 937218 Homo sapiens cDNA clone 565126 3'), R73372 (yl10g08.r1 Homo sapiens cDNA clone 157886 5' similar to SP F44B9.5 CE00552), T27033 (NIBT173E09R Infant brain, LLNL array of Dr. M. Soares 1NIB Homo sapiens cDNA clone LLAB173E09 5'end), and U41736 (Mus musculus ancient ubiquitous 46 kDa protein AUP1 precursor (Aup1) mRNA, complete cds). The predicted amino acid sequence disclosed herein for fy243_8 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted fy243_8 protein demonstrated at least some similarity to sequences identified as U41736 (ancient ubiquitous 46 kDa protein AUP46 precursor [Mus musculus]). Based upon sequence similarity, fy243_8 proteins and each similar protein or peptide may share at least some activity.

Clone "ga205_4"

A polynucleotide of the present invention has been identified as clone "ga205_4". ga205_4 was isolated from a human adult testes cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. ga205_4 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "ga205_4 protein").

The nucleotide sequence of ga205_4 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 ga205_4 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:21.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone ga205_4 should be approximately 1000 bp.

The nucleotide sequence disclosed herein for ga205_4 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. ga205_4 demonstrated at least some similarity with sequences identified as AA075247 (zm86e01.r1 Stratagene ovarian cancer (#937219) Homo sapiens cDNA clone 544824 5'), AA081273 (zn33e12.s1 Stratagene endothelial cell 937223 Homo sapiens cDNA clone 549262 3'), AA203476 (zx55e01.r1 Soares fetal liver spleen 1NFLS S1 Homo sapiens cDNA clone 446424 5' similar to contains element L1 repetitive element), T21011 (Human gene signature HUMGS02293), and U73030 (Rattus norvegicus pituitary tumor-specific transforming factor mRNA, complete cds). The predicted amino acid sequence disclosed herein for ga205_4 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted ga205_4 protein demonstrated at least some similarity to sequences identified as U73030 (PTTG gene product [Rattus norvegicus]). Based upon sequence similarity, ga205_4 proteins and each similar protein or peptide may share at least some activity.

Deposit of Clones

Clones bl209_10, cr1162_25, dh40_3, di39_9, dt674_2, eh61_1, fg265_1, fp273_10, fy243_8, and ga205_4 were deposited on March 28, 1997 with the American Type Culture Collection as an original deposit under the Budapest Treaty and were given the accession number ATCC 98379, 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), and the term of the deposit will comply with 37 C.F.R. § 1.806.

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
bl209_10	SEQ ID NO:22
cr1162_25	SEQ ID NO:23
dh40_3	SEQ ID NO:24
di39_9	SEQ ID NO:25
dt674_2	SEQ ID NO:26
eh61_1	SEQ ID NO:27

fg265_1	SEQ ID NO:28
fp273_10	SEQ ID NO:29
fy243_8	SEQ ID NO:30
ga205_4	SEQ ID NO:31

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

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

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

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

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

The bacterial culture containing the pool of full-length clones should preferably be thawed and 100 µl of the stock used to inoculate a sterile culture flask containing 25 ml of sterile L-broth containing ampicillin at 100 µg/ml. The culture should preferably be grown to saturation at 37°C, and the saturated culture should preferably be diluted in fresh L-broth. Aliquots of these dilutions should preferably be plated to determine the dilution and volume which will yield approximately 5000 distinct and well-separated colonies on solid bacteriological media containing L-broth containing ampicillin at 100 µg/ml and agar at 1.5% in a 150 mm petri dish when grown overnight at 37°C. Other known methods of obtaining distinct, well-separated colonies can also be employed.

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

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

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

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

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

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

Organisms that have enhanced, reduced, or modified expression of the gene(s) corresponding to the polynucleotide sequences disclosed herein are provided. The desired change in gene expression can be achieved through the use of antisense polynucleotides or ribozymes that bind and/or cleave the mRNA transcribed from the gene (Albert and Morris, 1994, *Trends Pharmacol. Sci.* **15(7)**: 250-254; Lavarosky *et al.*, 1997, *Biochem. Mol. Med.* **62(1)**: 11-22; and Hampel, 1998, *Prog. Nucleic Acid Res. Mol. Biol.* **58**: 1-39; all of which are incorporated by reference herein). Transgenic animals that have multiple copies of the gene(s) corresponding to the polynucleotide sequences disclosed herein, preferably produced by transformation of cells with genetic constructs that are stably maintained within the transformed cells and their progeny, are provided. Transgenic animals that have modified genetic control regions that increase or reduce gene expression levels, or that change temporal or spatial patterns of gene expression, are also provided (see European Patent No. 0 649 464 B1, incorporated by reference herein). In addition, organisms are provided in which the gene(s) corresponding to the polynucleotide sequences disclosed herein have been partially or completely inactivated, through insertion of extraneous sequences into the corresponding gene(s) or through deletion of all or part of the corresponding gene(s). Partial or complete gene inactivation can be accomplished through insertion, preferably followed by imprecise excision, of transposable elements (Plasterk, 1992, *Bioessays* **14(9)**: 629-633; Zwaal *et al.*, 1993, *Proc. Natl. Acad. Sci. USA* **90(16)**: 7431-7435; Clark *et al.*, 1994, *Proc. Natl. Acad. Sci. USA* **91(2)**: 719-722; all of which are incorporated by reference herein), or through homologous recombination,

preferably detected by positive/negative genetic selection strategies (Mansour *et al.*, 1988, *Nature* **336**: 348-352; U.S. Patent Nos. 5,464,764; 5,487,992; 5,627,059; 5,631,153; 5,614, 396; 5,616,491; and 5,679,523; all of which are incorporated by reference herein). These organisms with altered gene expression are preferably eukaryotes and more preferably are mammals. Such organisms are useful for the development of non-human models for the study of disorders involving the corresponding gene(s), and for the development of assay systems for the identification of molecules that interact with the protein product(s) of the corresponding gene(s).

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

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

Species homologues of the disclosed polynucleotides and proteins are also provided by the present invention. As used herein, a "species homologue" is a protein or polynucleotide with a different species of origin from that of a given protein or polynucleotide, but with significant sequence similarity to the given protein or polynucleotide. Preferably, polynucleotide species homologues have at least 60% sequence identity (more preferably, at least 75% identity; most preferably at least 90% identity) with the given polynucleotide, and protein species homologues have at least 30% sequence identity (more preferably, at least 45% identity; most preferably at least 60% identity) with

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

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

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

The present invention also includes polynucleotides capable of hybridizing under reduced stringency conditions, more preferably stringent conditions, and most preferably highly stringent conditions, to polynucleotides described herein. Examples of stringency

conditions are shown in the table below: highly stringent conditions are those that are at least as stringent as, for example, conditions A-F; stringent conditions are at least as stringent as, for example, conditions G-L; and reduced stringency conditions are at least as stringent as, for example, conditions M-R.

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

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

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

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

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

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

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

A number of types of cells may act as suitable host cells for expression of the protein. Mammalian host cells include, for example, monkey COS cells, Chinese Hamster Ovary (CHO) cells, human kidney 293 cells, human epidermal A431 cells, human Colo205 cells, 3T3 cells, CV-1 cells, other transformed primate cell lines, normal diploid cells, cell

strains derived from in vitro culture of primary tissue, primary explants, HeLa cells, mouse L cells, BHK, HL-60, U937, HaK or Jurkat cells.

Alternatively, it may be possible to produce the protein in lower eukaryotes such as yeast or in prokaryotes such as bacteria. Potentially suitable yeast strains include *Saccharomyces cerevisiae, Schizosaccharomyces pombe, Kluyveromyces* strains, *Candida,* or any yeast strain capable of expressing heterologous proteins. Potentially suitable bacterial strains include *Escherichia coli, Bacillus subtilis, Salmonella typhimurium,* or any bacterial strain capable of expressing heterologous proteins. If the protein is made in yeast or bacteria, it may be necessary to modify the protein produced therein, for example by phosphorylation or glycosylation of the appropriate sites, in order to obtain the functional protein. Such covalent attachments may be accomplished using known chemical or enzymatic methods.

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

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

Alternatively, the protein of the invention may also be expressed in a form which will facilitate purification. For example, it may be expressed as a fusion protein, such as those of maltose binding protein (MBP), glutathione-S-transferase (GST) or thioredoxin

(TRX). Kits for expression and purification of such fusion proteins are commercially available from New England BioLab (Beverly, MA), Pharmacia (Piscataway, NJ) and InVitrogen, respectively. The protein can also be tagged with an epitope and subsequently purified by using a specific antibody directed to such epitope. One such epitope ("Flag") is commercially available from Kodak (New Haven, CT).

Finally, one or more reverse-phase high performance liquid chromatography (RP-HPLC) steps employing hydrophobic RP-HPLC media, e.g., silica gel having pendant methyl or other aliphatic groups, can be employed to further purify the protein. Some or all of the foregoing purification steps, in various combinations, can also be employed to provide a substantially homogeneous isolated recombinant protein. The protein thus purified is substantially free of other mammalian proteins and is defined in accordance with the present invention as an "isolated protein."

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

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

The proteins provided herein also include proteins characterized by amino acid sequences similar to those of purified proteins but into which modification are naturally provided or deliberately engineered. For example, modifications in the peptide or DNA sequences can be made by those skilled in the art using known techniques. Modifications of interest in the protein sequences may include the alteration, substitution, replacement, insertion or deletion of a selected amino acid residue in the coding sequence. For example, one or more of the cysteine residues may be deleted or replaced with another amino acid to alter the conformation of the molecule. Techniques for such alteration, substitution, replacement, insertion or deletion are well known to those skilled in the art

(see, e.g., U.S. Patent No. 4,518,584). Preferably, such alteration, substitution, replacement, insertion or deletion retains the desired activity of the protein.

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

USES AND BIOLOGICAL ACTIVITY

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

Research Uses and Utilities

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

described in Gyuris *et al.*, 1993, *Cell* **75**: 791-803 and in Rossi *et al.*, 1997, *Proc. Natl. Acad. Sci. USA* **94**: 8405-8410, all of which are incorporated by reference herein) 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: 2355 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

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

CGCTTTTTTT TTTTTTTTTT TTCAGAAGGA GGAAGCTCAT TATGTTTGGA TCACCCACAG 60

CTATAGATTC TAAAAATATT TTGGCTTTTT TTGAGGTGCT TTAGTAAAAT ATAACCCCAA 120

ATGATTCACT TGGACAAGTG GTCTTAACAG CAAGGAAAAC AAACACTTTA TGAAAACAGC 180

TATAAGCCTT CTGTCTTTTA TCTTTACTAT TTTCTCCGAG TCTGGCATGA AACAGATACA 240

CAGCAGCCTC CACAGGGGGT TAAGTARAGA ACCATCCAAG CATCACAGAG TGTCATCCAG 300

AATTCTGATG ACTTCCATTC GTTGACTCTG ATGCACAATA TGCCTGGCTT GGGATGCAGC 360

GACCATGATG CCCCTCCCAG AACAGACACT TGCAGAGTGT TCCAGGAACA GCAGCTCCCT 420

CCAGCCCCCA GCACAAGATG CACACATCTC AGAACAAGCC TCCATCCTTT TCCTAGAGAA 480

CTGAGCATAA ATAACTTGTT CTATATCTGG CTCCAAGTCC ATTTCTGTTC TGTCTTGGAG 540

TAGAGTCTTA GCTCCCAGTT TGTTTTAGGT CAACTTTCAG CACCTACTTC AGCTCACTTG 600

TTTGATTTAC TAAGCTCTTG CTTCTGTATA TTATCAAATG TAGGGATGTA GGGAGAATAA 660

AAGGATCTAG ATACTTGCTT TTAGGAGAGA TTAGAACAAA GCTGAAGGTG GAGGCATTAG 720

TTCCTAGGTC TTCAGATCTC AGAGCAAAGG ACCCACTCTG GAGCCTAAAT TCTATGAGAG 780

ACCACAGAGC AGCCTGAAAT CCAAAGGAGT TTTACACAGG AAAAAAAAAA TACTGTGAGG 840

ACTTACACTA AATAATAATG TTGTTTTGAA TGGGGTTGTG GGTAATTCCT ATATTCTTCT 900

TTATAACTTT TGTACTTTTC AAATTCCCTA ATGTGAACTC ACTACTTAGT AGGTCTGTAA 960

GCTTAAACAT TACTATGGCT TGGAATCTCA TTTCAAAAAA TCTTTAAAAT GGGGACAAGA 1020

GTAAAAATTT CTTAGCTTCT ATGGAAGAAT AAAATGAAAT TATAATGATA CAGTGCCTGG 1080

CATGTTGTGG TCGCTCAATA AACACTGCTT TCCTCCCCAT TGTCCTCCTC TTTATTCTGT 1140

TTCATTACAA GGTCAGCAGA TTGAATCAGG ACCAGCTGGG AGGGCTACTT CTATGAGAGA 1200

AGATCTGTCC ACAGTCATGG TTTTCAATGT TTAGTGCACC AGAATCACCT TGAGGGTTTG 1260

TTAAAACAGA CTGCTGAACA TAACACATCT ATGAGAATGG CCAAAATCCA GAACACCAAA 1320

TGCTGGTGAG GATGTGGAGC AATAAAAACT CTCATTTATT GCTGATGGCA ATGCAAAATG 1380

GTACAGCCAC TTTGGAAGAC AATTTGCCAA ATTTTTACAA AACTAAGTGT ACTCTTACCA 1440

TACAATCTAG CAATCATGCT CCCTGGTATT TACCTAAAGG AGTTAAAAAC TTATGTCTAG 1500

ACAGAAACCT GCATATGAAT GTTTATAGCA GTTTTTTTCA TAATTGCTAA ACTTTGGAAG 1560

TAACCAAGAT GCCCTTCAGC AGGTGAATGG ACAAATAAAC TGCAGTAGAT GCAGACAGTG 1620

GAATATCATT CTAGGCCATG AAGGCCGAAT TCGGCCTTCA TGGCCTAATT AAAGAAAGTC 1680

AGGATAAAAA TTTTAAAAAG CAGGCCACTG TCAGCAAAGC CTGGAGAAGT GGGGCCGGAG 1740

GYTCCGCCCC CATCATGTGC CTGCCACCCC TTCCCAGTCA TCCCTTTAYT CTTACAGTAG 1800

CAAATAAGAC CCCTGTCTAA TGGGGGGAGA CAAATGTGTA GACCCTTAGC CACCTTGGCC 1860

AGGGCTGACT CCTTAAATTT CTGGATGATG ATGATTGTTA TTTAATAGCC AGAGGCTCAT 1920

ATAATTGGCC TCTTTGGAAG AGGCCTCATG GCCTCCTTAC TCTCACCAAA GCAATTTTTC 1980

CCTCAGGGGG GCTCCCATCT TCTTACACAG AGAGGCAGCT GAGGCAGGAC AGTGGGGCTA 2040

ACTGTAGACC AGGCGAGGGC ACGGGCTGCT GGGGTGGCCC TGCTTCCCCA GTGTACATAT 2100

TGTATCTGTG TAACATTTTG TATATTCCAG GGGTAGGGCC GCCCCCTGTA TCATACCTAG 2160

CAGAGGTTGG AGCTGGCACA TGGGGAGGAG GTTCTAATAA TTATTTGGGG CTGGGAAACT 2220

TATTTATTGA TAGCATAGGA CAGAGGAAGG AGGCGGGGAT GGGGTCGTGG CGCCCTGGTG 2280

ATGCGACTCC TGTTTATTTT GCTTTTTATT TCGGAATAAA TGGATTTAGC CATAAAAAAA 2340

AAAAAAAAAA AAAAA 2355

(2) INFORMATION FOR SEQ ID NO:2:

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

(ii) MOLECULE TYPE: protein

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

Met Lys Thr Ala Ile Ser Leu Leu Ser Phe Ile Phe Thr Ile Phe Ser
1 5 10 15

Glu Ser Gly Met Lys Gln Ile His Ser Ser Leu His Arg Gly Leu Ser
20 25 30

Xaa Glu Pro Ser Lys His His Arg Val Ser Ser Arg Ile Leu Met Thr
35 40 45

Ser Ile Arg
50

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2496 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

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

GCGCCCTTTC GGTCAACATC GTAGTCCACC CCCTCCCCAT CCCCAGCCCC CGGGGATTCA 60

GGCTCGCCAG CGCCCAGCCA GGGAGCCGGC CGGGAAGCGC GATGGGGGCC CCAGCCGCCT 120

CGCTCCTGCT CCTGCTCCTG CTGTTCGCCT GCTGCTGGGC GCCCGGCGGG GCCAACCTCT 180

CCCAGGACGA CAGCCAGCCC TGGACATCTG ATGAAACAGT GGTGGCTGGT GGCACCGTGG 240

TGCTCAAGTG CCAAGTGAAA GATCACGAGG ACTCATCCCT GCAATGGTCT AACCCTGCTC 300

AGCAGACTCT CTACTTTGGG GAGAAGAGAG CCCTTCGAGA TAATCGAATT CAGCTGGTTA 360

CCTCTACGCC CCACGAGCTC AGCATCAGCA TCAGCAATGT GGCCCTGGCA GACGAGGGCG 420

AGTACACCTG CTCAATCTTC ACTATGCCTG TGCGAACTGC CAAGTCCCTC GTCACTGTGC 480

TAGGAATTCC ACAGAAGCCC ATCATCACTG GTTATAAATC TTCATTACGG GAAAAAGACA 540

CAGCCACCCT AAACTGTCAG TCTTCTGGGA GCAAGCCTGC AGCCCGGCTC ACCTGGAGAA 600

AGGGTGACCA AGAACTCCAC GGAGAACCAA CCCGCATACA GGAAGATCCC AATGGTAAAA 660

CCTTCACTGT CAGCAGCTCG GTGACATTCC AGGTTACCCG GGAGGATGAT GGGGCGAGCA 720

TCGTGTGCTC TGTGAACCAT GAATCTCTAA AGGGAGCTGA CAGATCCACC TCTCAACGCA 780

TTGAAGTTTT ATACACACCA ACTGCGATGA TTAGGCCAGA CCCTCCCCAT CCTCGTGAGG 840

GCCAGAAGCT GTTGCTACAC TGTGAGGGTC GCGGCAATCC AGTCCCCCAG CAGTACCTAT 900

GGGAGAAGGA GGGCAGTGTG CCACCCCTGA AGATGACCCA GGAGAGTGCC CTGATCTTCC 960

CTTTCCTCAA CAAGAGTGAC AGTGGCACCT ACGGCTGCAC AGCCACCAGC AACATGGGCA 1020

GCTACAAGGC CTACTACACC CTCAATGTTA ATGACCCCAG TCCGGTGCCC TCCTCCTCCA 1080

GCACCTACCA CGCCATCATC GGTGGGATCG TGGCTTTCAT TGTCTTCCTG CTGCTCATCA 1140

TGCTCATCTT CCTCGGCCAC TACTTGATCC GGCACAAAGG AACCTACCTG ACACATGAGG 1200

CAAAAGGCTC CGACGATGCT CCAGACGCGG ACACGGCCAT CATCAATGCA GAAGGCGGGC 1260

AGTCAGGAGG GGACGACAAG AAGGAATATT TCATCTAGAG GCGCCTGCCC ACTTCCTGCG 1320

CCCCCCAGGG GCCCTGTGGG GACTGCTGGG GCCGTCACCA ACCCGGACTT GTACAGAGCA 1380

ACCGCAGGGC CGCCCCTCCC GCTTGCTCCC CAGCCCACCC ACCCCCCTGT ACAGAATGTC 1440

TGCTTTGGGT GCGGTTTTGT ACTCGGTTTG GAATGGGGAG GGAGGAGGGC GGGGGGAGGG 1500

GAGGGTTGCC CTCAGCCCTT TCCGTGGCTT CTCTGCATTT GGGTTATTAT TATTTTTGTA 1560

ACAATCCCAA ATCAAATCTG TCTCCAGGCT GGAGAGGCAG GAGCCCTGGG GTGAGAAAAG 1620

CAAAAAACAA ACAAAAAACA AAACCCTGGA GTGTTAGGAG GAGAGTGAAG GTAGAGGGGT 1680

GAGGAAGGGT AAGGGGCAGG GCTGGTTTCA GCTGGGGGCT CTCACCAGCC CTCCTTTCAG 1740

CCTCTACAAC AGAGCAGCTT CCCAGACTTC TCCAGGAACC CAGAAACGGG ATGGTTGTCG 1800

GCAAAGGTTG GGAGTGGCTT TTCCTCTGGT AGCCACACAC CTGAGCACTA CGGACAGGGA 1860

GGCAGGTGCC ACCTTGACAC CTCTCTTCCA TAGCAATGGG AAAGTGATGA GTGCGGGAGT 1920

CCTGAGGAGA TGTGGCCTGC AGACAACATG CAGCCATGCA GGGACCCAGG ACTGTAACCT 1980

GGGGAGGACG CGGGTCCCTG CAAGGAAGAG TAGATTTGGA GAGGAAGGAT GGAGGTGGAC 2040

TCTCACCCCA TTCCCCCCGG AAATGAACAA AGCCGGGCCC TTTCCATAGG AACTGCCCTT 2100

GGAGATAGCA GAGTGTGGCT GCCCCTCCTT GCTCCAGCAG CAGTGGGAGA GGCACTGCTC 2160

TGGGGCCTGA ACTGCCTCTG CTTCCCCCCC TGAGGGGCCC CTCACTCTTA CCCAAGACTC 2220

TGGATTGTTG CACGGCAACC ACTCCTCCCA TGGCATTGCT CAGCAACTAC TTCTCCCTTC 2280

CCGGCCACCC TGTGCCCCCT TCCTGGTCCC AACGCCAGCC CTTCATCCTT CCTCCCTCAG 2340

CAGCCAGGCA GACATAACAA CAAAACTACT AAAAGGAGCT TCAAAAAAAA AAAAAAAAAA 2400

AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 2460

AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAA 2496

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 398 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

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

```
Met Gly Ala Pro Ala Ala Ser Leu Leu Leu Leu Leu Leu Phe Ala
1               5                  10                  15

Cys Cys Trp Ala Pro Gly Gly Ala Asn Leu Ser Gln Asp Asp Ser Gln
            20                  25                  30

Pro Trp Thr Ser Asp Glu Thr Val Val Ala Gly Gly Thr Val Val Leu
        35                  40                  45

Lys Cys Gln Val Lys Asp His Glu Asp Ser Ser Leu Gln Trp Ser Asn
    50                  55                  60

Pro Ala Gln Gln Thr Leu Tyr Phe Gly Glu Lys Arg Ala Leu Arg Asp
65                  70                  75                  80

Asn Arg Ile Gln Leu Val Thr Ser Thr Pro His Glu Leu Ser Ile Ser
                85                  90                  95

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

Phe Thr Met Pro Val Arg Thr Ala Lys Ser Leu Val Thr Val Leu Gly
        115                 120                 125

Ile Pro Gln Lys Pro Ile Ile Thr Gly Tyr Lys Ser Ser Leu Arg Glu
    130                 135                 140

Lys Asp Thr Ala Thr Leu Asn Cys Gln Ser Ser Gly Ser Lys Pro Ala
145                 150                 155                 160

Ala Arg Leu Thr Trp Arg Lys Gly Asp Gln Glu Leu His Gly Glu Pro
                165                 170                 175

Thr Arg Ile Gln Glu Asp Pro Asn Gly Lys Thr Phe Thr Val Ser Ser
            180                 185                 190

Ser Val Thr Phe Gln Val Thr Arg Glu Asp Asp Gly Ala Ser Ile Val
        195                 200                 205

Cys Ser Val Asn His Glu Ser Leu Lys Gly Ala Asp Arg Ser Thr Ser
    210                 215                 220

Gln Arg Ile Glu Val Leu Tyr Thr Pro Thr Ala Met Ile Arg Pro Asp
225                 230                 235                 240

Pro Pro His Pro Arg Glu Gly Gln Lys Leu Leu Leu His Cys Glu Gly
                245                 250                 255

Arg Gly Asn Pro Val Pro Gln Gln Tyr Leu Trp Glu Lys Glu Gly Ser
            260                 265                 270
```

```
Val Pro Pro Leu Lys Met Thr Gln Glu Ser Ala Leu Ile Phe Pro Phe
        275                 280                 285

Leu Asn Lys Ser Asp Ser Gly Thr Tyr Gly Cys Thr Ala Thr Ser Asn
    290                 295                 300

Met Gly Ser Tyr Lys Ala Tyr Tyr Thr Leu Asn Val Asn Asp Pro Ser
305                 310                 315                 320

Pro Val Pro Ser Ser Ser Ser Thr Tyr His Ala Ile Ile Gly Gly Ile
                325                 330                 335

Val Ala Phe Ile Val Phe Leu Leu Leu Ile Met Leu Ile Phe Leu Gly
            340                 345                 350

His Tyr Leu Ile Arg His Lys Gly Thr Tyr Leu Thr His Glu Ala Lys
        355                 360                 365

Gly Ser Asp Asp Ala Pro Asp Ala Asp Thr Ala Ile Ile Asn Ala Glu
    370                 375                 380

Gly Gly Gln Ser Gly Gly Asp Asp Lys Lys Glu Tyr Phe Ile
385                 390                 395
```

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2764 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

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

```
GGGCCAAAGA GGCCTACCAG CTGCTGTTGA CCGCTGGACT CACAAACCTT TCTTTCTACT     60

CTTGTTTTTC ATTCACTTTG GGTCATTTTT CAGTGTTGAT GGGGACGTAA TAAAGCACGG    120

TAAGAAAATC CGTGAATTCC GTCAGAGCAG TCGTCCAGAG GGAAGGCGCG CCCGGCGTAG    180

GGAGGTCAGA GCTCATGTTA GCTATGAACA CAGGTCACAG GGGCGTACGG CGATGGGAAA    240

CACTGAGATG CTCAATATAT TGATTATTTA ATAGTGTTTA GCAAAATGGT CTTTTTTTAT    300

TCCTTAAATC AACTGAAACT CACTTCACGT CTCTTTCCTT GTAGAGCATC ATGCTTATTT    360

CTGGCTCACT CACATCTTTG TCTCGGGAGT TCTCTGCCGA GCCATTGCCC CCTACAGCAG    420

AGAGCACAGC TGGCTGCACT AGTGCTGAAG GAGCCAGCCC CAGAGCAGGG CATTTCCAGG    480
```

GGCTCTTGTC CCAGAGCGGC AGGCGTTGTG TGCAGAGAAC GCCCCTCCCA CGCAGCACAG 540

AGAACGCGGG GTGGGTGTGT GGCTCCGGGC CTGTGGGGCT TAGGCTGCCT GAACCACCGC 600

CGACTGGCAC CATGACTCGG CATTCCTGGA AGTGCCTTAC CAAGTTGTTG TTGTTGTTTT 660

GTTGTTTTTT AAGAGACGGG CTTGCTCTAT CATCCAGGCT CGAGTGCAAT GGCACAGTCA 720

CAGCTCACTG CAGCCTTGAA CTCGTGGGCT CAAGCCATCC TCCTGTGTCA GCCTCCCCAG 780

TACCTGGGAC TGTGGGCATG AGCACTGCGC CTGGCAGCTG TATCAGTGTT GACTCCACAT 840

TTTAATAGTT GCTTCTTGAA ATTAAAATGC TTTGATTCAG CCTTCAAGCC ATCAGGAAAG 900

TTTGCCCCTC TGAGTCACAC CTGGTGGTCT CCAGGGTTCC TGCCCCTCCC TCCTGAGCCA 960

GCTCCTCAGA GCGGATAGAG GCAGGACCCC CACCCAGGTC TTGAGACCCC CCTGCCCCGC 1020

ACTCCCCCGG AGACGGGCTA CCCCTGCAGA TGCAGATAGT CAAAGCTCAG GTTTCTTCCA 1080

AAGCTTTTAA AAAGATATTG TACCTTGAGC ACTTTAAAAA TGTCTTAAAA TTGCCATACA 1140

GGCTCTTAAA AGCTTATACG TTTAAACTGT TGATAGATGG GCCTTTACTA AAATGCATTC 1200

ATTTATTTTC CTAATCCCTT GGTTGTTAAA TAATTCTGGG GAAGGGCCCC GAGCACGACA 1260

GCCGCAGTCT CCACCCAGAA CCAGAGAGTC CCCCCCAACC CGGGATGTAC CCTCTGGCCA 1320

CACCAGGGAC CCTGCCAGAG GCCGCAGACT GGCAGCAGCA GCCTCCCCAC ACAGTGGGGG 1380

AAGGTCAGTG TGATGCCTTC AGGCCCCGTC TCCTGCCAGG GCTCTCCCTC CAGCCTACAT 1440

AGGGCCTCAG AGAAATGCAT TTTTAGTTCT GGCTTTGGCC CAGCCCAGGG CAAGGCAGGA 1500

AACTCTCCAG CGTGAGTCCG TGAGGGCCAA GAAGTCCCGC CCTGTTCTGG GGGAGGACCT 1560

GGCTTTTCTG GTGTCTCTGG TGCCCGAGAG CCCGGTGCTG CCATCTTTAG TGAAAGAGTA 1620

AATGGTGGCC GAGGGCTCCT TTTGTGAGGG ATGTGCCTTG GTGAAGAAGG CATGTTCCCT 1680

GCCGTGAAGA TACTTGGAAG CTCTGGGTGG AGAGGGAAAA GGGATACCCC TGGTGCTCCC 1740

TGGGCCTGGC GGAAGGCTAG GAGGAAGGAC AGCTGAGGTG AGGACTGAGT GGGGCAGGTA 1800

TCACCCTGAC AAACAGTTTG GGAAGATCAG GAAAGGCAGG TGAGACCTGG TGCAGAATCC 1860

AGGTTGGGTA ATAGATACAT CGTCGAAGAT GTAGCAAGCA AAGTAATATA CTCAACTCTG 1920

GAACATTGCA CAGAAGCTTT TAAAGCACTC TGTGACACTT TTTGTAATGA GGGATCTGAA 1980

GGAAACGGCC CCAGAGTCAC CCATCCCCAC GGGTCTGGTT GGCGGGGCTG GTGCCTTTCT 2040

TCTGCACTCA GTCACCATGG CTCCGTCTGT CAAACTCAAC TCTTTTTTTT TTTTTTTTTC 2100

TTCTCTTGGT GTGGTAATTT GTTTGAAGAG CCACTCCATC CCCAAATTCA AGATTAGAAA 2160

GATCCCTGAC TGCTTCTCAA GATCCAGAAC ATTCCTTGAC AGAGTATATT CACCATTTAG 2220

AAGTGATCCA GCAAAGATTG GGAGGGGTAC TACCAGATTC TACTTCAAAG AAATCCTGCC 2280

ACCCGATGAT TAAACAGTGA ATAAAATGTC ATGGCTCTTT CCTGCGACAA TTCTATTTGA 2340

GGAAAAGATT TGTTTTTCCC TTTTCCCAAG GAAGCTCGTG GGACAGCATG GGCACTACTC 2400

TTCATGTGCG GTGACACCAG CCCCCAGATG CCTTGAATTA AGTGTCCTCA CCTTTATGCA 2460

TGACTGCAAA GCCAGCTGGA GCATTTTCTA TGGAGCCTCC GTATGTTTTA GGCCCATGAC 2520

CTTCGTGAGG TGATGGGCAC TCACTCCCAT GAGCCCTGGC TGTGTGCTGT TGTGTGCCTA 2580

TCGGCAGATC CATCCTTCCT GCCTCCAAGG AGGATACACA GAGAATGGCT TCCTGTTGTT 2640

TTGTTTATTT TCTTAACGTG TACAGATGGA AACTTCATTT AAAAATAAAA ACAAAACAAY 2700

TCNAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 2760

AAAA 2764

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 164 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

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

Met Leu Ile Ser Gly Ser Leu Thr Ser Leu Ser Arg Glu Phe Ser Ala
1 5 10 15

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

Glu Gly Ala Ser Pro Arg Ala Gly His Phe Gln Gly Leu Leu Ser Gln
35 40 45

Ser Gly Arg Arg Cys Val Gln Arg Thr Pro Leu Pro Arg Ser Thr Glu
50 55 60

Asn Ala Gly Trp Val Cys Gly Ser Gly Pro Val Gly Leu Arg Leu Pro
65 70 75 80

Glu Pro Pro Pro Thr Gly Thr Met Thr Arg His Ser Trp Lys Cys Leu
85 90 95

Thr Lys Leu Leu Leu Leu Phe Cys Cys Phe Leu Arg Asp Gly Leu Ala
100 105 110

Leu Ser Ser Arg Leu Glu Cys Asn Gly Thr Val Thr Ala His Cys Ser
115 120 125

Leu Glu Leu Val Gly Ser Ser His Pro Pro Val Ser Ala Ser Pro Val
130 135 140

Pro Gly Thr Val Gly Met Ser Thr Ala Pro Gly Ser Cys Ile Ser Val
145 150 155 160

Asp Ser Thr Phe

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 3367 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

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

CAGAAGGGAG GTAGTCGCCC TCCGTCGTGG CCTGGCGTGG ATTCCGAGCG TTGGTGTCTG 60

GCGGTTTCCG ACCGTTGGTG TCTGGCACGC GCCACCCCGA TGTACCAGGT AAAGCCCTAT 120

CACGGGGTCG GCGCCCCTCT CCGTGTGGAG CCCACCTGCA TGTACTGGCT CCCCAACATG 180

CACGGCAGGA GCGGCGGCCC AGCACTCGGC ACTGGCCACT TGCAGACAAG AAGACAAGAA 240

AATGATTTGA GGACAGCTTC AATCGCGGTG TGAAGAAGAA AGCAACAAAA CGACCACTGA 300

AAACAATGCC GGTGGCAAAA CATCCAAAGA AAGGGTCCCA AGCGGTACAT CGTCATAGCT 360

GGAAACAGTC AGAGCCACCA GCCAATGATC TTTTCAATGC TGCGAAAGCT GCCAAAAGTG 420

ACATGCAGTG TGGCCATGAG GTCTGCCGGA AGTGACTTGT TGGTGTTATC TCCTGAGTTA 480

AAATGTGAAG GGATTTTTTT TTTTCAGATT ACTGAGAGTC TTCTGTTACT AGTTTGTCTT 540

TCCTAGATCC AGACACGGGG ACTGCAGAGA AAGGCTGTGT GCATCCGCTG TCTACTCCAC 600

TGTCTCCTCT GCAGAGGCGG ATTTCCCTGA CTGAAGACCA TGTTGCAGGC CCACAGCTGC 660

CTACAGAACC GTCCCAAAAT ATGGCAAAGA AACCTATTCT GAGGGTCTCA CCATGTTGCC 720

CAGGCTGGTC TTGAACTCCT GGACTCATCC TAAAGTGCTG GCCTCTCATT CCCTGTCTGT 780

GCACACCTCA CGGCAAGGGC CAGCCTGTTT CCTCCCGGTC ACCTCCAAAT CTTGCTGCTT 840

TTAATTCAAC TCAGAGGCCT AGCCAGGGTT GAGTTCTCAC CCACCTGTGC CGCCCTGCCT 900

TGTTACCTGG AAGCACAGCC TTGGGGACTG AGCAGGCCCT CACTGTCACT TTAAGAAGGG 960

AATCAGCCAC TTTGTGCTCA CCACCTCTGG GGAAGGTGTG AGAGGAGAGA AGGAAGTGGC 1020

TGTTTGGCTG CTGACAACAT GAAGACTTCC TGCGATGAGA ACAGAGGCAC AGGTGCCGGC 1080

CCTGCAGCCC CCAGAACCCG GACTGGAGGG GGCCATGGGG CGCCGGACCC TGGCCCTGCC 1140

CTGGGTGCTG CTGACCCTGC GTGTCACTGC AGGGACCCCG GAGGTGTGAG TACAAGTTCG 1200

GATGGAGGCC ACCGAGCTCT CGTCCTTCAC CATCCGTTGT GGGTTCCTGG AGTCTGGCTC 1260

CATCTCCCTG GTGACTGTGA GCTGGGGGGG CCCCGATGGT GCTGGGGGGA CCACGCTGGC 1320

TGTGTTGCAC CCGGAACTTG GCATCCAGCA ATGGGCCCCT GCTCGCCAGG CCCGCTGGGA 1380

AACCCAGAGC AGCGTCTCTC TTGCCCTGGA AGTCTCTGGG GCCAGCAGCC CCTGCACCAA 1440

CACCACCTTC TGCTGCAAGT TTGCGTCCTT CCCTGAGGGC TCCTGGGAGG CCTCTGGGAG 1500

CCTCCCGCCC AGCTCAGACC CAGGGCTCTC TGTCCCGCCG ACTCCTGCCC CCATTCTGCG 1560

GGCAGACCTG GCCGGGATCT TGGGGGTCTC AGGAGTCCTT CTCTTTGACT GTGGCTACCT 1620

CCTTCATCTG CTGTGCCGAC AGAAGCACCG CCCTGCCCCT AGGCTCCAGC CATCCCACAC 1680

CAGCTCCTAG GCACTGAGAG CACGAGCATG GGCACCCAGC CAGGCCTCCC AGGCTGCTCT 1740

CCACGTCCCT TATGCCACTA TCAACACCAG CTGCTGCCCA GCTACTTTGG ACACAGCTCA 1800

CCCCCGACAG GGGGCCGTCC TGTCGTTTCC TGCTGTGACT AAGTCAGCAA CACAGTTCCT 1860

CTGACATGGG CCTTGGCTGT GCTTCTTTGG GGGTGAAGAG ATTGGGGAGG AAGTCTCCAC 1920

CCCTGGGAGG CAGAAGCCAG GCATAGCGCG CTGGCTAGGA CTCCAGTACC GTGAAGGGAG 1980

GCAGTGAGAG CAGACATCTG TGTCTCATTC CTGATCTCAA GGGGAAAGCA AGAACAAGGG 2040

AGGCTTCCTC AGGATCTCAA ACCTGCGGAA GGAGGACCAG TCTGTGTACT TCTGCCAAGT 2100

CCAGCTGGAC ATACAGATCA GCCCTCAGGC AGCCCCTCCA CAGGACCCCT CTCCTGCCTG 2160

GACAGCTCTG CTGGTCTCCC CGTCCCCTGG AGAAGAACAA GGCCATGGGT CGGCCCCTGC 2220

TGCTGCCCCT GCTGCTCCTG CTGCAGCCGC CAGCATTTCT GCAGCCTGGT GGCTCCACAG 2280

GATCTGGTCC AAGCTACCTT TATGGGGTCA CTCAACCAAA ACACCTCTCA GCCTCCATGG 2340

GTGGCTCTGT GGAAATCCCC TTCTCCTTCT ATTACCCCTG GGAGTTAGCC ACAGCTCCCG 2400

ACGTGAGAAT ATCCTGGAGA CGGGGCCACT TCCACGGGCA GTCCTTCTAC AGCACAAGGC 2460

CGCCTTCCAT TCACAAGGAT TATGTGAACC GGCTCTTTCT GAACTGGACA GAGGGTCAGG 2520

AGAGCGGCTT CCTCAGGATC TCAAACCTGC GGAAGGAGGA CCAGTCTGTG TATTTCTGCC 2580

GAGTCGAGCT GGACACCCGG AGATCAGGGA GGCAGCAGTT GCAGTCCATC AAGGGGACCA 2640

AACTCACCAT CACCCAGGCT GTCACAACCA CCACCACCTG GACGCCCAGC AGCACAACCA 2700

CCATAGCCGG CCTCAGGGTC ACAGAAAGCA AAGGGCACTC AGAATCATGG CACCTAAGTC 2760

TGGACACTGC CATCAGGGTT GCATTGGCTG TCGCTGTGCT CAAAACTGTC ATTTTGGGAC 2820

TGCTGTGCCT CCTCCTGTGG TGGAGGAGAA GGAAAGGTAG CAGGGCGCCA AGCAGTGACT 2880

TCTGACCAAC AGAGTGTGGG GAGAAGGGAT GTGTATTAGC CCCGGAGGAC GTGATGTGAG 2940

ACCCGCTTGT GAGTCCTCCA CACTCGTTCC CCATTGGCAA GATACATGGA GAGCACCCTG 3000

AGGACCTTTA AAAGGCAAAG CCGCAAGGCA GAAGGAGGCT GGGTCCCTGA ATCACCGACT 3060

GGAGGAGAGT TACCTACAAG AGCCTTCATC CAGGAGCATC CACACTGCAA TGATATAGGA 3120

WTGAGGTCTG AACTCCACTG AATTAAACCA CTGGCATTTG GGGGCTGTTC ATTATAGCAG 3180

TGCAAAGAGT TCCTTTATCC TCCCCAAGGA TGGAAAATAC AATTTATTTT GCTTACCATA 3240

CACCCCTTTT CTCTTCGTCC ACATTTTCCA ATCTGTATGG TGGCTGTCTT CTATGGCAGA 3300

AGGTTTTGGG GAATAAATAG CGTGAAATGC TAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 3360

AAAAAAA 3367

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 226 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

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

```
Met Gly Arg Pro Leu Leu Leu Pro Leu Leu Leu Leu Gln Pro Pro
1               5                  10                  15

Ala Phe Leu Gln Pro Gly Gly Ser Thr Gly Ser Gly Pro Ser Tyr Leu
            20                  25                  30

Tyr Gly Val Thr Gln Pro Lys His Leu Ser Ala Ser Met Gly Gly Ser
        35                  40                  45
```

Val Glu Ile Pro Phe Ser Phe Tyr Tyr Pro Trp Glu Leu Ala Thr Ala
50 55 60

Pro Asp Val Arg Ile Ser Trp Arg Arg Gly His Phe His Gly Gln Ser
65 70 75 80

Phe Tyr Ser Thr Arg Pro Pro Ser Ile His Lys Asp Tyr Val Asn Arg
85 90 95

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

Ser Asn Leu Arg Lys Glu Asp Gln Ser Val Tyr Phe Cys Arg Val Glu
115 120 125

Leu Asp Thr Arg Arg Ser Gly Arg Gln Gln Leu Gln Ser Ile Lys Gly
130 135 140

Thr Lys Leu Thr Ile Thr Gln Ala Val Thr Thr Thr Thr Thr Trp Thr
145 150 155 160

Pro Ser Ser Thr Thr Thr Ile Ala Gly Leu Arg Val Thr Glu Ser Lys
165 170 175

Gly His Ser Glu Ser Trp His Leu Ser Leu Asp Thr Ala Ile Arg Val
180 185 190

Ala Leu Ala Val Ala Val Leu Lys Thr Val Ile Leu Gly Leu Leu Cys
195 200 205

Leu Leu Leu Trp Trp Arg Arg Arg Lys Gly Ser Arg Ala Pro Ser Ser
210 215 220

Asp Phe
225

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 3899 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

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

GGGAAGAGAT GGTGACTGAG GCAGAAGCTA ATAGGGAAGA TGATAGGAAA GAAATTTTAC 60

CCAAGGGAAT TAGATTTAGC AAGAGAGCGA AGGAAAGCTG AGAGGCCAAA AACATCTCTG 120

AGGAAAACTG ACTCTGAGAG AGAAGAGGTG ACAAGGGCAA ATGCACTCAA GGATGAAGAT 180

GCTTTTAAAG AAGAGCAAAA ACTTAAAGCG GAAGAAGGGG AAACAGAGAC AGAAGTWAGA 240

GCTGAGGAAG AGACAAAAGC TCCCCCAAAT GAAATGGGAT CTGATGCTGA RAACGAASCA 300

CCTGTGGAGG CTTCTGAGTT GTCTGACAAT CCAGGGCTTC TAGGAGAARA TTCACTAAAA 360

GAGACAGTGG TTCCCATATT TGAAGCAACG CCTGGATTTG AAAAGTCGCT GGAAAACATA 420

ACAGCTCTGA GGAAAGAAGG AGGAGGGGAA AGACTGAGTG AAGCCAGAGA CACAGAGCAC 480

AAAGACAGAG AAGAGCTGTC CAGCAGGGAG AATAGGGCCC TGAAGGAAGG GCACCGCCAA 540

GATGGAGAGG GGGCCTTAGC AGCTCCTGAA GCTGAGCCAG CAGGAAAGGT GCAGGCCCCT 600

GAGGGGCTGA TCCCAGCCAC AGGCCAGGCA GAGGAGCTAG CAGCCAAAGA TCACGACTCC 660

TGCGCAGGAC TGGAGGGGAG AGCTGAAGGG CAAGGAGGAG TGGATGTCGT GCTAAGGACC 720

CAGGAAGCTG TTGCTGAGGA AGATCCCATA WTGGCAGAAA AGTTCAGGGA GGAAGCGGTG 780

GATGAGGACC CAGAGGAGGA AGAGGACAAA GAGTGCAYTC TGGAGACAGA AGCGATGCAG 840

GACAGGAACT CGGAAGGGGA CGGGGACATG GAAGGAGAAG GAAACACACA AAAGAATGAG 900

GGCATGGGAG GAGGAAGGGT TGTGGCTGTG GAAGTTCTAC ACGGAGGTGG TGAAACGGCA 960

GAAACAGCCG CAGAGGAGAG GGAGGTGTTG GCAGGTTCGG AGACAGCCGA GGAGAAAACA 1020

ATAGCAAATA AAGCCTCCTC CTTTTCAGAT GTTGCTGAGG AAGAAACCTG GCACCAACAG 1080

GATGAGTTAG TAGGAAAAAC AGCAGCTGCA GGGAAGGTGG TGGTAGAGGA ATTAGCACGG 1140

AGTGGGGAGG AAGTGCCAGC AGCAGAGGAG ATGACAGTGA CATATACAAC AGAGGCTGGG 1200

GTGGGCACTC CAGGAGCCCT GGAGCGGAAG ACCTCAGGGC TAGGACAGGA GCAAGAGGAA 1260

GGGTCAGAGG GCCAGGAGGC AGCCACTGGG AGTGGCGATG GGAGGCAGGA GACAGGAGCA 1320

GCTGAAAAAT TCCGATTAGG ATTATCACGG GAGGGAGAGA GGGAATTGAG TCCGGAGAGT 1380

CTACAGGCGA TGGCAACACT TCCAGTGAAG CCTGATTTCA CTGAAACCCG AGAGAAGCAA 1440

CAGCATATGG TGCAAGGAGA AAGCGAGACT GCAGATGTTT CCCCCAACAA CATGCAGGTC 1500

TAGGAGACTT GCTGGCAGAC GGATAATTTA AAGATGTCTT CTGAAGATGT AAAGAGTGGA 1560

GAAAGATTCA CGCAAGCATC TCACCAGGAT TCTTGATTTT CTCTCTCTCC TCTTTAGTTG 1620

CTGGTTGCGC TTGTCTGAGA TGATTCCCAA TCTGTCAGCC CTGGTCAGTA GCTCAGTAAG 1680

CACCTTGAGA ATAGCTCAAG TAGATCTGTA GGACCCTTCT TAGAAGCAGT GGTTCCTCAT 1740

GGAGAAACTT GTGAGGCTGT TACACATTCT ACACACCTAA CATTATTTTC AAACAAAAAT 1800

GATAATTTTC AGATGCTTGA CTTTTACCAA AGATCACTGG AAGGCCCAGT CCTAATGTTA 1860

GGGGTTTGTT TAAAGTCCTT TTTATTTTAC AATACAGAGC CCCAGTCAAT TCCACAATCT 1920

CAATTTCATA CATGGGAATT TTATTTAAAA ATCTGTGGTT TGGGGCTTTA ATGAATTGGC 1980

CTGTGAAAAT GAGCTCTAAA TTTCCTCCCA CGTACACTCA AAACTCAAGA TTGCTCCAAA 2040

TCTCTAAGTT CTTCCAGCAA AAGATTTCTT GGCATGTATA TTCACTTATA CTTAGAAATA 2100

TTCATTCTTT TAATTTATGC CAGAATAACA AAGTGGAAAT CTTATTTCAA AATGCTCTTT 2160

GTTTTTTTGT GTGTGTTTCT GTAGTTCTGC TTTCTGGGGT AGACTAGTAA AATGGTAGCT 2220

TCCAGCATTT TGTCCCTGGG GCCTTCTTTA TAGGGCCACT CAAATTTAAA TAAAAGTAGT 2280

AAATAATTTA GCTAAGTGGA ATAAGTATAA TAATTATAGT GGTAAGCATA GCACATCAGC 2340

ATTATGCCAA CATTCTAGAC TCTTTAGTTG ATGTCATTAA ATGGAAAAGA AACTTGGATT 2400

AAATGAGTGT GCTGCTCACC TTCCCAAGTT CTGTTATTTC AAACCTGTGA ACTAACCTTG 2460

CAGTTCATTA TAAATCAACA GTAACAACTG CATTCTAAAT TACTCCCTGA TATTATTTTC 2520

TAGTTGTGTA TCAGCCTGTC TCCTAGGGGT TTTCATTTCC CTGAAGACAT ACAAGTGCCC 2580

CAGAGCGCAT GTATATGTCT ACCATTTCTC TATATGAGAA GGTAAAAAAA ATTTCCTTAA 2640

GCAGTGATTT TCCAGCCAGA ATATACATTA GATTTTCATG GGACGCTTTT ATAAATGACT 2700

CAACCCTTTT CCCCACCCCA GAGATTCAGA CTTAATTCGT TTTAGATGGA TCTACACATC 2760

AGTATATATA TATTTTTAAC TTTTCACTTG ATTCTTCTCT GTAGCCAAGG TTGAGAACCG 2820

CTGTTCTAAA TCATCATATA ATCCATGCTG GCCACATTAC ACTCAAGGTC CCTAGGGACC 2880

AGGCATATTA TCATAGTAGG TATCTTCCAT TTTAATGTGT AATGGAGCCA TTCAATGATC 2940

AAAAATACAC TGGACCAGAT AGTAGACTGG TCCCTTGATC AGAAGCATCA GCACATCAGC 3000

ATCACCTGGA AATTGTTCCC AGCCTTTGTC TCCTACCTAC TAAATTAGAA ACTCTTGGTG 3060

GGTTCCAGTA ATCCATAGCT TAACAAGCCC TGCAGTTAAT ACTGATGTAC ACTGATGTCC 3120

AAAAACTGCT GTCATGGACT ATTGATTGTA TTGAGGATTA GTCTCAGTTG GAAAGCCAAC 3180

TACAGAGGCA TTTTGAACTT TCTTTCTTTG CCTCTCTATG TCTCTCTGTC TTTTCCTGTC 3240

TTCTGATTTA TCTGTCTTTC TTTCTCTAGT AAATGGCACT CAATATAAAA GTGGTGGAGT 3300

CAATCTTAAA CTTATTTTTA TTATGATTGT ATTGATACAT GCACGAAGTC CCTCTGCCCT 3360

ACTCCCTATT CAAGGATATT ACTCACTGCA CATCATAAAT CTCCATCATC TGTCTTAAAG 3420

TTTTATGAGT AGATTTCATC TACATTATAT TCAAGTTCAT TTATTACTGA GCTGTATTAC 3480

TGTGGAGCTC TAACAGTATT TGTTTCCTGA TTTCAAACTC AATGCTACAG AGCACTTTGA 3540

ATACATCACA CCTTATAGGA AAGATAGTAA ATGTATTAAT CCCATTGAAA AATTAGTTTT 3600

GTACAATGTG CTAAATAGTA TTGCATTGGA TTACTTTTAT ATTTAACACA CTCCATCAAA 3660

ACATCCCATA ACATAATTTT ACAATCTGCA TGTGAATTTA ACTGTGAAAT TCAGTATTGT 3720

GATATTTTGA ATAAGTGAAT TCTTTCTCTG CAAATACTAT GTTGATAAAA TTACTTGTAT 3780

GTTCCCCTGA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 3840

AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAGA 3899

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 487 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

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

Met Ile Gly Lys Lys Phe Tyr Pro Arg Glu Leu Asp Leu Ala Arg Glu
1 5 10 15

Arg Arg Lys Ala Glu Arg Pro Lys Thr Ser Leu Arg Lys Thr Asp Ser
20 25 30

Glu Arg Glu Glu Val Thr Arg Ala Asn Ala Leu Lys Asp Glu Asp Ala
35 40 45

Phe Lys Glu Glu Gln Lys Leu Lys Ala Glu Glu Gly Glu Thr Glu Thr
50 55 60

Glu Val Arg Ala Glu Glu Glu Thr Lys Ala Pro Pro Asn Glu Met Gly
65 70 75 80

Ser Asp Ala Glu Asn Glu Xaa Pro Val Glu Ala Ser Glu Leu Ser Asp
85 90 95

Asn Pro Gly Leu Leu Gly Glu Xaa Ser Leu Lys Glu Thr Val Val Pro
100 105 110

Ile Phe Glu Ala Thr Pro Gly Phe Glu Lys Ser Leu Glu Asn Ile Thr
115 120 125

Ala Leu Arg Lys Glu Gly Gly Gly Glu Arg Leu Ser Glu Ala Arg Asp
130 135 140

```
Thr Glu His Lys Asp Arg Glu Glu Leu Ser Ser Arg Glu Asn Arg Ala
145                 150                 155                 160

Leu Lys Glu Gly His Arg Gln Asp Gly Glu Gly Ala Leu Ala Ala Pro
                165                 170                 175

Glu Ala Glu Pro Ala Gly Lys Val Gln Ala Pro Glu Gly Leu Ile Pro
            180                 185                 190

Ala Thr Gly Gln Ala Glu Glu Leu Ala Ala Lys Asp His Asp Ser Cys
        195                 200                 205

Ala Gly Leu Glu Gly Arg Ala Glu Gly Gln Gly Gly Val Asp Val Val
    210                 215                 220

Leu Arg Thr Gln Glu Ala Val Ala Glu Glu Asp Pro Ile Xaa Ala Glu
225                 230                 235                 240

Lys Phe Arg Glu Glu Ala Val Asp Glu Asp Pro Glu Glu Glu Glu Asp
                245                 250                 255

Lys Glu Cys Xaa Leu Glu Thr Glu Ala Met Gln Asp Arg Asn Ser Glu
            260                 265                 270

Gly Asp Gly Asp Met Glu Gly Glu Gly Asn Thr Gln Lys Asn Glu Gly
        275                 280                 285

Met Gly Gly Gly Arg Val Val Ala Val Glu Val Leu His Gly Gly Gly
    290                 295                 300

Glu Thr Ala Glu Thr Ala Ala Glu Glu Arg Glu Val Leu Ala Gly Ser
305                 310                 315                 320

Glu Thr Ala Glu Glu Lys Thr Ile Ala Asn Lys Ala Ser Ser Phe Ser
                325                 330                 335

Asp Val Ala Glu Glu Glu Thr Trp His Gln Gln Asp Glu Leu Val Gly
            340                 345                 350

Lys Thr Ala Ala Ala Gly Lys Val Val Val Glu Glu Leu Ala Arg Ser
        355                 360                 365

Gly Glu Glu Val Pro Ala Ala Glu Glu Met Thr Val Thr Tyr Thr Thr
    370                 375                 380

Glu Ala Gly Val Gly Thr Pro Gly Ala Leu Glu Arg Lys Thr Ser Gly
385                 390                 395                 400

Leu Gly Gln Glu Gln Glu Glu Gly Ser Glu Gly Gln Glu Ala Ala Thr
                405                 410                 415

Gly Ser Gly Asp Gly Arg Gln Glu Thr Gly Ala Ala Glu Lys Phe Arg
            420                 425                 430

Leu Gly Leu Ser Arg Glu Gly Glu Arg Glu Leu Ser Pro Glu Ser Leu
```

435 440 445

Gln Ala Met Ala Thr Leu Pro Val Lys Pro Asp Phe Thr Glu Thr Arg
450 455 460

Glu Lys Gln Gln His Met Val Gln Gly Glu Ser Glu Thr Ala Asp Val
465 470 475 480

Ser Pro Asn Asn Met Gln Val
485

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 483 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

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

CATTGCTAGA CAGACTCTCT TGCTTGGATG GTACTCCACC ACTTTCTTGG CACATGAGAT 60

GCAAGATTGC TCAGGGTGCA GCTAATGGCA TCAATTTTCT ACATGAAAAT CATCATATTC 120

ATAGAGATAT TAAAAGTGCA AATATCTTAC TGGATGAAGC TTTTACTGCT AAAATATCTG 180

ACTTTGGCCT TGCACGGGCT TCTGAGAAGT TTTGCCCAGA CAGTCATGAC TAGCAGAATT 240

GTGGGAACAA CAGCTTATAT GGCACCAGAA GCTTTGCGTG GAGAAATAAC ACCCAAATCT 300

GATATTTACA GCTTTGGTGT GGTTTTACTA GAAATAATAA CTGGACTTCC AGCTGTGGAT 360

GAACACCGTG AACCTCAGTT ATTGCTAGAT ATTAAAGAAG AAATTGAAGA TGAAGAAAAG 420

ACATTGAAGA TTATATTGAT AAAAAGATGA ATGATGCTGA TTCCACTTCA GTTGAAGCTA 480

TGT 483

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 121 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

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

```
Met Ala Ser Ile Phe Tyr Met Lys Ile Ile Ile Phe Ile Glu Ile Leu
1               5                  10                  15

Lys Val Gln Ile Ser Tyr Trp Met Lys Leu Leu Leu Leu Lys Tyr Leu
            20                  25                  30

Thr Leu Ala Leu His Gly Leu Leu Arg Ser Phe Ala Gln Thr Val Met
        35                  40                  45

Thr Ser Arg Ile Val Gly Thr Thr Ala Tyr Met Ala Pro Glu Ala Leu
    50                  55                  60

Arg Gly Glu Ile Thr Pro Lys Ser Asp Ile Tyr Ser Phe Gly Val Val
65                  70                  75                  80

Leu Leu Glu Ile Ile Thr Gly Leu Pro Ala Val Asp Glu His Arg Glu
                85                  90                  95

Pro Gln Leu Leu Leu Asp Ile Lys Glu Glu Ile Glu Asp Glu Glu Lys
            100                 105                 110

Thr Leu Lys Ile Ile Leu Ile Lys Arg
        115                 120
```

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 493 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

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

```
AATCTGAGTC AGCTTAGAAG ATANTCCAAG CTTCAGATGA TAACCACAGC CTGGGCTGAC     60

ACCTGGATTT CAGCTTTGCA TGATCCTCAG TATGAGAATC TATCTGTTCT GTGCTGGACT    120

TCTAATATAT AGAACTGTGA GATAATGGGT CACATTGGCT GGATGTGGTG GCTCATACCT    180

GTAAATCCCA GCACTTTGGG AGGCCGAGGC AGGCAGATCA CCTGAGGTCA GGAGTTCAAG    240

ACCGGCCTGG CCAGCATGGT GAAGCCCCGT CTTTACTAGA AATACAAAAA TTAGACGAGC    300

GTGGTGGTGG ACACCTGTGT TCCCAGCTAC TTGGGAGGCT GAGGCAGGAG ACTGGCTGGA    360

ACCAGGGAGG TAGAGGTTGC AGTGAGCTGA GATCGTGCCA CTGCACTCCA GCCTGGGTGA    420
```

CAGAGTGAGA CTCCATCATA AATAAATAAA TAAATAAATG GGTCACATTA AGCCTTTAAA 480

AAAAAAAAAA AAA 493

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2682 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

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

GGTTCCCAGA AGAGTTTGCG ACGTGGTAAA GAAATAAGGC GAGTACACAA GCGAAGACTT 60

TCCAGCTCAG AGAGTGAAGA GAGCTATTTG TCCAAGAACT CTGAAGATGA TGAGCTAGCT 120

AAAGAATCAA AGCGGTCAGT TCGAAAGCGG GGCCGAAGCA CAGACGAGTA TTCAGAAGCA 180

GATGAGGAGG AGGAGGAAGA RGAAGGCAAA CCATCCCGCA AACGGCTACA CCGGATTGAG 240

ACGGATGAGG ARGAGAGTTG TGACAATGCT CATGGAGATG CAAATCAGCC TGCCCGTGAC 300

AGCCAGCCTA GGGTCCTGCC CTCAGAACAA GAGAGCACCA AGAAGCCCTA CCGGATAGAA 360

AGTGATGAGG AAGAGGACTT TGAAAATGTA GGCAAAGTGG GGAGCCCATT GGACTATAGC 420

TTAGTGGACT TACCTTCAAC CAATGGACAG AGCCCTGGCA AAGCCATTGA GAACTTGATT 480

GGCAAGCCTA CTGAGAAGTC TCAGACCCCC AAGGACAACA GCACAGCCAG TGCAAGCCTA 540

GCYTCCCAAT GGGACAAGTG GTGGGCAGGA GGCAGGAGCA CCAGAAGAGG AGGAAGATGA 600

GCTTTTGAGA GTGACTGACC TTGTTGATTA TGTCTGTAAC AGTGAACAGT TATAAGACTT 660

TTTTTCCATT TTTGTGCTAA TTTATTCCAC GGTAGCTCTC ACACCAGCGG GCCAGTTATT 720

AAAAGCTGTT TAATTTTTCC TAGAAAACTC CACTACAGAA TGACTTTTAG AAGAAAAATT 780

TCAACAAATC CTGAAGTCTT TCTGTGAAGT GACCAGTTCT GAACTTTGAA GATAAATAAT 840

TGCTGTAAAT TCCTTTTGAT TTTCTTTTTC CAGGTTCATG GTCCTTGGTA ATTTCATTCA 900

TGGAAAAAAA TCTTATTATA ATAACAACAA AGATTTGTAT ATTTTTGACT TTATATTTCC 960

TGAGCTCTCC TGACTTTGTG AAAAAGGGTG ATGAAAATGC ATTCCGAATC TGTGAGGGCC 1020

CAAAACAGAA TTTAGGGGTG GGTGAAAGCA CTTGTGCTTT AGCTTTTTCA TATTAAATAT 1080

ATATTATATT TAAACATTCA TGGCATAGAT GATGATTTAC AGACAATTTA AAAGTTCAAG 1140

TCTGTACTGT TACAGTTTGA GAATTGTAGA TAACATCATA CATAAGTCAT TTAGTAACAG 1200

CCTTTGTGAA ATGAACTTGT TTACTATTGG AGATAACCAC ACTTAATAAA GAAGAGACAG 1260

TGAAAGTACC ATCATAATTA ACCTAAATTT TTGTTATAGC AGAGTTTCTT GTTTAAAAAA 1320

AAAWAAAAWG CRKCYGMAAA GCATTTGTAC AGTAAAATGT ATAATGAAGC TTTGCCAACC 1380

AGACTGTGCT AGCAACAAAT TTTTTTAAAT AAGCTTTATG CAGTGGTAAT AAGGTGGCCT 1440

CAAATATATT GTGTCTGATG GAGAGTTATT AGTGAAATGA ATGTGGTCTT TCTTAAGGCC 1500

TGGGTGGACT GTAAACTTTG CCAATAGTAT AACTCTTGTC TTCTGGCCAC TTGATGTTTA 1560

AATATCTGAA ATATCATTTT GAAAAAAATA CATCTATATA TAACATACAT GAAGAGATGC 1620

TAAGCTGACA GTGATATTTT AGCACATTTG AAGACTGGGA AGAGATTTTC AGGTGAATTT 1680

TAACTGGTCT ATTCTTGCCC TTAGTATCTA CTTCAAATTG AAGTCTACAA ACAAAGCAGT 1740

TCCTTTGGGA GGTTTTTAGT TTGAGTTTTA GCGTGTGTGT GTGTTTGTGT GTGTGCGTGT 1800

GCGTGTGTGT GTGTGTTGGA ATTTCCTATC TGCCTGGATA TATTAGCAGA GTTTGAATGT 1860

AGTTTTGGCC TTTGGCCATT AGACTTCTAT TAAAATTCAT TAATAGTCAT ACAACCAACA 1920

TAGAGTTGAA TGAGAACTGC CGATGTAATT AATAGGCATG ACATCCATTT CAAACATCTC 1980

AACACTTTAA AGAAAAGCCC TTTGTTTCAA GAAAAAAGGG TTTGTAACTA ACTAAATACC 2040

TAACATGTAA TTGACACTAA AATATGAACT TTGTCTTATT TAGTTTCTGT TATAGCTGTA 2100

AAATTTCAGG CAGAGCCATA ACATTGTACA GAGTGTAGCA CTTGTGATTA AACCTAGCCT 2160

GTTAAATCCT GAAACCTTCA ACCATTACTT CTGTGAATAC TTTAGCCCTG GGATTTGGGT 2220

TTTTCTGTTC CGGTGTTGTG TCTGTTGCCG GCAATGGACA CACCATATCT GCTGCTGGCC 2280

CAAGGAACGT CATTAATTTT TCTTTCCAAA TTAAGTATTA TGTGCTAGTC AGTGTATAGT 2340

AAAGCACTTC TCTTTTTTAT TACTAAAAAG CTGGCATTAG ATTTGCATTA TAAATACCTC 2400

TCTAGGAACT TTATACTCCT TTTCCTTCTT CAACAGGTAT TGCCCTTAAA TCTTATCTTT 2460

TGGCCTTGAA AGTTTATAGC TATTGTTTTT CAGTTGTTCG TTGTTTTGTT TTGTTTCACT 2520

TTAGTTCTGT AGTACCTGCC CATTAATATT TTTGCTTTGA TTCTAGCAAT GTGTATGTAT 2580

CTGTATAAAA AATAAAATAA TGAAAGCAAC CTAAAAATAG GATGCACCAA TTAAAAAAAA 2640

AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AA 2682

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 58 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

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

Met Glu Lys Asn Leu Ile Ile Ile Thr Thr Lys Ile Cys Ile Phe Leu
1 5 10 15

Thr Leu Tyr Phe Leu Ser Ser Pro Asp Phe Val Lys Lys Gly Asp Glu
20 25 30

Asn Ala Phe Arg Ile Cys Glu Gly Pro Lys Gln Asn Leu Gly Val Gly
35 40 45

Glu Ser Thr Cys Ala Leu Ala Phe Ser Tyr
50 55

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2522 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

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

GCCGAGCGCC CGCGCCGCCG CTGCCTCTGT CCTCCGCGCG CTGCTCAGCT GAAGGCGCAC 60

AGGATTCAAT TACTGGACTT GTCAACTCTG CCAGTGTACG TGCCATTTCT CTTCCACTAT 120

GAGAGGACCG ATTGTATTGC ACATTTGTCT GGCTTTCTGT AGCCTTCTGC TTTTCAGCGT 180

TGCCACACAA TGTCTGGCCT TCCCCAAAAT AGAAAGGAGG AGGGAGATAG CACATGTTCA 240

TGCGGAAAAA GGGCAGTCCG ATAAGATGAA CACCGATGAC CTAGAAAATA GCTCTGTTAC 300

CTCAAAGCAG ACTCCCCAAC TGGTGGTCTC TGAAGATCCA ATGATGATGT CAGCAGTACC 360

ATCGGCAACA TCATTAAATA AAGCATTCTC GATTAACAAA GAAACCCAGC CTGGACAAGC 420

TGGGCTCATG CAAACAGAAC GCCCTGGTGT TTCCACACYT ACTGAGTCAG GTGTCCCCTC 480

AGCTGAAGAA GTATTTGGTT CCAGCCAGCC AGAGAGAATA TCTCCTGAAA GTGGACTTGC 540

CAAGGCCATG TTAACCATTG CTATCACTGC GACTCCTTCT CTGACTGTTG ATGAAAAGGA 600

GGAACTCCTT ACAAGCACTA ACTTTCAGCC CATTGTAGAA GAGATCACAG AAACCACAAA 660

AGGTTTTCTG AAGTATATGG ATAATCAATC ATTTGCAACT GAAAGTCAGG AAGGAGTTGG 720

TTTGGGACAT TCACCTTCAT CCTATGTGAA TACTAAGGAA ATGCTAACCA CCAATCCAAA 780

GACTGAGAAA TTTGAAGCAG ACACAGACCA CAGGACAACT TCTTTTCCTG GTGCTGAGTC 840

CACAGCAGGC AGTGAGCCTG GAAGCCTCAC CCCTGATAAG GAGAAGCCTT CGCAGATGAC 900

AGCTGATAAC ACCCAGGCTG CTGCCACCAA GCAACCACTC GAAACTTCCG AGTACACCCT 960

GAGTGTTGAG CCAGAAACTG ATAGTCTGCT GGGAGCCCCA GAAGTCACAG TGAGTGTCAG 1020

CACAGCTGTT CCAGCTGCCT CTGCCTTAAG TGATGAGTGG GATGACACCA AATTAGAGAG 1080

TGTAAGCCGG ATAAGGACCC CCAAGCTTGG AGACAATGAA GAGACTCAGG TGAGAACGGA 1140

GATGTCTCAG ACAGCACAAG TAAGCCATGA GGGTATGGAA GGAGGCCAGC CTTGGACAGA 1200

GGCTGCACAG GTGGCTCTGG GGCTGCCTGA AGGGGAAACA CACACGGGCA CAGCCCTGCT 1260

AATAGCGCAT GGGAATGAGA GATCACCTGC TTTCACTGAT CAAAGTTCCT TTACCCCCAC 1320

AAGTCTGATG GAAGACATGA AAGTTTCCAT TGTGAACTTG CTCCAAAGTA CGGGAGACTT 1380

CACGGAATCC ACCAAGGAAA ACGATGCCCT GTTTTTCTTA GAAACCACTG TTTCTGTCTC 1440

TGTATATGAG TCTGAGGCAG ACCAACTGTT GGGAAATACA ATGAAAGACA TCATCACTCA 1500

AGAGATGACA ACAGCTGTTC AAGAGCCAGA TGCCACTTTA TCCATGGTGA CACAAGAGCA 1560

GGTTGCTACC CTCGAGCTTA TCAGAGACAG TGGCAAGACT GAGGAAGAAA AGGAGGACCC 1620

CTCTCCTGTG TCTGACGTTC CTGGTGTTAC TCAGCTGTCA AGAAGATGGG AGCCTCTGGC 1680

CACTACAATT TCAACTACAG TCGTCCCTTT GTCTTTTGAA GTTACTCCCA CTGTGGAAGA 1740

ACAAATGGAC ACAGTCACAG GGCCAAATGA GGAGTTCACA CCAGTTCTGG GATCTCCAGT 1800

GACACCTCCT GGAATAATGG TGGGGGAACC CAGCATTTCC CCTGCACTTC CTGCTTTGGA 1860

GGCATCCTCT GAGAGAAGAA CTGTTGTTCC ATCTATTACT CGTGTTAATA CAGCTGCCTC 1920

ATATGGCCTG GACCAACTTG AATCTGAAGA GGGACAAGAA GATGAGGATG AAGAGGATGA 1980

AGAAGATGAA GATGAAGAAG AGGAAGATGA GGAAGAAGAT GAGGAAGATA AAGATGCAGA 2040

CTCGCTGGAT GAGGGCTTGG ATGGTGACAC TGAGCTGCCA GGTTTTACCC TCCCTGGTAT 2100

CACATCCCAG GAACCAGGCT TAGAGGAGGG AAACATGGAC CTGTTGGAGG GAGCTACCTA 2160

CCAGGTGCCA GATGCCYTCG AGTGGGAACA GCAGAATCAA GGCCTGGTGA GAAGCTGGAT 2220

GGAAAAATTM AAAGACAAGG CTGGTTACAT GTCTGGGATG CTGGTGCCTG TAGGGGTTGG 2280

GATAGCTGGA GCCTTGTTCA TCTTGGGAGC CCTCTACAGC ATTAAGGTTA TGAATCGCCG 2340

AAGGAGAAAT GGCTTCAAAA GGCATAAAAG AAAGCAGAGA GAATTCAACA GCATGCAAGA 2400

TCGAGTAATG CTCTTAGCCG ACAGCTCTGA AGATGAATTT TGAATTGGAC TGGGTTTTAA 2460

TTGGGATATT CAACGATGCT ACTATTCTAA TTTTTATTTT GGAGCAGAAA AAAAAAAAAA 2520

AA 2522

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 774 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

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

Met Arg Gly Pro Ile Val Leu His Ile Cys Leu Ala Phe Cys Ser Leu
1 5 10 15

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

Arg Arg Arg Glu Ile Ala His Val His Ala Glu Lys Gly Gln Ser Asp
35 40 45

Lys Met Asn Thr Asp Asp Leu Glu Asn Ser Ser Val Thr Ser Lys Gln
50 55 60

Thr Pro Gln Leu Val Val Ser Glu Asp Pro Met Met Met Ser Ala Val
65 70 75 80

Pro Ser Ala Thr Ser Leu Asn Lys Ala Phe Ser Ile Asn Lys Glu Thr
85 90 95

Gln Pro Gly Gln Ala Gly Leu Met Gln Thr Glu Arg Pro Gly Val Ser
100 105 110

Thr Xaa Thr Glu Ser Gly Val Pro Ser Ala Glu Glu Val Phe Gly Ser
115 120 125

Ser Gln Pro Glu Arg Ile Ser Pro Glu Ser Gly Leu Ala Lys Ala Met
130 135 140

```
Leu Thr Ile Ala Ile Thr Ala Thr Pro Ser Leu Thr Val Asp Glu Lys
145                 150                 155                 160

Glu Glu Leu Leu Thr Ser Thr Asn Phe Gln Pro Ile Val Glu Glu Ile
                165                 170                 175

Thr Glu Thr Thr Lys Gly Phe Leu Lys Tyr Met Asp Asn Gln Ser Phe
            180                 185                 190

Ala Thr Glu Ser Gln Glu Gly Val Gly Leu Gly His Ser Pro Ser Ser
        195                 200                 205

Tyr Val Asn Thr Lys Glu Met Leu Thr Thr Asn Pro Lys Thr Glu Lys
    210                 215                 220

Phe Glu Ala Asp Thr Asp His Arg Thr Thr Ser Phe Pro Gly Ala Glu
225                 230                 235                 240

Ser Thr Ala Gly Ser Glu Pro Gly Ser Leu Thr Pro Asp Lys Glu Lys
                245                 250                 255

Pro Ser Gln Met Thr Ala Asp Asn Thr Gln Ala Ala Ala Thr Lys Gln
            260                 265                 270

Pro Leu Glu Thr Ser Glu Tyr Thr Leu Ser Val Glu Pro Glu Thr Asp
        275                 280                 285

Ser Leu Leu Gly Ala Pro Glu Val Thr Val Ser Val Ser Thr Ala Val
    290                 295                 300

Pro Ala Ala Ser Ala Leu Ser Asp Glu Trp Asp Asp Thr Lys Leu Glu
305                 310                 315                 320

Ser Val Ser Arg Ile Arg Thr Pro Lys Leu Gly Asp Asn Glu Glu Thr
                325                 330                 335

Gln Val Arg Thr Glu Met Ser Gln Thr Ala Gln Val Ser His Glu Gly
            340                 345                 350

Met Glu Gly Gly Gln Pro Trp Thr Glu Ala Ala Gln Val Ala Leu Gly
        355                 360                 365

Leu Pro Glu Gly Glu Thr His Thr Gly Thr Ala Leu Leu Ile Ala His
    370                 375                 380

Gly Asn Glu Arg Ser Pro Ala Phe Thr Asp Gln Ser Ser Phe Thr Pro
385                 390                 395                 400

Thr Ser Leu Met Glu Asp Met Lys Val Ser Ile Val Asn Leu Leu Gln
                405                 410                 415

Ser Thr Gly Asp Phe Thr Glu Ser Thr Lys Glu Asn Asp Ala Leu Phe
            420                 425                 430

Phe Leu Glu Thr Thr Val Ser Val Ser Val Tyr Glu Ser Glu Ala Asp
```

```
        435                 440                 445

Gln Leu Leu Gly Asn Thr Met Lys Asp Ile Ile Thr Gln Glu Met Thr
    450                 455                 460

Thr Ala Val Gln Glu Pro Asp Ala Thr Leu Ser Met Val Thr Gln Glu
465                 470                 475                 480

Gln Val Ala Thr Leu Glu Leu Ile Arg Asp Ser Gly Lys Thr Glu Glu
                485                 490                 495

Glu Lys Glu Asp Pro Ser Pro Val Ser Asp Val Pro Gly Val Thr Gln
            500                 505                 510

Leu Ser Arg Arg Trp Glu Pro Leu Ala Thr Thr Ile Ser Thr Thr Val
        515                 520                 525

Val Pro Leu Ser Phe Glu Val Thr Pro Thr Val Glu Glu Gln Met Asp
    530                 535                 540

Thr Val Thr Gly Pro Asn Glu Glu Phe Thr Pro Val Leu Gly Ser Pro
545                 550                 555                 560

Val Thr Pro Pro Gly Ile Met Val Gly Glu Pro Ser Ile Ser Pro Ala
                565                 570                 575

Leu Pro Ala Leu Glu Ala Ser Ser Glu Arg Arg Thr Val Val Pro Ser
            580                 585                 590

Ile Thr Arg Val Asn Thr Ala Ala Ser Tyr Gly Leu Asp Gln Leu Glu
        595                 600                 605

Ser Glu Glu Gly Gln Glu Asp Glu Asp Glu Glu Asp Glu Glu Asp Glu
    610                 615                 620

Asp Glu Glu Glu Glu Asp Glu Glu Glu Asp Glu Glu Asp Lys Asp Ala
625                 630                 635                 640

Asp Ser Leu Asp Glu Gly Leu Asp Gly Asp Thr Glu Leu Pro Gly Phe
                645                 650                 655

Thr Leu Pro Gly Ile Thr Ser Gln Glu Pro Gly Leu Glu Glu Gly Asn
            660                 665                 670

Met Asp Leu Leu Glu Gly Ala Thr Tyr Gln Val Pro Asp Ala Xaa Glu
        675                 680                 685

Trp Glu Gln Gln Asn Gln Gly Leu Val Arg Ser Trp Met Glu Lys Xaa
    690                 695                 700

Lys Asp Lys Ala Gly Tyr Met Ser Gly Met Leu Val Pro Val Gly Val
705                 710                 715                 720

Gly Ile Ala Gly Ala Leu Phe Ile Leu Gly Ala Leu Tyr Ser Ile Lys
                725                 730                 735
```

Val Met Asn Arg Arg Arg Arg Asn Gly Phe Lys Arg His Lys Arg Lys
740 745 750

Gln Arg Glu Phe Asn Ser Met Gln Asp Arg Val Met Leu Leu Ala Asp
755 760 765

Ser Ser Glu Asp Glu Phe
770

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2002 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

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

GGCACGCCGG TACCTGAAGT CCTTCAGAAG TGCACGCCGG GACCAGGATT CCGGGAGGCC 60

GACTCCTCCC TGCCCCACGA ATGCCGGGAA TTGTGGTCTC CGCCGGACGC GAGTTGTGAG 120

ACGGCCCAAG GGGCCGCGGG GTATGCTGGG ACCGCTAGCC CTTCCGGCGC GCCTCAGGAC 180

TTCGGGTCCC CTCACCCCGG GCGGATGCCC AAAGACTCCG CCTTCCCAAG AGCCCCTGCG 240

GCCGGGCGCG AAAATGGCGG CGGCGGCGAC GGCCGGGCGC TCCTGAAGCA GCAGTTATGG 300

AGCTTCCCTC AGGGCCGGGG CCGGAGCGGC TCTTTGACTC GCACCGGCTT CCGGGTGACT 360

GCTTCCTACT GCTCGTGCTG CTGCTCTACG CGCCAGTCGG GTTCTGCCTC CTCGTCCTGC 420

GCCTGTTTCT CGGGATCCAC GTCTTCCTGG TCAGCTGCGC GCTGCCAGAC AGCGTCCTTC 480

GCAGATTCGT AGTGCGGACC ATGTGTGCGG TGCTAGGGCT CGTGGCCCGG CAGGAGGACT 540

CCGGACTCCG GGATCACAGT GTCAGGGTCC TCATTTCCAA CCATGTGACA CCTTTCGACC 600

ACAACATAGT CAATTTGCTT ACCACCTGTA GCACCGTGAG TGAGAGCGAG GCCGAGAGCG 660

CCACGGGGCG GTTCCCTGGG GCCCAGCTGA AGGCCCCCCT GTCCCCACTC GCGTTCCCCA 720

TGGAGGATAC TGAGCCTTAC CCCTAACCCC GATCCTCTAC CCAACATGTC AGTTTTTTTT 780

TTCATTTTCC TCAATATTTT TCTTCTTGCT TTCTCTTCTC CTGGTTCCCA GCCTCTACTC 840

AATAGTCCCC CCAGCTTTGT GTGCTGGTCT CGGGGCTTCA TGGAGATGAA TGGGCGGGGG 900

GAGTTGGTGG AGTCACTCAA GAGATTCTGT GCTTCCACGA GGCTTCCCCC CACTCCTCTG 960

CTGCTATTCC CTGAGGAAGA GGCCACCAAT GGCCGGGAGG GGCTCCTGCG CTTCAGAGTT 1020

TGACAGTTGC CTGTTATAAG GCAGGTGTGA GCTGCTGACT AGGCTGGCTG GATTCCCATC 1080

CTACTTTCTC CTTCCTCTTC TAGTTCCTGG CCATTTTCTA TCCAAGATGT GGTACAACCT 1140

CTTACCCTGC AAGTTCAGAG ACCCCTGGTC TCTGTGACGG TGTCAGATGC CTCCTGGGTC 1200

TCAGAACTGC TGTGGTCACT TTTCGTCCCT TTCACGGTGT ATCAAGTGGC TTCGTCCTGT 1260

TCATCGCCAA CTAGGGGAAG CGAATGAGGA GTTTGCACTC CGTGTACAAC AGCTGGTGGC 1320

CAAGGAATTG GGCCAGACAG GGACACGGCT CACTCCAGCT GACAAAGCAG AGCACATGAA 1380

GCGACAAAGA CACCCCAGAT TGCGCCCCCA GTCAGCCCAG TCTTCTTTCC CTCCCTCCCC 1440

TGGTCCTTCT CCTGATGTGC AACTGGCAAC TCTGGCTCAG AGAGTCAAGG AAGTTTTGCC 1500

CCATGTGCCA TTTGGTGTCA TCCAGAGAGA CCTGGCCAAG ACTGGCTGTG TAGACTTGAC 1560

TATCACTAAT CTGCTTGAGG GGGCCGTAGC TTTCATGCCT GAAGACATCA CCAAGGGAAC 1620

TCAGTCCCTA CCCACAGCCT CTGCCTCCAA GTTTCCCAGC TCTGGCCCGG TGACCCCTCA 1680

GCCAACAGCC CTAACATTTG CCAAGTCTTC CTGGGCCCGG CAGGAGAGCC TGCAGGAGCG 1740

CAAGCAAGCA CTATATGAAT ACGCAAGAAG GAGATTCACA GAGAGACGAG CCCAGGAGGC 1800

TGACTGAGCT CAAAGGAACA GGATGGCACC CAGAGCCGCA GGACGGAGAC TGGGGGCAGC 1860

CCTCACCCAA CTCACAACAG GCTGGATGGG TGGGTGGTAA AAAGGGAAGG ATGAGGCTCC 1920

CCCAATGTCA CATTAAATTC ATGGTTTTCA TTCAAGGVAA AAAAAAAAAA AAAAAAAAAA 1980

AAAAAAAAAA AAAAAAAAAA AA 2002

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 206 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

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

Met Pro Pro Gly Ser Gln Asn Cys Cys Gly His Phe Ser Ser Leu Ser
1 5 10 15

Arg Cys Ile Lys Trp Leu Arg Pro Val His Arg Gln Leu Gly Glu Ala

```
                20                  25                  30

Asn Glu Glu Phe Ala Leu Arg Val Gln Gln Leu Val Ala Lys Glu Leu
        35                  40                  45

Gly Gln Thr Gly Thr Arg Leu Thr Pro Ala Asp Lys Ala Glu His Met
    50                  55                  60

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

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

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

Gln Arg Asp Leu Ala Lys Thr Gly Cys Val Asp Leu Thr Ile Thr Asn
        115                 120                 125

Leu Leu Glu Gly Ala Val Ala Phe Met Pro Glu Asp Ile Thr Lys Gly
    130                 135                 140

Thr Gln Ser Leu Pro Thr Ala Ser Ala Ser Lys Phe Pro Ser Ser Gly
145                 150                 155                 160

Pro Val Thr Pro Gln Pro Thr Ala Leu Thr Phe Ala Lys Ser Ser Trp
                165                 170                 175

Ala Arg Gln Glu Ser Leu Gln Glu Arg Lys Gln Ala Leu Tyr Glu Tyr
            180                 185                 190

Ala Arg Arg Arg Phe Thr Glu Arg Arg Ala Gln Glu Ala Asp
        195                 200                 205
```

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 819 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

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

```
CAATTGGGCC GCGAGTTGTG GTTTAAACCA GGAGTGCGCC GCGTCCGTTC ACCGCGGCCT     60

CAGATGAATG CGGCTGTTAA GACCTGCAAT AATCCAGAAT GGCTACTCTG ATCTATGTTG     120

ATAAGGAAAA TGGAGAACCA GGCACCCGTG TGGTTGCTAA GGATGGGCTG AAGCTGGGGT     180
```

CTGGACCTTC AATCAAAGCC TTAGATGGGA GATCTCAAGT TTCAACACCA CGTTTTGGCA 240

AAACGTTCGA TGCCCCACCA GCCTTACCTA AAGCTACTAG AAAGGCTTTG GGAACTGTCA 300

ACAGAGCTAC AGAAAAGTCT GTAAAGACCA AGGGACCCCT CAAACAAAAA CAGCCAAGCT 360

TTTCTGCCAA AAAGATGACT GAGAAGACTG TTAAAGCAAA AAGCTCTGTT CCTGCCTCAG 420

ATGATGCCTA TCCAGAAATA GAAAAATTCT TTCCCTTCAA TCCTCTAGAC TTTGAGAGTT 480

TTGACCTGCC TGAAGAGCAC CAGATTGCGC ACCTCCCCTT GAGTGGAGTG CCTCTCWTGA 540

TCCTTGACGA GGAGAGAGAG CTTGAAAAGC TGTTTCAGCT GGGCCCCCCT TCACCTGTGA 600

AGATGCCCTC TCCACCATGG GAATCCAATC TGTTGCAGTC TCCTTCAAGC ATTCTGTCGA 660

CCCTGGATGT TGAATTGCCA CCTGTTTGCT GTGACATAGA TATTTAAATT TCTTAGTGCT 720

TCAGAGTTTG TGTGTATTTG TATTAATAAA GCATTCTTTA ACAGAAAAAA AAAAAAAAAA 780

AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAA 819

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 146 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

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

Met Ala Thr Leu Ile Tyr Val Asp Lys Glu Asn Gly Glu Pro Gly Thr
1 5 10 15

Arg Val Val Ala Lys Asp Gly Leu Lys Leu Gly Ser Gly Pro Ser Ile
20 25 30

Lys Ala Leu Asp Gly Arg Ser Gln Val Ser Thr Pro Arg Phe Gly Lys
35 40 45

Thr Phe Asp Ala Pro Pro Ala Leu Pro Lys Ala Thr Arg Lys Ala Leu
50 55 60

Gly Thr Val Asn Arg Ala Thr Glu Lys Ser Val Lys Thr Lys Gly Pro
65 70 75 80

Leu Lys Gln Lys Gln Pro Ser Phe Ser Ala Lys Lys Met Thr Glu Lys
85 90 95

Thr Val Lys Ala Lys Ser Ser Val Pro Ala Ser Asp Asp Ala Tyr Pro
100 105 110

Glu Ile Glu Lys Phe Phe Pro Phe Asn Pro Leu Asp Phe Glu Ser Phe
115 120 125

Asp Leu Pro Glu Glu His Gln Ile Ala His Leu Pro Leu Ser Gly Val
130 135 140

Pro Leu
145

(2) INFORMATION FOR SEQ ID NO:22:

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

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

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

TNTCCTGCCTC AGCTGCCTCT CTGTGTAA 29

(2) INFORMATION FOR SEQ ID NO:23:

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

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

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

CNCACTGCCCT CCTTCTCCCA TAGGTACT 29

(2) INFORMATION FOR SEQ ID NO:24:

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

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

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

GNAATAAGCAT GATGCTCTAC AAGGAAAG 29

(2) INFORMATION FOR SEQ ID NO:25:

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

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

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

TNGGTGCCATG ATTCTGAGTG CCCTTTGC 29

(2) INFORMATION FOR SEQ ID NO:26:

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

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

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

GNATATGTCAC TGTCATCTCC TCTGCTGC 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:

ANAAGCTTCAT CCAGTAAGAT ATTTGCAC 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:

ANTTCAGAACT GGTCACTTCA CAGAAAGA 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:

GNATTCACATA GGATGAAGGT GAATGTCC 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:

ANTAGAGGCTG GGAACCAGGA GAAGAGAA 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:

TNTTGCAGGTC TTAACAGCCG CATTCATC 29

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 113 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

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

```
Met Glu Leu Pro Ser Gly Pro Gly Pro Glu Arg Leu Phe Asp Ser His
1               5                  10                  15

Arg Leu Pro Gly Asp Cys Phe Leu Leu Leu Val Leu Leu Leu Tyr Ala
            20                  25                  30

Pro Val Gly Phe Cys Leu Leu Val Leu Arg Leu Phe Leu Gly Ile His
        35                  40                  45

Val Phe Leu Val Ser Cys Ala Leu Pro Asp Ser Val Leu Arg Arg Phe
    50                  55                  60

Val Val Arg Thr Met Cys Ala Val Leu Gly Leu Val Ala Arg Gln Glu
65                  70                  75                  80
```

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

Val Thr Pro Phe Asp His Asn Ile Val Asn Leu Leu Thr Thr Cys Ser
100 105 110

Thr

(2) INFORMATION FOR SEQ ID NO:33:

(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:33:

Ser Gln Pro Leu Leu Asn Ser Pro Pro Ser Phe Val Cys Trp Ser Arg
1 5 10 15

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

Arg Phe Cys Ala Ser Thr Arg Leu Pro Pro Thr Pro Leu Leu Leu Phe
35 40 45

Pro Glu Glu Glu Ala Thr Asn Gly Arg Glu Gly Leu Leu Arg Phe
50 55 60

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 49 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

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

Ser Ser Trp Pro Phe Ser Ile Gln Asp Val Val Gln Pro Leu Thr Leu
1 5 10 15

Gln Val Gln Arg Pro Leu Val Ser Val Thr Val Ser Asp Ala Ser Trp

20 25 30

Val Ser Glu Leu Leu Trp Ser Leu Phe Val Pro Phe Thr Val Tyr Gln
35 40 45

Val

What is claimed is:

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

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

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 170 to nucleotide 322;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 218 to nucleotide 322;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 1814 to nucleotide 2355;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone bl209_10 deposited under accession number ATCC 98379;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone bl209_10 deposited under accession number ATCC 98379;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone bl209_10 deposited under accession number ATCC 98379;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone bl209_10 deposited under accession number ATCC 98379;

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

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:2 having biological activity, the fragment comprising the amino acid sequence from amino acid 20 to amino acid 29 of SEQ ID NO:2;

(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).

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

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

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

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

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

(b) purifying said protein from the culture.

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

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

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

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

(b) fragments of the amino acid sequence of SEQ ID NO:2 comprising the amino acid sequence from amino acid 20 to amino acid 29 of SEQ ID NO:2; and

(c) the amino acid sequence encoded by the cDNA insert of clone bl209_10 deposited under accession number ATCC 98379;

the protein being substantially free from other mammalian proteins.

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

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

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

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

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

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

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 102 to nucleotide 1295;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 162 to nucleotide 1295;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 804 to nucleotide 1184;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone cr1162_25 deposited under accession number ATCC 98379;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone cr1162_25 deposited under accession number ATCC 98379;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone cr1162_25 deposited under accession number ATCC 98379;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone cr1162_25 deposited under accession number ATCC 98379;

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

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:4 having biological activity, the fragment comprising the amino acid sequence from amino acid 194 to amino acid 203 of SEQ ID NO:4;

(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).

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

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

(b) the amino acid sequence of SEQ ID NO:4 from amino acid 236 to amino acid 361;

(c) fragments of the amino acid sequence of SEQ ID NO:4 comprising the amino acid sequence from amino acid 194 to amino acid 203 of SEQ ID NO:4; and

(d) the amino acid sequence encoded by the cDNA insert of clone cr1162_25 deposited under accession number ATCC 98379;

the protein being substantially free from other mammalian proteins.

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

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

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

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 351 to nucleotide 842;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 687 to nucleotide 842;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 1 to nucleotide 689;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone dh40_3 deposited under accession number ATCC 98379;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone dh40_3 deposited under accession number ATCC 98379;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone dh40_3 deposited under accession number ATCC 98379;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone dh40_3 deposited under accession number ATCC 98379;

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

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:6 having biological activity, the fragment comprising the amino acid sequence from amino acid 77 to amino acid 86 of SEQ ID NO:6;

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

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

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

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

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

(b) the amino acid sequence of SEQ ID NO:6 from amino acid 1 to amino acid 113;

(c) fragments of the amino acid sequence of SEQ ID NO:6 comprising the amino acid sequence from amino acid 77 to amino acid 86 of SEQ ID NO:6; and

(d) the amino acid sequence encoded by the cDNA insert of clone dh40_3 deposited under accession number ATCC 98379;

the protein being substantially free from other mammalian proteins.

18. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:5.

19. 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 2205 to nucleotide 2882;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 2262 to nucleotide 2882;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 2494 to nucleotide 3120;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone di39_9 deposited under accession number ATCC 98379;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone di39_9 deposited under accession number ATCC 98379;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone di39_9 deposited under accession number ATCC 98379;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone di39_9 deposited under accession number ATCC 98379;

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

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:8 having biological activity, the fragment comprising the amino acid sequence from amino acid 108 to amino acid 117 of SEQ ID NO:8;

(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).

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

(a) the amino acid sequence of SEQ ID NO:8;

(b) fragments of the amino acid sequence of SEQ ID NO:8 comprising the amino acid sequence from amino acid 108 to amino acid 117 of SEQ ID NO:8; and

(c) the amino acid sequence encoded by the cDNA insert of clone di39_9 deposited under accession number ATCC 98379;

the protein being substantially free from other mammalian proteins.

21. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:7.

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

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9 from nucleotide 40 to nucleotide 1503;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9 from nucleotide 863 to nucleotide 1377;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone dt674_2 deposited under accession number ATCC 98379;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone dt674_2 deposited under accession number ATCC 98379;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone dt674_2 deposited under accession number ATCC 98379;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone dt674_2 deposited under accession number ATCC 98379;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:10;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:10 having biological activity, the fragment comprising the amino acid sequence from amino acid 238 to amino acid 247 of SEQ ID NO:10;

(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).

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

(a) the amino acid sequence of SEQ ID NO:10;

(b) the amino acid sequence of SEQ ID NO:10 from amino acid 277 to amino acid 446;

(c) fragments of the amino acid sequence of SEQ ID NO:10 comprising the amino acid sequence from amino acid 238 to amino acid 247 of SEQ ID NO:10; and

(d) the amino acid sequence encoded by the cDNA insert of clone dt674_2 deposited under accession number ATCC 98379;

the protein being substantially free from other mammalian proteins.

24. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:9.

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

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11 from nucleotide 85 to nucleotide 450;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11 from nucleotide 217 to nucleotide 450;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone eh61_1 deposited under accession number ATCC 98379;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone eh61_1 deposited under accession number ATCC 98379;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone eh61_1 deposited under accession number ATCC 98379;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone eh61_1 deposited under accession number ATCC 98379;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:12;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:12 having biological activity, the fragment comprising the amino acid sequence from amino acid 55 to amino acid 64 of SEQ ID NO:12;

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

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

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

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

(a) the amino acid sequence of SEQ ID NO:12;

(b) the amino acid sequence of SEQ ID NO:12 from amino acid 9 to amino acid 94;

(c) fragments of the amino acid sequence of SEQ ID NO:12 comprising the amino acid sequence from amino acid 55 to amino acid 64 of SEQ ID NO:12; and

(d) the amino acid sequence encoded by the cDNA insert of clone eh61_1 deposited under accession number ATCC 98379;

the protein being substantially free from other mammalian proteins.

27. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:11 and SEQ ID NO:13.

28. 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 900 to nucleotide 1073;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:14 from nucleotide 544 to nucleotide 1022;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone fg265_1 deposited under accession number ATCC 98379;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone fg265_1 deposited under accession number ATCC 98379;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone fg265_1 deposited under accession number ATCC 98379;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone fg265_1 deposited under accession number ATCC 98379;

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

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:15 having biological activity, the fragment comprising the amino acid sequence from amino acid 24 to amino acid 33 of SEQ ID NO:15;

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

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

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

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

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

(b) the amino acid sequence of SEQ ID NO:15 from amino acid 1 to amino acid 41;

(c) fragments of the amino acid sequence of SEQ ID NO:15 comprising the amino acid sequence from amino acid 24 to amino acid 33 of SEQ ID NO:15; and

(d) the amino acid sequence encoded by the cDNA insert of clone fg265_1 deposited under accession number ATCC 98379;

the protein being substantially free from other mammalian proteins.

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

31. 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 119 to nucleotide 2440;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:16 from nucleotide 200 to nucleotide 2440;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:16 from nucleotide 460 to nucleotide 1153;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone fp273_10 deposited under accession number ATCC 98379;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone fp273_10 deposited under accession number ATCC 98379;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone fp273_10 deposited under accession number ATCC 98379;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone fp273_10 deposited under accession number ATCC 98379;

(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, the fragment comprising the amino acid sequence from amino acid 382 to amino acid 391 of SEQ ID NO:17;

(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).

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

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

(b) the amino acid sequence of SEQ ID NO:17 from amino acid 115 to amino acid 345;

(c) fragments of the amino acid sequence of SEQ ID NO:17 comprising the amino acid sequence from amino acid 382 to amino acid 391 of SEQ ID NO:17; and

(d) the amino acid sequence encoded by the cDNA insert of clone fp273_10 deposited under accession number ATCC 98379;

the protein being substantially free from other mammalian proteins.

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

34. 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 1187 to nucleotide 1804;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:18 from nucleotide 674 to nucleotide 1014;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone fy243_8 deposited under accession number ATCC 98379;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone fy243_8 deposited under accession number ATCC 98379;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone fy243_8 deposited under accession number ATCC 98379;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone fy243_8 deposited under accession number ATCC 98379;

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

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:19 having biological activity, the fragment comprising the amino acid sequence from amino acid 98 to amino acid 107 of SEQ ID NO:19;

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

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

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

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

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

(b) the amino acid sequence of SEQ ID NO:19 from amino acid 21 to amino acid 69;

(c) fragments of the amino acid sequence of SEQ ID NO:19 comprising the amino acid sequence from amino acid 98 to amino acid 107 of SEQ ID NO:19; and

(d) the amino acid sequence encoded by the cDNA insert of clone fy243_8 deposited under accession number ATCC 98379;

the protein being substantially free from other mammalian proteins.

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

37. 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 99 to nucleotide 536;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:20 from nucleotide 1 to nucleotide 370;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone ga205_4 deposited under accession number ATCC 98379;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone ga205_4 deposited under accession number ATCC 98379;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone ga205_4 deposited under accession number ATCC 98379;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone ga205_4 deposited under accession number ATCC 98379;

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

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:21 having biological activity, the fragment comprising the amino acid sequence from amino acid 68 to amino acid 77 of SEQ ID NO:21;

(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).

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

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

(b) the amino acid sequence of SEQ ID NO:21 from amino acid 1 to amino acid 90;

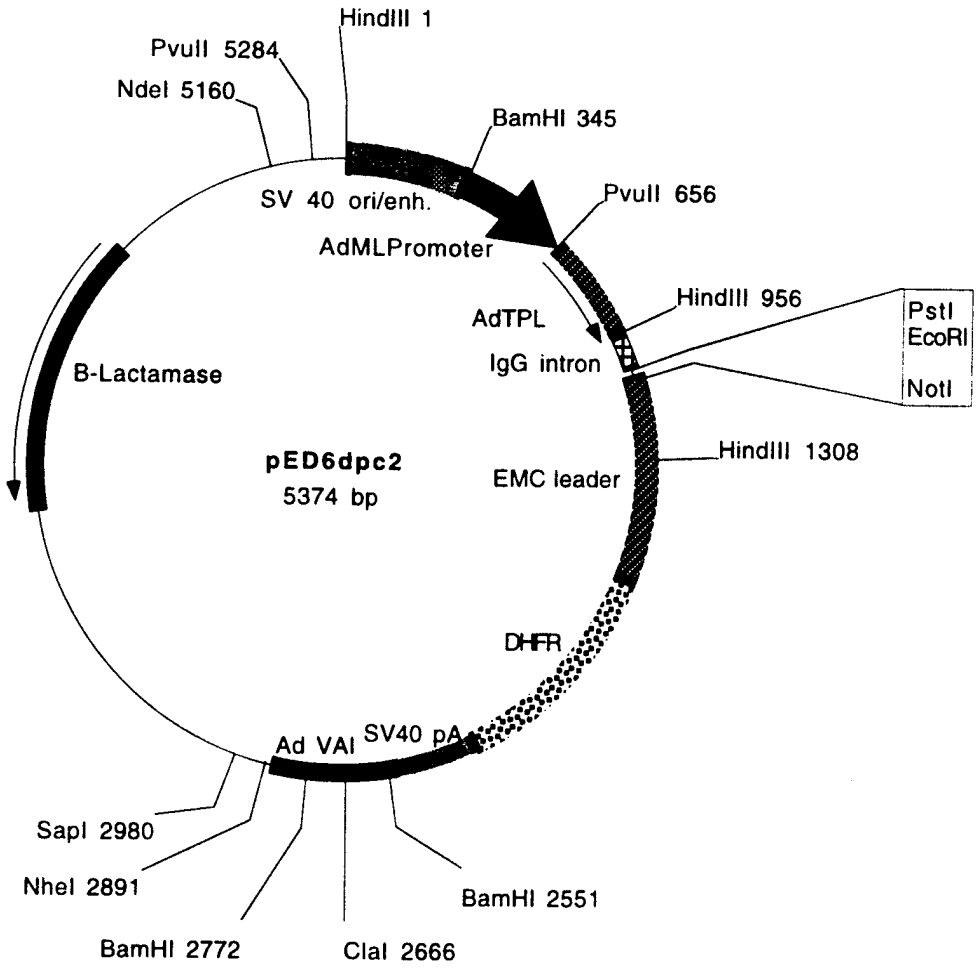
(c) fragments of the amino acid sequence of SEQ ID NO:21 comprising the amino acid sequence from amino acid 68 to amino acid 77 of SEQ ID NO:21; and

(d) the amino acid sequence encoded by the cDNA insert of clone ga205_4 deposited under accession number ATCC 98379;

the protein being substantially free from other mammalian proteins.

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

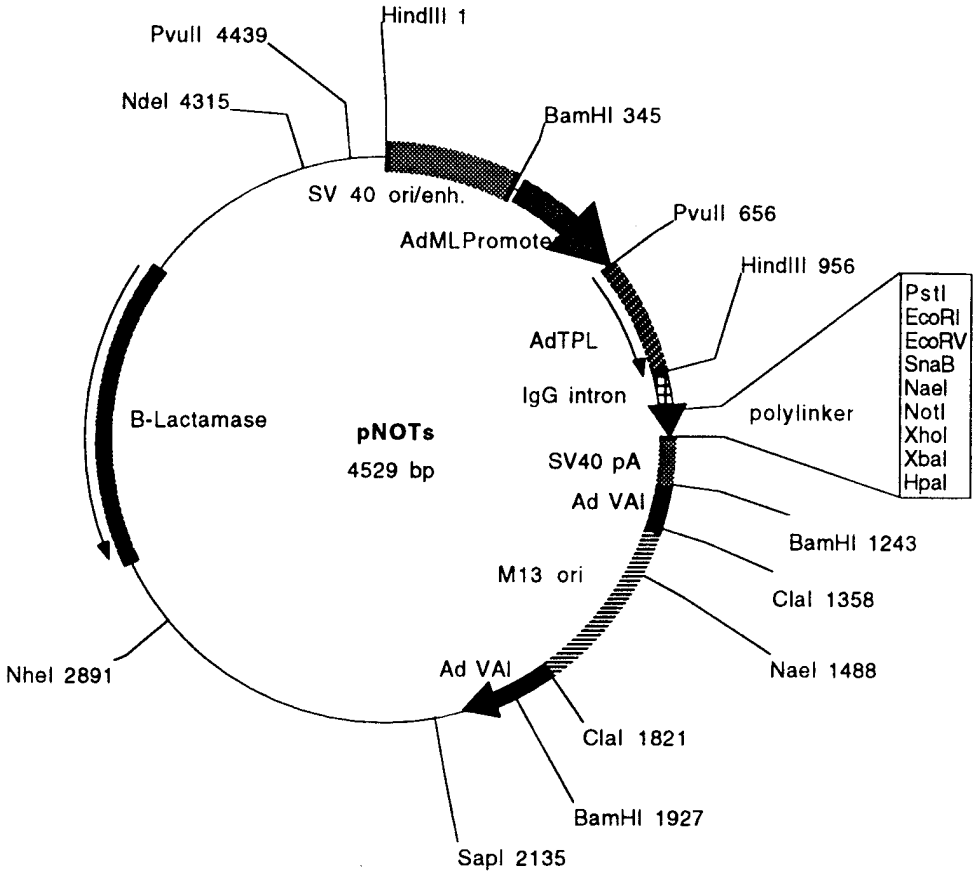
FIGURE 1A



Plasmid name: pED6dpc2
Plasmid size: 5374 bp

Comments/References: pED6dpc2 is derived from pED6dpc1 by insertion of a new polylinker to facilitate cDNA cloning. SST cDNAs are cloned between EcoRI and NotI. pED vectors are described in Kaufman et al.(1991), NAR 19: 4485-4490.

FIGURE 1B



Plasmid name: pNOTs
Plasmid size: 4529 bp

Comments/References: pNOTs is a derivative of pMT2 (Kaufman et al,1989. Mol.Cell.Biol.9:1741-1750). DHFR was deleted and a new polylinker was inserted between EcoRI and HpaI. M13 origin of replication was inserted in the ClaI site. SST cDNAs are cloned between EcoRI and NotI

INTERNATIONAL SEARCH REPORT

Inter nal Application No
PCT/US 98/06176

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/12 C07K14/47 A61K38/17

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N C07K A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Database EMBL, entry HS622277, Accession number N41622, 27 January 1996 98% identity to Seq.ID:1 nt.1718-2236 XP002067589 cited in the application see the whole document ---	1,12
X	Database EMBL, entry HS981289, Accession number N52981, 31 January 1997 98% identity to Seq.ID:1 nt.1888-2344 reverse orientation XP002067590 see the whole document --- -/--	1,12

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
11 June 1998	16. 09. 1998

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Macchia, G

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Database EMBL, entry HS620267, Accession number N29620, 6 January 1996 99% identity to Seq.ID:1 nt.1948-2333 reverse orientation XP002067591 cited in the application see the whole document	1,12
X	Database EMBL, entry HS172310, Accession number N80172, 5 April 1996 100% identity to Seq.ID:1 nt.1951-2347 reverse orientation XP002067592 cited in the application see the whole document	1,12
X	Database EMBL, entry HSN93579, Accession number N93579, 26 August 1996 100% identity to Seq.ID:1 nt.2211-2355 reverse orientation XP002067593 see the whole document	1,12
A	WO 97 07198 A (GENETICS INSTITUTE INC.) 27 February 1997	
A	US 5 536 637 A (JACOBS KENNETH) 16 July 1996 cited in the application	

INTERNATIONAL SEARCH REPORT

In. .national application No.

PCT/US 98/ 06176

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

Remark: Although claim 11 is directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

2. [ ] Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see further information sheet

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [X] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1-12

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: 1-12

   Polynucleotide comprising the nucleotide sequence of Seq.ID:1 and encoding a polypeptide of Seq.ID:2 or fragments. Polynucleotide fragments, variants, homologues, gene thereof. Host cell transformed with said polynucleotide. Protein comprising an amino acid sequence of Seq.ID:2 or fragments thereof. Method for making said recombinant protein. Application of said protein in therapy.

2. Claims: 13-15

   Polynucleotide comprising the nucleotide sequence of Seq.ID:3 and encoding a polypeptide of Seq.ID:4 or fragments. Polynucleotide fragments, variants, homologues, gene thereof. Protein comprising an amino acid sequence of Seq.ID:4 or fragments thereof.

3. Claims: 16-18

   As invention 2 but concerning Seq.ID:5 and 6.

4. Claims: 19-21

   As invention 2 but concerning Seq.ID:7 and 8.

5. Claims: 22-24

   As invention 2 but concerning Seq.ID:9 and 10.

6. Claims: 25-27

   As invention 2 but concerning Seq.ID:11, 12 and 13.

7. Claims: 28-30

   As invention 2 but concerning Seq.ID:14 and 15.

8. Claims: 31-33

   As invention 2 but concerning Seq.ID:16 and 17.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

9. Claims: 34-36

As invention 2 but concerning Seq.ID:18 and 19.

10. Claims: 37-39

As invention 2 but concerning Seq.ID:20 and 21.

INTERNATIONAL SEARCH REPORT

...ormation on patent family members

Inte onal Application No
PCT/US 98/06176

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9707198 A	27-02-97	US 5707829 A	13-01-98
		AU 6712396 A	18-02-97
		AU 6768596 A	12-03-97
		EP 0839196 A	06-05-98
		EP 0851875 A	08-07-98
		WO 9704097 A	06-02-97
US 5536637 A	16-07-96	US 5712116 A	27-01-98