PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification [6] **:**
C07K 14/00

A1

(11) International Publication Number: **WO 99/19348**

(43) International Publication Date: 22 April 1999 (22.04.99)

(21) International Application Number: PCT/US98/21391

(22) International Filing Date: 8 October 1998 (08.10.98)

(30) Priority Data:
60/061,609 10 October 1997 (10.10.97) US

(71) Applicant: THE GENERAL HOSPITAL CORPORATION [US/US]; 55 Fruit Street, Boston, MA 02110 (US).

(72) Inventors: BURGESON, Robert, E.; 84 Naugus Avenue, Marblehead, MA 01945 (US). CHAMPLIAUD, Marie–France; 23 Sheafe Street #3F, Boston, MA 02113 (US). OLSON, Pamela; 127 Wallis Road #3F, Brookline, MA 02167 (US). KOCH, Manuel; 64 Garfield Street, Cambridge, MA 02138 (US). BRUNKEN, William; Dept. of Neurobiology, 140 Commonwealth Avenue, Chestnut Hill, MA 02167 (US).

(74) Agents: MYERS, Paul, Louis et al.; Fish & Richardson P.C., 225 Franklin Street, Boston, MA 02110 (US).

(81) Designated States: AU, CA, JP, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

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: LAMININS AND USES THEREOF

(57) Abstract

The invention is drawn to a purified laminin 12 polypeptide that includes an $\alpha 1$ subunit, a $\beta 2$ subunit and a $\gamma 3$ subunit. The invention is also drawn to isolated laminin $\beta 4$ and $\gamma 3$ subunits.

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.

Code	State	Code	State	Code	State	Code	State
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	**ML**	Mali	**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**MN**	Mongolia	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MR**	Mauritania	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MW**	Malawi	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MX**	Mexico	**US**	United States of America
CA	Canada	**IT**	Italy	**NE**	Niger	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NL**	Netherlands	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NO**	Norway	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NZ**	New Zealand	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**PL**	Poland		
CM	Cameroon	**KR**	Republic of Korea	**PT**	Portugal		
CN	China	**KZ**	Kazakstan	**RO**	Romania		
CU	Cuba	**LC**	Saint Lucia	**RU**	Russian Federation		
CZ	Czech Republic	**LI**	Liechtenstein	**SD**	Sudan		
DE	Germany	**LK**	Sri Lanka	**SE**	Sweden		
DK	Denmark	**LR**	Liberia	**SG**	Singapore		
EE	Estonia						

5

LAMININS AND USES THEREOF

BACKGROUND OF THE INVENTION

The invention relates to the laminin 12, laminin subunit γ3, and laminin subunit β1,
10 and methods of making and using these molecules.

SUMMARY OF THE INVENTION

The present invention is based, in part, on the discovery of a novel member of the laminin family, laminin 12. Accordingly, the present invention features a purified or isolated
15 preparation or a recombinant preparation of laminin 12 which includes an α2 subunit, a β1 subunit and a γ3 subunit.

In a preferred embodiment, the α2 subunit has at least 60% to about 70%, more preferably at least about 80%, even more preferably at least about 90% to about 95%, and most preferably at least about 99% sequence identity with human α2 subunit, e.g., the human
20 α2 subunit of SEQ ID NO:7. The α2 subunit can be identical to a human α2 sequence, e.g., that of SEQ ID NO:7. In another embodiment, the α2 subunit is encoded by a nucleic acid molecule which hybridizes under stringent conditions to a nucleic acid molecule of the nucleic acid sequence shown in SEQ ID NO:8. In addition, the α2 subunit can have substantially the same electrophoretic mobility as human α2 subunit, e.g., it appears as a 205
25 kDa electrophoretic band on reducing gels. Yet another preferred embodiment of the invention features an α2 subunit which is reactive with an α2-specific antibody, e.g., an antibody which binds to the epitope recognized by mAb 5H2. α2 specific antibodies can be made by methods known in the art.

Another preferred embodiment of the invention features a β1 subunit having at least
30 60% to about 70%, more preferably at least about 80%, even more preferably at least about 90% to about 95%, and most preferably at least about 99% sequence identity with human β1 subunit, e.g., the human β1 subunit of SEQ ID NO:9. Preferably, the β1 subunit has the identical amino acid sequence of human β1 subunit, e.g., that of SEQ ID NO:9. In another embodiment, the β1 subunit is encoded by a nucleic acid molecule which hybridizes under
35 stringent conditions to a nucleic acid molecule of the nucleic acid sequence shown in SEQ ID NO:10. In addition, the β1 subunit can have substantially the same electrophoretic mobility as human β1 subunit, e.g., it appears as a 185 kDa electrophoretic band on reducing gels. Yet another preferred embodiment of the invention features an β1 subunit which is reactive with an β1-specific antibody, e.g., an antibody which binds to the epitope recognized by mAb 545.
40 β1-specific antibodies can be made by methods known in the art.

In yet another preferred embodiment, the γ3 subunit of laminin 12 has at least 60% to about 70%, more preferably at least about 80%, even more preferably at least about 90% to about 95%, and most preferably at least about 99% sequence identity with human γ3 subunit,

5 e.g., the γ3 subunit of SEQ ID NO:3. The γ3 subunit can be identical to a naturally occuring human γ3 subunit, e.g., that of SEQ ID NO:3. In another embodiment, the γ3 subunit is encoded by a nucleic acid molecule which hybridizes under stringent conditions to a nucleic acid molecule of the nucleic acid sequence shown in SEQ ID NO:4. In addition, the γ3 subunit can have substantially the same electrophoretic mobility as human γ3 subunit, e.g., it
10 appears as a 170 kDa electrophoretic band on reducing gels. Yet another preferred embodiment of the invention features an γ3 subunit which is reactive with an γ3-specific antibody. γ3-specific antibodies can be made by methods known in the art and taught herein.

In a preferred embodiment, the laminin 12 is a trimer which can be found in, or can be isolated from human placental chorionic villi. In another embodiment, the laminin 12 is
15 expressed by a recombinant cell, e.g., a bacterial cell, a cultured cell (e.g., a cultured eukaryotic cell) or a cell of a non-human transgenic animal. Cultured cells can include CHO cells or SF8 cells. Expression of laminin 12 in a transgenic animal can be general or can be under the control of a tissue specific promoter. Preferably, one or more sequences which encode subunits of the laminin 12 trimer are expressed in a preferred cell-type by a tissue
20 specific promoter, e.g., a milk specific promoter.

The present invention is also based, in part, on the discovery of a novel laminin subunit, γ3. Accordingly, the invention features a recombinant or substantially pure or isolated preparation of a γ3 polypeptide.

In a preferred embodiment, the γ3 polypeptide has the following biological acitivities:
25 1) it promotes adhesion between tissue elements; 2) provides a site for insertion of nerves into the basement membrane. In other preferred embodiments: the γ3 polypeptide includes an amino acid sequence with at least 60%, 80%, 90%, 95%, 98%, or 99% sequence identity to an amino acid sequence from SEQ ID NO:3; the γ3 polypeptide includes an amino acid sequence essentially the same as the amino acid sequence in SEQ ID NO:3; the γ3
30 polypeptide is at least 5, 10, 20, 50, 100, or 150 amino acids in length; the γ3 polypeptide includes at least 5, preferably at least 10, more preferably at least 20, most preferably at least 50, 100, or 150 contiguous amino acids from SEQ ID NO:3; the γ3 polypeptide is either, an agonist or an antagonist, of a biological activity of a naturally occurring γ3 subunit; the γ3 polypeptide is a vertebrate, e.g., a mammalian, e.g. a primate, e.g., a human, γ3 polypeptide.
35 In a preferred embodiment, the invention includes a γ3 polypeptide encoded by a DNA insert of a plasmid deposited with ATCC as Accession No: 209357. In another embodiment, the γ3 polypeptide is a polypeptide encoded by nucleotide sequences of the overlapping DNA inserts of more than one, preferably all seven of the plasmids deposited with ATCC as Accession No:209357.
40 In preferred embodiments: the γ3 polypeptide is encoded by the nucleic acid in SEQ ID NO:4, or by a nucleic acid having at least about 85%, more preferably at least about 90% to about 95%, and most preferably at least about 99% sequence identity with the nucleic acid from SEQ ID NO: 4.

5 In preferred embodiments, the γ3 polypeptide includes a nidogen-binding domain. Generally, the nidogen-binding domain is at least 5 residues in length and preferably, has about 70, 80, 90, or 95% sequence identity with the nidogen-binding domain of the protein shown in SEQ ID NO: 3 (amino acid residues 750-755). In another embodiment, the γ3 polypeptide includes at least 5, preferably 6 to 7, and most preferably 8 of the cysteins found 10 in native γ3 protein. In yet another embodiment of the invention features a γ3 polypeptide that does not include or has an inactivated nidogen-binding domain which serves as an antagonist to γ3 biological activities. Furthermore, a γ3 polypeptide which has antagonist activity can have inactivated or excluded regions which comprise at least one cystein found in native γ3 protein.

15 In a preferred embodiment, the γ3 polypeptide differs in amino acid sequence at up to 1, 2, 3, 5, or 10 residues, from a sequence in SEQ ID NO: 3. In other preferred embodiments, the γ3 polypeptide differs in amino acid sequence at up to 1, 2, 3, 5, or 10 % of the residues from a sequence in SEQ ID NO: 3. Preferably, the differences are such that: the γ3 polypeptide exhibits a γ3 biological activity, e.g., the γ3 polypeptide retains a biological 20 activity of a naturally occurring γ3 subunit.

In preferred embodiments the γ3 polypeptide includes a γ3 subunit sequence described herein as well as other N-terminal and/or C-terminal amino acid sequence.

In preferred embodiments, the γ3 polypeptide includes all or a fragment of an amino acid sequence from SEQ ID NO: 3, fused, in reading frame, to additional amino acid residues, 25 preferably to residues encoded by genomic DNA 5' to the genomic DNA which encodes a sequence from SEQ ID NO: 3.

In yet other preferred embodiments, the γ3 polypeptide is a recombinant fusion protein having a first γ3 portion and a second polypeptide portion, e.g., a second polypeptide portion having an amino acid sequence unrelated to γ3. The second polypeptide portion can 30 be, e.g., any of glutathione-S-transferase, a DNA binding domain, or a polymerase activating domain. In preferred embodiment the fusion protein can be used in a two-hybrid assay.

In a preferred embodiment the γ3 polypeptide includes amino acid residues 750-755 of SEQ ID NO:3. In another embodiment, the γ3 polypeptide encodes domains IV-VI of the γ3 subunit.

35 In preferred embodiments the γ3 polypeptide has antagonistic activity, and is capable of: inhibiting adhesion between connective tissues.

In a preferred embodiment, the γ3 polypeptide is a fragment of a naturally occurring γ 3 which inhibits connective tissue adhesion.

Polypeptides of the invention include those which arise as a result of the existence of 40 multiple genes, alternative transcription events, alternative RNA splicing events, and alternative translational and postranslational events. The γ3 polypeptide can be expressed in systems, e.g., cultured cells, which result in substantially the same postranslational modifications present when expressed γ3 is expressed in a native cell, or in systems which

5 result in the omission of postranslational modifications present when expressed in a native cell.

The invention includes an immunogen which includes a γ3 polypeptide in an immunogenic preparation, the immunogen being capable of eliciting an immune response specific for the γ3 polypeptide, e.g., a humoral response, an antibody response, or a cellular
10 response. In preferred embodiments, the immunogen comprising an antigenic determinant, e.g., a unique determinant, from a protein represented by SEQ ID NO: 3.

The present invention also includes an antibody preparation specifically reactive with an epitope of the γ3 immunogen or generally of a γ3 polypeptide, preferably an epitope which consists all or in part of residues from the the amino acid sequence of SEQ ID NO:3, or an
15 epitope, which when bound to an antibody, results in the modulation of a biological activity.

In preferred embodiments the γ3-like polypeptide, as expressed in the cells in which it is normally expressed or in other eukaryotic cells, has a molecular weight of 170 kDa as determined by SDS-PAGE.

In another embodiment, the γ3 polypeptide comprises amino acid residues 100-1761
20 of SEQ ID NO: 3.

In a preferred embodiment, the γ3 polypeptide has one or more of the following characteristics:

(i) it has the ability to promote adhesion between connective tissues;

(ii) it has a molecular weight, amino acid composition or other physical
25 characteristic of γ3 subunit of SEQ ID NO:3;

(iii) it has an overall sequence similarity of at least 50%, preferably at least 60%, more preferably at least 70, 80, 90, or 95%, with a γ3 polypeptide of SEQ ID NO:3;

(iv) it can be isolated from human placenta chorionic villi;

(v) it has a nidogen-binding domain which is preferably about 70%, 80%,
30 90% or 95% with amino acid residues 750-755 of SEQ ID NO:3;

(vi) it can colocalize with protein ubiquitin carboxy terminal hydroxylase I;

(vii) it has at least 5, preferably 6 or 7, and most preferably 8 of the cysteins found amino acid sequence of native γ3.

Also included in the invention is a composition which includes a γ3 polypeptide (or a
35 nucleic acid which encodes it) and one or more additional components, e.g., a carrier, diluent, or solvent. The additional component can be one which renders the composition useful for *in vitro* and *in vivo* pharmaceutical or veterinary use.

In another aspect, the invention provides an isolated or substantially pure nucleic acid having or comprising a nucleotide sequence which encodes a γ3 polypeptide, e.g., a γ3
40 polypeptide described herein.

A preferred embodiment of the invention features a nucleic acid molecule having a nucleotide sequence at least about 85% sequence identity to a nucleotide sequence of SEQ ID NO:4. In other preferred embodiments, the γ3 polypeptide is encoded by a nucleic acid

5 molecule having a nucleotide sequence with at least about 90% to about 95%, and more preferably about 98% to about 99% sequence identity to the nucleotide sequence from SEQ ID NO:4. In another preferred embodiment, the γ3 polypeptide is encoded by the nulceic acid molecule of SEQ ID NO:4.

In prefered embodiments, the isolated nucleic acid molecule includes the nucleotide 10 sequence of at least one and preferably all of the DNA inserts of the plasmids deposited with ATCC as Accession No: 209357.

In preferred embodiments, the subject γ3 nucleic acid will include a transcriptional regulatory sequence, e.g. at least one of a transcriptional promoter or transcriptional enhancer sequence, operably linked to the γ3 gene sequence (also referred to as LAMG3), e.g., to 15 render the γ3 gene sequence suitable for use as an expression vector.

In yet a further preferred embodiment, the nucleic acid which encodes a γ3 polypeptide of the invention, hybridizes under stringent conditions to a nucleic acid probe corresponding to at least 12 consecutive nucleotides of SEQ ID NO:4. More preferably, the nucleic acid probe corresponds to at least 20 consecutive nucleotides from SEQ ID NO: 4.

20 The invention also provides a probe or primer which includes or comprises a substantially purified oligonucleotide. The oligonucleotide includes a region of nucleotide sequence which hybridizes under stringent conditions to at least 10 consecutive nucleotides of sense or antisense sequence from SEQ ID NO: 4, or naturally occurring mutants thereof. In preferred embodiments, the probe or primer further includes a label group attached thereto. 25 The label group can be, e.g., a radioisotope, a fluorescent compound, an enzyme, and/or an enzyme co-factor. Preferably the oligonucleotide is at least 10 and less than 20, 30, 50, 100, or 150 nucleotides in length.

The invention involves nucleic acids, e.g., RNA or DNA, encoding a γ3 polypeptide of the invention. This includes double stranded nucleic acids as well as coding and antisense 30 single strands.

In another aspect, the invention features a cell or purified preparation of cells which include a γ3 subunit transgene, or which otherwise misexpress a γ3 gene. The cell preparation can consist of human or non human cells, e.g., rodent cells, e.g., mouse or rat cells, rabbit cells, or pig cells. In preferred embodiments, the cell or cells include a γ3 35 transgene, e.g., a heterologous form of a γ3 gene, e.g., a gene derived from humans (in the case of a non-human cell). The γ3 transgene can be misexpressed, e.g., overexpressed or underexpressed. In other preferred embodiments, the cell or cells include a gene which misexpress an endogenous γ3 gene, e.g., a gene the expression of which is disrupted, e.g., a knockout. Such cells can serve as a model for studying disorders which are related to 40 mutated or mis-expressed γ3 alleles or for use in drug screening.

In another aspect, the invention features a transgenic γ3 animal, e.g., a rodent, e.g., a mouse or a rat, a rabbit, a pig, a goat, or a cow. In preferred embodiments, the transgenic animal includes (and preferably express) a heterologous form of a γ3 gene, e.g., a gene

5 derived from humans. In a further embodiment, the γ3 transgene includes a tissue specific promoter, e.g., a milk-specific promoter. In other preferred embodiments, the animal has an endogenous γ3 gene which is misexpressed, e.g., a knockout. Such a transgenic animal can serve as a model for studying disorders which are related to mutated or mis-expressed γ3 alleles or for use in drug screening.

10 The invention is also based, in part, on the discovery of a novel laminin subunit, β4. Accordingly, the invention features a recombinant or substantially pure preparation of a β4 polypeptide.

In preferred embodiment, the β4 polypeptide has the following biological activities: 1) it promotes adhesion between tissue elements; 2) it aids in wound healing. In other preferred 15 embodiments: the β4 polypeptide includes an amino acid sequence with at least 65%, 80%, 90%, 95%, 98%, or 99% sequence identity to an amino acid sequence from SEQ ID NO:1; the β4 polypeptide includes an amino acid sequence essentially the same as an amino acid sequence in SEQ ID NO: 1; the β4 polypeptide is at least 5, 10, 20, 50, 100, or 150 amino acids in length; the β4 polypeptide includes at least 5, preferably at least 10, more preferably 20 at least 20, most preferably at least 50, 100, or 150 contiguous amino acids from SEQ ID NO:1; the β4 polypeptide is either, an agonist or an antagonist, of a biological activity of a naturally occurring β4 subunit; the β4 polypeptide is a vertebrate, e.g., a mammalian, e.g. a primate, e.g., a human, β4 polypeptide.

In preferred embodiments: the β4 polypeptide is encoded by the nucleic acid in SEQ 25 ID NO:2, or by a nucleic acid having at least about 65% to about 70%, more preferably at least 80%, even more preferably at least about 90% to about 95%, and most preferably about 99% sequence identity with the nucleic acid from SEQ ID NO: 2.

In preferred embodiments, the β4 polypeptide includes domains VI and V found in native β4 subunits. Amino acid residues from about 221-262 and 263-535 of SEQ ID NO: 1 30 are exemplary of domains VI and V, respectively, of β4. Generally, domain VI is at least 33 residues in length and has preferably at least about 60%, more preferably about 70% to about 80%, and most preferably about 90% to about 95% sequence identity with the amino acid residues 221-262 of the β4 protein shown in SEQ ID NO: 1. Domain V is at least 272 residues in length and has preferably at least about 60%, more preferably about 70% to about 35 80%, and most preferably about 90% to about 95% sequence identity with the amino acid residues 263-535 of the β4 protein shown in SEQ ID NO: 1. In another embodiment, the β4 polypeptide has at least 5, preferably 6 or 7, and most preferably 8 cysteins as found in native β4. In yet another embodiment, a β4 polypeptide which has antagonist activity has inactivated or excluded regions which comprise at least one of the cysteins found in native β4 40 protein.

In a preferred embodiment, the β4 polypeptide differs in amino acid sequence at up to 1, 2, 3, 5, or 10 residues, from a sequence in SEQ ID NO: 1. In other preferred embodiments, the β4 polypeptide differs in amino acid sequence at up to 1, 2, 3, 5, or 10 % of the residues

from a sequence in SEQ ID NO: 1. Preferably, the differences are such that: the β4 polypeptide exhibits a β4 biological activity, e.g., the β4 polypeptide retains a biological activity of a naturally occurring β4 subunit.

In preferred embodiments the β4 polypeptide includes a β4 sequence described herein as well as other N-terminal and/or C-terminal amino acid sequence.

In preferred embodiments, the β4 polypeptide includes all or a fragment of an amino acid sequence from SEQ ID NO:1, fused, in reading frame, to additional amino acid residues, preferably to residues encoded by genomic DNA 5' to the genomic DNA which encodes a sequence from SEQ ID NO:1.

In yet other preferred embodiments, the β4 polypeptide is a recombinant fusion protein having a first β4 portion and a second polypeptide portion, e.g., a second polypeptide portion having an amino acid sequence unrelated to β4. The second polypeptide portion can be, e.g., any of glutathione-S-transferase, a DNA binding domain, or a polymerase activating domain. In preferred embodiment the fusion protein can be used in a two-hybrid assay.

In preferred embodiments the β4 polypeptide has antagonistic activity, and is capable of: inhibiting the adhesion of connective tissues.

Preferably, the β4 polypeptide is a fragment of a naturally occurring β4 which inhibits connective tissue adhesion.

Polypeptides of the invention include those which arise as a result of the existence of multiple genes, alternative transcription events, alternative RNA splicing events, and alternative translational and postranslational events. In one aspect of the invention, the β4 polypeptide is a splice variant of the β4 subunit. In another preferred embodiment, the β4 splice variant is encoded by a nucleic acid molecule identical to the nucleotide sequence of SEQ ID NO:6. The polypeptide can be expressed in systems, e.g., cultured cells, which result in substantially the same postranslational modifications present when expressed β4 is expressed in a native cell, or in systems which result in the omission of postranslational modifications present when expressed in a native cell.

The invention includes an immunogen which includes a β4 polypeptide in an immunogenic preparation, the immunogen being capable of eliciting an immune response specific for the β4 polypeptide, e.g., a humoral response, an antibody response, or a cellular response. In preferred embodiments, the immunogen comprising an antigenic determinant, e.g., a unique determinant, from a protein represented by SEQ ID NO: 1.

The present invention also includes an antibody preparation specifically reactive with an epitope of the β4 immunogen or generally of a β4 polypeptide, preferably an epitope which consists all or in part of residues from the amino acid sequence of SEQ ID NO:1, or an epitope, which when bound to an antibody, results in the modulation of a biological activity.

In preferred embodiments the β4-like polypeptide, as expressed in the cells in which it is normally expressed or in other eukaryotic cells, has an estimated molecular weight of 200 kDa as determined by SDS-PAGE.

5 In a preferred embodiment, the β4 polypeptide has one or more of the following characteristics:

(i) it has the ability to promote adhesion between connective tissues;

(ii) it has a molecular weight, amino acid composition or other physical characteristic of β4 subunit of SEQ ID NO:1;

10 (iii) it has an overall sequence similarity of at least 50%, preferably at least 65%, more preferably at least 70, 80, 90, or 95%, with a β4 polypeptide of SEQ ID NO:1;

(iv) it can be isolated from human placenta chorionic villi;

(v) it can associate with α3 or γ2 subunits;

(vi) it has coiled coils in domains I and II.

15 (vii) it has at least 5, preferably 6 or 7, and most preferably 8 of the cysteins found in native β4 sequence.

Also included in the invention is a composition which includes a β4 polypeptide (or a nucleic acid which encodes it) and one or more additional components, e.g., a carrier, diluent, or solvent. The additional component can be one which renders the composition for *in vitro*
20 and *in vivo* pharmaceutical or veterinary use. Such uses can include aiding in wound healing or promotion of the adhesion of dermal and epidermal cells.

In another aspect, the invention provides an isolated or substantially pure nucleic acid having or comprising a nucleotide sequence which encodes a β4 polypeptide, e.g., a β4 polypeptide described herein.

25 A preferred embodiment of the invention features a nucleic acid molecule having a nucleotide sequence at least about 65% sequence identity to a nucleotide sequence of SEQ ID NO:2. In other preferred embodiments, the β4 polypeptide is encoded by a nucleic acid molecule having a nucleotide sequence with at least 70%, preferably 80%, more preferably about 90% to about 95%, and even more preferably about 99% sequence identity to the
30 nucleotide sequence from SEQ ID NO:2. In another preferred embodiment, the β4 polypeptide is encoded by the nulceic acid molecule of SEQ ID NO:2.

In preferred embodiments, the subject β4 nucleic acid will include a transcriptional regulatory sequence, e.g. at least one of a transcriptional promoter or transcriptional enhancer sequence, operably linked to the β4 gene sequence (also referred to as LAMB4), e.g., to
35 render the β4 gene sequence suitable for use as an expression vector.

In yet a further preferred embodiment, the nucleic acid which encodes a β4 polypeptide of the invention, hybridizes under stringent conditions to a nucleic acid probe corresponding to at least 12 consecutive nucleotides from SEQ ID NO:2, more preferably to at least 20 consecutive nucleotides from SEQ ID NO:2.

40 In a preferred embodiment, the nucleic acid differs by at least one nucleotide from a nucleotide sequence of SEQ ID NO:2, nucleotides 4686-5870.

The invention also provides a probe or primer which includes or comprises a substantially purified oligonucleotide. The oligonucleotide includes a region of nucleotide

5 sequence which hybridizes under stringent conditions to at least 10 consecutive nucleotides of sense or antisense sequence from SEQ ID NO: 2, or naturally occurring mutants thereof. In preferred embodiments, the probe or primer further includes a label group attached thereto. The label group can be, e.g., a radioisotope, a fluorescent compound, an enzyme, and/or an enzyme co-factor. Preferably the oligonucleotide is at least 10 and less than 20, 30, 50, 100,
10 or 150 nucleotides in length.

The invention involves nucleic acids, e.g., RNA or DNA, encoding a β4 polypeptide of the invention. This includes double stranded nucleic acids as well as coding and antisense single strands.

In another aspect, the invention features a cell or purified preparation of cells which
15 include a β4 transgene, or which otherwise misexpress a β4 gene. The cell preparation can consist of human or non human cells, e.g., rodent cells, e.g., mouse or rat cells, rabbit cells, or pig cells. In preferred embodiments, the cell or cells include a β4 transgene, e.g., a heterologous form of a β4 gene, e.g., a gene derived from humans (in the case of a non-human cell). The β4 transgene can be misexpressed, e.g., overexpressed or underexpressed.
20 In other preferred embodiments, the cell or cells include a gene which misexpress an endogenous β4 gene, e.g., a gene the expression of which is disrupted, e.g., a knockout. Such cells can serve as a model for studying disorders which are related to mutated or mis-expressed β4 alleles or for use in drug screening.

In another aspect, the invention features a transgenic β4 animal, e.g., a rodent, e.g., a
25 mouse or a rat, a rabbit, a pig, a goat, or a cow. In preferred embodiments, the transgenic animal includes (and preferably express) a heterologous form of a β4 gene, e.g., a gene derived from humans. In a further embodiment, the β4 transgene includes a tissue specific promoter, e.g., a milk-specific promoter. In other preferred embodiments, the animal has an endogenous β4 gene which is misexpressed, e.g., a knockout. Such a transgenic animal can
30 serve as a model for studying disorders which are related to mutated or mis-expressed β4 alleles or for use in drug screening.

In another aspect, the invention features, a method of promoting adhesion of a first tissue element to a second tissue element. The method includes contacting one or both of the first tissue element and the second tissue element with an amount of a laminin molecule
35 described herein, e.g., laminin 12, or γ3 (or a laminin trimer which inlcudes γ3), sufficient to promote adhesion. The method can be performed in vivo, or in vitro. In in vivo methods the laminin is administered to the subject. The administration can be directed to the site where adhesion is desired, e.g., by topical appication or by injection, or administered in a systemic fashion.

40 A tissue element can be a cell or a multi-cellular on acellular structure. Examples of tissue elements include, skin cells, e.g., epidermal or dermal cells, neuronal cells, e.g., nerve cells, retinal cells, central or pereipheral nervous system components, basement membrane or components of the basement membrane, or any cell or structure which in normal, non-

5 traumatized, or non-diseased tissue is adjacent or adhered to a specific tissue element recited herein.

In preferred embodiments the molecule is exogenous (e.g., administered to a subject) or is recombinant.

In preferred embodiments the method is an vivo method. In vivo methods can be
10 autologous, allogeneic, or xenogeneic. In autologous methods, adhesion between two tissue elements from the subject is promoted. In allogeneic methods, adhesion between a recipient tissue element and a donor tissue element from an allogeneic donor is promoted. In xenogeneic methods, adhesion between a recipient tissue element and a donor tissue element from a xenogeneic donor is promoted. Thus, one element can be a donor tissue element
15 which is implanted into a recipient subject.

In preferred embodiments the first tissue is healthy tissue, e.g., skin tissue, and the second tissue is wounded, e.g., burned, diseased, traumatized, cut, and the tissue, or is a wound bed. For example, the first tissue is skin tissue, from the subject or from a donor, and the second tissue is wounded, e.g., burned or abraided tissue.

20 In preferred embodiments the first tissue and second tissue element are normally adhered but have become detached from one another due to trauma, burn or other physical injury, disease, or age.

In preferred embodiments: the first tissue element is a dermal cell and the second tissue element is an epidermal cell; the first tissue element is a nerve cell or nerve and the
25 second tissue element is a cell or structure which in normal, non-traumatized, or non-diseased tissue is adjacent or adhered to the nerve cell or nerve; the first tissue element is a retinal cell or retina tissue and the second tissue element is a cell or structure which in normal, non-traumatized, or non-diseased tissue is adjacent or adhered to the a retinal cell or retina tissue, the first tissue is a nerve and the second tissue is basement membrane.

30 The administration of laminin can be repeated.

In another aspect, the invention features a method of promoting wound healing in a subject. The method includes administering an amount of a laminin molecule described herein, e.g., laminin 12, $\gamma3$ (or a laminin trimer which inlcudes $\gamma3$), sufficient to promote healing to the wound. The administration can be directed to the site where healing is desired,
35 e.g., by topical appication or by injection, or administered in a systemic fashion.

The wound can be in any tissue, but preferably ina tissue in which the laminin normally occurs. Examples skin, central or peripheral nervous tissue, tissues of the eye, e.g., the retinal, the basement membrane, or any tissue which in normal, non-traumatized, or non-diseased tissue is adjacent or adhered thereto.

40 In preferred embodiments the molecule is exogenous (e.g., administered to a subject) or is recombinant.

In preferred embodiments the wound tissue is burned, diseased, traumatized, cut, the subject of immune attack, e.g, autoimmune attack, or abraided.

5 The administration of laminin can be repeated.

In another aspect, the invention features a method of promoting nerve growth or regeneration in a subject. The method includes administering an amount of a laminin molecule described herein, e.g., laminin 12, or γ3 (or a laminin trimer which inlcudes γ3), sufficient to promote nerve growth or regeneration. The administration can be directed to the 10 site where nerve growth or regeneration is desired, e.g., by topical appication or by injection, or administered in a systemic fashion.

In preferred embodiments the molecule is exogenous (e.g., administered to a subject) or is recombinant.

In preferred embodiments the nerve growth or regeneration is promoted at a wound 15 site.

The administration of laminin can be repeated.

In another aspect, the invention provides, a method of determining if a subject is at risk for a disorder related to a lesion in or the misexpression of a gene which encodes a laminin described herein, e.g., γ3 or laminin 12.

20 Such disorders include, e.g., a disorder associated with the misexpression of a laminin, e.g., laminin 12, or misexpression of the γ3 subunit; a disorder of the central or peripheral nervous system; a disorder associated with a genetic lesion at chromosome 9, region q31-34; Fukuyama-type muscular dystrophy; muscle-eye-brain disease; Walker-Warburg Syndrome (hydrocephalus, ageria, and retinal displasia); a retinal disorder, e.g, retinitis pigmentosa-deafness syndrome (which may be a 25 subtype of Walker-Warburg Syndrome); a disorder associated with abnormal levels, e.g., abnormally low levels, of adhesion between tissues; a disorder associated with the basement membrane; a skin disorder, e.g., an epidermal or dermal, disorder; a disorder associated with the testis, spleen, placenta, thymus, ovary, small intestine, lung, or liver.

The method includes one or more of the following:

30 detecting, in a tissue of the subject, the presence or absence of a mutation which affects the expression of the γ3 gene, or other gene which encodes a subunit of laminin 12, e.g., detecting the presence or absence of a mutation in a region which controls the expression of the gene, e.g., a mutation in the 5' control region;

detecting, in a tissue of the subject, the presence or absence of a mutation which alters 35 the structure of the γ3 gene, or other gene which encodes a subunit of laminin 12;

detecting, in a tissue of the subject, the misexpression of theγ3 gene, or other gene which encodes a subunit of laminin 12 at the mRNA level, e.g., detecting a non-wild type level of a γ3, or an other laminin 12 subunit mRNA ;

detecting, in a tissue of the subject, the misexpression of the γ3 gene, or other gene 40 which encodes a subunit of laminin 12, at the protein level, e.g., detecting a non-wild type level of a γ3, or an other laminin 12 subunit polypeptide.

In preferred embodiments the method includes: ascertaining the existence of at least one of: a deletion of one or more nucleotides from the γ3 gene, or other gene which encodes a

5 subunit of laminin 12; an insertion of one or more nucleotides into the gene, a point mutation, e.g., a substitution of one or more nucleotides of the gene, a gross chromosomal rearrangement of the gene, e.g., a translocation, inversion, or deletion.

For example, detecting the genetic lesion can include: (i) providing a probe/primer including an oligonucleotide containing a region of nucleotide sequence which hybridizes to a 10 sense or antisense sequence from SEQ ID NO:4, or naturally occurring mutants thereof or 5' or 3' flanking sequences naturally associated with the LAMG3 gene; (ii) exposing the probe/primer to nucleic acid of the tissue; and detecting, by hybridization, e.g., *in situ* hybridization, of the probe/primer to the nucleic acid, the presence or absence of the genetic lesion.

15 In preferred embodiments detecting the misexpression includes ascertaining the existence of at least one of: an alteration in the level of a messenger RNA transcript of the γ3 gene, or other gene which encodes a subunit of laminin 12; the presence of a non-wild type splicing pattern of a messenger RNA transcript of the γ3 gene, or other gene which encodes a subunit of laminin 12; or a non-wild type level of γ3, or other subunit of laminin 12.

20 Methods of the invention can be used prenatally or to determine if a subject's offspring will be at risk for a disorder.

In preferred embodiments the method includes determining the structure of a γ3 gene, or other gene which encodes a subunit of laminin 12, an abnormal structure being indicative of risk for the disorder.

25 In preferred embodiments the method includes contacting a sample form the subject with an antibody to the laminin protein or a nucleic acid which hybridizes specifically with the γ3 gene, or other gene which encodes a subunit of laminin 12.

In another aspect, the invention features, a method of promoting adhesion of a first tissue element to a second tissue element. The method includes contacting one or both of the 30 first tissue element and the second tissue element with an amount of a laminin molecule described herein, e.g., β4, sufficient to promote adhesion. The method can be performed in vivo, or in vitro. In in vivo methods the laminin is administered to the subject. The administration can be directed to the site where adhesion is desired, e.g., by topical application or by injection, or administered in a systemic fashion.

35 A tissue element can be a cell or a multi-cellular on acellular structure. Examples of tissue elements include, skin cells, e.g., epidermal or dermal cells, neuronal cells, e.g., nerve cells, retinal cells, central or pereipheral nervous system components, basement membrane or components of the basement membrane, or any cell or structure which in normal, non-traumatized, or non-diseased tissue is adjascent or adhered to a specific tissue element recited 40 herein.

In preferred embodiments the molecule is exogenous (e.g., administered to a subject) or is recombinant.

5 In preferred embodiments the method is an vivo method. In vivo methods can be autologous, allogeneic, or xenogeneic. In autologous methods, adhesion between two tissue elements from the subject is promoted. In allogeneic methods, adhesion between a recipient tissue element and a donor tissue element from an allogeneic donor is promoted. In xenogeneic methods, adhesion between a recipient tissue element and a donor tissue element
10 from a xenogeneic donor is promoted. Thus, one element can be a donor tissue element which is implanted into a recipient subject.

In preferred embodiments the first tissue is healthy tissue, e.g., skin tissue, and the second tissue is wounded, e.g., burned, diseased, traumatized, cut, and the tissue, or is a wound bed. For example, the first tissue is skin tissue, from the subject or from a donor, and
15 the second tissue is wounded, e.g., burned or abraided tissue.

In preferred embodiments: the first tissue element is a dermal cell and the second tissue element is an epidermal cell; the first tissue element is a nerve cell or nerve and the second tissue element is a cell or structure which in normal, non-traumatized, or non-diseased tissue is adjascent or adhered to the nerve cell or nerve; the first tissue is a nerve and the
20 second tissue is basement membrane.

The administration of laminin can be repeated.

In another aspect, the invention features a method of promoting wound healing in a subject. The method includes administering an amount of a laminin molecule described herein, e.g., β4, sufficient to promote healing to the wound. The administration can be
25 directed to the site where healing is desired, e.g., by topical appication or by injection, or administered in a systemic fashion.

The wound can be in any tissue, but preferably in a tissue in which the laminin normally occurs in fetal or adult life. Examples examples include skin the basement membrane.

30 In preferred embodiments the molecule is exogenous (e.g., administered to a subject) or is recombinant.

In preferred embodiments the wound tissue is burned, diseased, traumatized, cut, the subject of immune attack, e.g, autoimmune attack, or abraded.

The administration of laminin can be repeated.

35 In another aspect, the invention features a method of promoting tissue growth, development, or regeneration in a subject. The method includes administering an amount of a laminin molecule described herein, e.g., β4, sufficient to promote tissue growth, development, or regeneration in a subject. The administration can be directed to the site where nerve growth or regeneration is desired, e.g., by topical appication or by injection, or
40 administered in a systemic fashion.

In preferred embodiments the molecule is exogenous (e.g., administered to a subject) or is recombinant.

5 In preferred embodiments the nerve growth or regeneration is promoted at a wound site.

The administration of laminin can be repeated.

In another aspect, the invention provides, a method of determining if a subject is at risk for a disorder related to a lesion in or the misexpression of a laminin molecule described herein, e.g., β4.

10 Such disorders include, e.g., a disorder associated with the misexpression of a laminin, e.g., β 4; a disorder associated with a genetic lesion at chromosome region 7q22-q31.2; a developmetnal disorder; a disorder associated with abnormal levels, e.g., abnormally low levels, of adhesion between tissues; a disorder associated with the basement membrane; a skin disorder, e.g., an epidermal or dermal, disorder.

15 The method includes one or more of the following:

detecting, in a tissue of the subject, the presence or absence of a mutation which affects the expression of the β4 gene, e.g, detecting the presence or absence of a mutation in a region which controls the expression of the gene, e.g., a mutation in the 5' control region;

detecting, in a tissue of the subject, the presence or absence of a mutation which alters 20 the structure of the β4 gene;

detecting, in a tissue of the subject, the misexpression of the β4 gene, e.g., detecting a non-wild type level of a β4 mRNA ;

detecting, in a tissue of the subject, the misexpression of the β4, at the protein level, e.g., detecting a non-wild type level of a β4 polypeptide.

25 In preferred embodiments the method includes: ascertaining the existence of at least one of: a deletion of one or more nucleotides from the β4; an insertion of one or more nucleotides into the gene, a point mutation, e.g., a substitution of one or more nucleotides of the β4 gene, a gross chromosomal rearrangement of the β4 gene, e.g., a translocation, inversion, or deletion.

30 For example, detecting the genetic lesion can include: (i) providing a probe/primer including an oligonucleotide containing a region of nucleotide sequence which hybridizes to a sense or antisense sequence from SEQ ID NO:2, or naturally occurring mutants thereof or 5' or 3' flanking sequences naturally associated with the LAMB4 gene; (ii) exposing the probe/primer to nucleic acid of the tissue; and detecting, by hybridization, e.g., *in situ* 35 hybridization, of the probe/primer to the nucleic acid, the presence or absence of the genetic lesion.

In preferred embodiments: detecting the misexpression includes ascertaining the existence of at least one of: an alteration in the level of a messenger RNA transcript of the β 4; the presence of a non-wild type splicing pattern of a messenger RNA transcript of the β4; 40 or a non-wild type level of β4.

Methods of the invention can be used prenatally or to determine if a subject's offspring will be at risk for a disorder.

5 In preferred embodiments the method includes determining the structure of the a β4, an abnormal structure being indicative of risk for the disorder.

In preferred embodiments the method includes contacting a sample form the subject with an antibody to the β4 protein or a nucleic acid which hybridizes specifically with the β4.

In another aspect, the invention features, a method of evaluating a compound for the
10 ability to interact with, e.g., bind, a subject laminin polypeptide, e.g., laminin 12, γ3, a laminin trimer which inlcudes γ3, β4, or a laminin trimer which includes β4. The method includes: contacting the compound with the subject laminin polypeptide; and evaluating ability of the compound to interact with, e.g., to bind or form a complex with the subject laminin polypeptide. This method can be performed in vitro, e.g., in a cell free system, or in
15 vivo, e.g., in a two-hybrid interaction trap assay. This method can be used to identify naturally occurring molecules which interact with subject laminin polypeptide. It can also be used to find natural or synthetic inhibitors of subject laminin polypeptide.

In another aspect, the invention features, a method of evaluating a compound, e.g., a polypeptide, e.g., a naturally occurring ligand of or a naturally occuring substrate to which
20 binds a subject laminin polypeptide, e.g., of laminin 12, γ3, a laminin trimer which inlcudes γ 3, β4, or a laminin trimer which includes β4, for the ability to bind a subject laminin polypeptide. The method includes: contacting the compound with the subject laminin polypeptide; and evaluating the ability of the compound to interact with, e.g., to bind or form a complex with the subject laminin polypeptide, e.g., the ability of the compound to inhibit a
25 subject laminin polypeptide/ligand interaction. This method can be performed in vitro, e.g., in a cell free system, or in vivo, e.g., in a two-hybrid interaction trap assay. This method can be used to identify compounds, e.g., fragments or analogs of a subject laminin polypeptide, which are agonists or antagonists of a subject laminin polypeptide.

In another aspect, the invention features, a method of evaluating a first compound,
30 e.g., a subject laminin polypeptide, e.g., laminin 12, γ3, a laminin trimer which inlcudes γ3, β 4, or a laminin trimer which includes β4, for the ability to bind a second compound, e.g., a second polypeptide, e.g., a naturally occurring ligand of or substrate to which binds a subject laminin polypeptide. The method includes: contacting the first compound with the second compound; and evaluating the ability of the first compound to form a complex with the
35 second compound. This method can be performed in vitro, e.g., in a cell free system, or in vivo, e.g., in a two-hybrid interaction trap assay. This method can be used to identify compounds, e.g., fragments or analogs of a subject laminin polypeptide, which are agonists or antagonists of a subject laminin polypeptide.

In yet another aspect, the invention features a method for evaluating a compound, e.g.,
40 for the ability to modulate an interaction, e.g., the ability to inhibit an interaction of a subject laminin polypeptide, e.g., of laminin 12, γ3, a laminin trimer which inlcudes γ3, β4, or a laminin trimer which includes β4, with a second polypeptide, e.g., a polypeptide, e.g., a natural ligand of the of or a substrate wo which binds a subject laminin polypeptide, or a

5 fragment thereof. The method includes the steps of (i) combining the second polypeptide (or preferably a purified preparation thereof), a subject laminin polypeptide, (or preferably a purified preparation thereof), and a compound, e.g., under conditions wherein in the absence of the compound, the second polypeptide, and the subject laminin polypeptide, are able to interact, e.g., to bind or form a complex; and (ii) detecting the interaction, e.g., detecting the
10 formation (or dissolution) of a complex which includes the second polypeptide, and the subject laminin polypeptide. A change, e.g., a decrease or increase, in the formation of the complex in the presence of a compound (relative to what is seen in the absence of the compound) is indicative of a modulation, e.g., an inhibition or promotion, of the interaction between the second polypeptide, and the subject laminin polypeptide. In preferred
15 embodiments: the second polypeptide, and the subject laminin polypeptide, are combined in a cell-free system and contacted with the compound; the cell-free system is selected from a group consisting of a cell lysate and a reconstituted protein mixture; the subject laminin polypeptide, and the second polypeptide are simultaneously expressed in a cell, and the cell is contacted with the compound, e.g. in an interaction trap assay (e.g., a two-hybrid assay).
20 In yet another aspect, the invention features a two-phase method (e.g., a method having an in vitro, e.g., in a cell free system, and an in vivo phase) for evaluating a compound, e.g., for the ability to modulate, e.g., to inhibit or promote, an interaction of a subject laminin polypeptide subject laminin polypeptide, e.g., of laminin 12, γ3, a laminin trimer which inlcudes γ3, β4, or a laminin trimer which includes β4, with a second
25 compound, e.g., a second polypeptide, e.g., a naturally occurring ligand of or a substrate to which binds a subject laminin polypeptide, or a fragment thereof. The method includes steps (i) and (ii) of the method described immediately above performed in vitro, and further includes: (iii) determining if the compound modulates the interaction in vitro, e.g., in a cell free system, and if so; (iv) administering the compound to a cell or animal; and (v)
30 evaluating the in vivo effect of the compound on an interaction, e.g., inhibition, of a subject laminin polypeptide, with a second polypeptide.

In another aspect, the invention features, a method of evaluating a compound for the ability to bind a nucleic acid encoding a subject laminin polypeptide, e.g., a laminin 12, γ3, a laminin trimer which inlcudes γ3, β4, or a laminin trimer which includes β4 polypeptide
35 regulatory sequence. The method includes: contacting the compound with the nucleic acid; and evaluating ability of the compound to form a complex with the nucleic acid.

In another aspect, the invention features a method of making a γ3 or β4 polypeptide, e.g., a peptide having a non-wild type activity, e.g., an antagonist, agonist, or super agonist of a naturally occurring γ3 or β4 polypeptide, e.g., a naturally occurring γ3 or β4 polypeptide.
40 The method includes: altering the sequence of a γ3 or β4 polypeptide, e.g., altering the sequence , e.g., by substitution or deletion of one or more residues of a non-conserved region, a domain or residue disclosed herein, and testing the altered polypeptide for the desired activity.

5 In another aspect, the invention features a method of making a fragment or analog of a γ3 or β4 polypeptide having a biological activity of a naturally occurring γ3 or β4 polypeptide. The method includes: altering the sequence, e.g., by substitution or deletion of one or more residues, of a γ3 or β4 polypeptide, e.g., altering the sequence of a non-conserved region, or a domain or residue described herein, and testing the altered polypeptide
10 for the desired activity.

In another aspect, the invention features, a human cell, e.g., a hematopoietic stem cell, transformed with nucleic acid which encodes a subject laminin polypeptide, e.g., a laminin 12, γ3, a laminin trimer which inlcudes γ3, β4, or a laminin trimer which includes β4.

In another aspect, the invention includes: a γ3, β4 nucleic acid, e.g., a γ3, β4 nucleic
15 acid inserted into a vector; a cell transformed with a γ3, β4 nucleic acid; a γ3, β4 made by culturing a cell transformed with a γ3, β4 nucleic acid; and a method of making a γ3, β4 polypeptide including culturing a a cell transformed with a γ3, β4 nucleic acid.

The inventors have shown that γ3 forms laminin 12 in association with α2 and β1. However, we are unsure of the chain associations of γ3 within other tissues. It is very likely
20 that γ3 can also associate with γ3, α3, α4, and α5; with β2, β3, β4 and β5. Therefore, our results predict 25 new laminins: laminins 12-37. γ3 and β4 polypetides of the invention can be expressed with, assembled with, or administered with other laminin subunits in any of the methods described herein. E.g., γ3 can be assembled with an α and a β subunit to form a laminin trimer. β4 can be assembled with an α and a β subunit to form a laminin trimer.
25 In any treatment or therapeutic application which administers γ3, a β2 subunit can also be administered.

A "heterologous promoter", as used herein is a promoter which is not naturally associated with a gene or a purified nucleic acid.

A "purified"or "substantially pure" or isolated "preparation" of a polypeptide, as used
30 herein, means a polypeptide that has been separated from other proteins, lipids, and nucleic acids with which it naturally occurs. Preferably, the polypeptide is also separated from substances, e.g., antibodies or gel matrix, e.g., polyacrylamide, which are used to purify it. Preferably, the polypeptide constitutes at least 10, 20, 50 70, 80 or 95% dry weight of the purified preparation. Preferably, the preparation contains: sufficient polypeptide to allow
35 protein sequencing; at least 1, 10, or 100 μg of the polypeptide; at least 1, 10, or 100 mg of the polypeptide.

A "purified preparation of cells", as used herein, refers to, in the case of plant or animal cells, an in vitro preparation of cells and not an entire intact plant or animal. In the case of cultured cells or microbial cells, it consists of a preparation of at least 10% and more
40 preferably 50% of the subject cells.

A "treatment", as used herein, includes any therapeutic treatment, e.g., the administration of a therapeutic agent or substance, e.g., a drug.

An "isolated" or " pure nucleic acid", e.g., a substantially pure DNA, is a nucleic acid which is one or both of: not immediately contiguous with either one or both of the sequences, e.g., coding sequences, with which it is immediately contiguous (i.e., one at the 5' end and one at the 3' end) in the naturally-occurring genome of the organism from which the nucleic acid is derived; or which is substantially free of a nucleic acid sequence with which it occurs in the organism from which the nucleic acid is derived. The term includes, for example, a recombinant DNA which is incorporated into a vector, e.g., into an autonomously replicating plasmid or virus, or into the genomic DNA of a prokaryote or eukaryote, or which exists as a separate molecule (e.g., a cDNA or a genomic DNA fragment produced by PCR or restriction endonuclease treatment) independent of other DNA sequences. Substantially pure DNA can also includes a recombinant DNA which is part of a hybrid gene encoding sequence.

"Sequence identity or homology", as used herein, refers to the sequence similarity between two polypeptide molecules or between two nucleic acid molecules. When a position in both of the two compared sequences is occupied by the same base or amino acid monomer subunit, e.g., if a position in each of two DNA molecules is occupied by adenine, then the molecules are homologous or sequence identical at that position. The percent of homology or sequence identity between two sequences is a function of the number of matching or homologous identical positions shared by the two sequences divided by the number of positions compared x 100. For example, if 6 of 10, of the positions in two sequences are the same then the two sequences are 60% homologous or have 60% sequence identity. By way of example, the DNA sequences ATTGCC and TATGGC share 50% homology or sequence identity. Generally, a comparison is made when two sequences are aligned to give maximum homology.

The terms "peptides", "proteins", and "polypeptides" are used interchangeably herein.

As used herein, the term "transgene" means a nucleic acid sequence (encoding, e.g., one or more subject laminin polypeptides), which is partly or entirely heterologous, i.e., foreign, to the transgenic animal or cell into which it is introduced, or, is homologous to an endogenous gene of the transgenic animal or cell into which it is introduced, but which is designed to be inserted, or is inserted, into the animal's genome in such a way as to alter the genome of the cell into which it is inserted (e.g., it is inserted at a location which differs from that of the natural gene or its insertion results in a knockout). A transgene can include one or more transcriptional regulatory sequences and any other nucleic acid, such as introns, that may be necessary for optimal expression of the selected nucleic acid, all operably linked to the selected nucleic acid, and may include an enhancer sequence.

As used herein, the term "transgenic cell" refers to a cell containing a transgene.

As used herein, a "transgenic animal" is any animal in which one or more, and preferably essentially all, of the cells of the animal includes a transgene. The transgene can be introduced into the cell, directly or indirectly by introduction into a precursor of the cell, by way of deliberate genetic manipulation, such as by microinjection or by infection with a

5 recombinant virus. This molecule may be integrated within a chromosome, or it may be extrachromosomally replicating DNA.

As used herein, the term "tissue-specific promoter" means a DNA sequence that serves as a promoter, i.e., regulates expression of a selected DNA sequence operably linked to the promoter, and which effects expression of the selected DNA sequence in
10 specific cells of a tissue, such as mammary tissue. The term also covers so-called "leaky" promoters, which regulate expression of a selected DNA primarily in one tissue, but cause expression in other tissues as well.

"Unrelated to a γ3 or β4 amino acid or nucleic acid sequence" means having less than 30% sequence identity, less than 20% sequence identity, or, preferably, less than 10%
15 homology with a naturally occuring γ3 or β4 sequence disclosed herein.

A polypeptide has γ3 biological activity if it has one or more of the properties of γ3 disclosed herein. A polypeptide has biological activity if it is an antagonist, agonist, or super-agonist of a polypeptide having one of the properties of γ3 disclosed herein.

A polypeptide has β4 biological activity if it has one or more of the properties of β4
20 disclosed herein. A polypeptide has biological activity if it is an antagonist, agonist, or super-agonist of a polypeptide having one of the properties of β4 disclosed herein.

"Misexpression", as used herein, refers to a non-wild type pattern of gene expression, at the RNA or protein level. It includes: expression at non-wild type levels, i.e., over or under expression; a pattern of expression that differs from wild type in terms of the time or stage at
25 which the gene is expressed, e.g., increased or decreased expression (as compared with wild type) at a predetermined developmental period or stage; a pattern of expression that differs from wild type in terms of decreased expression (as compared with wild type) in a predetermined cell type or tissue type; a pattern of expression that differs from wild type in terms of the splicing size, amino acid sequence, post-transitional modification, or biological
30 activity of the expressed polypeptide; a pattern of expression that differs from wild type in terms of the effect of an environmental stimulus or extracellular stimulus on expression of the gene, e.g., a pattern of increased or decreased expression (as compared with wild type) in the presence of an increase or decrease in the strength of the stimulus.

Subject, as used herein, can refer to a mammal, e.g., a human, or to an experimental or
35 animal or disease model. The subject can also be a non-human animal, e.g., a horse, cow, goat, or other domestic animal.

As described herein, one aspect of the invention features a substantially pure (or recombinant) nucleic acid which includes a nucleotide sequence encoding a γ3 or β4 polypeptide and/or equivalents of such nucleic acids. The term nucleic acid as used herein
40 can include fragments and equivalents. The term equivalent refers to nucleotide sequences encoding functionally equivalent polypeptides. Equivalent nucleotide sequences will include sequences that differ by one or more nucleotide substitutions, additions or deletions, such as

5 allelic variants, and include sequences that differ from the nucleotide sequences disclosed herein by degeneracy of the genetic code.

The practice of the present invention will employ, unless otherwise indicated, conventional techniques of cell biology, cell culture, molecular biology, transgenic biology, microbiology, recombinant DNA, and immunology, which are within the skill of the art.
10 Such techniques are described in the literature. See, for example, *Molecular Cloning A Laboratory Manual*, 2nd Ed., ed. by Sambrook, Fritsch and Maniatis (Cold Spring Harbor Laboratory Press: 1989); *DNA Cloning*, Volumes I and II (D. N. Glover ed., 1985); *Oligonucleotide Synthesis* (M. J. Gait ed., 1984); Mullis et al. U.S. Patent No: 4,683,195; *Nucleic Acid Hybridization* (B. D. Hames & S. J. Higgins eds. 1984); *Transcription And*
15 *Translation* (B. D. Hames & S. J. Higgins eds. 1984); *Culture Of Animal Cells* (R. I. Freshney, Alan R. Liss, Inc., 1987); *Immobilized Cells And Enzymes* (IRL Press, 1986); B. Perbal, *A Practical Guide To Molecular Cloning* (1984); the treatise, *Methods In Enzymology* (Academic Press, Inc., N.Y.); *Gene Transfer Vectors For Mammalian Cells* (J. H. Miller and M. P. Calos eds., 1987, Cold Spring Harbor Laboratory); *Methods In Enzymology*, Vols. 154
20 and 155 (Wu et al. eds.), *Immunochemical Methods In Cell And Molecular Biology* (Mayer and Walker, eds., Academic Press, London, 1987); *Handbook Of Experimental Immunology*, Volumes I-IV (D. M. Weir and C. C. Blackwell, eds., 1986); *Manipulating the Mouse Embryo*, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1986).

Other features and advantages of the invention will be apparent from the following
25 detailed description, and from the claims.

DETAILED DESCRIPTION

The drawings are briefly described.

30

Figure 1 depicts the cDNA sequence for human α2 subunit.

Figure 2 depicts the predicted amino acid sequence for human α2 subunit.

35 *Figure 3* depicts the cDNA sequence for human β4 subunit.

Figure 4 depicts the predicted amino acid sequence for human β4 subunit.

Figure 5 depicts an alignment of the amino acid sequence of human β4 of SEQ ID
40 NO: 1 and β4 splice varient of SEQ ID NO:5 and laminin β1, β2, and β3 subunits.

Figure 6 provides a comparision of the similarities of laminin β4 domains with the domains of other known laminin β subunits.

5

Isolation of laminin 12

Laminin 12 was isolated from human placental chorionic villi. Briefly, human chorionic placental villi were frozen in liquid nitrogen, ground in a Waring blender and washed in 1 M NaCl. The final tissue pellet (200g, wet weight) was suspended in 1 L of
10 extraction buffer (50 mM Tris-HCl 50 mM, pH=7.8; NaCl 0.5M, EDTA 10mM, 625 mg/l of N-ethylmaleimide, 150 mg/l of phenylmethylsulphonyl fluoride. The suspension was incubated at 4°C with stirring for 48 h. Unless otherwise noted, all subsequent steps were performed at 4°C. The soluble fraction was collected following centrifugation (30000 x g, 60 min) and precipitated by 300g/l of Ammonium Sulfate. The precipitated proteins were
15 collected by centrifugation (30000 x g, 60 min) and redissolved into chromatography buffer (2M Urea, 25 mM NaCl, 5 mM EDTA, and 50 mM Tris-HCl, pH=7.8). The sample was then dialyzed against the same buffer. Following dialysis, 0.5 volumes of buffer equilibrated DEAE-cellulose (DE-52, Whatman) was added and the mixture shaken overnight. Material not bound to DEAE-cellulose was collected by filtration on a Buchner funnel (Whatman filter
20 4) and precipitated by addition of 300g/l of ammonium sulfate. The proteins were collected by centrifugation (30000 x g, 60 min), redissolved in the Concanavalin-A buffer (0.5 M NaCl, 5 mM $CaCl_2$, 5 mM $MgCl_2$, and Tris-HCl 50 mM, pH=7.8) and dialyzed against the same buffer overnight. The fraction was applied to a 2.5 x 5 cm Concanavalin-A sepharose column (Pharmacia), and unbound material was removed by extensive washing. Bound
25 proteins were first eluted with 10 mM α-D- Mannopyrannoside (Sigma, St. Louis, MO) and secondly with 1 M α-D-Glucopyrannoside (Sigma, St. Louis, MO). A third elution with 1M α-D-Manno-pyrannoside (Sigma, St. Louis, MO) allowed the recovery of the proteins of interest. Each fraction was independently concentrated to 10 ml on a Amicon™ concentrator (30 kDa membrane) and applied to a 2.5 x 100 cm Sephacryl S-500 column in a 0.5 M NaCl,
30 50 mM Tris-HCl, pH=7.8 buffer. The fractions of interest were pooled, dialyzed against Mono-Q buffer (0.1 M NaCl, 25 mM Tris-HCl, pH=7.8) and applied to the 1 x 5 cm Mono-Q column (Pharmacia). Elution was achieved with a 60 ml 0.1-0.5 M NaCl gradient.

The final fraction of interest resulting from the above protocol contains multiple laminins. The laminin 12 was resolved from this mixture by SDS-PAGE (3-5%
35 polyacrylamide) under non-reducing conditions. Six band were resolved. Only the bands at approximately 560 kDa and at the top of the gel were shown to be reactive with polyclonal anti-laminin antiserum (Sigma, St. Louis, MO).

Isolation of α2, β1,γ3 subunits from laminin 12

40 Laminin 12 was excised, equilibrated and reduced in 10% 2-me SDS-PAGE sample buffer, and resolved by 5% SDS-PAGE. Three bands were resolved, which were approximately 205 kDa, 185 kDa, and 170 kDa. The band at 185 kDa reacted with monoclonal antibody 545, specific to the laminin β1 subunit. Each of the three bands were

5 digested with trypsin and the peptides were resolved by HPLC. The selected resolves were subject to peptide sequencing.

Sequencing of the α2, β1 subunits of laminin 12

Protein sequencing was done according to Aebersold et al. (1987). The complex
10 laminin 5-laminin 7 was run on a polyacrylamide gel in the presence of 2-mercaptoethanol and blotted onto a nitrocellulose membrane (Biorad). The 190 kDa band of β2 and the 165 kDa α3 band were separately excised and digested by protease trypsin. The digested product was separated by HPLC and one fragment was sequenced on an Applied Biosystems sequenator (Applied Biosystems, Foster City, CA). The 205 kDa chain
15 contained a sequence identical to human laminin α2, and was thus identified as human laminin α2 subunit. The 185 kDa produced two peptides identical to human β1, and was thus identified as human laminin β1 subunit. The band at 170 kDa contained three sequences not contained in any known laminin chain. A N-terminal sequence of the 170 kDa chain was also determined. In addition, the N-terminal sequence was not identical to
20 any known laminin sequence.

Identification of the γ3 subunit

The cDNA sequences of human γ1 and γ2 were used to probe the National Center for Biomedical Information (NCBI) dBest™ data base by BLAST search and a clone was
25 isolated that was homologous, but not identical to γ1 and γ2. This clone was extended by PCR at the 5' end using Marathon cDNA from human placenta from Clonetech (Palo Alto, CA). The resulting sequence was determined to be 100% identical to all three of the 170 kDa band peptide sequences.

Comparison of the nucleotide sequence of the isolated γ3 subunit to γ1,
30 demonstrated about 80% sequence identity.

Structural Analysis of γ3 encoding DNA

The human cDNA encoding γ3, which is approximately 4710 nucleotides in length, encodes a protein having an estimated molecular weight of approximately 146 kDa (including
35 post-translational modifications) and which is approximately 1570 amino acid residues in length. The human γ3 protein contains a nidogen-binding domain, which can be found, for example, from about amino acids 750-755 of SEQ ID NO:3. The γ3 amino acid sequence and the nucleotide sequence encoding human laminin γ3 is shown in SEQ ID NO:3 and SEQ ID NO:4, respectively.

40 By Northern analysis the size of the γ3 mRNA is approximately 5 kb, which is consistent with other laminin γ subunits. The γ3 mRNA transcript is expressed in human tissues including spleen, testis, brain, placenta, lung, and possibly liver. Chromosomal mapping using the γ3 cDNA sequence indicates that the human γ3 gene is located on

5 chromosome 9q31-34. The location of γ3 on chromosome 9 was confirmed by FISH analysis using a 1.3 kb γ3 cDNA probe within the predicted domains I and II, which are the regions of the least sequence identity among γ subunits. Four human genes associated with Walker-Walburg syndrome, Fukuyama muscular dystrophy, retinitis pigmentosa-deafness syndrome and Eye, Muscle, Brain disease have also been mapped to chromosome 9q31-34.

10

Production of a γ3 specific antibody and tissue localization of γ3

The 170 kDa (γ3) chain was excised from the reducing SDS-PAGE gel described above and injected into a rabbit for antibody production. The resulting serum (rabbit 16) was evaluated by Western analysis and shown to react with the 170 kDa γ3 chain, and showed 15 minor crossreactivity with other laminin chains.

Using immunofluorescence, this antiserum shows localization of γ3 to the following tissue areas: 1) sites of insertions of nerves into the dermal-epidermal junction basement membrane of human skin; 2) the inner nuclear layers, outer nuclear layers, and outer limiting membranes of human, mouse and rat neural retina; 3) the Purkinje cells, and molecular 20 layers, and (perhaps) the glial cells of the mouse and rat cerebellum; 4) the neuromuscular junctions of skeletal muscle; and, 5) the taste buds of the cow tongue.

The γ3 was also shown to colocalize with protein ubiquitin carboxy terminal hydrolase I using antibody pGp 9.5. The γ3 subunit also appears to colocalize with the α2 subunit in the same tissue sections.

25

Isolation and Sequencing of cDNA encoding β4

The initial 350 bp fragment of human laminin β4 cDNA was amplified by touchdown RT-PCR from cultured human keratinocyte total RNA using nested primers made from the published chicken laminin β x 503 bp cDNA sequence (as described in Ybot-30 Gonzalez et al. (1995)). Subsequent cDNA clones were isolated by nested PCR directly from a human placenta cDNA library packaged in lambda-gt11 (Clontech, Palo Alto, CA) or by nested PCR directly from human placenta Marathon-Ready cDNA (Clontech, Palo Alto, CA). The 5' end of the cDNA was cloned using the 5'-RACE technique from human placenta total RNA. The Expanded Long Template PCR System (Boehringer Mannheim Biochemicals, 35 Indianapolis, IN) was used for all PCR reactions. The PCR products were ligated into the pCR2.1 vector (Invitrogen, San Diego, CA) and recombinant plasmids purified for sequencing using the QIAprep™ kit (Qiagen). The DNA sequence was determined using either the Sequenase version 2.0 DNA Sequencing Kit (Amersham) and ^{35}S-dATP or the Thermo Sequenase Radiolabeled Terminator Cycle Sequencing kit (Amersham) and ^{33}P-40 ddNTPs. At least two independent cDNA subclones were sequenced to rule out Taq polymerase-generated nucleotide substitutions. In some cases, PCR product bands were sequenced directly by cycle sequencing after excision from a TAE-EtBr agarose gel and purification using QIAquick Gel Extraction kit (Qiagen).

5

Structural Analysis of DNA encoding β4

The human cDNA encoding a long form β4, which is approximately 5.87 kb, encodes a protein having an estimated molecular weight of approximately 200 kDa and which is approximately 1761 amino acid residues in length. The human β4 protein retains the highest 10 amino acid sequence identity with domains VI and V, which can be found, for example, from about amino acids 221-262 and about 263-535 of SEQ ID NO:1. In addition, a short form, splice variant of β4, which is approximately 3.84 kb and an estimated molecular weight of 120 kDa, was also isolated. The splice variant has 132 nucleotide sequence identical to the long form of β4, with the sequence diverging at nucleotide 3375 and spliced into a unique 3' 15 untranslated region. The short form cDNA encodes a truncated β4 subunit which contains only the short arm of the β4 subunit and is missing the domains necessary for heterodimerization. The β4 amino acid sequence and the nucleotide sequence encoding human laminin β4 is shown in SEQ ID NO:1 and SEQ ID NO:2, respectively.

Northern analysis was performed using total RNA prepared from JAR cell, cultured 20 human keratinocytes and human placenta using either Trizol (Gibco BRL, Bethesda, MD) or RNeasy™ (Qiagen) which was denatured, separated on a formaldehyde agarose gel and blotted onto nitrocellulose according to standard protocols (Sambrook, et al., 1989). In addition, A human multiple tissue northern blot (Clontech, Palo Alto, CA) and Human Northern Territory normal tissue blots and custom fetal skin northern blot (Invitrogen, San 25 Diego, CA) were used. Hybridization and washing were performed using NorthernMAX™ buffer system (Ambion) by manufacturer's recommended protocols. 32P-dCTP-labelled probes were generated from gel-purified restriction fragments using Rediprime™ random primer labeling kit (Amersham). 32P-UTP-labelled antisense RNA probes were generated using the RNA transcription kit (Stratagene, La Jolla, CA) from cDNAs subcloned into 30 Bluescript II KS+ (Stratagene, La Jolla, CA).

Northern blotting showed that human laminin β4 is expressed in JAR cells, derived from undeveloped chronic villi and in placenta. By RT-PCR, it is also expressed in cultured keratinocytes. Using a northern blot of human fetal skin developmental progression, β4 subunit (long form) demonstrates strong expression at week twelve of fetal development and 35 persists until birth, but expression is barely detectable in adult skin. The β4 splice variant, however, is expressed in various tissues including adult heart, brain, lung, liver, skeletal muscle, kidney, spleen, stomach, esophagus, intestine, colon, uterus, bladder, adipose tissue and pancreas. Chromosomal mapping with a β4 cDNA probe indicates that the human β4 subunit is located at locus 7q22-q31.2. The gene encoding β1 is located near, but not on, this 40 position of chromosome 7. Statistical analysis of the mapping data using markers for β1 and β4 suggest that the gene encoding β1 is linked to both ends of the gene encoding β4. In addition, neonatal cutis laxa with manifold phenotype has been mapped near, but not in the same position, as the gene encoding β4.

5 In situ hybridization to wounded human skin grafted into nude mice suggests that laminin β x is expressed in the dermis underneath the migrating epidermal tongues during wound closure.

A GenBank™ search using the human nucleotide sequence encoding β4 as shown in SEQ ID NO:3 revealed an EST, which corresponds to nucleotides 4686-5870 of the human 10 nucleotide sequence encoding β4 depicted in SEQ ID NO:3. Alignment of cDNA encoding β4 with the genes encoding human laminin β1 and laminin β2 shows 61% and 59% sequence identity, respectively, as shown in Figure 5.

Production of a β4 specific antibody and tissue localization of β4

15 Antibodies were raised in rabbits against a 26 kDa bacterial fusion protein which corresponds to the 175 amino acid residues of domain VI (e.g., from about amino acid residues 221-262) of SEQ ID NO:1. Briefly the fusion protein was made by PCR amplification of nucleotides 302-785 of the cDNA encoding β4 using adapter primers and cloned in-frame into the NdeI and SacII sites of pET-15b (Novagen). The fusion protein 20 construct was confirmed by restriction mapping and DNA sequencing. Expression of the fusion protein was induced and separated from *E. coli* proteins using reducing SDS-PAGE. Bands corresponding to the fusion protein were excised from the gel, equilibrated and homogenized using Freud's adjuvant. The same fusion protein was also western blotted on nitrocellulose, dissolved in DMSO and used to immunize mice for monoclonal antibody 25 production.

The polyclonal antisera raised in mice against the fusion protein reacted well with β4, as well as, β1 and β2 polypeptides.

Structural Analysis of the β4 subunit and the β4 splice variant

30 The β4 subunit contains six domains, and α interruption and a signal peptide. The signal peptide and domain VI can be found, for example, at about amino acid residues 1-262 of SEQ ID NO:1. Domain V can be found, for example, at about amino acid residues 263-535 of SEQ ID NO:1. Domains IV and III can be found, for example, at about amino acid residues 536-767 and 768-1178 of SEQ ID NO:1, respectively. Domain I can be found, for 35 example, at about amino acid residues 1409-1761 of SEQ ID NO:1.

The β4 subunit (long form) is most similar in size and domain structure to laminin β1 with an amino acid sequence identity of 42.5%. β4 retains the highest levels of amino acid identity with the other laminin β subunits in domains VI and V, and the lowest levels in domains I and II, as shown in Figure 6. Using the Multicoil™ program, it was determined 40 that only domains I and II of β4 have a high probability of forming coiled coil structures. Domains I and II of β4 look most similar to human β3. Both β4 and β3 are epithelial and the coiled coil structures in domains I and II dictate the α and γ subunits with which the β

5 subunits are associated. Thus, it is likely that β4 associates with α3 and γ2, as does the laminin β3 subunit.

The cDNA encoding the splice variant of β4 contains only the short arm of the β4 subunit, and is missing the EGF repeat of domain III, as shown in Figure 5. Thus, the β4 polypeptide encoded by the β4 c DNA splice variant is missing the coiled coil structures in
10 domains I and II, rendering the short subunit unable to associate into a laminin heterotrimer. PCR amplification of human genomic DNA suggest that the exon which encodes the alternative short form 3' untranslated region is located downstream from the carboxyl-most common exon, exon 23, and is splices out of the β4 subunit, long form, by exon skipping.

15 Analogs of γ3 and β4

Analogs can differ from naturally occurring γ3 or β4 in amino acid sequence or in ways that do not involve sequence, or both. Non-sequence modifications include in vivo or in vitro chemical derivatization of γ3 or β4. Non-sequence modifications include changes in acetylation, methylation, phosphorylation, carboxylation, or glycosylation.

20 Preferred analogs include γ3 or β4 (or biologically active fragments thereof) whose sequences differ from the wild-type sequence by one or more conservative amino acid substitutions or by one or more non-conservative amino acid substitutions, deletions, or insertions which do not abolish the γ3 or β4 biological activity. Conservative substitutions typically include the substitution of one amino acid for another with similar characteristics,
25 e.g., substitutions within the following groups: valine, glycine; glycine, alanine; valine, isoleucine, leucine; aspartic acid, glutamic acid; asparagine, glutamine; serine, threonine; lysine, arginine; and phenylalanine, tyrosine. Other conservative substitutions can be taken from the table below.

5 TABLE 1

CONSERVATIVE AMINO ACID REPLACEMENTS

For Amino Acid	Code	Replace with any of
Alanine	A	D-Ala, Gly, beta-Ala, L-Cys, D-Cys
Arginine	R	D-Arg, Lys, D-Lys, homo-Arg, D-homo-Arg, Met, Ile, D-Met, D-Ile, Orn, D-Orn
Asparagine	N	D-Asn, Asp, D-Asp, Glu, D-Glu, Gln, D-Gln
Aspartic Acid	D	D-Asp, D-Asn, Asn, Glu, D-Glu, Gln, D-Gln
Cysteine	C	D-Cys, S-Me-Cys, Met, D-Met, Thr, D-Thr
Glutamine	Q	D-Gln, Asn, D-Asn, Glu, D-Glu, Asp, D-Asp
Glutamic Acid	E	D-Glu, D-Asp, Asp, Asn, D-Asn, Gln, D-Gln
Glycine	G	Ala, D-Ala, Pro, D-Pro, β-Ala, Acp
Isoleucine	I	D-Ile, Val, D-Val, Leu, D-Leu, Met, D-Met
Leucine	L	D-Leu, Val, D-Val, Leu, D-Leu, Met, D-Met
Lysine	K	D-Lys, Arg, D-Arg, homo-Arg, D-homo-Arg, Met, D-Met, Ile, D-Ile, Orn, D-Orn
Methionine	M	D-Met, S-Me-Cys, Ile, D-Ile, Leu, D-Leu, Val, D-Val
Phenylalanine	F	D-Phe, Tyr, D-Thr, L-Dopa, His, D-His, Trp, D-Trp, Trans-3,4, or 5-phenylproline, cis-3,4, or 5-phenylproline
Proline	P	D-Pro, L-I-thioazolidine-4-carboxylic acid, D-or L-1-oxazolidine-4-carboxylic acid
Serine	S	D-Ser, Thr, D-Thr, allo-Thr, Met, D-Met, Met(O), D-Met(O), L-Cys, D-Cys
Threonine	T	D-Thr, Ser, D-Ser, allo-Thr, Met, D-Met, Met(O), D-Met(O), Val, D-Val
Tyrosine	Y	D-Tyr, Phe, D-Phe, L-Dopa, His, D-His
Valine	V	D-Val, Leu, D-Leu, Ile, D-Ile, Met, D-Met

Other analogs within the invention are those with modifications which increase
10 peptide stability; such analogs may contain, for example, one or more non-peptide bonds (which replace the peptide bonds) in the peptide sequence. Also included are: analogs that include residues other than naturally occurring L-amino acids, e.g., D-amino acids or non-naturally occurring or synthetic amino acids, e.g., β or γ amino acids; and cyclic analogs.

15 Gene Therapy

5 The gene constructs of the invention can also be used as a part of a gene therapy protocol to deliver nucleic acids encoding either an agonistic or antagonistic form of a γ3 or β 4 polypeptide. The invention features expression vectors for *in vivo* transfection and expression of a γ3 or β4 polypeptide in particular cell types so as to reconstitute the function of, or alternatively, antagonize the function of γ3 or β4 polypeptide in a cell in which that
10 polypeptide is misexpressed. Expression constructs of γ3 or β4 polypeptides, may be administered in any biologically effective carrier, e.g. any formulation or composition capable of effectively delivering the γ3 or β4 gene to cells *in vivo*. Approaches include insertion of the subject gene in viral vectors including recombinant retroviruses, adenovirus, adeno-associated virus, and herpes simplex virus-1, or recombinant bacterial or eukaryotic
15 plasmids. Viral vectors transfect cells directly; plasmid DNA can be delivered with the help of, for example, cationic liposomes (lipofectin) or derivatized (e.g. antibody conjugated), polylysine conjugates, gramacidin S, artificial viral envelopes or other such intracellular carriers, as well as direct injection of the gene construct or $CaPO_4$ precipitation carried out *in vivo*.

20 A preferred approach for *in vivo* introduction of nucleic acid into a cell is by use of a viral vector containing nucleic acid, e.g. a cDNA, encoding a γ3 or β4 polypeptide. Infection of cells with a viral vector has the advantage that a large proportion of the targeted cells can receive the nucleic acid. Additionally, molecules encoded within the viral vector, e.g., by a cDNA contained in the viral vector, are expressed efficiently in cells which have taken up
25 viral vector nucleic acid.

Retrovirus vectors and adeno-associated virus vectors can be used as a recombinant gene delivery system for the transfer of exogenous genes *in vivo*, particularly into humans. These vectors provide efficient delivery of genes into cells, and the transferred nucleic acids are stably integrated into the chromosomal DNA of the host. The development of specialized
30 cell lines (termed "packaging cells") which produce only replication-defective retroviruses has increased the utility of retroviruses for gene therapy, and defective retroviruses are characterized for use in gene transfer for gene therapy purposes (for a review see Miller, A.D. (1990) *Blood* 76:271). A replication defective retrovirus can be packaged into virions which can be used to infect a target cell through the use of a helper virus by standard techniques.
35 Protocols for producing recombinant retroviruses and for infecting cells *in vitro* or *in vivo* with such viruses can be found in Current Protocols in Molecular Biology, Ausubel, F.M. et al. (eds.) Greene Publishing Associates, (1989), Sections 9.10-9.14 and other standard laboratory manuals. Examples of suitable retroviruses include pLJ, pZIP, pWE and pEM which are known to those skilled in the art. Examples of suitable packaging virus lines for
40 preparing both ecotropic and amphotropic retroviral systems include ψCrip, ψCre, ψ2 and ψ Am. Retroviruses have been used to introduce a variety of genes into many different cell types, including epithelial cells, *in vitro* and/or *in vivo* (see for example Eglitis, et al. (1985) *Science* 230:1395-1398; Danos and Mulligan (1988) *Proc. Natl. Acad. Sci. USA* 85:6460-

5 6464; Wilson et al. (1988) *Proc. Natl. Acad. Sci. USA* 85:3014-3018; Armentano et al. (1990) *Proc. Natl. Acad. Sci. USA* 87:6141-6145; Huber et al. (1991) *Proc. Natl. Acad. Sci. USA* 88:8039-8043; Ferry et al. (1991) *Proc. Natl. Acad. Sci. USA* 88:8377-8381; Chowdhury et al. (1991) *Science* 254:1802-1805; van Beusechem et al. (1992) *Proc. Natl. Acad. Sci. USA* 89:7640-7644; Kay et al. (1992) *Human Gene Therapy* 3:641-647; Dai et al. (1992) *Proc.*
10 *Natl. Acad. Sci. USA* 89:10892-10895; Hwu et al. (1993) *J. Immunol.* 150:4104-4115; U.S. Patent No. 4,868,116; U.S. Patent No. 4,980,286; PCT Application WO 89/07136; PCT Application WO 89/02468; PCT Application WO 89/05345; and PCT Application WO 92/07573).

Another viral gene delivery system useful in the present invention utilizes adenovirus-
15 derived vectors. The genome of an adenovirus can be manipulated such that it encodes and expresses a gene product of interest but is inactivated in terms of its ability to replicate in a normal lytic viral life cycle. See, for example, Berkner et al. (1988) *BioTechniques* 6:616; Rosenfeld et al. (1991) *Science* 252:431-434; and Rosenfeld et al. (1992) *Cell* 68:143-155. Suitable adenoviral vectors derived from the adenovirus strain Ad type 5 dl324 or other
20 strains of adenovirus (e.g., Ad2, Ad3, Ad7 etc.) are known to those skilled in the art. Recombinant adenoviruses can be advantageous in certain circumstances in that they are not capable of infecting nondividing cells and can be used to infect a wide variety of cell types, including epithelial cells (Rosenfeld et al. (1992) cited *supra*). Furthermore, the virus particle is relatively stable and amenable to purification and concentration, and as above, can be
25 modified so as to affect the spectrum of infectivity. Additionally, introduced adenoviral DNA (and foreign DNA contained therein) is not integrated into the genome of a host cell but remains episomal, thereby avoiding potential problems that can occur as a result of insertional mutagenesis in situations where introduced DNA becomes integrated into the host genome (e.g., retroviral DNA). Moreover, the carrying capacity of the adenoviral genome for foreign
30 DNA is large (up to 8 kilobases) relative to other gene delivery vectors (Berkner et al. cited *supra*; Haj-Ahmand and Graham (1986) *J. Virol.* 57:267).

Yet another viral vector system useful for delivery of the subject gene is the adeno-associated virus (AAV). Adeno-associated virus is a naturally occurring defective virus that requires another virus, such as an adenovirus or a herpes virus, as a helper virus for efficient
35 replication and a productive life cycle. (For a review see Muzyczka et al. *Curr. Topics in Micro. and Immunol.* (1992) 158:97-129). It is also one of the few viruses that may integrate its DNA into non-dividing cells, and exhibits a high frequency of stable integration (see for example Flotte et al. (1992) *Am. J. Respir. Cell. Mol. Biol.* 7:349-356; Samulski et al. (1989) *J. Virol.* 63:3822-3828; and McLaughlin et al. (1989) *J. Virol.* 62:1963-1973). Vectors
40 containing as little as 300 base pairs of AAV can be packaged and can integrate. Space for exogenous DNA is limited to about 4.5 kb. An AAV vector such as that described in Tratschin et al. (1985) *Mol. Cell. Biol.* 5:3251-3260 can be used to introduce DNA into cells. A variety of nucleic acids have been introduced into different cell types using AAV vectors

(see for example Hermonat et al. (1984) *Proc. Natl. Acad. Sci. USA* 81:6466-6470; Tratschin et al. (1985) *Mol. Cell. Biol.* 4:2072-2081; Wondisford et al. (1988) *Mol. Endocrinol.* 2:32-39; Tratschin et al. (1984) *J. Virol.* 51:611-619; and Flotte et al. (1993) *J. Biol. Chem.* 268:3781-3790).

In addition to viral transfer methods, such as those illustrated above, non-viral methods can also be employed to cause expression of a γ3 or β4 polypeptide in the tissue of an animal. Most nonviral methods of gene transfer rely on normal mechanisms used by mammalian cells for the uptake and intracellular transport of macromolecules. In preferred embodiments, non-viral gene delivery systems of the present invention rely on endocytic pathways for the uptake of the subject γ3 or β4 gene by the targeted cell. Exemplary gene delivery systems of this type include liposomal derived systems, poly-lysine conjugates, and artificial viral envelopes.

In a representative embodiment, a gene encoding a γ3 or β4 polypeptide can be entrapped in liposomes bearing positive charges on their surface (e.g., lipofectins) and (optionally) which are tagged with antibodies against cell surface antigens of the target tissue (Mizuno et al. (1992) *No Shinkei Geka* 20:547-551; PCT publication WO91/06309; Japanese patent application 1047381; and European patent publication EP-A-43075).

In clinical settings, the gene delivery systems for the therapeutic γ3 or β4 gene can be introduced into a patient by any of a number of methods, each of which is familiar in the art. For instance, a pharmaceutical preparation of the gene delivery system can be introduced systemically, e.g. by intravenous injection, and specific transduction of the protein in the target cells occurs predominantly from specificity of transfection provided by the gene delivery vehicle, cell-type or tissue-type expression due to the transcriptional regulatory sequences controlling expression of the receptor gene, or a combination thereof. In other embodiments, initial delivery of the recombinant gene is more limited with introduction into the animal being quite localized. For example, the gene delivery vehicle can be introduced by catheter (see U.S. Patent 5,328,470) or by Stereotactic injection (e.g. Chen et al. (1994) *PNAS* 91: 3054-3057).

The pharmaceutical preparation of the gene therapy construct can consist essentially of the gene delivery system in an acceptable diluent, or can comprise a slow release matrix in which the gene delivery vehicle is imbedded. Alternatively, where the complete gene delivery system can be produced in tact from recombinant cells, e.g. retroviral vectors, the pharmaceutical preparation can comprise one or more cells which produce the gene delivery system.

Transgenic Animals

The invention includes transgenic animals which include cells (of that animal) which contain a γ3 or β4 transgene and which preferably (though optionally) express (or misexpress) an endogenous or exogenous γ3 or β4 gene in one or more cells in the animal.

5 The γ3 or β4 transgene can encode the wild-type form of the protein, or can encode homologs thereof, including both agonists and antagonists, as well as antisense constructs. In preferred embodiments, the expression of the transgene is restricted to specific subsets of cells, or tissues utilizing, for example, cis-acting sequences that control expression in the desired pattern. Tissue-specific regulatory sequences and conditional regulatory sequences can be
10 used to control expression of the transgene in certain spatial patterns, e.g., to restrict production to the milk or other secreted product of the animal.

Production of Fragments and Analogs

Generation of Fragments

15 Fragments of a protein can be produced in several ways, e.g., recombinantly, by proteolytic digestion, or by chemical synthesis. Internal or terminal fragments of a polypeptide can be generated by removing one or more nucleotides from one end (for a terminal fragment) or both ends (for an internal fragment) of a nucleic acid which encodes the polypeptide. Expression of the mutagenized DNA produces polypeptide fragments.
20 Digestion with "end-nibbling" endonucleases can thus generate DNA's which encode an array of fragments. DNA's which encode fragments of a protein can also be generated by random shearing, restriction digestion or a combination of the above-discussed methods.

Fragments can also be chemically synthesized using techniques known in the art such as conventional Merrifield solid phase f-Moc or t-Boc chemistry. For example, peptides of
25 the present invention may be arbitrarily divided into fragments of desired length with no overlap of the fragments, or divided into overlapping fragments of a desired length.

Generation of Analogs: Production of Altered DNA and Peptide Sequences by Random Methods

30 Amino acid sequence variants of a protein can be prepared by random mutagenesis of DNA which encodes a protein or a particular domain or region of a protein. Useful methods include PCR mutagenesis and saturation mutagenesis. A library of random amino acid sequence variants can also be generated by the synthesis of a set of degenerate oligonucleotide sequences. (Methods for screening proteins in a library of variants are
35 elsewhere herein.)

PCR Mutagenesis

In PCR mutagenesis, reduced Taq polymerase fidelity is used to introduce random mutations into a cloned fragment of DNA (Leung et al., 1989, *Technique* 1:11-15). This is a
40 very powerful and relatively rapid method of introducing random mutations. The DNA region to be mutagenized is amplified using the polymerase chain reaction (PCR) under conditions that reduce the fidelity of DNA synthesis by Taq DNA polymerase, e.g., by using a dGTP/dATP ratio of five and adding Mn^{2+} to the PCR reaction. The pool of amplified

5 DNA fragments are inserted into appropriate cloning vectors to provide random mutant libraries.

Saturation Mutagenesis

Saturation mutagenesis allows for the rapid introduction of a large number of single 10 base substitutions into cloned DNA fragments (Mayers et al., 1985, *Science* 229:242). This technique includes generation of mutations, e.g., by chemical treatment or irradiation of single-stranded DNA *in vitro,* and synthesis of a complimentary DNA strand. The mutation frequency can be modulated by modulating the severity of the treatment, and essentially all possible base substitutions can be obtained. Because this procedure does not involve a 15 genetic selection for mutant fragments both neutral substitutions, as well as those that alter function, are obtained. The distribution of point mutations is not biased toward conserved sequence elements.

Degenerate Oligonucleotides

20 A library of homologs can also be generated from a set of degenerate oligonucleotide sequences. Chemical synthesis of a degenerate sequences can be carried out in an automatic DNA synthesizer, and the synthetic genes then ligated into an appropriate expression vector. The synthesis of degenerate oligonucleotides is known in the art (see for example, Narang, SA (1983) *Tetrahedron* 39:3; Itakura et al. (1981) *Recombinant DNA, Proc 3rd Cleveland* 25 *Sympos. Macromolecules*, ed. AG Walton, Amsterdam: Elsevier pp273-289; Itakura et al. (1984) *Annu. Rev. Biochem.* 53:323; Itakura et al. (1984) *Science* 198:1056; Ike et al. (1983) *Nucleic Acid Res.* 11:477. Such techniques have been employed in the directed evolution of other proteins (see, for example, Scott et al. (1990) *Science* 249:386-390; Roberts et al. (1992) *PNAS* 89:2429-2433; Devlin et al. (1990) *Science* 249: 404-406; Cwirla et al. (1990) 30 *PNAS* 87: 6378-6382; as well as U.S. Patents Nos. 5,223,409, 5,198,346, and 5,096,815).

Generation of Analogs: Production of Altered DNA and Peptide Sequences by Directed Mutagenesis

Non-random or directed, mutagenesis techniques can be used to provide specific 35 sequences or mutations in specific regions. These techniques can be used to create variants which include, e.g., deletions, insertions, or substitutions, of residues of the known amino acid sequence of a protein. The sites for mutation can be modified individually or in series, e.g., by (1) substituting first with conserved amino acids and then with more radical choices depending upon results achieved, (2) deleting the target residue, or (3) inserting residues of 40 the same or a different class adjacent to the located site, or combinations of options 1-3.

Alanine Scanning Mutagenesis

5 Alanine scanning mutagenesis is a useful method for identification of certain residues or regions of the desired protein that are preferred locations or domains for mutagenesis, Cunningham and Wells (*Science* 244:1081-1085, 1989). In alanine scanning, a residue or group of target residues are identified (e.g., charged residues such as Arg, Asp, His, Lys, and Glu) and replaced by a neutral or negatively charged amino acid (most preferably alanine or
10 polyalanine). Replacement of an amino acid can affect the interaction of the amino acids with the surrounding aqueous environment in or outside the cell. Those domains demonstrating functional sensitivity to the substitutions are then refined by introducing further or other variants at or for the sites of substitution. Thus, while the site for introducing an amino acid sequence variation is predetermined, the nature of the mutation per se need not
15 be predetermined. For example, to optimize the performance of a mutation at a given site, alanine scanning or random mutagenesis may be conducted at the target codon or region and the expressed desired protein subunit variants are screened for the optimal combination of desired activity.

20 Oligonucleotide-Mediated Mutagenesis

Oligonucleotide-mediated mutagenesis is a useful method for preparing substitution, deletion, and insertion variants of DNA, see, e.g., Adelman et al., (*DNA* 2:183, 1983). Briefly, the desired DNA is altered by hybridizing an oligonucleotide encoding a mutation to a DNA template, where the template is the single-stranded form of a plasmid or
25 bacteriophage containing the unaltered or native DNA sequence of the desired protein. After hybridization, a DNA polymerase is used to synthesize an entire second complementary strand of the template that will thus incorporate the oligonucleotide primer, and will code for the selected alteration in the desired protein DNA. Generally, oligonucleotides of at least 25 nucleotides in length are used. An optimal oligonucleotide will have 12 to 15 nucleotides
30 that are completely complementary to the template on either side of the nucleotide(s) coding for the mutation. This ensures that the oligonucleotide will hybridize properly to the single-stranded DNA template molecule. The oligonucleotides are readily synthesized using techniques known in the art such as that described by Crea et al. (*Proc. Natl. Acad. Sci.* USA, 75: 5765[1978]).

35

Cassette Mutagenesis

Another method for preparing variants, cassette mutagenesis, is based on the technique described by Wells et al. (*Gene*, 34:315[1985]). The starting material is a plasmid (or other vector) which includes the protein subunit DNA to be mutated. The codon(s) in the
40 protein subunit DNA to be mutated are identified. There must be a unique restriction endonuclease site on each side of the identified mutation site(s). If no such restriction sites exist, they may be generated using the above-described oligonucleotide-mediated mutagenesis method to introduce them at appropriate locations in the desired protein subunit

5 DNA. After the restriction sites have been introduced into the plasmid, the plasmid is cut at these sites to linearize it. A double-stranded oligonucleotide encoding the sequence of the DNA between the restriction sites but containing the desired mutation(s) is synthesized using standard procedures. The two strands are synthesized separately and then hybridized together using standard techniques. This double-stranded oligonucleotide is referred to as the cassette.
10 This cassette is designed to have 3' and 5' ends that are comparable with the ends of the linearized plasmid, such that it can be directly ligated to the plasmid. This plasmid now contains the mutated desired protein subunit DNA sequence.

Combinatorial Mutagenesis

15 Combinatorial mutagenesis can also be used to generate mutants. E.g., the amino acid sequences for a group of homologs or other related proteins are aligned, preferably to promote the highest homology possible. All of the amino acids which appear at a given position of the aligned sequences can be selected to create a degenerate set of combinatorial sequences. The variegated library of variants is generated by combinatorial mutagenesis at
20 the nucleic acid level, and is encoded by a variegated gene library. For example, a mixture of synthetic oligonucleotides can be enzymatically ligated into gene sequences such that the degenerate set of potential sequences are expressible as individual peptides, or alternatively, as a set of larger fusion proteins containing the set of degenerate sequences.

25 <u>Primary High-Through-Put Methods for Screening Libraries of Peptide Fragments or Homologs</u>

Various techniques are known in the art for screening generated mutant gene products. Techniques for screening large gene libraries often include cloning the gene library into replicable expression vectors, transforming appropriate cells with the resulting library of
30 vectors, and expressing the genes under conditions in which detection of a desired activity, e.g., in this case, binding to other laminin subunits, assembly into a trimeric laminin molecules, binding to natural ligands or substrates, facilitates relatively easy isolation of the vector encoding the gene whose product was detected. Each of the techniques described below is amenable to high through-put analysis for screening large numbers of sequences
35 created, e.g., by random mutagenesis techniques.

Two Hybrid Systems

Two hybrid assays such as the system described above (as with the other screening methods described herein), can be used to identify fragments or analogs. These may include
40 agonists, superagonists, and antagonists. (The subject protein and a protein it interacts with are used as the bait protein and fish proteins.) .

Display Libraries

In one approach to screening assays, the candidate peptides are displayed on the surface of a cell or viral particle, and the ability of particular cells or viral particles to bind an appropriate receptor protein via the displayed product is detected in a "panning assay". For example, the gene library can be cloned into the gene for a surface membrane protein of a bacterial cell, and the resulting fusion protein detected by panning (Ladner et al., WO 88/06630; Fuchs et al. (1991) *Bio/Technology* 9:1370-1371; and Goward et al. (1992) *TIBS* 18:136-140). In a similar fashion, a detectably labeled ligand can be used to score for potentially functional peptide homologs. Fluorescently labeled ligands, e.g., receptors, can be used to detect homolog which retain ligand-binding activity. The use of fluorescently labeled ligands, allows cells to be visually inspected and separated under a fluorescence microscope, or, where the morphology of the cell permits, to be separated by a fluorescence-activated cell sorter.

A gene library can be expressed as a fusion protein on the surface of a viral particle. For instance, in the filamentous phage system, foreign peptide sequences can be expressed on the surface of infectious phage, thereby conferring two significant benefits. First, since these phage can be applied to affinity matrices at concentrations well over 10^{13} phage per milliliter, a large number of phage can be screened at one time. Second, since each infectious phage displays a gene product on its surface, if a particular phage is recovered from an affinity matrix in low yield, the phage can be amplified by another round of infection. The group of almost identical *E. coli* filamentous phages M13, fd., and f1 are most often used in phage display libraries. Either of the phage gIII or gVIII coat proteins can be used to generate fusion proteins without disrupting the ultimate packaging of the viral particle. Foreign epitopes can be expressed at the NH_2-terminal end of pIII and phage bearing such epitopes recovered from a large excess of phage lacking this epitope (Ladner et al. PCT publication WO 90/02909; Garrard et al., PCT publication WO 92/09690; Marks et al. (1992) *J. Biol. Chem.* 267:16007-16010; Griffiths et al. (1993) *EMBO J* 12:725-734; Clackson et al. (1991) *Nature* 352:624-628; and Barbas et al. (1992) *PNAS* 89:4457-4461).

A common approach uses the maltose receptor of *E. coli* (the outer membrane protein, LamB) as a peptide fusion partner (Charbit et al. (1986) *EMBO* 5, 3029-3037). Oligonucleotides have been inserted into plasmids encoding the LamB gene to produce peptides fused into one of the extracellular loops of the protein. These peptides are available for binding to ligands, e.g., to antibodies, and can elicit an immune response when the cells are administered to animals. Other cell surface proteins, e.g., OmpA (Schorr et al. (1991) *Vaccines 91*, pp. 387-392), PhoE (Agterberg, et al. (1990) *Gene* 88, 37-45), and PAL (Fuchs et al. (1991) *Bio/Tech* 9, 1369-1372), as well as large bacterial surface structures have served as vehicles for peptide display. Peptides can be fused to pilin, a protein which polymerizes to form the pilus-a conduit for interbacterial exchange of genetic information (Thiry et al. (1989) *Appl. Environ. Microbiol.* 55, 984-993). Because of its role in interacting with other cells, the pilus provides a useful support for the presentation of peptides to the extracellular

5 environment. Another large surface structure used for peptide display is the bacterial motive organ, the flagellum. Fusion of peptides to the subunit protein flagellin offers a dense array of may peptides copies on the host cells (Kuwajima et al. (1988) *Bio/Tech.* 6, 1080-1083). Surface proteins of other bacterial species have also served as peptide fusion partners. Examples include the *Staphylococcus* protein A and the outer membrane protease IgA of
10 *Neisseria* (Hansson et al. (1992) *J. Bacteriol.* 174, 4239-4245 and Klauser et al. (1990) *EMBO J.* 9, 1991-1999).

In the filamentous phage systems and the LamB system described above, the physical link between the peptide and its encoding DNA occurs by the containment of the DNA within a particle (cell or phage) that carries the peptide on its surface. Capturing the peptide captures
15 the particle and the DNA within. An alternative scheme uses the DNA-binding protein LacI to form a link between peptide and DNA (Cull *et al.* (1992) *PNAS USA* 89:1865-1869). This system uses a plasmid containing the LacI gene with an oligonucleotide cloning site at its 3'-end. Under the controlled induction by arabinose, a LacI-peptide fusion protein is produced. This fusion retains the natural ability of LacI to bind to a short DNA sequence known as
20 LacO operator (LacO). By installing two copies of LacO on the expression plasmid, the LacI-peptide fusion binds tightly to the plasmid that encoded it. Because the plasmids in each cell contain only a single oligonucleotide sequence and each cell expresses only a single peptide sequence, the peptides become specifically and stably associated with the DNA sequence that directed its synthesis. The cells of the library are gently lysed and the peptide-
25 DNA complexes are exposed to a matrix of immobilized receptor to recover the complexes containing active peptides. The associated plasmid DNA is then reintroduced into cells for amplification and DNA sequencing to determine the identity of the peptide ligands. As a demonstration of the practical utility of the method, a large random library of dodecapeptides was made and selected on a monoclonal antibody raised against the opioid peptide dynorphin
30 B. A cohort of peptides was recovered, all related by a consensus sequence corresponding to a six-residue portion of dynorphin B. (Cull et al. (1992) *Proc. Natl. Acad. Sci. U.S.A.* 89-1869)

This scheme, sometimes referred to as peptides-on-plasmids, differs in two important ways from the phage display methods. First, the peptides are attached to the C-terminus of
35 the fusion protein, resulting in the display of the library members as peptides having free carboxy termini. Both of the filamentous phage coat proteins, pIII and pVIII, are anchored to the phage through their C-termini, and the guest peptides are placed into the outward-extending N-terminal domains. In some designs, the phage-displayed peptides are presented right at the amino terminus of the fusion protein. (Cwirla, et al. (1990) *Proc. Natl. Acad. Sci.*
40 *U.S.A.* 87, 6378-6382) A second difference is the set of biological biases affecting the population of peptides actually present in the libraries. The LacI fusion molecules are confined to the cytoplasm of the host cells. The phage coat fusions are exposed briefly to the cytoplasm during translation but are rapidly secreted through the inner membrane into the

5 periplasmic compartment, remaining anchored in the membrane by their C-terminal hydrophobic domains, with the N-termini, containing the peptides, protruding into the periplasm while awaiting assembly into phage particles. The peptides in the LacI and phage libraries may differ significantly as a result of their exposure to different proteolytic activities. The phage coat proteins require transport across the inner membrane and signal peptidase
10 processing as a prelude to incorporation into phage. Certain peptides exert a deleterious effect on these processes and are underrepresented in the libraries (Gallop et al. (1994) *J. Med. Chem.* 37(9):1233-1251). These particular biases are not a factor in the LacI display system.

The number of small peptides available in recombinant random libraries is enormous.
15 Libraries of 10^7-10^9 independent clones are routinely prepared. Libraries as large as 10^{11} recombinants have been created, but this size approaches the practical limit for clone libraries. This limitation in library size occurs at the step of transforming the DNA containing randomized segments into the host bacterial cells. To circumvent this limitation, an *in vitro* system based on the display of nascent peptides in polysome complexes has
20 recently been developed. This display library method has the potential of producing libraries 3-6 orders of magnitude larger than the currently available phage/phagemid or plasmid libraries. Furthermore, the construction of the libraries, expression of the peptides, and screening, is done in an entirely cell-free format.

In one application of this method (Gallop et al. (1994) *J. Med. Chem.* 37(9):1233-
25 1251), a molecular DNA library encoding 10^{12} decapeptides was constructed and the library expressed in an *E. coli* S30 *in vitro* coupled transcription/translation system. Conditions were chosen to stall the ribosomes on the mRNA, causing the accumulation of a substantial proportion of the RNA in polysomes and yielding complexes containing nascent peptides still linked to their encoding RNA. The polysomes are sufficiently robust to be affinity purified
30 on immobilized receptors in much the same way as the more conventional recombinant peptide display libraries are screened. RNA from the bound complexes is recovered, converted to cDNA, and amplified by PCR to produce a template for the next round of synthesis and screening. The polysome display method can be coupled to the phage display system. Following several rounds of screening, cDNA from the enriched pool of polysomes
35 was cloned into a phagemid vector. This vector serves as both a peptide expression vector, displaying peptides fused to the coat proteins, and as a DNA sequencing vector for peptide identification. By expressing the polysome-derived peptides on phage, one can either continue the affinity selection procedure in this format or assay the peptides on individual clones for binding activity in a phage ELISA, or for binding specificity in a completion phage
40 ELISA (Barret, et al. (1992) *Anal. Biochem* 204,357-364). To identify the sequences of the active peptides one sequences the DNA produced by the phagemid host.

Secondary Screens

5 The high through-put assays described above can be followed by secondary screens in order to identify further biological activities which will, e.g., allow one skilled in the art to differentiate agonists from antagonists. The type of a secondary screen used will depend on the desired activity that needs to be tested. For example, an assay can be developed in which the ability to inhibit an interaction between a protein of interest and its respective ligand can 10 be used to identify antagonists from a group of peptide fragments isolated though one of the primary screens described above.

Therefore, methods for generating fragments and analogs and testing them for activity are known in the art. Once the core sequence of interest is identified, it is routine to perform for one skilled in the art to obtain analogs and fragments.

15

Peptide Mimetics

The invention also provides for reduction of the protein binding domains of the subject γ3 or β4 polypeptides to generate mimetics, e.g. peptide or non-peptide agents. See, for example, "Peptide inhibitors of human papillomavirus protein binding to retinoblastoma 20 gene protein" European patent applications EP-412,762A and EP-B31,080A.

Non-hydrolyzable peptide analogs of critical residues can be generated using benzodiazepine (e.g., see Freidinger et al. in *Peptides: Chemistry and Biology*, G.R. Marshall ed., ESCOM Publisher: Leiden, Netherlands, 1988), azepine (e.g., see Huffman et al. in 25 *Peptides: Chemistry and Biology*, G.R. Marshall ed., ESCOM Publisher: Leiden, Netherlands, 1988), substituted gama lactam rings (Garvey et al. in *Peptides: Chemistry and Biology*, G.R. Marshall ed., ESCOM Publisher: Leiden, Netherlands, 1988), keto-methylene pseudopeptides (Ewenson et al. (1986) *J Med Chem* 29:295; and Ewenson et al. in *Peptides: Structure and Function* (Proceedings of the 9th American Peptide Symposium) Pierce 30 Chemical Co. Rockland, IL, 1985), β-turn dipeptide cores (Nagai et al. (1985) *Tetrahedron Lett* 26:647; and Sato et al. (1986) *J Chem Soc Perkin Trans* 1:1231), and β-aminoalcohols (Gordon et al. (1985) *Biochem Biophys Res Commun*126:419; and Dann et al. (1986) *Biochem Biophys Res Commun* 134:71).

35 Antibodies

The invention also includes antibodies specifically reactive with a subject γ3 or β4 polypeptides. Anti-protein/anti-peptide antisera or monoclonal antibodies can be made by standard protocols (See, for example, *Antibodies: A Laboratory Manual* ed. by Harlow and Lane (Cold Spring Harbor Press: 1988)).

40 Antibodies which specifically bind γ3 or β4 epitopes can also be used in immunohistochemical staining of tissue samples in order to evaluate the abundance and pattern of expression of γ3 or β4. Anti γ3 or β4 antibodies can be used diagnostically in

5 immuno-precipitation and immuno-blotting to detect and evaluate γ3 or β4 levels in tissue or bodily fluid as part of a clinical testing procedure.

Another application of antibodies of the present invention is in the immunological screening of cDNA libraries constructed in expression vectors such as λgt11, λgt18-23, λZAP, and λORF8. Messenger libraries of this type, having coding sequences inserted in the
10 correct reading frame and orientation, can produce fusion proteins. For instance, λgt11 will produce fusion proteins whose amino termini consist of ß-galactosidase amino acid sequences and whose carboxy termini consist of a foreign polypeptide. Antigenic epitopes of a subject polypeptide can then be detected with antibodies, as, for example, reacting nitrocellulose filters lifted from infected plates with antibodies of the invention. Phage, scored by this
15 assay, can then be isolated from the infected plate. Thus, the presence of homologs can be detected and cloned from other animals, and alternate isoforms (including splicing variants) can be detected and cloned from human sources.

20 *Other Embodiments*

Included in the invention are: allelic variations; natural mutants; induced mutants; proteins encoded by DNA that hybridizes under high or low stringency conditions to a nucleic acid which encodes a polypeptide of SEQ ID NO:1 or SEQ ID NO:3 (for definitions
25 of high and low stringency see Current Protocols in Molecular Biology, John Wiley & Sons, New York, 1989, 6.3.1 - 6.3.6, hereby incorporated by reference); and, polypeptides specifically bound by antisera to γ3 or β4.

Nucleic acids and polypeptides of the invention includes those that differ from the sequences discolosed herein by virtue of sequencing errors in the disclosed sequences.
30 The invention also includes fragments, preferably biologically active fragments, or analogs of γ3 or β4. A biologically active fragment or analog is one having any in vivo or in vitro activity which is characteristic of theγ3 or β4 shown in SEQ ID NO:3 and SEQ ID NO:1, respectively, or of other naturally occurring γ3 or β4, e.g., one or more of the biological activities described above. Especially preferred are fragments which exist in vivo,
35 e.g., fragments which arise from post transcriptional processing or which arise from translation of alternatively spliced RNA's. Fragments include those expressed in native or endogenous cells, e.g., as a result of post-translational processing, e.g., as the result of the removal of an amino-terminal signal sequence, as well as those made in expression systems, e.g., in CHO cells. Particularly preferred fragments are fragments, e.g., active fragments,
40 which are generated by proteolytic cleavage or alternative splicing events.

5 Other embodiments are within the following claims.

What is claimed is:

5

1. An isolated laminin 12 which includes an $\alpha 2$ subunit, a $\beta 1$ subunit and a $\gamma 3$ subunit.

2. An isolated $\gamma 3$ subunit.

10

3. An isolated $\beta 4$ subunit.

cDNA sequence encoding laminin α2 subunit

1 cagcgactcc tctggctccc gagaagtgga tccggtcgcg gccactacga tgccgggagc
61 cgccggggtc ctcctccttc tgctgctctc cggaggcctc gggggcgtac aggcgcagcg
121 gccgcagcag cagcggcagt cacaggcaca tcagcaaaga ggtttattcc ctgctgtcct
181 gaatcttgct tctaatgctc ttatcacgac caatgcaaca tgtggagaaa aaggacctga
241 aatgtactgc aaattggtag aacatgtccc tgggcagcct gtgaggaacc cgcagtgtcg
301 aatctgcaat caaaacagca gcaatccaaa ccagagacac ccgattacaa atgctattga
361 tggaaagaac acttggtggc agagtcccag tattaagaat ggaatcgaat accattatgt
421 gacaattaca ctggatttac agcaggtgtt ccagatcgcg tatgtgattg tgaaggcagc
481 taactccccc cggcctggaa actggatttt ggaacgctct cttgatgatg ttgaatacaa
541 gccctggcag tatcatgctg tgacagacac ggagtgccta acgctttaca atatttatcc
601 ccgcactggg ccaccgtcat atgccaaaga tgatgaggtc atctgcactt cattttactc
661 caagatacac cccttagaaa atggagagat tcacatctct ttaatcaatg ggagaccaag
721 tgccgatgat ccttctccag aactgctaga atttacctcc gctcgctata ttcgcctgag
781 atttcagagg atccgcacac tgaatgctga cttgatgatg tttgctcaca aagacccaag
841 agaaattgac cccattgtca ccagaagata ttactactcg gtcaaggata tttcagttgg
901 agggatgtgc atctgctatg gtcatgccag ggcttgtcca cttgatccag cgacaaataa
961 atctcgctgt gagtgtgagc ataacacatg tggcgatagc tgtgatcagt gctgtccagg
1021 attccatcag aaaccctgga gagctggaac ttttctaact aaaactgaat gtgaagcatg
1081 caattgtcat ggaaaagctg aagaatgcta ttatgatgaa aatgttgcca gaagaaatct
1141 gagtttgaat atacgtggaa agtacattgg agggggtgtc tgcattaatt gtacccaaaa
1201 cactgctggt ataaactgcg agacatgtac agatggcttc ttcagaccca aagggggtatc
1261 tccaaattat ccaaggccat gccagccatg tcattgcgat ccaattggtt ccttaaatga
1321 agtctgtgtc aaggatgaga aacatgctcg acgaggtttg gcacctggat cctgtcattg
1381 caaaactggt tttggaggtg tgagctgtga tcggtgtgcc aggggctaca ctggctaccc
1441 ggactgcaaa gcctgtaact gcagtgggtt agggagcaaa aatgaggatc cttgttttgg
1501 cccctgtatc tgcaaggaaa atgttgaagg aggagactgt agtcgttgca aatccggctt
1561 cttcaatttg caagaggata attggaaagg ctgcgatgag tgtttctgtt caggggtttc
1621 aaacagatgt cagagttcct actggaccta tggcaaaata caagatatga gtggctggta
1681 tctgactgac cttcctggcc gcattcgagt ggctccccag caggacgact tggactcacc
1741 tcagcagatc agcatcagta acgcggaggc ccggcaagcc ctgccgcaca gctactactg
1801 gagcgcgccg gctccctatc tgggaaacaa actcccagca gtaggaggac agttgacatt
1861 taccatatca tatgaccttg aagaagagga agaagataca gaacgtgttc tccagcttat
1921 gattatctta gagggtaatg acttgagcat cagcacagcc caagatgagg tgtacctgca
1981 cccatctgaa gaacatacta atgtattgtt acttaaagaa gaatcattta ccatacatgg
2041 cacacatttt ccagtccgta gaaaggaatt tatgacagtg cttgcgaatt tgaagagagt
2101 cctcctacaa atcacataca gctttgggat ggatgccatc ttcaggttga gctctgttaa
2161 ccttgaatcc gctgtctcct atcctactga tggaagcatt gcagcagctg tagaagtgtg
2221 tcagtgccca ccagggtata ctggctcctc ttgtgaatct tgttggccta ggcacaggcg
2281 agttaacggc actatttttg gtggcatctg tgagccatgt cagtgctttg gtcatgcgga
2341 gtcctgtgat gacgtcactg gagaatgcct gaactgtaag gatcacacag gtggcccata
2401 ttgtgataaa tgtcttcctg gtttctatgg cgagcctact aaaggaacct ctgaagactg
2461 tcaaccctgt gcctgtccac tcaatatccc atccaataac tttagcccaa cgtgccattt
2521 agaccggagt cttggattga tctgtgatgg atgccctgtc gggtacacag gaccacgctg
2581 tgagaggtgt gcagaaggct attttggaca accctctgta cctggaggat catgtcagcc
2641 atgccaatgc aatgacaacc ttgacttctc catccctggc agctgtgaca gcttgtctgg
2701 ctcctgtctg atatgtaaac caggtacaac aggccggtac tgtgagctct gtgctgatgg

FIGURE 1

2761 atattttgga gatgcagttg atgcgaagaa ctgtcagccc tgtcgctgta atgccggtgg
2821 ctctttctct gaggtttgcc acagtcaaac tggacagtgt gagtgcagag ccaacgttca
2881 gggtcagaga tgtgacaaat gcaaggctgg gacctttggc ctacaatcag caaggggctg
2941 tgttccctgc aactgcaatt cttttgggtc taagtcattc gactgtgaag agagtggaca
3001 atgttggtgc caacctggag tcacagggaa gaaatgtgac cgctgtgccc acggctattt
3061 caacttccaa gaaggaggct gcacagcttg tgaatgttct catctgggta ataattgtga
3121 cccaaagact gggcgatgca tttgcccacc caataccatt ggagagaaat gttctaaatg
3181 tgcacccaat acctggggcc acagcattac cactggttgt aaggcttgta actgcagcac
3241 agtgggatcc ttggatttcc aatgcaatgt aaatacaggc caatgcaact gtcatccaaa
3301 attctctggt gcaaaatgta cagagtgcag tcgaggtcac tggaactacc ctcgctgcaa
3361 tctctgtgac tgcttcctcc ctgggacaga tgccacaacc tgtgattcag agactaaaaa
3421 atgctcctgt agtgatcaaa ctgggcagtg cacttgtaag gtgaatgtgg aaggcatcca
3481 ctgtgacaga tgccggcctg gcaaattcgg actcgatgcc aagaatccac ttggctgcag
3541 cagctgctat tgcttcggca ctactaccca gtgctctgaa gcaaaaggac tgatccggac
3601 gtgggtgact ctgaaggctg agcagaccat tctacccctg gtagatgagg ctctgcagca
3661 cacgaccacc aagggcattg tttttcaaca tccagagatt gttgcccaca tggacctgat
3721 gagagaagat ctccatttgg aacctttttа ttggaaactt ccagaacaat ttgaaggaaa
3781 gaagttgatg gcctatgggg gcaaactcaa gtatgcaatc tatttcgagg ctcgggaaga
3841 aacaggtttc tctacatata atcctcaagt gatcattcga ggtgggacac ctactcatgc
3901 tagaattatc gtcaggcata tggctgctcc tctgattggc caattgacaa ggcatgaaat
3961 tgaaatgaca gagaaagaat ggaaatatta tggggatgat cctcgagtcc atagaactgt
4021 gacccgagaa gacttcttgg atatactata tgatattcat tacattctta tcaaagctac
4081 ttatggaaat ttcatgcgac aaagcaggat ttctgaaatc tcaatggagg tagctgaaca
4141 aggacgtgga acaacaatga ctcctccagc tgacttgatt gaaaaatgtg attgtcccct
4201 gggctattct ggcctgtcct gtgaggcatg cttgccggga ttttatcgac tgcgttctca
4261 accaggtggc cgcacccctg gaccaaccct gggcacctgt gttccatgtc aatgtaatgg
4321 acacagcagc ctgtgtgacc ctgaaacatc gatatgccag aattgtcaac atcacactgc
4381 tggtgacttc tgtgaacgat gtgctcttgg atactatgga attgtcaagg gattgccaaa
4441 tgactgtcag caatgtgcct gccctctgat ttcttccagt aacaatttca gcccctcttg
4501 tgtcgcagaa ggacttgacg actaccgctg cacggcttgt ccacggggat atgaaggcca
4561 gtactgtgaa aggtgtgccc ctggctatac tggcagtcca ggcaaccctg gaggctcctg
4621 ccaagaatgt gagtgtgatc cctatggctc actgcctgtg ccctgtgacc ctgtcacagg
4681 attctgcacg tgccgacctg gagccacggg aaggaagtgt gacggctgca agcactggca
4741 tgcacgcgag ggctgggagt gtgttttttg tggagatgag tgcactggcc ttcttctcgg
4801 tgacttggct cgcctggagc agatggtcat gagcatcaac ctcactggtc cgctgcctgc
4861 gccatataaa atgctgtatg gtcttgaaaa tatgactcag gagctaaagc acttgctgtc
4921 acctcagcgg gccccagaga ggcttattca gctggcagag ggcaatctga atacactcgt
4981 gaccgaaatg aacgagctgc tgaccagggc taccaaagtg acagcagatg gcgagcagac
5041 cggacaggat gctgagagga ccaacacaag agcaaagtcc ctgggagaat tcattaagga
5101 gcttgcccgg gatgcagaag ctgtaaatga aaaagctata aaactaaatg aaactctagg
5161 aactcgagac gaggcctttg agagaaattt ggaagggctt cagaaagaga ttgaccagat
5221 gattaaagaa ctgaggagga aaaatctaga gacacaaaag gaaattgctg aagatgagtt
5281 ggtagctgca gaagcccttc tgaaaaaagt gaagaagctg tttggagagt cccgggggga
5341 aaatgaagaa atggagaagg atctccggga aaaactggct gactacaaaa acaaagttga
5401 tgatgcttgg gacctttga gagaagccac agataaaatc agagaagcta atcgcctatt
5461 tgcagtaaat cagaaaaaca tgactgcatt ggagaaaaag aaggaggctg ttgagagcgg
5521 caaacgacaa attgagaaca ctttaaaaga aggcaatgac atactcgatg aagccaaccg
5581 tcttgcagat gaaatcaact ccatcataga ctatgttgaa gacatccaaa ctaaattgcc

FIGURE 1 (continued)

5641 acctatgtct gaggagctta atgataaaat agatgacctc tcccaagaaa taaaggacag
5701 gaagcttgct gagaaggtgt cccaggctga gagccacgca gctcagttga atgactcatc
5761 tgctgtcctt gatggaatcc ttgatgaggc taaaaacatc tccttcaatg ccactgcagc
5821 cttcaaagct tacagcaata ttaaggacta tattgatgaa gctgagaaag ttgccaaaga
5881 agccaaagat cttgcacatg aagctacaaa actggcaaca ggtcctcggg gtttattaaa
5941 ggaagatgcc aaaggctgtc ttcagaaaag cttcaggatt cttaacgaag ccaagaagtt
6001 agcaaatgat gtaaaagaaa atgaagacca tctaaatggc ttaaaaacca ggatagaaaa
6061 tgctgatgct agaaatgggg atctcttgag aactttgaat gacactttgg gaaagttatc
6121 agctattcca aatgatacag ctgctaaact gcaagctgtt aaggacaaag ccagacaagc
6181 caacgacaca gctaaagatg tactggcaca gattacagag ctccaccaga acctcgatgg
6241 cctgaagaag aattacaata aactagcaga cagcgtcgcc aaaacgaatg ctgtggttaa
6301 agatccttcc aagaacaaaa tcattgccga tgcagatgcc actgtcaaaa atttagaaca
6361 ggaagctgac cggctaatag ataaactcaa acccatcaag gaacttgagg ataacctaaa
6421 gaaaaacatc tctgagataa aggaattgat aaaccaagct cggaaacaag ccaattctat
6481 caaagtatct gtgtcttcag gaggtgactg cattcgaaca tacaaaccag aaatcaagaa
6541 aggaagttac aataatattg ttgtcaacgt aaagacagct gttgctgata acctcctctt
6601 ttatcttgga agtgccaaat ttattgactt tctggctata gaaatgcgta aaggcaaagt
6661 cagcttcctc tgggatgttg gatctggagt tggacgtgta gagtacccag atttgactat
6721 tgatgactca tattggtacc gtatcgtagc atcaagaact gggagaaatg gaactatttc
6781 tgtgagagcc ctggatggac ccaaagccag cattgtgccc agcacacacc attcgacgtc
6841 tcctccaggg tacacgattc tagatgtgga tgcaaatgca atgctgtttg ttggtggcct
6901 gactgggaaa ttaaagaagg ctgatgctgt acgtgtgatt acattcactg gctgcatggg
6961 agaaacatac tttgacaaca aacctatagg tttgtggaat ttccgagaaa aagaaggtga
7021 ctgcaaagga tgcactgtca gtcctcaggt ggaagatagt gaggggacta ttcaatttga
7081 tggagaaggt tatgcattgg tcagccgtcc cattcgctgg taccccaaca tctccactgt
7141 catgttcaag ttcagaacat ttcttcgag tgctcttctg atgtatcttg ccacacgaga
7201 cctgagagat ttcatgagtg tggagctcac tgatgggcac ataaaagtca gttacgatct
7261 gggctcagga atggcttccg ttgtcagcaa tcaaaaccat aatgatggga aatggaaatc
7321 attcactctg tcaagaattc aaaaacaagc caatatatca attgtagata tagatactaa
7381 tcaggaggag aatatagcaa cttcgtcttc tggaaacaac tttggtcttg acttgaaagc
7441 agatgacaaa atatattttg gtggcctgcc aacgctgaga aacttgagta tgaaagcaag
7501 gccagaagta aatctgaaga aatattccgg ctgcctcaaa gatattgaaa tttcaagaac
7561 tccgtacaat atactcagta gtcccgatta tgttggtgtt accaaaggat gttccctgga
7621 gaatgtttac acagttagct ttcctaagcc tggttttgtg gagctctccc ctgtgccaat
7681 tgatgtagga acagaaatca acctgtcatt cagcaccaag aatgagtccg gcatcattct
7741 tttgggaagt ggagggacac cagcaccacc taggagaaaa cgaaggcaga ctggacaggc
7801 ctattatgta atactcctca acaggggccg tctggaagtg catctctcca caggggcacg
7861 aacaatgagg aaaattgtca tcagaccaga gccgaatctg tttcatgatg gaagagaaca
7921 ttccgttcat gtagagcgaa ctagaggcat ctttacagtt caagtggatg aaaacagaag
7981 atacatgcaa aacctgacag ttgaacagcc tatcgaagtt aaaaagcttt tcgttggggg
8041 tgctccacct gaatttcaac cttccccact cagaaatatt cctccttttg aaggctgcat
8101 atggaatctt gttattaact ctgtccccat ggactttgca aggcctgtgt ccttcaaaaa
8161 tgctgacatt ggtcgctgtg cccatcagaa actccgtgaa gatgaagatg gagcagctcc
8221 agctgaaata gttatccagc ctgagccagt tcccaccccca gcctttccta cgcccacccc
8281 agttctgaca catggtcctt gtgctgcaga atcagaacca gctcttttga tagggagcaa
8341 gcagttcggg ctttcaagaa acagtcacat tgcaattgca tttgatgaca ccaaagttaa
8401 aaaccgtctc acaattgagt tggaagtaag aaccgaagct gaatccggct tgctttttta
8461 catggctgcg atcaatcatg ctgattttgc aacagttcag ctgagaaatg gattgcccta

FIGURE 1 (continued)

8521 cttcagctat gacttgggga gtggggacac ccacaccatg atccccacca aaatcaatga
8581 tggccagtgg cacaagatta agataatgag aagtaagcaa gaaggaattc tttatgtaga
8641 tggggcttcc aacagaacca tcagtcccaa aaaagccgac atcctggatg tcgtgggaat
8701 gctgtatgtt ggtgggttac ccatcaacta cactacccga agaattggtc cagtgaccta
8761 tagcattgat ggctgcgtca ggaatctcca catggcagag gcccctgccg atctggaaca
8821 acccacctcc agcttccatg ttgggacatg ttttgcaaat gctcagaggg gaacatattt
8881 tgacggaacc ggttttgcca aagcagttgg tggattcaaa gtgggattgg accttcttgt
8941 agaatttgaa ttcgcgacaa ctacaacgac tggagttctt ctggggatca gtagtcaaaa
9001 aatggatgga atgggtattg aaatgattga tgaaaagttg atgtttcatg tggacaatgg
9061 tgcgggcaga ttcactgctg tctatgatgc tggggttcca gggcatttgt gtgatggaca
9121 atggcataaa gtcactgcca acaagatcaa acaccgcatt gagctcacag tcgatgggaa
9181 ccaggtggaa gcccaaagcc caaacccagc atctacatca gctgacacaa atgaccctgt
9241 gtttgttgga ggcttcccag atgacctcaa gcagtttggc ctaacaacca gtattccgtt
9301 ccgaggttgc atcagatccc tgaagctcac caaaggcaca gcaagccact ggaggttaat
9361 tttgccaagg ccctggaact gaggggcgtt caacctgtat catgcccagc caactaataa
9421 aaataagtgt aaccccagga agagtctgtc aaaacaagta tatcaagtaa aacaaacaaa
9481 tatattttac ctatatatgt taattaaact aatttgtgca tgtacataga attc

FIGURE 1 (continued)

amino acid sequence of laminin α2 subunit

MPGAAGVLLLLLLLSGGLGGVQAQRPQQQRQSQAHQQRGLFPAVL

NLASNALITTNATCGEKGPEMYCKLVEHVPGQPVRNPQCRICNQNSSNPNQRHPITNA

IDGKNTWWQSPSIKNGIEYHYVTITLDLQQVFQIAYVIVKAANSPRPGNWILERSLDD

VEYKPWQYHAVTDTECLTLYNIYPRTGPPSYAKDDEVICTSFYSKIHPLENGEIHISL

INGRPSADDPSPELLEFTSARYIRLRFQRIRTLNADLMMFAHKDPREIDPIVTRRYYY

SVKDISVGGMCICYGHARACPLDPATNKSRCECEHNTCGDSCDQCCPGFHQKPWRAGT

FLTKTECEACNCHGKAEECYYDENVARRNLSLNIRGKYIGGGVCINCTQNTAGINCET

CTDGFFRPKGVSPNYPRPCQPCHCDPIGSLNEVCVKDEKHARRGLAPGSCHCKTGFGG

VSCDRCARGYTGYPDCKACNCSGLGSKNEDPCFGPCICKENVEGGDCSRCKSGFFNLQ

EDNWKGCDECFCSGVSNRCQSSYWTYGKIQDMSGWYLTDLPGRIRVAPQQDDLDSPQQ

ISISNAEARQALPHSYYWSAPAPYLGNKLPAVGGQLTFTISYDLEEEEEDTERVLQLM

IILEGNDLSISTAQDEVYLHPSEEHTNVLLLKEESFTIHGTHFPVRRKEFMTVLANLK

RVLLQITYSFGMDAIFRLSSVNLESAVSYPTDGSIAAAVEVCQCPPGYTGSSCESCWP

RHRRVNGTIFGGICEPCQCFGHAESCDDVTGECLNCKDHTGGPYCDKCLPGFYGEPTK

GTSEDCQPCACPLNIPSNNFSPTCHLDRSLGLICDGCPVGYTGPRCERCAEGYFGQPS

VPGGSCQPCQCNDNLDFSIPGSCDSLSGSCLICKPGTTGRYCELCADGYFGDAVDAKN

CQPCRCNAGGSFSEVCHSQTGQCECRANVQGQRCDKCKAGTFGLQSARGCVPCNCNSF

GSKSFDCEESGQCWCQPGVTGKKCDRCAHGYFNFQEGGCTACECSHLGNNCDPKTGRC

ICPPNTIGEKCSKCAPNTWGHSITTGCKACNCSTVGSLDFQCNVNTGQCNCHPKFSGA

KCTECSRGHWNYPRCNLCDCFLPGTDATTCDSETKKCSCSDQTGQCTCKVNVEGIHCD

RCRPGKFGLDAKNPLGCSSCYCFGTTTQCSEAKGLIRTWVTLKAEQTILPLVDEALQH

TTTKGIVFQHPEIVAHMDLMREDLHLEPFYWKLPEQFEGKKLMAYGGKLKYAIYFEAR

EETGFSTYNPQVIIRGGTPTHARIIVRHMAAPLIGQLTRHEIEMTEKEWKYYGDDPRV

FIGURE 2

HRTVTREDFLDILYDIHYILIKATYGNFMRQSRISEISMEVAEQGRGTTMTPPADLIE

KCDCPLGYSGLSCEACLPGFYRLRSQPGGRTPGPTLGTCVPCQCNGHSSLCDPETSIC

QNCQHHTAGDFCERCALGYYGIVKGLPNDCQQCACPLISSSNNFSPSCVAEGLDDYRC

TACPRGYEGQYCERCAPGYTGSPGNPGGSCQECECDPYGSLPVPCDPVTGFCTCRPGA

TGRKCDGCKHWHAREGWECVFCGDECTGLLLGDLARLEQMVMSINLTGPLPAPYKMLY

GLENMTQELKHLLSPQRAPERLIQLAEGNLNTLVTEMNELLTRATKVTADGEQTGQDA

ERTNTRAKSLGEFIKELARDAEAVNEKAIKLNETLGTRDEAFERNLEGLQKEIDQMIK

ELRRKNLETQKEIAEDELVAAEALLKKVKKLFGESRGENEEMEKDLREKLADYKNKVD

DAWDLLREATDKIREANRLFAVNQKNMTALEKKKEAVESGKRQIENTLKEGNDILDEA

NRLADEINSIIDYVEDIQTKLPPMSEELNDKIDDLSQEIKDRKLAEKVSQAESHAAQL

NDSSAVLDGILDEAKNISFNATAAFKAYSNIKDYIDEAEKVAKEAKDLAHEATKLATG

PRGLLKEDAKGCLQKSFRILNEAKKLANDVKENEDHLNGLKTRIENADARNGDLLRTL

NDTLGKLSAIPNDTAAKLQAVKDKARQANDTAKDVLAQITELHQNLDGLKKNYNKLAD

SVAKTNAVVKDPSKNKIIADADATVKNLEQEADRLIDKLKPIKELEDNLKKNISEIKE

LINQARKQANSIKVSVSSGGDCIRTYKPEIKKGSYNNIVVNVKTAVADNLLFYLGSAK

FIDFLAIEMRKGKVSFLWDVGSGVGRVEYPDLTIDDSYWYRIVASRTGRNGTISVRAL

DGPKASIVPSTHHSTSPPGYTILDVDANAMLFVGGLTGKLKKADAVRVITFTGCMGET

YFDNKPIGLWNFREKEGDCKGCTVSPQVEDSEGTIQFDGEGYALVSRPIRWYPNISTV

MFKFRTFSSSALLMYLATRDLRDFMSVELTDGHIKVSYDLGSGMASVVSNQNHNDGKW

KSFTLSRIQKQANISIVDIDTNQEENIATSSSGNNFGLDLKADDKIYFGGLPTLRNLS

MKARPEVNLKKYSGCLKDIEISRTPYNILSSPDYVGVTKGCSLENVYTVSFPKPGFVE

LSPVPIDVGTEINLSFSTKNESGIILLGSGGTPAPPRRKRRQTGQAYYVILLNRGRLE

VHLSTGARTMRKIVIRPEPNLFHDGREHSVHVERTRGIFTVQVDENRRYMQNLTVEQP

IEVKKLFVGGAPPEFQPSPLRNIPPFEGCIWNLVINSVPMDFARPVSFKNADIGRCAH

FIGURE 2 (continued)

QKLREDEDGAAPAEIVIQPEPVPTPAFPTPTPVLTHGPCAAESEPALLIGSKQFGLSR

NSHIAIAFDDTKVKNRLTIELEVRTEAESGLLFYMAAINHADFATVQLRNGLPYFSYD

LGSGDTHTMIPTKINDGQWHKIKIMRSKQEGILYVDGASNRTISPKKADILDVVGMLY

VGGLPINYTTRRIGPVTYSIDGCVRNLHMAEAPADLEQPTSSFHVGTCFANAQRGTYF

DGTGFAKAVGGFKVGLDLLVEFEFATTTTTGVLLGISSQKMDGMGIEMIDEKLMFHVD

NGAGRFTAVYDAGVPGHLCDGQWHKVTANKIKHRIELTVDGNQVEAQSPNPASTSADT
NDPVFVGGFPDDLKQFGLTTSIPFRGCIRSLKLTKGTASHWRLILPRPWN

FIGURE 2 (continued)

cDNA sequence encoding laminin β1 subunit

1 cccggagcag ggcgagagct cgcgtcgccg gaaaggaaga cgggaagaaa gggcaggcgg
61 ctcggcgggc gtcttctcca ctcctctgcc gcgtccccgt ggctgcaggg agccggcatg
121 gggcttctcc agttgctagc tttcagtttc ttagccctgt gcagagcccg agtgcgcgct
181 caggaacccg agttcagcta cggctgcgca gaaggcagct gctatcccgc cacgggcgac
241 cttctcatcg gccgagcaca gaagctttcg gtgacctcga cgtgcgggct gcacaagccc
301 gaaccctact gtatcgtcag ccacttgcag gaggacaaaa aatgcttcat atgcaattcc
361 caagatcctt atcatgagac cctgaatcct gacagccatc tcattgaaaa tgtggtcact
421 acatttgctc caaaccgcct taagatttgg tggcaatctg aaaatggtgt ggaaaatgta
481 actatccaac tggatttgga agcagaattc cattttactc atctcataat gactttcaag
541 acattccgtc cagctgctat gctgatagaa cgatcgtccg actttgggaa aacctggggt
601 gtgtatagat acttcgccta tgactgtgag gcctcgtttc caggcatttc aactggcccc
661 atgaaaaaag tcgatgacat aatttgtgat tctcgatatt ctgacattga accctcaact
721 gaaggagagg tgatatttcg tgctttagat cctgctttca aaatagaaga tccttatagc
781 ccaaggatac agaatttatt aaaaattacc aacttgagaa tcaagtttgt gaaactgcat
841 actttgggag ataaccttct ggattccagg atggaaatca gagaaaagta ttattatgca
901 gtttatgata tggtggttcg aggaaattgc ttctgctatg gtcatgccag cgaatgtgcc
961 cctgtggatg gattcaatga agaagtggaa ggaatggttc acggacactg catgtgcagg
1021 cataacacca agggcttaaa ctgtgaactc tgcatggatt tctaccatga tttaccttgg
1081 agacctgctg aaggccgaaa cagcaacgcc tgtaaaaaat gtaactgcaa tgaacattcc
1141 atctcttgtc actttgacat ggctgtttac ctggccacgg ggaacgtcag cggaggcgtg
1201 tgtgatgact gtcagcacaa caccatgggg cgcaactgtg agcagtgcaa gccgttttac
1261 taccagcacc cagagaggga catccgagat cctaatttct gtgaacgatg tacgtgtgac
1321 ccagctggct ctcaaaatga gggaatttgt gacagctata ctgatttttc tactggtctc
1381 attgctggcc agtgtcggtg taaattaaat gtggaaggag aacattgtga tgtttgcaaa
1441 gaaggcttct atgatttaag cagtgaagat ccatttggtt gtaaatcttg tgcttgcaat
1501 cctctgggaa caattcctgg agggaatcct tgtgattccg agacaggtca ctgctactgc
1561 aagcgtctgg tgacaggaca gcattgtgac cagtgcctgc cagagcactg ggggcttaagc
1621 aatgatttgg atggatgtcg accatgtgac tgtgaccttg ggggagcctt aaacaacagt
1681 tgctttgcgg agtcaggcca gtgctcatgc cggcctcaca tgattggacg tcagtgcaac
1741 gaagtggaac ctggttacta ctttgccacc ctggatcact acctctatga agcggaggaa
1801 gccaacttgg ggcctggggt tagcatagtg gagcggcaat atatccagga ccggattccc
1861 tcctggactg gagccggctt cgtccgagtg cctgaagggg cttatttgga gtttttcatt
1921 gacaacatac catattccat ggagtacgac atcctaattc gctacgagcc acagctaccc
1981 gaccactggg aaaaagctgt catcacagtg cagcgacctg gaaggattcc aaccagcagc
2041 cgatgtggta ataccatccc cgatgatgac aaccaggtgg tgtcattatc accaggctca
2101 agatatgtcg tccttcctcg gccggtgtgc tttgagaagg gaacaaacta cacggtgagg
2161 ttggagctgc ctcagtacac ctcctctgat agcgacgtgg agagccccta cacgctgatc
2221 gattctcttg ttctcatgcc atactgtaaa tcactggaca tcttcaccgt gggaggttca
2281 ggagatgggg tggtcaccaa cagtgcctgg gaaacctttc agagataccg atgtctagag
2341 aacagcagaa gcgttgtgaa aacaccgatg acagatgttt gcagaaacat catctttagc
2401 atttctgccc tgttacacca gacaggcctg gcttgtgaat gcgaccctca gggttcgtta
2461 agttccgtgt gtgatcccaa cggaggccag tgccagtgcc ggcccaacgt ggttggaaga
2521 acctgcaaca gatgtgcacc tggaactttt ggctttggcc ccagtggatg caaaccttgt
2581 gagtgccatc tgcaaggatc tgtcaatgcc ttctgcaatc ccgtcactgg ccagtgccac
2641 tgtttccagg gagtgtatgc tcggcagtgt gatcggtgct tacctgggca ctggggcttt
2701 ccaagttgcc agccctgcca gtgcaatggc cacgccgatg actgcgaccc agtgactggg

2761 gagtgcttga actgccagga ctacaccatg ggtcataact gtgaaaggtg cttggctggt
2821 tactatggcg accccatcat tgggtcaggt gatcactgcc gcccttgccc ttgcccagat
2881 ggtcccgaca gtggacgcca gtttgccagg agctgctacc aagatcctgt tactttacag
2941 cttgcctgtg tttgtgatcc tggatacatt ggttccagat gtgacgactg tgcctcagga
3001 tactttggca atccatcaga agttgggggg tcgtgtcagc cttgccagtg tcacaacaac
3061 attgacacga cagacccaga agcctgtgac aaggagactg ggaggtgtct caagtgcctg
3121 taccacacgg aaggggaaca ctgtcagttc tgccggtttg gatactatgg tgatgccctc
3181 cggcaggact gtcgaaagtg tgtctgtaat tacctgggca ccgtgcaaga gcactgtaac
3241 ggctctgact gccagtgcga caaagccact ggtcagtgct tgtgtcttcc taatgtgatc
3301 gggcagaact gtgaccgctg tgcgcccaat acctggcagc tggccagtgg cactggctgt
3361 gacccatgca actgcaatgc tgctcattcc ttcgggccat cttgcaatga gttcacgggg
3421 cagtgccagt gcatgcctgg gtttggaggc cgcacctgca gcgagtgcca ggaactcttc
3481 tggggagacc ccgacgtgga gtgccgagcc tgtgactgtg accccagggg cattgagacg
3541 ccacagtgtg accagtccac gggccagtgt gtctgcgttg agggtgttga gggtccacgc
3601 tgtgacaagt gcacgcgagg gtactcgggg gtcttccctg actgcacacc ctgccaccag
3661 tgctttgctc tctgggatgt gatcattgcc gagctgacca acaggacaca cagattcctg
3721 gagaaagcca aggccttgaa gatcagtggt gtgatcgggc cttaccgtga gactgtggac
3781 tcggtggaga ggaaagtcag cgagataaaa gacatcctgg cgcagagccc cgcagcagag
3841 ccactgaaaa acattgggaa tctctttgag gaagcagaga aactgattaa agatgttaca
3901 gaaatgatgg ctcaagtaga agtgaaatta tctgacacaa cttcccaaag caacagcaca
3961 gccaaagaac tggattctct acagacagaa gccgaaagcc tagacaacac tgtgaaagaa
4021 cttgctgaac aactggaatt tatcaaaaac tcagatattc ggggtgcctt ggatagcatt
4081 accaagtatt tccagatgtc tcttgaggca gaggagaggg tgaatgcctc caccacagaa
4141 cccaacagca ctgtggagca gtcagccctc atgagagaca gagtagaaga cgtgatgatg
4201 gagcgagaat cccagttcaa ggaaaaacaa gaggagcagg ctcgcctcct tgatgaactg
4261 gcaggcaagc tacaaagcct agacctttca gccgctgccg aaatgacctg tggaacaccc
4321 ccaggggcct cctgttccga gactgaatgt ggcgggccaa actgcagaac tgacgaagga
4381 gagaggaagt gtggggggcc tggctgtggt ggtctggtta ctgttgcaca caacgcctgg
4441 cagaaagcca tggacttgga ccaagatgtc ctgagtgccc tggctgaagt ggaacagctc
4501 tccaagatgg tctctgaagc aaaactgagg gcagatgagg caaaacaaag tgctgaagac
4561 attctgttga agacaaatgc taccaaagaa aaaatggaca agagcaatga ggagctgaga
4621 aatctaatca agcaaatcag aaactttttg acccaggata gtgctgattt ggacagcatt
4681 gaagcagttg ctaatgaagt attgaaaatg gagatgccta gcacccccaca gcagttacag
4741 aacttgacag aagatatacg tgaacgagtt gaaagccttt ctcaagtaga ggttattctt
4801 cagcatagtg ctgctgacat tgccagagct gagatgttgt tagaagaagc taaaagagca
4861 agcaaaagtg caacagatgt taaagtcact gcagatatgg taaaggaagc tctggaagaa
4921 gcagaaaagg cccaggtcgc agcagagaag gcaattaaac aagcagatga agacattcaa
4981 ggaacccaga acctgttaac ttcgattgag tctgaaacag cagcttctga ggaaaccttg
5041 ttcaacgcgt cccagcgcat cagcgagtta gagaggaatg tggaagaact taagcggaaa
5101 gctgcccaaa actccgggga ggcagaatat attgaaaaag tagtatatac tgtgaagcaa
5161 agtgcagaag atgttaagaa gactttagat ggtgaacttg atgaaaagta taaaaaagta
5221 gaaaatttaa ttgccaaaaa aactgaagag tcagctgatg ccagaaggaa agccgaaatg
5281 ctacaaaatg aagcaaaaac tcttttagct caagcaaata gcaagctgca actgctcaaa
5341 gatttagaaa gaaaatatga agacaatcaa agatacttag aagataaagc tcaagaatta
5401 gcaagactgg aaggagaagt ccgttcactc ctaaaggata taagccagaa agttgctgtg
5461 tatagcacat gcttgtaaca gaggagaata aaaaatggct gaggtgaaca aggtaaaaca
5521 actacatttt aaaaactgac ttaatgctct tcaaaataaa acatcaccta tttaatgttt
5581 ttaatcacat tttgtatgag ttaaataaag ccc

FIGURE 3 (continued)

amino acid sequence of laminin β1 subunit

MGLLQLLAFSFLALCRARVRAQEPEFSYGCAEGSCYPATGDLLI

GRAQKLSVTSTCGLHKPEPYCIVSHLQEDKKCFICNSQDPYHETLNPDSHLIENVVTT

FAPNRLKIWWQSENGVENVTIQLDLEAEFHFTHLIMTFKTFRPAAMLIERSSDFGKTW

GVYRYFAYDCEASFPGISTGPMKKVDDIICDSRYSDIEPSTEGEVIFRALDPAFKIED

PYSPRIQNLLKITNLRIKFVKLHTLGDNLLDSRMEIREKYYYAVYDMVVRGNCFCYGH

ASECAPVDGFNEEVEGMVHGHCMCRHNTKGLNCELCMDFYHDLPWRPAEGRNSNACKK

CNCNEHSISCHFDMAVYLATGNVSGGVCDDCQHNTMGRNCEQCKPFYYQHPERDIRDP

NFCERCTCDPAGSQNEGICDSYTDFSTGLIAGQCRCKLNVEGEHCDVCKEGFYDLSSE

DPFGCKSCACNPLGTIPGGNPCDSETGHCYCKRLVTGQHCDQCLPEHWGLSNDLDGCR

PCDCDLGGALNNSCFAESGQCSCRPHMIGRQCNEVEPGYYFATLDHYLYEAEEANLGP

GVSIVERQYIQDRIPSWTGAGFVRVPEGAYLEFFIDNIPYSMEYDILIRYEPQLPDHW

EKAVITVQRPGRIPTSSRCGNTIPDDDNQVVSLSPGSRYVVLPRPVCFEKGTNYTVRL

ELPQYTSSDSDVESPYTLIDSLVLMPYCKSLDIFTVGGSGDGVVTNSAWETFQRYRCL

ENSRSVVKTPMTDVCRNIIFSISALLHQTGLACECDPQGSLSSVCDPNGGQCQCRPNV

VGRTCNRCAPGTFGFGPSGCKPCECHLQGSVNAFCNPVTGQCHCFQGVYARQCDRCLP

GHWGFPSCQPCQCNGHADDCDPVTGECLNCQDYTMGHNCERCLAGYYGDPIIGSGDHC

RPCPCPDGPDSGRQFARSCYQDPVTLQLACVCDPGYIGSRCDDCASGYFGNPSEVGGS

CQPCQCHNNIDTTDPEACDKETGRCLKCLYHTEGEHCQFCRFGYYGDALRQDCRKCVC

NYLGTVQEHCNGSDCQCDKATGQCLCLPNVIGQNCDRCAPNTWQLASGTGCDPCNCNA

AHSFGPSCNEFTGQCQCMPGFGGRTCSECQELFWGDPDVECRACDCDPRGIETPQCDQ

STGQCVCVEGVEGPRCDKCTRGYSGVFPDCTPCHQCFALWDVIIAELTNRTHRFLEKA

KALKISGVIGPYRETVDSVERKVSEIKDILAQSPAAEPLKNIGNLFEEAEKLIKDVTE

MMAQVEVKLSDTTSQSNSTAKELDSLQTEAESLDNTVKELAEQLEFIKNSDIRGALDS

FIGURE 4

ITKYFQMSLEAEERVNASTTEPNSTVEQSALMRDRVEDVMMERESQFKEKQEEQARLL

DELAGKLQSLDLSAAAEMTCGTPPGASCSETECGGPNCRTDEGERKCGGPGCGGLVTV

AHNAWQKAMDLDQDVLSALAEVEQLSKMVSEAKLRADEAKQSAEDILLKTNATKEKMD

KSNEELRNLIKQIRNFLTQDSADLDSIEAVANEVLKMEMPSTPQQLQNLTEDIRERVE

SLSQVEVILQHSAADIARAEMLLEEAKRASKSATDVKVTADMVKEALEEAEKAQVAAE

KAIKQADEDIQGTQNLLTSIESETAASEETLFNASQRISELERNVEELKRKAAQNSGE

AEYIEKVVYTVKQSAEDVKKTLDGELDEKYKKVENLIAKKTEESADARRKAEMLQNEA

KTLLAQANSKLQLLKDLERKYEDNQRYLEDKAQELARLEGEVRSLLKDISQKVAVYST

FIGURE 4 (continued)

```
Helix II:~/humlam> more beta.prty
 LINEUP of: beta  from: 1  to: 1841   May 7, 1997 12:31  ..

           1                                                     50
 shortb4                         MQFQLTLF LHLGWLSYSK AQDDCNRGAC
   b4pep                         MQFQLTLF LHLGWLSYSK AQDDCNRGAC
   b3pep   .......... .....MRPFF LLCFALPGL. ...LHAQQ.. ...ACSRGAC
   b2pep   MELTSTERGR GQPLPWELRL PLLLSVLAAT LAQAPAPDVP ...GCSRGSC
   b1pep   .......... .....MGLLQ LLAFSFLALC RARVRAQEPE FSYGCAEGSC
    beta   .......... .....m.l.. llm..ql.lf l.lgwa.ysk ....C.rGaC

           51                                                   100
 shortb4   HPTTGDLLVG RNTQLMASST CGLSRAQKYC ILSYLEG.EQ KCSICDSRFP
   b4pep   HPTTGDLLVG RNTQLMASST CGLSRAQKYC ILSYLEG.EQ KCSICDSRFP
   b3pep   YPPVGDLLVG RTRFLRASST CGLTKPETYC ..TQYGEWQM KCCKCDSRQP
   b2pep   YPATADLLVG RADRLTASST CGLNGRQPYC IVSHLQD.EK KCFLCDSRRP
   b1pep   YPATGDLLIG RAQKLSVTST CGLHKPEPYC IVSHLQE.DK KCFICNSQDP
    beta   yP.tgDLLvG R.tqlmasST CGLs..q.YC i.s.l...e. KC.iCdSrfP

           101                                                  150
 shortb4   YDPYDQPNSH TIENVTVSFE PDREKKWWQS ENGLDHVSIR LDLEALFRFS
   b4pep   YDPYDQPNSH TIENVTVSFE PDREKKWWQS ENGLDHVSIR LDLEALFRFS
   b3pep   H....NYYSH RVENVASSSG PMR...WWQS QNDVNPVSLQ LDLDRRFQLQ
   b2pep   FSARDNPHTH RIQNVVTSFA PQRRAAWWQS QNGIPAVTIQ LDLEAEFHFT
   b1pep   YHETLNPDSH LIENVVTTFA PNRLKIWWQS ENGVENVTIQ LDLEAEFHFT
    beta   ydpydnpnsH .ieNV..sf. PdRekkWWQS eNg.dhVsiq LDLea.F.f.

           151                                                  200
 shortb4   HLILTFKTFR PAAMLVERST DYGHNWKVFK YFAKDCATSF PNITSGQAQG
   b4pep   HLILTFKTFR PAAMLVERST DYGHNWKVFK YFAKDCATSF PNITSGQAQG
   b3pep   EVMMEFRGPM PAGMLIERSS DFGKTWRVYQ YLAADCTSTF PRVRQGRPQS
   b2pep   HLIMTFKTFR PAAMLVERSA DFGRTWHVYR YFSYHCGADF PGVPLAPPRH
   b1pep   HLIMTFKTFR PAAMLIERSS DFGKTWGVYR YFAYDCEASF PGISTGPMKK
    beta   hlimtFktfr PAaMLvERS. DfG.tWkVy. Yfa.dCa.sF P.itsg..qg

           201                                                  250
 shortb4   VGDIVCDS.K YSDIEPSTGG EVVLKVLDPS FEIENPYSPY IQDLVTLTNL
   b4pep   VGDIVCDS.K YSDIEPSTGG EVVLKVLDPS FEIENPYSPY IQDLVTLTNL
   b3pep   WQDVRCQSLP QRPNARLNGG KVQLNLMDLV SGIPATQSQK IQEVGEITNL
   b2pep   WDDVVCES.R YSEIEPSTEG EVIYRVLDPA IPIPDPYSSR IQNLLKITNL
   b1pep   VDDIICDS.R YSDIEPSTEG EVIFRALDPA FKIEDPYSPR IQNLLKITNL
    beta   v.DivCdS.. ysdiepstgG eV.l.vlDp. feIe.pySp. IQ.l..iTNL

           251                                                  300
 shortb4   RINFTKLHTL GDALLGRRQN DSLDKYYYAL YEMIVRGSCF CNGHASECRP
   b4pep   RINFTKLHTL GDALLGRRQN DSLDKYYYAL YEMIVRGSCF CNGHASECRP
   b3pep   RVNFTRLAPV PQRGYHPPS. .....AYYAV SQLRLQGSCF CHGHADRCAP
   b2pep   RVNLTRLHTL GDNLLDPRR. EIREKYYYAL YELVVRGNCF CYGHASECAP
   b1pep   RIKFVKLHTL GDNLLDSRM. EIREKYYYAV YDMVVRGNCF CYGHASECAP
    beta   RinftkLhtl gd.ll..rq. ....kyYYAl yem.vrGsCF C.GHAseCaP

           301                                                  350
 shortb4   MQKMRGDVFS PPGMVHGQCV CQHNTDGPNC ERCKDFFQDA PWRPAADLQD
   b4pep   MQKMRGDVFS PPGMVHGQCV CQHNTDGPNC ERCKDFFQDA PWRPAADLQD
   b3pep   KPGASAGSTA V..QVHDVCV CQHNTAGPNC ERCAPFYNNR PWRPAEGQDA
   b2pep   APGAPAHA.. .EGMVHGACI CKHNTRGLNC EQCQDFYRDL PWRPAEDGHS
   b1pep   VDGFNEEV.. .EGMVHGHCM CRHNTKGLNC ELCMDFYHDL PWRPAEGRNS
    beta   m.g.r.dv.. ..gmVHgqCv CqHNTdGpNC ErCkdFyqd. PWRPAedlq.

           351                                                  400
 shortb4   NACRSCSCNS HSSRCHFDMT TYLASGGLSG GVCEDCQHNT EGQHCDRCRP
   b4pep   NACRSCSCNS HSSRCHFDMT TYLASGGLSG GVCEDCQHNT EGQHCDRCRP
   b3pep   HECQRCDCNG HSETCHFDPA VFAASQGAYG GVCDNCRDHT EGKNCERCQL
   b2pep   HACRKCDRHG HTHSCHFDMA VYLGSGNVSG GVCDGCQHNT AWRHCELCRP
   b1pep   NACKKCNCNE HSISCHFDMA VYLATGNVSG GVCDDCQHNT MGRNCEQCKP
    beta   naCr.C.cn. Hss.CHFDma vylasgg.sG GVCddCqhnT eg.hCerCrp

           401                                                  450
 shortb4   LFYRDPLKTI SDPYACIPCE CDPDGTISGG ICVSHSDPAL GSVAGQCLCK
   b4pep   LFYRDPLKTI SDPYACIPCE CDPDGTISGG ICVSHSDPAL GSVAGQCLCK
   b3pep   HYFRNRRPGA SIQETCISCE CDPDGAVAGA PCDP...... ..VTGQCVCK
   b2pep   FFYRDPTKDL RDPAVCRSCD CDPMGSQDGG RCDSHDDPAL GLVSGQCRCK
   b1pep   FYYQHPERDI RDPNFCERCT CDPAGSQNEG ICDSYTDFST GLIAGQCRCK
    beta   .fyrdplk.i sdpyaCi.Ce CDPdG..sgg iCdshsdpal g.vaGQC.CK

           451                                                  500
 shortb4   ENVEGAKCDQ CKPNHYGLSA TDPLGCQPCD CNPLGSLPFL T.CDVDTGQC
   b4pep   ENVEGAKCDQ CKPNHYGLSA TDPLGCQPCD CNPLGSLPFL T.CDVDTGQC
   b3pep   EHVQGERCDL CKPGFTGLTY ANPRRCHRCD CNIL...... ..........
   b2pep   EHVVGTRCQQ CRDGFFGLSI SDPSGCRRCQ CNARGTVPGS TPCDPNSGSC
```

FIGURE 5

```
   b1pep  LNVEGEHCDV CKEGFYDLSS EDPFGCKSCA CNPLGTIPGG NPCDSETGHC
    beta  enVeG..Cdq CkpgfygLsa tdPlgCq.Cd CNplg.lp.l t.cdvdtgqc

          501                                                550
 shortb4  LCLSYVTGAH CEECTVGYWG LGNHLHGCSP CDCDIGGAYS NVCSPKNGQC
   b4pep  LCLSYVTGAH CEECTVGYWG LGNHLHGCSP CDCDIGGAYS NVCSPKNGQC
   b3pep  .......... .......... .......... .......... ..........
   b2pep  YCKRLVTGRG CDRCLPGHWG LSLDLLGCRP CDCDVGGALD PQCDEGTGQC
   b1pep  YCKRLVTGQH CDQCLPEHWG LSNDLDGCRP CDCDLGGALN NSCFAESGQC
    beta  .c...vtgah c.ec..g.wg l.n.lhgc.p cdcdigga.s nvcspkngqc

          551                                                600
 shortb4  ECRPHVTGRS|CSEPAPGYFF APLNFYLYEA EEATTLQGLA PLGSETFGQS
   b4pep  ECRPHVTGRS|CSEPAPGYFF APLNFYLYEA EEATTLQGLA PLGSETFGQS
   b3pep  ..........|.......... .......... .......... ..........
   b2pep  HCRQHMVGRR|CEQVQPGYFR PFLDHLIWEA ENTR...... .........G
   b1pep  SCRPHMIGRQ|CNEVEPGYYF ATLDHYLYEA EEANL..... .........G
    beta  ecrph.tgrs|cse.apgyff apl..ylyea eeat...... ..........

          601                                                650
 shortb4  PAVHVVLGEP VPGNPVTWTG PGFARVLPGA GLRFAVNNIP FPVDFTIAIH
   b4pep  PAVHVVLGEP VPGNPVTWTG PGFARVLPGA GLRFAVNNIP FPVDFTIAIH
   b3pep  .......... .......... .......... .......... ..........
   b2pep  QVLDVVERLV TPGETPSWTG SGFVR[illegible]Q TLEFLVASVP NAMDYDLLLR
   b1pep  PGVSIVERQY IQDRIPSWTG AGFVRVPEGA YLEFFIDNIP YSMEYDILIR
    beta  pavhvv..ep vpgnp..wtg pgf.rvl.ga gl.favnnip fp.d..i.i.

          651                                                700
 shortb4  YETQSAADWT VQIV.VNPPG G...SEHCIP KTLQSKPQSF ALPAATRIML
   b4pep  YETQSAADWT VQIV.VNPPG G...SEHCIP KTLQSKPQSF ALPAATRIML
   b3pep  .......... .......... ...GSR.... .......... .........E
   b2pep  LEPQVPEQWA ELELIVQRPG PVPAHSLCGH LVPRDDRIQG TLQPHARYLI
   b1pep  YEPQLPDHWE KAVITVQRPG RIPTSSRCGN TIPDDDNQVV SLSPGSRYVV
    beta  ye.qs.adwt vqiv.v..pg g...s.hc.p kt.q..pqsf alp.atr.ml

          701                                                750
 shortb4  LPTPICLEPD VQYSIDVYFS QPLQGESHAH S.....HVLV DSLGLIPQIN
   b4pep  LPTPICLEPD VQYSIDVYFS QPLQGESHAH S.....HVLV DSLGLIPQIN
   b3pep  MP........ .......... .......... .......... ..........
   b2pep  FPNPVCLEPG ISYKLHLKLV R.TGGSAQPE TPYSGPGLLI DSLVLLPRVL
   b1pep  LPRPVCFEKG TNYTVRLELP QYTSSDSDVE SPYT....LI DSLVLMPYCK
    beta  lPtp.clep. vqysid.y.s q..qgesha. s.......l. dsl.lipqin

          751                                                800
 shortb4  SLENF..... ....CSKQDL DEYQLHNCVE IASAMGPQVL PGACERLIIS
   b4pep  SLENF..... ....CSKQDL DEYQLHNCVE IASAMGPQVL PGACERLIIS
   b3pep  .......... .......... .......... .......... ..........
   b2pep  VLEMF....S GGDAAALERQ ATFERYQCHE EGLVPSKTSP SEACAPLLIS
   b1pep  SLDIFTVGGS GDGVVTNSAW ETFQRYRCLE NSRSVVKTPM TDVCRNIIFS
    beta  slenf..... ....cskqdl d..q..ncve iasamg..vl pgacerliis

          801                                                850
 shortb4  MSAKLHDGAV|ACKCHPQGSV GSSCSRLGGQ CQCKPLVVGR CCDRCSTGSY
   b4pep  MSAKLHDGAV|ACKCHPQGSV GSSCSRLGGQ CQCKPLVVGR CCDRCSTGSY
   b3pep  ..........|.......... .......... .......... ..........
   b2pep  LSTLIYNGAL|PCQCNPQGSL SSECNPHGGQ CLCKPGVVGR RCDTCAPGYY
   b1pep  ISALLHQTGL|ACECDPQGSL SSVCDPNGGQ CQCRPNVVGR TCNRCAPGTF
    beta  msa.lhdga.|ackchpqgs. .sscs.lggq cqckplvvgr ccdrc..gsy

          851                                                900
 shortb4  DLGHHGCHPC HCHPQGSKDT VCDQVTGQCP CHGEVSGRRC DRCLAGYFGF
   b4pep  DLGHHGCHPC HCHPQGSKDT VCDQVTGQCP CHGEVSGRRC DRCLAGYFGF
   b3pep  .......... .......... .......... .......... ..........
   b2pep  GFGPTGCQAC QCSPRGALSS LCERTSGQCL CRTGAFGLRC DACQRGQWGF
   b1pep  GFGPSGCKPC ECHLQGSVNA FCNPVTGQCH CFQGVYARQC DRCLPGHWGF
    beta  ..g.hgchpc hchpqgskdt vcdqvtgqcp chg.vsgrrc drclagy.gf

          901                                                950
 shortb4  PSCHPCPCNR FAELCDPETG SCFNCGGFTT GRNCERCIDG YYGNP..SSG
   b4pep  PSCHPCPCNR FAELCDPETG SCFNCGGFTT GRNCERCIDG YYGNP..SSC
   b3pep  .......... .......... .......... .......... ..........
   b2pep  PSCRPCVCNG HADECNTHTC ACLGCRDLTG GEHCERCIAG FHGDPRLPYG
   b1pep  PSCQPCQCNG HADDCDPVTG ECLNCQDYTM GHNCERCLAG YYGDPIIGSG
    beta  pschpcpcn. .a.lcdpetg sc.ncg.ftt grncerci.g yyg.p..ssg

          951                                               1000
 shortb4  QPCRPCLCPD DPSSNQYFAH SCYQNLWSSD VICNCLQGYT GTQCGECSTG
   b4pep  QPCRPCLCPD DPSSNQYFAH SCYQNLWSSD VICNCLQGYT GTQCGECSTG
   h3pep  .......... .......... .......... .......... ..........
   b2pep  AQCRPCPCPE GPGSQRHFAT SCHQDEYSQQ IVCHCRAGYT GLRCEACAPG
```

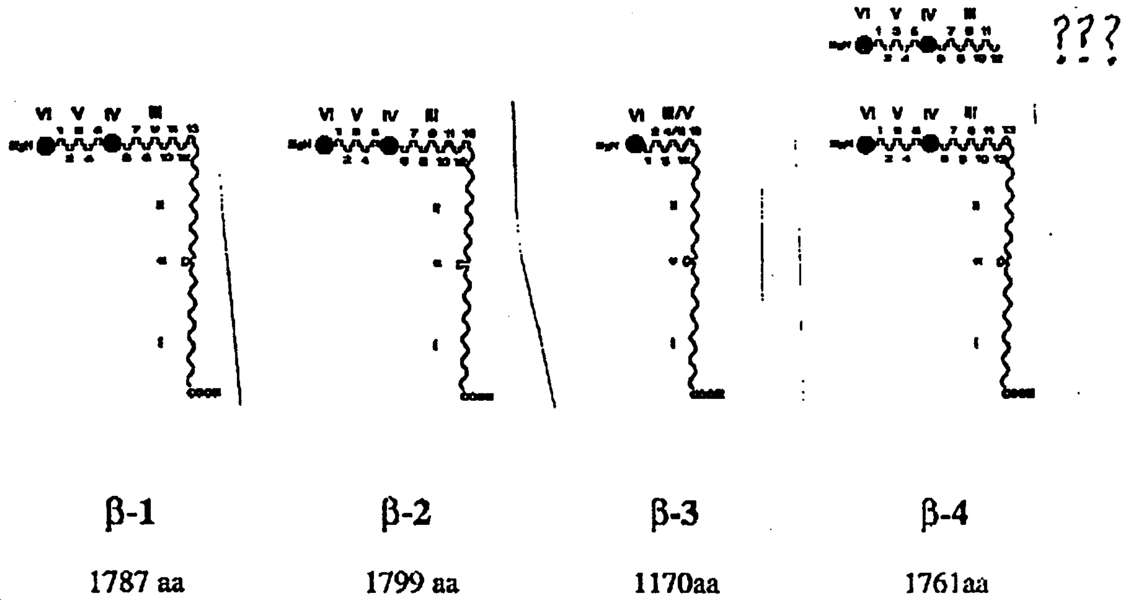
FIGURE 5 (continued)

```
   b1pep  DHCRPCPCPD G[illegible]QFAR SCYQDPVTLQ LACVCDPG[illegible]CDDCASG
    beta  qpcrpc.cpd .pssn.yfah scyq.lwss. vicnclqgyt gt.cgec.tg

          1001                                               1050
 shortb4  FYGNPRISGA PCQPCACNNN IDVTDPESCS RVTGECLRCL HNTQGANCQL
   b4pep  FYGNPRISGA PCQPCACNNN IDVTDPESCS RVTGECLRCL HNTQGANCQL
   b3pep  .......... .......... .......... .......... ..........
   b2pep  QFGDPSRPGG RCQLCECSGN IDPMDPDACD PHPGQCLRCL HHTEGPHCAH
   b1pep  YFGNPSEVGG SCQPCQCHNN IDTTDPEACD KETGRCLKCL YHTEGEHCQF
    beta  f.gnp.isg. pcqpcacnnn idvtdpe.c. rvtgeclrcl h.t.ga.cql

          1051                                               1100
 shortb4  CKPGHYGSAL NQTCRRCSCH ASGVSPMECP PGGGACLCDP VTGACPCLPN
   b4pep  CKPGHYGSAL NQTCRRCSCH ASGVSPMECP PGGGACLCDP VTGACPCLPN
   b3pep  .......... .......... .......... .......CDE ESGRCLCLPN
   b2pep  SKPGFHGQAA RQSCHRCTCN LLGTNPQQCP SPD.QCHCDP SSGQCPCLPN
   b1pep  CRFGYYGDAL RQDCRKCVCN YLGTVQEHCN GSD..CQCDK ATGQCLCLPN
    beta  ckpghygsal .qtcrrcsc. a.g.spmecp pg...clCDp vtG.CpCLPN

          1101                                               1150
 shortb4  VTGLACDRCA DGYWNLVPGR GCQSCDCDPR TSQSSHCDQA RYFKAY
   b4pep  VTGLACDRCA DGYWNLVPGR GCQSCDCDPR TSQSSHCDQL TGQCPCKLGY
   b3pep  VVGPKCDQCA PYHWKLASGQ GCEPCACDPH NSLSPQCNQF TGQCPCREGF
   b2pep  VQALAVDRCA PNFWNLTSGH GCQPCACLPS PEEGPTCNEF TGQCHCLCGF
   b1pep  VIGQNCDRCA PNTWQLASGT GCDPCNCNAA HSFGPSCNEF TGQCQCMPGF
    beta  VtglacDrCA p.yWnL.sGr GCqpC.Cdpr tsqsphCnqf tgqcpc..Gf

          1151                                               1200
   b4pep  GGKRCS.... .ECQENYYGD PPGRCIPCDC NRAGTQKPIC DPDTGMCRCR
   b3pep  GGLMCSAAAI RQCPDRTYGD VATGCRACDC DFRGTEGPGC DKASGRCLCR
   b2pep  GGRTCS.... .ECQELHWGD PGLQCHACDC DSRGIDTPQC HRFTGHCTCR
   b1pep  GGRTCS.... .ECQELFWGD PDVECRACDC DPRGIETPQC DQSTGQCVCV
    beta  GGrtCS.... .eCqel..GD p...CraCDC d.rG.etPqC d..tG.C.Cr

          1201                                               1250
   b4pep  EGVSGQRCDR CARGHSQEFP TCLQCHLCFD QWDHTISSLS KAVQGLMRLA
   b3pep  PGITGPRCDQ CQRGYCNRYP VCVACHPCFQ TYDADLREQA LRFGRLPNAT
   b2pep  PGVSGVRCDQ CARGFSGIFP ACHPCHACFG DWDRVVQDLA ARTQRLEQRA
   b1pep  EGVEGPRCDK CTRGYSGVFP DCTPCHQCFA LWDVIIAELT NRTHRFLEKA
    beta  .GvsGpRCDq CaRGysg.fP .C.pCH.CF. .wD..i.ela .rtqrl...a

          1251                                               1300
   b4pep  ANMEDKRETL PVCEADFKDL RGNVSEIERI L......KHP VFPSGK.FLK
   b3pep  ASLWSGPGLE D.RGL.ASRI LDAKSKIEQI RAVLSSPAVT EQEVAQVASA
   b2pep  QELQQTGVLG AFESS.FWHM QEKLGIVQGI VGARNTSAAS TAQLVEATEE
   b1pep  KALKISGVIG PYRET.VDSV ERKVSEIKDI L.AQSPAAEP LKNIGNLFEE
    beta  a.l...gvlg p.re..f... ..kvseie.I l.a.s..a.p ....g..fee

          1301                                               1350
   b4pep  VKDYHDSVRR QIMQLNEQLK .AVYEFQDLK DTIERAKNEA DLLLEDLQEE
   b3pep  ILSLRRTLQG L..QLDLPLE EET...LSLP RDLESLDRSF NGLLTMYQRK
   b2pep  LRREIGEATE HLTQLEADLT DVQDENFNAN HALSGLERDR LALNLTLRQL
   b1pep  AEKLIKDVTE MMAQVEVKLS DTTSQSNSTA KELDSLQTEA ESLDNTVKEL
    beta  ...li..vte ...Qle..L. d.t.e..sl. ..lesl.rea ..Ll.tlqel

          1351                                               1400
   b4pep  IDLQSSVLNA SIADSSENIK KYYHISSSAE KKIN....ET SSTINTSANT
   b3pep  REQFEKISSA DPSGAFRMLS TAYEQSAQAA QQVS...... .....DSSRL
   b2pep  DQHLDLLKHS NFLGAYDSIR HAHSQSAEAE RRANTSALAV PSPVSNSASA
   b1pep  AEQLEFIKNS DIRGALDSIT KYFQMSLEAE ERVNASTTEP NSTVEQSALM
    beta  .eqle.ikn. di.ga.dsi. k.y.qSaeAe .rvn....e. .stv..Sa..

          1401                                               1450
   b4pep  RNDLLTIL.. ......DTLT SKGNLSLERL KQIKIPDIQI LNEKV.....
   b3pep  LDQLRDSRRE AERLVRQAGG GGGTGSPKLV A..LRLEMSS LPDLTPTFNK
   b2pep  RHRTEALMDA QKEDFNSKHM ANQRALGKLS AHTHTLSLTD INELV.....
   b1pep  RDRVEDVMME RESQFKEKQE EQARLLDELA GKLQSLDLSA AAEMT.....
    beta  rdrled.m.e .e..f..k.. ..grl...l. a....ldls. lnel......

          1451                                               1500
   b4pep  .CGDPGNVPC VPLPCGGALC TGRKGHRKCR GPGCHGSLTL STNALQKAQE
   b3pep  LCGNSRQMAC TPISCPGELC PQDNG.TAC. ASRCRGVLPR AGGAFLMAGQ
   b2pep  .CGAQGLHHD RTSPCGGAGC RDEDGQPRCG GLSCNGAAAT ADLALGRARH
   b1pep  .CGTPPGASC SETECGGPNC RTDEGERKCG GPGCGGLVTV AHNANQKAMD
    beta  .CG.pg...c .p.pCgGalC r.d.G.rkCg gpgC.G.lt. a.nAlqkA..

          1501                                               1550
   b4pep  AKSIIRNLDK QVRGLKNQIE STSEQAEVSK NNALQ[illegible]KL GNIRNQSDSE
   b3pep  VAEQLRGFNA QLQRTRQMIR AAEESASQIQ SSAQRLETQV SASRSQMEED
   b2pep  TQAELQRALA EGGSILSRVA ETRRQASEAQ QRAQAALDKA NASRGQVEQA
```

FIGURE 5 (continued)

```
b1pep  LDQDVLSALA EVEQLSKMVS EAKLRADEAK QSAEDILLKT NATKEKMDKS
 beta  ....lr.ala .v..l..m.. ea.eqAsea. qsAq.ll.k. nasr.qm...

       1551                                              1600
b4pep  EENINLFIKK VKNFLLEENV PPEDIEKVAN GVLDIHLPIP SQNI/TDELVK
b3pep  VRRTRLLIQQ VRDFLTDPDT DAATIQEVSE AVLALWLPTD SATVLQKMNE
b2pep  NQE[illegible]LIQS VKDFLNQEGA DPDSIEMVAT RVLELSIPAS AEQIQHLAGA
b1pep  NEKLRNLIKQ IRNFLTQDSA DLDSIEAVAN EVLKMEMPST PQQLQNLTED
 beta  neelrllI.q v..FLtqe.a dpdsIe.Van .VL.l.lP.. sqqlq.l...

       1601                                              1650
b4pep  IQKHMQLCED YRTDENRSNE EADGAQKLLV KAKAAEKAA. NILLNLDKTL
b3pep  IQAIAARLPN VDLVLSQTKQ DIARARRLQA EAEEARSRAH AVEGQVEDVV
b2pep  IAERVRSLAD VDAILARTVG DVRRAEQLLQ DARRARSWAE DEKQKAETVQ
b1pep  IRERVESLSQ VEVILQHSAA DIARAEMLLE EAKRASKSAT DVKVTADMVK
 beta  Iqerv.sl.d vd.il.r... diarAe.Ll. eAkrAr..A. dvk..a..v.

       1651                                              1700
b4pep  NQLQQAQITQ GRANSTITQL TANITKIKKN VLQAENQTRE MKSELELAKQ
b3pep  GNLRQGTVAL QEAQDTMQGT SRSLRLIQDR VAEVQQVLRP AEKLVTSMTK
b2pep  AALEEAQRAQ GIAQGAIRGA VADTRDTEQT LYQVQERMAG AERALSSAGE
b1pep  EALEEAEKAQ VAAEKAIKQA DEDIQGTQNL LTSIESETAA SEETLFNASQ
 beta  .aLe.aq.aq g.Aq..i..a .adir..q.. ..qv...t.. ae..l.sa.q

       1701                                              1750
b4pep  R.SGLEDGLS LLQTKLQRHQ DHAVNAKVQA ESAQHQAGSL EKEF.VELKK
b3pep  QLGDFWTRME ELRHQARQQG AEAVQAQQLA EGASEQALSA QEGFE.RIKQ
b2pep  RARQLDALLE ALKLKRAGNS LAASTAEETA GSAQGRAQEA EQL[illegible]PLGD
b1pep  RISELERNVE ELKRKAAQNS GEAEYIEKVV YTVKQSAEDV KKTLDGELDE
 beta  r.s.le..le eLk.kaaqns .eAv.ae..a esaq.qA.sa ek...gelk.

       1751                                              1800
b4pep  QYAILQRK.T STTGLTKETL GKVKQLKDAA EKLAGDTEAK IRRITDLERK
b3pep  KYAELKDRLG QSSMLGEQGA R.IQSVKTEA EELFGETMEM MDRMKDMELE
b2pep  QYQTVKALAE RKAQGVLAAQ ARAEQLPDEA RDLLQAAQDK LQRLQELEGT
b1pep  KYKKVENLIA KKTEESADAR RKAEMLQNEA KTLLAQANSK LQLLKDLERK
 beta  .Ya..k.l.. .kt.l...a. rkaeqlkdeA e.Llg....k lqrlkdlErk

       1801                                 1841
b4pep  IQDLNLSRQA KADQLRILED QVVAIKNEIV EQEKKYARCY S
b3pep  LLRGSQAIML RSADLTGLEK RVEQIRDHIN GRVLYYATCK
b2pep  YEENERALES KAAQLDGLEA RMRSVLQAIN LQVQIYNTCQ
b1pep  YEDNQRYLED KAQELARLEG EVRSLLKDIS QKVAVYSTCL
 beta  yedn.rale. kaaqL.gLE. rvrsil..In .qv..YatC. S
```

FIGURE 5 (continued)



β-1	β-2	β-3	β-4
1787 aa	1799 aa	1170aa	1761aa

β-4	I	α	II	III	IV	V	VI
β-1	24.6 34.0	48.5 48.5	23.0 37.7	53.3 59.2	35.0 44.9	55.8 61.7	59.4 68.6
β-2	31.9 22.8	41.2 44.1	19.3 31.3	49.9 55.5	29.3 43.9	55.0 59.5	56.9 67.8
β-3	21.7 30.6	36.0 40.0	16.3 28.6	39.7 44.9	——	48.4 54.0	45.4 56.0

Comparison of the similarity of laminin beta-4 chain domains to the domains of other laminin beta chains

The percentage amino acid identity and percentage amino acid similarity are given for each domain compared.

FIGURE 6

5 SEQUENCE LISTING

(1) GENERAL INFORMATION:

10 (i) APPLICANT: Burgeson, Robert, et al.

(ii) TITLE OF INVENTION: DNA Sequences

(iii) NUMBER OF SEQUENCES: 6

15

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: LAHIVE & COCKFIELD
(B) STREET: 28 State Street
(C) CITY: Boston
20 (D) STATE: Massachusetts
(E) COUNTRY: USA
(F) ZIP: 02109

(v) COMPUTER READABLE FORM:
25 (A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

30 (vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: US 000000
(B) FILING DATE: 13-APR-1994
(C) CLASSIFICATION:

35 (vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US 08/111,111
(B) FILING DATE: 12-DEC-1909

(viii) ATTORNEY/AGENT INFORMATION:
40 (A) NAME: Attorney, Name Init
(B) REGISTRATION NUMBER: 000000
(C) REFERENCE/DOCKET NUMBER: oe

(ix) TELECOMMUNICATION INFORMATION:
45 (A) TELEPHONE: (617)227-7400
(B) TELEFAX: (617)742-4214

(2) INFORMATION FOR SEQ ID NO:1:
50

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1761 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear
55

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

60

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

5 Met Gln Phe Gln Leu Thr Leu Phe Leu His Leu Gly Trp Leu Ser Tyr
1 5 10 15

Ser Lys Ala Gln Asp Asp Cys Asn Arg Gly Ala Cys His Pro Thr Thr
10 20 25 30

Gly Asp Leu Leu Val Gly Arg Asn Thr Gln Leu Met Ala Ser Ser Thr
35 40 45

Cys Gly Leu Ser Arg Ala Gln Lys Tyr Cys Ile Leu Ser Tyr Leu Glu
15 50 55 60

Gly Glu Gln Lys Cys Ser Ile Cys Asp Ser Arg Phe Pro Tyr Asp Pro
65 70 75 80

20 Tyr Asp Gln Pro Asn Ser His Thr Ile Glu Asn Val Thr Val Ser Phe
85 90 95

Glu Pro Asp Arg Glu Lys Lys Trp Trp Gln Ser Glu Asn Gly Leu Asp
25 100 105 110

His Val Ser Ile Arg Leu Asp Leu Glu Ala Leu Phe Arg Phe Ser His
115 120 125

Leu Ile Leu Thr Phe Lys Thr Phe Arg Pro Ala Ala Met Leu Val Glu
30 130 135 140

Arg Ser Thr Asp Tyr Gly His Asn Trp Lys Val Phe Lys Tyr Phe Ala
145 150 155 160

35 Lys Asp Cys Ala Thr Ser Phe Pro Asn Ile Thr Ser Gly Gln Ala Gln
165 170 175

Gly Val Gly Asp Ile Val Cys Asp Ser Lys Tyr Ser Asp Ile Glu Pro
40 180 185 190

Ser Thr Gly Gly Glu Val Val Leu Lys Val Leu Asp Pro Ser Phe Glu
195 200 205

Ile Glu Asn Pro Tyr Ser Pro Tyr Ile Gln Asp Leu Val Thr Leu Thr
45 210 215 220

Asn Leu Arg Ile Asn Phe Thr Lys Leu His Thr Leu Gly Asp Ala Leu
225 230 235 240

50 Leu Gly Arg Arg Gln Asn Asp Ser Leu Asp Lys Tyr Tyr Tyr Ala Leu
245 250 255

Tyr Glu Met Ile Val Arg Gly Ser Cys Phe Cys Asn Gly His Ala Ser
55 260 265 270

Glu Cys Arg Pro Met Gln Lys Met Arg Gly Asp Val Phe Ser Pro Pro
275 280 285

Gly Met Val His Gly Gln Cys Val Cys Gln His Asn Thr Asp Gly Pro
60 290 295 300

Asn Cys Glu Arg Cys Lys Asp Phe Phe Gln Asp Ala Pro Trp Arg Pro
305 310 315 320

5

Ala Ala Asp Leu Gln Asp Asn Ala Cys Arg Ser Cys Ser Cys Asn Ser
325 330 335

His Ser Ser Arg Cys His Phe Asp Met Thr Thr Tyr Leu Ala Ser Gly
10 340 345 350

Gly Leu Ser Gly Gly Val Cys Glu Asp Cys Gln His Asn Thr Glu Gly
355 360 365

15 Gln His Cys Asp Arg Cys Arg Pro Leu Phe Tyr Arg Asp Pro Leu Lys
370 375 380

Thr Ile Ser Asp Pro Tyr Ala Cys Ile Pro Cys Glu Cys Asp Pro Asp
385 390 395 400

20

Gly Thr Ile Ser Gly Gly Ile Cys Val Ser His Ser Asp Pro Ala Leu
405 410 415

Gly Ser Val Ala Gly Gln Cys Leu Cys Lys Glu Asn Val Glu Gly Ala
25 420 425 430

Lys Cys Asp Gln Cys Lys Pro Asn His Tyr Gly Leu Ser Ala Thr Asp
435 440 445

30 Pro Leu Gly Cys Gln Pro Cys Asp Cys Asn Pro Leu Gly Ser Leu Pro
450 455 460

Phe Leu Thr Cys Asp Val Asp Thr Gly Gln Cys Leu Cys Leu Ser Tyr
465 470 475 480

35

Val Thr Gly Ala His Cys Glu Glu Cys Thr Val Gly Tyr Trp Gly Leu
485 490 495

Gly Asn His Leu His Gly Cys Ser Pro Cys Asp Cys Asp Ile Gly Gly
40 500 505 510

Ala Tyr Ser Asn Val Cys Ser Pro Lys Asn Gly Gln Cys Glu Cys Arg
515 520 525

45 Pro His Val Thr Gly Arg Ser Cys Ser Glu Pro Ala Pro Gly Tyr Phe
530 535 540

Phe Ala Pro Leu Asn Phe Tyr Leu Tyr Glu Ala Glu Glu Ala Thr Thr
545 550 555 560

50

Leu Gln Gly Leu Ala Pro Leu Gly Ser Glu Thr Phe Gly Gln Ser Pro
565 570 575

Ala Val His Val Val Leu Gly Glu Pro Val Pro Gly Asn Pro Val Thr
55 580 585 590

Trp Thr Gly Pro Gly Phe Ala Arg Val Leu Pro Gly Ala Gly Leu Arg
595 600 605

60 Phe Ala Val Asn Asn Ile Pro Phe Pro Val Asp Phe Thr Ile Ala Ile
610 615 620

His Tyr Glu Thr Gln Ser Ala Ala Asp Trp Thr Val Gln Ile Val Val

5 625 630 635 640

Asn Pro Pro Gly Gly Ser Glu His Cys Ile Pro Lys Thr Leu Gln Ser
645 650 655

10 Lys Pro Gln Ser Phe Ala Leu Pro Ala Ala Thr Arg Ile Met Leu Leu
660 665 670

Pro Thr Pro Ile Cys Leu Glu Pro Asp Val Gln Tyr Ser Ile Asp Val
675 680 685

15

Tyr Phe Ser Gln Pro Leu Gln Gly Glu Ser His Ala His Ser His Val
690 695 700

Leu Val Asp Ser Leu Gly Leu Ile Pro Gln Ile Asn Ser Leu Glu Asn
20 705 710 715 720

Phe Cys Ser Lys Gln Asp Leu Asp Glu Tyr Gln Leu His Asn Cys Val
725 730 735

25 Glu Ile Ala Ser Ala Met Gly Pro Gln Val Leu Pro Gly Ala Cys Glu
740 745 750

Arg Leu Ile Ile Ser Met Ser Ala Lys Leu His Asp Gly Ala Val Ala
755 760 765

30

Cys Lys Cys His Pro Gln Gly Ser Val Gly Ser Ser Cys Ser Arg Leu
770 775 780

Gly Gly Gln Cys Gln Cys Lys Pro Leu Val Val Gly Arg Cys Cys Asp
35 785 790 795 800

Arg Cys Ser Thr Gly Ser Tyr Asp Leu Gly His His Gly Cys His Pro
805 810 815

40 Cys His Cys His Pro Gln Gly Ser Lys Asp Thr Val Cys Asp Gln Val
820 825 830

Thr Gly Gln Cys Pro Cys His Gly Glu Val Ser Gly Arg Arg Cys Asp
835 840 845

45

Arg Cys Leu Ala Gly Tyr Phe Gly Phe Pro Ser Cys His Pro Cys Pro
850 855 860

Cys Asn Arg Phe Ala Glu Leu Cys Asp Pro Glu Thr Gly Ser Cys Phe
50 865 870 875 880

Asn Cys Gly Gly Phe Thr Thr Gly Arg Asn Cys Glu Arg Cys Ile Asp
885 890 895

55 Gly Tyr Tyr Gly Asn Pro Ser Ser Gly Gln Pro Cys Arg Pro Cys Leu
900 905 910

Cys Pro Asp Asp Pro Ser Ser Asn Gln Tyr Phe Ala His Ser Cys Tyr
915 920 925

60

Gln Asn Leu Trp Ser Ser Asp Val Ile Cys Asn Cys Leu Gln Gly Tyr
930 935 940

5 Thr Gly Thr Gln Cys Gly Glu Cys Ser Thr Gly Phe Tyr Gly Asn Pro
945 950 955 960

Arg Ile Ser Gly Ala Pro Cys Gln Pro Cys Ala Cys Asn Asn Asn Ile
965 970 975

10

Asp Val Thr Asp Pro Glu Ser Cys Ser Arg Val Thr Gly Glu Cys Leu
980 985 990

Arg Cys Leu His Asn Thr Gln Gly Ala Asn Cys Gln Leu Cys Lys Pro
15 995 1000 1005

Gly His Tyr Gly Ser Ala Leu Asn Gln Thr Cys Arg Arg Cys Ser Cys
1010 1015 1020

20 His Ala Ser Gly Val Ser Pro Met Glu Cys Pro Pro Gly Gly Gly Ala
1025 1030 1035 1040

Cys Leu Cys Asp Pro Val Thr Gly Ala Cys Pro Cys Leu Pro Asn Val
1045 1050 1055

25

Thr Gly Leu Ala Cys Asp Arg Cys Ala Asp Gly Tyr Trp Asn Leu Val
1060 1065 1070

Pro Gly Arg Gly Cys Gln Ser Cys Asp Cys Asp Pro Arg Thr Ser Gln
30 1075 1080 1085

Ser Ser His Cys Asp Gln Leu Thr Gly Gln Cys Pro Cys Lys Leu Gly
1090 1095 1100

35 Tyr Gly Gly Lys Arg Cys Ser Glu Cys Gln Glu Asn Tyr Tyr Gly Asp
1105 1110 1115 1120

Pro Pro Gly Arg Cys Ile Pro Cys Asp Cys Asn Arg Ala Gly Thr Gln
1125 1130 1135

40

Lys Pro Ile Cys Asp Pro Asp Thr Gly Met Cys Arg Cys Arg Glu Gly
1140 1145 1150

Val Ser Gly Gln Arg Cys Asp Arg Cys Ala Arg Gly His Ser Gln Glu
45 1155 1160 1165

Phe Pro Thr Cys Leu Gln Cys His Leu Cys Phe Asp Gln Trp Asp His
1170 1175 1180

50 Thr Ile Ser Ser Leu Ser Lys Ala Val Gln Gly Leu Met Arg Leu Ala
1185 1190 1195 1200

Ala Asn Met Glu Asp Lys Arg Glu Thr Leu Pro Val Cys Glu Ala Asp
1205 1210 1215

55

Phe Lys Asp Leu Arg Gly Asn Val Ser Glu Ile Glu Arg Ile Leu Lys
1220 1225 1230

His Pro Val Phe Pro Ser Gly Lys Phe Leu Lys Val Lys Asp Tyr His
60 1235 1240 1245

Asp Ser Val Arg Arg Gln Ile Met Gln Leu Asn Glu Gln Leu Lys Ala
1250 1255 1260

5

```
Val Tyr Glu Phe Gln Asp Leu Lys Asp Thr Ile Glu Arg Ala Lys Asn
1265                1270                1275                1280

Glu Ala Asp Leu Leu Leu Glu Asp Leu Gln Glu Glu Ile Asp Leu Gln
                1285                1290                1295

Ser Ser Val Leu Asn Ala Ser Ile Ala Asp Ser Ser Glu Asn Ile Lys
            1300                1305                1310

Lys Tyr Tyr His Ile Ser Ser Ser Ala Glu Lys Lys Ile Asn Glu Thr
        1315                1320                1325

Ser Ser Thr Ile Asn Thr Ser Ala Asn Thr Arg Asn Asp Leu Leu Thr
    1330                1335                1340

Ile Leu Asp Thr Leu Thr Ser Lys Gly Asn Leu Ser Leu Glu Arg Leu
1345                1350                1355                1360

Lys Gln Ile Lys Ile Pro Asp Ile Gln Ile Leu Asn Glu Lys Val Cys
                1365                1370                1375

Gly Asp Pro Gly Asn Val Pro Cys Val Pro Leu Pro Cys Gly Gly Ala
            1380                1385                1390

Leu Cys Thr Gly Arg Lys Gly His Arg Lys Cys Arg Gly Pro Gly Cys
        1395                1400                1405

His Gly Ser Leu Thr Leu Ser Thr Asn Ala Leu Gln Lys Ala Gln Glu
    1410                1415                1420

Ala Lys Ser Ile Ile Arg Asn Leu Asp Lys Gln Val Arg Gly Leu Lys
1425                1430                1435                1440

Asn Gln Ile Glu Ser Ile Ser Glu Gln Ala Glu Val Ser Lys Asn Asn
                1445                1450                1455

Ala Leu Gln Leu Arg Glu Lys Leu Gly Asn Ile Arg Asn Gln Ser Asp
            1460                1465                1470

Ser Glu Glu Glu Asn Ile Asn Leu Phe Ile Lys Lys Val Lys Asn Phe
        1475                1480                1485

Leu Leu Glu Glu Asn Val Pro Pro Glu Asp Ile Glu Lys Val Ala Asn
    1490                1495                1500

Gly Val Leu Asp Ile His Leu Pro Ile Pro Ser Gln Asn Leu Thr Asp
1505                1510                1515                1520

Glu Leu Val Lys Ile Gln Lys His Met Gln Leu Cys Glu Asp Tyr Arg
                1525                1530                1535

Thr Asp Glu Asn Arg Ser Asn Glu Glu Ala Asp Gly Ala Gln Lys Leu
            1540                1545                1550

Leu Val Lys Ala Lys Ala Ala Glu Lys Ala Ala Asn Ile Leu Leu Asn
        1555                1560                1565

Leu Asp Lys Thr Leu Asn Gln Leu Gln Gln Ala Gln Ile Thr Gln Gly
```

10 15 20 25 30 35 40 45 50 55 60

```
        1570                1575                1580

Arg Ala Asn Ser Thr Ile Thr Gln Leu Thr Ala Asn Ile Thr Lys Ile
1585                1590                1595                1600

Lys Lys Asn Val Leu Gln Ala Glu Asn Gln Thr Arg Glu Met Lys Ser
                1605                1610                1615

Glu Leu Glu Leu Ala Lys Gln Arg Ser Gly Leu Glu Asp Gly Leu Ser
            1620                1625                1630

Leu Leu Gln Thr Lys Leu Gln Arg His Gln Asp His Ala Val Asn Ala
        1635                1640                1645

Lys Val Gln Ala Glu Ser Ala Gln His Gln Ala Gly Ser Leu Glu Lys
    1650                1655                1660

Glu Phe Val Glu Leu Lys Lys Gln Tyr Ala Ile Leu Gln Arg Lys Thr
1665                1670                1675                1680

Ser Thr Thr Gly Leu Thr Lys Glu Thr Leu Gly Lys Val Lys Gln Leu
                1685                1690                1695

Lys Asp Ala Ala Glu Lys Leu Ala Gly Asp Thr Glu Ala Lys Ile Arg
            1700                1705                1710

Arg Ile Thr Asp Leu Glu Arg Lys Ile Gln Asp Leu Asn Leu Ser Arg
        1715                1720                1725

Gln Ala Lys Ala Asp Gln Leu Arg Ile Leu Glu Asp Gln Val Val Ala
    1730                1735                1740

Ile Lys Asn Glu Ile Val Glu Gln Glu Lys Lys Tyr Ala Arg Cys Tyr
1745                1750                1755                1760

Ser
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 5874 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

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

```
ACATGCCCCG TTTGCTGCCT GAACCTCTCC ACAAAGACTC CCAGATCCTG AATTGAATTT      60

AATCATCTCC TGACAAAAGA ATGCAATTTC AACTGACCCT TTTTTTGCAC CTTGGGTGGC     120

TCAGTTACTC AAAAGCTCAA GATGACTGCA ACAGGGGTGC CTGTCATCCC ACCACTGGTG     180

ATCTCCTGGT GGGCAGGAAC ACGCAGCTTA TGGCTTCTTC TACCTGTGGG CTGAGCAGAG     240
```

5
CCCAGAAATA CTGCATCCTC AGTTACCTGG AGGGGGAACA AAAATGCTCC ATCTGTGACT 300

CTAGATTTCC ATATGATCCG TATGACCAAC CCAACAGCCA CACCATTGAG AATGTCACTG 360

10 TAAGTTTTGA ACCAGACAGA GAAAAGAAAT GGTGGCAATC TGAAAATGGT CTTGATCATG 420

TCAGCATCAG ACTGGACTTA GAGGCATTAT TTCGGTTCAG CCACCTTATC CTGACCTTTA 480

AGACTTTTCG GCCTGCTGCA ATGTTAGTTG AACGTTCCAC AGACTATGGA CACAACTGGA 540

15
AAGTGTTCAA ATATTTTGCA AAAGACTGTG CCACTTCCTT TCCTAACATC ACATCTGGCC 600

AGGCCCAGGG AGTGGGAGAC ATTGTTTGTG ACTCCAAATA CTCGGATATT GAACCCTCAA 660

20 CAGGTGGAGA GGTTGTTTTA AAAGTTTTGG ATCCCAGTTT TGAAATTGAA AACCCTTATA 720

GCCCCTACAT CCAAGACCTT GTGACATTGA CAAACCTGAG GATAAACTTT ACCAAGCTCC 780

ACACCCTTGG GGATGCTTTG CTTGGAAGGA GGCAAAATGA TTCCCTTGAT AAATACTACT 840

25
ATGCTCTGTA CGAGATGATT GTTCGGGGAA GCTGCTTTTG CAATGGCCAT GCTAGCGAAT 900

GTCGCCCTAT GCAGAAGATG CGGGGAGATG TTTTCAGCCC TCCTGGAATG GTTCACGGTC 960

30 AGTGTGTGTG TCAGCACAAT ACAGATGGTC CGAACTGTGA GAGATGCAAG GACTTCTTCC 1020

AGGATGCTCC TTGGAGGCCA GCTGCAGACC TCCAGGACAA CGCTTGCAGA TCGTGCAGCT 1080

GTAATAGCCA CTCCAGCCGC TGTCACTTTG ACATGACTAC GTACCTGGCA AGCGGTGGCC 1140

35
TCAGCGGGGG CGTGTGTGAA GACTGCCAGC ACAACACTGA GGGGCAGCAC TGCGACCGCT 1200

GCAGACCCCT CTTCTACAGG GACCCGCTCA AGACCATCTC AGATCCCTAC GCGTGCATTC 1260

40 CTTGTGAATG TGACCCCGAT GGGACCATAT CTGGTGGCAT TTGTGTGAGC CACTCTGATC 1320

CTGCCTTAGG GTCTGTGGCC GGCCAGTGCC TTTGTAAAGA GAACGTGGAA GGAGCCAAAT 1380

GCGACCAGTG CAAACCCAAC CACTACGGAC TAAGCGCCAC CGACCCCCTG GGCTGCCAGC 1440

45
CCTGCGACTG TAACCCCCTT GGGAGTCTGC CATTCTTGAC CTGTGATGTG GATACAGGCC 1500

AATGCTTGTG CCTGTCATAT GTCACCGGAG CACACTGCGA AGAATGCACT GTTGGATACT 1560

50 GGGGCCTGGG AAATCATCTC CATGGGTGTT CTCCCTGTGA CTGTGATATT GGAGGTGCTT 1620

ATTCTAACGT GTGCTCACCC AAGAATGGGC AGTGTGAATG CCGCCCACAT GTCACTGGCC 1680

GTAGCTGCTC TGAACCAGCC CCTGGCTACT TCTTTGCTCC TTTGAATTTC TATCTCTACG 1740

55
AGGCAGAGGA AGCCACAACA CTCCAAGGAC TGGCGCCTTT GGGCTCGGAG ACGTTTGGCC 1800

AGAGTCCTGC TGTTCACGTT GTTTTAGGAG AGCCAGTTCC TGGGAACCCT GTTACATGGA 1860

60 CTGGACCTGG ATTTGCCAGG GTTCTCCCTG GGGCTGGCTT GAGATTTGCT GTCAACAACA 1920

TTCCCTTTCC TGTGGACTTC ACCATTGCCA TTCACTATGA AACCCAGTCT GCAGCTGACT 1980

5 GGACTGTCCA GATTGTGGTG AACCCCCCTG GAGGGAGTGA GCACTGCATA CCCAAGACTC 2040

TACAGTCAAA GCCTCAGTCT TTTGCCTTAC CAGCGGCTAC GAGAATCATG CTGCTTCCCA 2100

CACCCATCTG TTTAGAACCA GATGTACAAT ATTCCATAGA TGTCTATTTT TCTCAGCCTT 2160

10 TGCAAGGAGA GTCCCACGCT CATTCACATG TCCTGGTGGA CTCTCTTGGC CTTATTCCCC 2220

AAATCAATTC ATTGGAGAAT TTCTGCAGCA AGCAGGACTT AGATGAGTAT CAGCTTCACA 2280

15 ACTGTGTTGA AATTGCCTCA GCAATGGGAC CTCAAGTGCT CCCGGGTGCC TGTGAAAGGC 2340

TGATCATCAG CATGTCTGCC AAGCTGCATG ATGGGGCTGT GGCCTGCAAG TGTCACCCCC 2400

AGGGCTCAGT CGGATCCAGC TGCAGCCGAC TTGGAGGCCA GTGCCAGTGT AAACCTCTTG 2460

20 TGGTCGGGCG CTGCTGTGAC AGGTGCTCAA CTGGAAGCTA TGATTTGGGG CATCACGGCT 2520

GTCACCCATG TCACTGCCAT CCTCAAGGAT CAAAGGACAC TGTATGTGAC CAAGTAACAG 2580

25 GACAGTGCCC CTGCCATGGA GAGGTGTCTG GCCGCCGCTG TGATCGCTGC CTGGCAGGCT 2640

ACTTTGGATT TCCCAGCTGC CACCCTTGCC CTTGTAATAG GTTTGCTGAA CTTTGTGATC 2700

CTGAGACAGG GTCATGCTTC AATTGTGGAG GCTTTACAAC TGGCAGAAAC TGTGAAAGGT 2760

30 GTATTGATGG TTACTATGGA AATCCTTCTT CAGGACAGCC CTGTCGTCCT TGCCTGTGTC 2820

CAGATGATCC CTCAAGCAAT CAGTATTTTG CCCATTCCTG TTATCAGAAT CTGTGGAGCT 2880

35 CAGATGTAAT CTGCAATTGT CTTCAAGGTT ATACGGGTAC TCAGTGTGGA GAATGCTCTA 2940

CTGGTTTCTA TGGAAATCCA AGAATTTCAG GAGCACCTTG CCAACCATGT GCCTGCAACA 3000

ACAACATAGA TGTAACCGAT CCAGAGTCCT GCAGCCGGGT AACAGGGGAG TGCCTTCGAT 3060

40 GTTTGCACAA CACTCAGGGC GCAAACTGCC AGCTCTGCAA ACCAGGTCAC TATGGATCAG 3120

CCCTCAATCA GACCTGCAGA AGATGCTCCT GCCATGCTTC CGGCGTGAGT CCCATGGAGT 3180

45 GTCCCCCTGG TGGGGGAGCT TGCCTCTGTG ACCCTGTCAC TGGTGCATGT CCTTGTCTGC 3240

CGAATGTCAC AGGCCTGGCC TGTGACCGTT GTGCTGATGG ATACTGGAAT CTGGTCCCTG 3300

GCAGAGGATG TCAGTCATGT GACTGTGACC CTAGGACCTC TCAAAGTAGC CACTGTGACC 3360

50 AGCTTACAGG CCAGTGTCCG TGTAAATTAG GTTACGGCGG GAAACGTTGC AGTGAGTGCC 3420

AGGAAAATTA TTATGGTGAT CCACCTGGGC GATGCATTCC ATGTGATTGT AACAGGGCAG 3480

55 GTACCCAGAA GCCCATCTGT GATCCAGACA CAGGCATGTG CCGCTGCCGG GAGGGTGTCA 3540

GCGGCCAGAG ATGTGATCGC TGTGCCCGGG GACACAGCCA GGAATTCCCT ACTTGTCTTC 3600

AATGTCACTT GTGCTTTGAT CAATGGGACC ACACCATTTC TTCCCTCTCC AAAGCGGTGC 3660

60 AAGGGTTAAT GAGACTGGCT GCTAACATGG AAGATAAAAG AGAGACCCTG CCTGTCTGTG 3720

AGGCAGACTT CAAAGACCTC AGAGGGAACG TGTCTGAAAT AGAAAGGATT TTGAAACATC 3780

```
5
CTGTTTTCCC ATCTGGGAAA TTCTTAAAAG TCAAGGATTA TCATGACTCT GTTAGAAGAC 3840
AAATCATGCA GCTAAATGAA CAACTGAAAG CAGTGTATGA ATTTCAAGAT CTGAAAGATA 3900
10 CAATAGAAAG AGCAAAGAAT GAAGCAGACC TCTTACTTGA AGACCTTCAG GAAGAAATTG 3960
ATTTGCAATC CAGTGTCCTT AATGCAAGCA TTGCGGACTC CTCAGAAAAC ATCAAGAAAT 4020
ATTATCACAT ATCATCATCT GCTGAAAAGA AAATTAATGA AACTAGTTCC ACCATTAATA 4080
15
CCTCTGCAAA TACAAGGAAT GACTTACTTA CCATCTTAGA TACACTAACC TCAAAAGGAA 4140
ACTTGTCATT GGAAAGATTA AAGCAGATTA AGATACCAGA TATCCAAATA TTGAATGAAA 4200
20 AGGTGTGCGG AGATCCAGGA AATGTGCCAT GTGTGCCCTT GCCCTGTGGC GGTGCTCTCT 4260
GCACGGGCCG GAAGGGGCAC AGGAAGTGTA GGGGTCCCGG CTGTCACGGC TCCCTGACCC 4320
TCTCAACGAA TGCCCTCCAA AAAGCCCAGG AAGCAAAATC CATTATTCGT AATTTGGACA 4380
25
AACAGGTTCG TGGGTTGAAA AATCAGATCG AAAGTATAAG TGAACAGGCA GAAGTCTCCA 4440
AAAACAATGC CTTACAGCTG AGGGAAAAAC TGGGAAATAT AAGAAACCAA AGTGACTCTG 4500
30 AAGAAGAAAA CATCAATCTT TTCATCAAAA AAGTGAAAAA CTTTTTGTTA GAGGAAAACG 4560
TGCCTCCAGA AGACATCGAG AAGGTTGCGA ATGGTGTGCT TGACATTCAC CTACCAATTC 4620
CATCCCAAAA TCTAACCGAT GAACTTGTCA AAATACAGAA ACATATGCAA CTCTGTGAGG 4680
35
ATTACAGGAC AGATGAAAAC AGGTCAAATG AAGAAGCAGA TGGAGCCCAA AAGCTTTTGG 4740
TGAAGGCCAA AGCAGCTGAG AAAGCAGCAA ATATTCTATT AAATCTTGAC AAAACATTGA 4800
40 ACCAGTTACA ACAAGCTCAA ATCACTCAAG GACGGGCAAA CTCTACCATT ACACAGCTGA 4860
CTGCCAATAT AACAAAAATA AAAAAGAATG TGCTGCAGGC TGAAAATCAA ACCAGGGAAA 4920
TGAAGAGTGA GCTGGAGTTA GCAAAGCAGC GATCAGGGCT GGAGGATGGA CTTTCCCTGC 4980
45
TGCAGACCAA GTTGCAAAGG CATCAAGACC ACGCTGTCAA TGCGAAAGTT CAGGCTGAAT 5040
CTGCCCAACA CCAGGCTGGG AGTCTTGAGA AGGAATTTGT TGAGCTGAAA AAACAATATG 5100
50 CTATTCTCCA ACGTAAGACA AGCACTACAG GACTAACAAA GGAGACATTA GGAAAAGTTA 5160
AACAGCTAAA AGATGCGGCA GAAAAATTGG CTGGAGATAC AGAGGCCAAG ATAAGAAGAA 5220
TAACAGATTT AGAAAGGAAA ATCCAAGATT TGAATCTAAG TAGACAAGCA AAAGCTGATC 5280
55
AACTGAGAAT ATTGGAAGAT CAAGTTGTTG CCATTAAAAA TGAAATTGTT GAACAAGAAA 5340
AAAAATATGC TAGGTGCTAT AGCTAGGCAG AGTTAAAGAG CAAAAGCTTG TGCCTTTGTT 5400
60 TCTGGTTTCT GATGTACAAG CCCCTGGGGC TCTGTTGAAC CTGTGAAATA CTGACAATGT 5460
CTTCTACCTT CCTTCCCCAC ACCCTGTCCT TATTAGACAC CTGCTCAGTG TGGCTGGAGG 5520
```

5 TTGAAATGCC ACCAGGAAAA TGCCACTTCA TAATTGAAAG GGGAAAGTAA TGAAATTGTC 5580

TCTGGTTTCA GAAACTTTTC CTCTTACCTT CCTTTCTCTT TCCTAACTTA AAAATAACAG 5640

TTTCCATATA ACAAGTAGAA ATTTAAGTAA GTACTCTACT AACTAATAAT CATTTCAGTC 5700

10 AGATAAACCT AAACATTAAA TAAATATCTC CAATATTAGG ATGGAATACA TATGTATGGC 5760

ATGTACTAGA TTGTCCTATA TTTTATGTTT ATTTGGATTT GCTTTTATTT GTAAAATTAT 5820

15 TCTTTTCTGA ATAAACTGCA TACAATTCAA AATGGAAAAA AAAAAAAAAA AAAA 5874

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
20 (A) LENGTH: 1524 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
25
(v) FRAGMENT TYPE: internal

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

Ala Ala Gly Ala Gly Ala His Cys Gln Arg Cys Asp Ala Ala Asp Pro
1 5 10 15

35 Gln Arg His His Asn Ala Ser Tyr Leu Thr Asp Phe His Ser Gln Asp
20 25 30

Glu Ser Thr Trp Trp Gln Ser Pro Ser Met Ala Phe Gly Val Gln Tyr
35 40 45

40
Pro Thr Ser Val Asn Ile Thr Leu Xaa Arg Leu Gly Lys Ala Tyr Glu
50 55 60

Ile Thr Tyr Val Arg Leu Lys Phe His Thr Ser Arg Pro Glu Ser Phe
45 65 70 75 80

Ala Ile Tyr Lys Arg Ser Arg Ala Asp Gly Pro Trp Glu Pro Tyr Gln
85 90 95

50 Phe Tyr Ser Ala Ser Cys Gln Lys Thr Tyr Gly Arg Pro Glu Gly Gln
100 105 110

Tyr Leu Arg Pro Gly Glu Asp Glu Arg Val Ala Phe Cys Thr Ser Glu
115 120 125
55

Phe Ser Asp Ile Ser Pro Leu Ser Gly Gly Asn Val Ala Phe Ser Thr
130 135 140

Leu Glu Gly Arg Pro Ser Ala Tyr Asn Phe Glu Glu Ser Pro Gly Leu
60 145 150 155 160

Gln Glu Trp Val Thr Ser Thr Glu Leu Leu Ile Ser Leu Asp Arg Leu
165 170 175

```
5
        Asn Thr Phe Gly Asp Asp Ile Phe Lys Asp Pro Lys Val Leu Gln Ser
                    180                 185                 190

        Tyr Tyr Tyr Ala Val Ser Asp Phe Ser Val Gly Gly Arg Cys Lys Cys
10              195                 200                 205

        Asn Gly His Ala Ser Glu Cys Gly Pro Asp Val Ala Gly Gln Leu Ala
            210                 215                 220

15      Cys Arg Cys Gln His Asn Thr Thr Gly Thr Asp Cys Glu Arg Cys Leu
        225                 230                 235                 240

        Pro Phe Phe Gln Asp Arg Pro Trp Ala Arg Gly Thr Ala Glu Ala Ala
                        245                 250                 255
20
        His Glu Cys Leu Pro Cys Asn Cys Ser Gly Arg Ser Glu Glu Cys Thr
                    260                 265                 270

        Phe Asp Arg Glu Leu Phe Arg Ser Thr Gly His Gly Gly Arg Cys His
25              275                 280                 285

        His Cys Arg Asp His Thr Ala Gly Pro His Cys Glu Arg Cys Gln Glu
            290                 295                 300

30      Asn Phe Tyr His Trp Asp Pro Arg Met Pro Cys Gln Pro Cys Asp Cys
        305                 310                 315                 320

        Gln Ser Ala Gly Ser Leu His Leu Gln Cys Asp Asp Thr Gly Thr Cys
                        325                 330                 335
35
        Ala Cys Lys Pro Thr Val Thr Gly Trp Lys Cys Asp Arg Cys Leu Pro
                    340                 345                 350

        Gly Phe His Ser Leu Ser Glu Gly Gly Cys Arg Pro Cys Thr Cys Asn
40              355                 360                 365

        Pro Ala Gly Ser Leu Asp Thr Cys Asp Pro Arg Ser Gly Arg Cys Pro
            370                 375                 380

45      Cys Lys Glu Asn Val Glu Gly Asn Leu Cys Asp Arg Cys Arg Pro Gly
        385                 390                 395                 400

        Thr Phe Asn Leu Gln Pro His Asn Pro Ala Gly Cys Ser Ser Cys Phe
                        405                 410                 415
50
        Cys Tyr Gly His Ser Lys Val Cys Ala Ser Thr Ala Gln Phe Gln Val
                    420                 425                 430

        His His Ile Leu Ser Asp Phe His Gln Gly Ala Glu Gly Trp Trp Ala
55              435                 440                 445

        Arg Ser Val Gly Gly Ser Glu His Ser Pro Gln Trp Ser Pro Asn Gly
            450                 455                 460

60      Val Leu Leu Ser Pro Glu Asp Glu Glu Glu Leu Thr Ala Pro Gly Lys
        465                 470                 475                 480

        Phe Leu Gly Asp Gln Arg Phe Ser Tyr Gly Gln Pro Leu Ile Leu Thr
```

```
5                     485                     490                     495

          Phe Arg Val Pro Pro Gly Asp Ser Pro Leu Pro Val Gln Leu Arg Leu
                      500                 505                 510

10        Glu Gly Thr Gly Leu Ala Leu Ser Leu Arg His Ser Ser Leu Ser Gly
                  515                 520                 525

          Pro Gln Asp Ala Arg Ala Ser Gln Gly Gly Arg Ala Gln Val Pro Leu
              530                 535                 540
15
          Gln Glu Thr Ser Glu Asp Val Ala Pro Pro Leu Pro Pro Phe His Phe
          545                 550                 555                 560

          Gln Arg Leu Leu Ala Asn Leu Thr Ser Leu Arg Leu Arg Val Ser Pro
20                            565                 570                 575

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

25        Ser Ala Arg Pro Gly Leu Ser Pro Pro Ala Ser Trp Val Glu Ile Cys
                  595                 600                 605

          Ser Cys Pro Thr Gly Tyr Thr Gly Gln Phe Cys Glu Ser Cys Ala Pro
              610                 615                 620
30
          Gly Tyr Lys Arg Glu Met Pro Gln Gly Gly Pro Tyr Ala Ser Cys Val
          625                 630                 635                 640

          Pro Cys Thr Cys Asn Gln His Gly Thr Cys Asp Pro Asn Thr Gly Ile
35                            645                 650                 655

          Cys Val Cys Ser His His Thr Glu Gly Pro Ser Cys Glu Arg Cys Leu
                      660                 665                 670

40        Pro Gly Phe Tyr Gly Asn Pro Phe Ala Gly Gln Ala Asp Asp Cys Gln
                  675                 680                 685

          Pro Cys Pro Cys Pro Gly Gln Ser Ala Cys Thr Thr Ile Pro Glu Ser
              690                 695                 700
45
          Gly Glu Val Val Cys Thr His Cys Pro Pro Gly Gln Arg Gly Arg Arg
          705                 710                 715                 720

          Cys Glu Val Cys Asp Asp Gly Phe Phe Gly Asp Pro Leu Gly Leu Phe
50                            725                 730                 735

          Gly His Pro Gln Pro Cys His Gln Cys Gln Cys Ser Gly Asn Val Asp
                      740                 745                 750

55        Pro Asn Ala Val Gly Asn Cys Asp Pro Leu Ser Gly His Cys Leu Arg
                  755                 760                 765

          Cys Leu His Asn Thr Thr Gly Asp His Cys Glu His Cys Gln Glu Gly
              770                 775                 780
60
          Phe Tyr Gly Ser Ala Leu Ala Pro Arg Pro Ala Asp Lys Cys Met Pro
          785                 790                 795                 800
```

Cys Ser Cys His Pro Gln Gly Ser Val Ser Glu Gln Met Pro Cys Asp
805 810 815

Pro Val Thr Gly Gln Cys Ser Cys Leu Pro His Val Thr Ala Arg Asp
820 825 830

Cys Ser Arg Cys Tyr Pro Gly Phe Phe Asp Leu Gln Pro Gly Arg Gly
835 840 845

Cys Arg Ser Cys Lys Cys His Pro Leu Gly Ser Gln Glu Asp Gln Cys
850 855 860

His Pro Lys Thr Gly Gln Cys Thr Cys Arg Pro Gly Val Thr Gly Gln
865 870 875 880

Ala Cys Asp Arg Cys Gln Leu Gly Phe Phe Gly Ser Ser Ile Lys Gly
885 890 895

Cys Arg Ala Cys Arg Cys Ser Pro Leu Gly Ala Ala Ser Ala Gln Cys
900 905 910

His Tyr Asn Gly Thr Cys Val Cys Arg Pro Gly Phe Glu Gly Tyr Lys
915 920 925

Cys Asp Arg Cys His Tyr Asn Phe Phe Leu Thr Ala Asp Gly Thr His
930 935 940

Cys Gln Gln Cys Pro Ser Cys Tyr Ala Leu Val Lys Glu Glu Xaa Ala
945 950 955 960

Lys Leu Lys Ala Arg Leu Thr Leu Thr Glu Gly Trp Leu Gln Gly Ser
965 970 975

Asp Cys Gly Ser Pro Trp Gly Pro Leu Asp Ile Leu Leu Gly Glu Ala
980 985 990

Pro Arg Xaa Asp Val Tyr Gln Gly His His Leu Leu Pro Gly Ala Arg
995 1000 1005

Glu Ala Phe Leu Glu Gln Met Met Gly Leu Glu Gly Ala Val Lys Ala
1010 1015 1020

Ala Arg Glu Gln Leu Gln Arg Leu Asn Lys Gly Ala Arg Cys Ala Gln
1025 1030 1035 1040

Ala Gly Ser Gln Lys Thr Cys Thr Gln Leu Ala Asp Leu Glu Ala Val
1045 1050 1055

Leu Glu Ser Ser Glu Glu Glu Ile Leu His Ala Ala Ala Ile Leu Ala
1060 1065 1070

Ser Leu Glu Ile Pro Gln Glu Gly Pro Ser Gln Pro Thr Lys Trp Ser
1075 1080 1085

His Leu Ala Ile Glu Ala Arg Ala Leu Ala Arg Ser His Arg Asp Thr
1090 1095 1100

Ala Thr Lys Ile Ala Ala Thr Ala Trp Arg Ala Leu Leu Ala Ser Asn
1105 1110 1115 1120

5

Thr Ser Tyr Ala Leu Leu Trp Asn Leu Leu Glu Gly Arg Val Ala Leu
1125 1130 1135

Glu Thr Gln Arg Asp Leu Glu Asp Arg Tyr Gln Glu Val Gln Ala Ala
1140 1145 1150

10

Gln Lys Ala Leu Arg Thr Ala Val Ala Glu Val Leu Pro Glu Ala Xaa
1155 1160 1165

15

Lys Arg Val Gly His Arg Ala Ala Ser Trp Arg Arg Tyr Ser Pro Val
1170 1175 1180

Pro Gly Leu Ala Gly Phe Pro Gly Ser Ser Ala Ser Xaa Lys Ser Arg
1185 1190 1195 1200

20

Ala Glu Asp Leu Gly Leu Lys Ala Lys Ala Leu Glu Lys Thr Val Ala
1205 1210 1215

Ser Trp Gln His Met Ala Thr Glu Ala Ala Arg Thr Leu Gln Thr Ala
1220 1225 1230

25

Ala Gln Ala Thr Leu Arg Gln Thr Glu Pro Leu Thr Met Ala Arg Ser
1235 1240 1245

30

Arg Leu Thr Ala Thr Phe Ala Ser Gln Leu His Gln Gly Ala Arg Ala
1250 1255 1260

Ala Leu Thr Gln Ala Ser Ser Ser Val Gln Ala Ala Thr Val Thr Val
1265 1270 1275 1280

35

Met Gly Ala Arg Thr Leu Leu Ala Asp Leu Glu Gly Met Lys Leu Gln
1285 1290 1295

Phe Pro Arg Pro Lys Asp Gln Ala Ala Leu Gln Arg Lys Ala Asp Ser
1300 1305 1310

40

Val Ser Asp Arg Leu Leu Ala Asp Thr Arg Lys Lys Thr Lys Gln Ala
1315 1320 1325

45

Glu Arg Met Leu Gly Asn Ala Ala Pro Leu Ser Ser Ser Ala Lys Lys
1330 1335 1340

Lys Gly Arg Glu Ala Glu Val Leu Ala Lys Asp Ser Ala Lys Leu Ala
1345 1350 1355 1360

50

Lys Ala Leu Leu Arg Glu Arg Lys Gln Ala His Arg Arg Ala Ser Arg
1365 1370 1375

Leu Thr Ser Gln Xaa Leu Gln Ala Thr Leu Gln Gln Ala Ser Gln Gln
1380 1385 1390

55

Val Leu Ala Ser Glu Ala Arg Arg Gln Glu Leu Glu Glu Ala Glu Arg
1395 1400 1405

60

Val Gly Ala Gly Leu Ser Glu Met Glu Gln Gln Ile Arg Glu Ser Arg
1410 1415 1420

Ile Ser Leu Glu Lys Asp Ile Glu Thr Leu Ser Glu Leu Leu Ala Arg

5 1425 1430 1435 1440

Leu Gly Ser Leu Asp Thr His Gln Ala Pro Ala Gln Ala Leu Asn Glu
1445 1450 1455

10 Thr Gln Trp Ala Leu Glu Arg Leu Arg Leu Gln Leu Gly Ser Pro Gly
1460 1465 1470

Ser Leu Gln Arg Lys Leu Ser Leu Leu Glu Gln Glu Ser Gln Gln Gln
1475 1480 1485

15

Glu Leu Gln Ile Gln Gly Phe Glu Ser Asp Leu Ala Glu Ile Arg Ala
1490 1495 1500

Asp Lys Gln Asn Leu Glu Ala Ile Leu His Ser Leu Pro Glu Asn Cys
20 1505 1510 1515 1520

Ala Ser Trp Gln

25 (2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 4890 base pairs
(B) TYPE: nucleic acid
30 (C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

35

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

GCCGCGGGCG CGGGGGCTCA TTGCCAGCGC TGCGACGCCG CCGACCCCCA GCGCCACCAC 60
40
AACGCCTCCT ACCTCACCGA CTTCCACAGC CAGGACGAGA GCACCTGGTG GCAGAGCCCG 120

TCCATGGCCT TCGGCGTGCA GTACCCCACC TCGGTCAACA TCACCCTCCC GCCTAGGGAA 180

45 GGCTTATGAG ATCACGTATG TGAGGCTGAA GTTCCACACC AGTCGCCCTG AGAGCTTTGC 240

CATCTACAAG CGCAGCCGCG CCGACGGCCC ATGGGAGCCC TACCAGTTCT ACAGCGCCTC 300

CTGCCAGAAG ACCTACGGCC GGCCCGAGGG CCAGTACCTG CGCCCCGGCG AGGACGAGCG 360
50
CGTGGCCTTC TGCACCTCTG AGTTCAGCGA CATCTCCCCG CTGAGTGGCG GCAACGTGGC 420

CTTCTCCACC CTGGAGGGCC GGCCCAGCGC CTACAACTTC GAGGAGAGCC CTGGGCTGCA 480

55 GGAGTGGGTC ACCAGCACCG AACTCCTCAT CTCTCTAGAC CGGCTCAACA CGTTTGGGGA 540

CGACATCTTC AAGGACCCCA AGGTGCTCCA GTCCTACTAT TATGCCGTGT CCGACTTCTC 600

TGTGGGCGGC AGGTGCAAGT GCAACGGGCA TGCCAGCGAG TGCGGCCCCG ACGTGGCAGG 660
60
CCAGTTGGCC TGCCGGTGCC AGCACAACAC CACCGGCACA GACTGTGAGC GCTGCCTGCC 720

CTTCTTCCAG GACCGCCCGT GGGCCCGGGG CACCGCCGAG GCTGCCCACG AGTGTCTGCC 780

5
CTGCAACTGC AGTGGCCGCT CCGAGGAATG CACGTTTGAT CGGGAGCTCT TCCGCAGCAC 840
AGGCCACGGC GGGCGCTGTC ACCACTGCCG TGACCACACA GCTGGGCCAC ACTGTGAGCG 900
10 CTGTCAGGAG AATTTCTATC ACTGGGACCC GCGGATGCCA TGCCAGCCCT GTGACTGCCA 960
GTCGGCAGGC TCCCTACACC TCCAGTGCGA TGACACAGGC ACCTGCGCCT GCAAGCCCAC 1020
AGTGACTGGC TGGAAGTGTG ACCGCTGTCT GCCCGGGTTC CACTCGCTCA GTGAGGGAGG 1080
15
CTGCAGACCC TGCACTTGCA ATCCCGCTGG CAGCCTGGAC ACCTGTGACC CCCGCAGTGG 1140
GCGCTGCCCC TGCAAAGAGA ATGTGGAAGG CAACCTATGT GACAGATGTC GCCCGGGGAC 1200
20 CTTTAACCTG CAGCCCCACA ATCCAGCTGG CTGCAGCAGC TGTTTCTGCT ATGGCCACTC 1260
CAAGGTGTGC GCGTCCACTG CCCAGTTCCA GGTGCATCAC ATCCTCAGCG ATTTCCACCA 1320
GGGAGCCGAA GGCTGGTGGG CCAGAAGTGT GGGGGGCTCT GAGCACTCCC CACAATGGAG 1380
25
CCCAAATGGG GTCCTCCTGA GCCCAGAAGA CGAGGAGGAG CTCACAGCAC CAGGGAAGTT 1440
CCTGGGAGAC CAGCGGTTCA GCTATGGGCA GCCCCTCATA CTGACCTTCC GGGTGCCCCC 1500
30 CGGGGACTCC CCACTCCCTG TACAGCTGAG GCTGGAAGGG ACAGGCTTGG CCCTGTCCCT 1560
GAGGCACTCT AGCCTGTCTG GCCCCCAGGA TGCCAGGGCA TCCCAGGGAG GTAGAGCTCA 1620
GGTTCCACTG CAGGAGACCT CCGAGGACGT GGCCCCTCCA CTGCCCCCCT TCCACTTCCA 1680
35
GCGGCTCCTC GCCAACCTGA CCAGCCTCCG CCTCCGCGTC AGTCCCGGCC CCAGCCCTGC 1740
CGGTCCAGTG TTCCTGACTG AGGTCCGGCT CACATCCGCC CGGCCAGGGC TTTCCCCGCC 1800
40 AGCCTCCTGG GTGGAGATTT GTTCATGTCC CACTGGCTAC ACGGGCCAGT TCTGTGAATC 1860
CTGTGCTCCG GGATACAAGA GGGAGATGCC ACAGGGGGGT CCCTATGCCA GCTGTGTCCC 1920
CTGCACCTGT AACCAGCATG GCACCTGTGA CCCCAACACA GGGATCTGTG TCTGCAGCCA 1980
45
CCATACCGAG GGCCCATCCT GTGAACGCTG TTTGCCAGGT TTCTATGGCA ACCCTTTCGC 2040
GGGCCAAGCC GACGACTGCC AGCCCTGTCC CTGCCCTGGC CAGTCGGCCT GTACGACCAT 2100
50 CCCAGAGAGC GGGGAGGTGG TGTGTACCCA CTGCCCCCCG GGCCAGAGAG GGCGGCGCTG 2160
TGAGGTCTGT GATGATGGCT TTTTTGGGGA CCCGCTGGGG CTCTTTGGGC ACCCCCAGCC 2220
CTGCCACCAG TGCCAGTGTA GCGGGAACGT GGACCCCAAT GCCGTGGGCA ACTGTGACCC 2280
55
CCTGTCTGGC CACTGCCTGC GCTGCCTGCA CAACACCACG GGTGACCACT GTGAGCACTG 2340
TCAGGAAGGC TTCTACGGGA GCGCCCTGGC CCCTCGACCC GCAGACAAAT GCATGCCTTG 2400
60 CAGCTGTCAC CCACAGGGCT CGGTCAGTGA GCAGATGCCC TGCGACCCAG TGACAGGCCA 2460
ATGCTCCTGC CTGCCTCATG TGACTGCACG GGACTGCAGC CGCTGCTACC CTGGCTTCTT 2520

```
5   CGACCTCCAG CCTGGGAGGG GCTGCCGGAG CTGCAAGTGT CACCCACTGG GCTCCCAGGA   2580
    GGACCAGTGC CATCCCAAGA CTGGACAGTG CACCTGCCGC CCAGGTGTCA CAGGCCAGGC   2640
    CTGTGACAGG TGCCAGCTGG GTTTCTTCGG CTCCTCAATC AAGGGCTGCC GGGCCTGCAG   2700
10
    GTGCTCCCCA CTGGGCGCTG CCTCGGCCCA GTGCCACTAT AACGGCACAT GCGTGTGCAG   2760
    GCCTGGCTTC GAGGGCTACA AATGTGACCG CTGCCACTAC AACTTCTTCC TCACGGCAGA   2820
15  CGGCACACAC TGCCAGCAAT GTCCGTCCTG CTACGCCCTG GTGAAGGAGG AGCAGCCAAG   2880
    CTGAAGGCCA GACTGACTTT GACGGAGGGG TGGCTCCAAG GGTCCGACTG TGGCAGTCCC   2940
    TGGGGACCAC TAGACATTCT GCTGGGAGAG GCCCCAAGGG GGACGTCTAC CAGGGCCATC   3000
20
    ACCTGCTTCC AGGGGCTCGG GAAGCCTTCC TGGAGCAGAT GATGGGCCTC GAGGGTGCTG   3060
    TCAAGGCCGC CCGGGAGCAG CTGCAGAGGC TGAACAAGGG TGCCCGCTGT GCCCAGGCCG   3120
25  GATCCCAGAA GACCTGCACC CAGCTGGCAG ACCTGGAGGC AGTGCTGGAG TCCTCGGAAG   3180
    AGGAGATTCT GCATGCAGCT GCCATTCTCG CGTCTCTGGA GATTCCTCAG GAAGGTCCCA   3240
    GTCAGCCGAC CAAATGGAGC CACCTGGCCA TAGAGGCCCG TGCCCTCGCC AGGAGCCACA   3300
30
    GAGACACCGC CACCAAGATC GCAGCCACTG CTTGGAGGGC CCTGCTCGCC TCCAACACCA   3360
    GCTACGCGCT TCTCTGGAAT CTGCTGGAGG GAAGGGTGGC CCTAGAGACC CAGCGGGACC   3420
35  TGGAGGACAG GTACCAGGAG GTCCAGGCGG CCCAGAAAGC ACTGAGGACG GCTGTGGCAG   3480
    AGGTGCTGCC TGAAGCGGAA AGCGTGTTGG CCACCGTGCA GCAAGTTGGC GCAGATACAG   3540
    CCCCGTACCT GGCCTTGCTG GCTTCCCCGG GAGCTCTGCC TCAGAAGTCC CGGGCTGAAG   3600
40
    ACCTGGGCCT GAAGGCGAAG GCCCTGGAGA AGACAGTTGC ATCATGGCAG CACATGGCCA   3660
    CTGAGGCTGC CCGAACCCTC CAGACTGCTG CCCAGGCGAC GCTACGGCAA ACAGAACCCC   3720
45  TCACAATGGC GCGATCTCGG CTCACTGCAA CCTTTGCCTC CCAGCTGCAC CAGGGGGCCA   3780
    GAGCCGCCCT GACCCAGGCT TCCTCATCTG TCCAGGCTGC GACAGTGACT GTCATGGGAG   3840
    CCAGGACTCT GCTGGCTGAT CTGGAAGGAA TGAAGCTGCA GTTTCCCCGG CCCAAGGACC   3900
50
    AGGCGGCATT GCAGAGGAAG GCAGACTCCG TCAGTGACAG ACTCCTTGCA GACACGAGAA   3960
    AGAAGACCAA GCAGGCGGAG AGGATGCTGG GAAACGCGGC CCCTCTTTCC TCCAGTGCCA   4020
55  AGAAGAAGGG CAGAGAAGCA GAGGTGTTGG CCAAGGACAG TGCCAAGCTT GCCAAGGCCT   4080
    TGCTGAGGGA GCGGAAACAG GCGCACCGCC GTGCCAGCAG GCTCACCAGC CAGACTGCAA   4140
    GCCACGCTCC AACAGGCGTC CCAGCAGGTG CTGGCGTCTG AAGCACGCAG ACAGGAGCTG   4200
60
    GAGGAAGCTG AGCGGGTGGG TGCTGGGCTG AGCGAGATGG AGCAGCAGAT CCGGGAATCG   4260
    CGTATCTCAC TGGAGAAGGA CATCGAGACC TTGTCAGAGC TGCTTGCCAG GCTGGGGTCG   4320
```

5

CTGGACACCC ATCAAGCCCC AGCCCAGGCC CTGAACGAGA CTCAGTGGGC ACTAGAACGC 4380

CTGAGGCTGC AGCTGGGCTC CCCGGGGTCC TTGCAGAGGA AACTCAGTCT GCTGGAGCAG 4440

10 GAATCCCAGC AGCAGGAGCT GCAGATCCAG GGCTTCGAGA GTGACCTCGC CGAGATCCGC 4500

GCCGACAAAC AGAACCTGGA GGCCATTCTG CACAGCCTGC CCGAGAACTG TGCCAGCTGG 4560

CAGTGAGGGC TGCCCAGATC CCCGGCACAC ACTCCCCCAC CTGCTGTTTA CATGACCCAG 4620

15

GGGGTGCACA CTACCCCACA GGTGTGCCCA TACAGACATT CCCCGGAGCC GGCTGCTGTG 4680

AACTCGACCC CGTGTGGATA GTCACACTCC CTGCCGATTC TGTCTGTGGC TTCTTCCCTG 4740

20 CCAGCAGGAC TGAGTGTGCG TACCCAGTTC ACCTGGACAT GAGTGCACAC TCTCACCCCT 4800

GCACATGCAT AAACGGGCAC ACCCCAGTGT CAATAACATA CACACGTGAG GGTGCATGTC 4860

TGTGTGTATG ACCCAAATAA AAAAAAAAAA 4890

25

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1105 amino acids
30 (B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

35 (v) FRAGMENT TYPE: internal

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

40

Met Gln Phe Gln Leu Thr Leu Phe Leu His Leu Gly Trp Leu Ser Tyr
1 5 10 15

45 Ser Lys Ala Gln Asp Asp Cys Asn Arg Gly Ala Cys His Pro Thr Thr
20 25 30

Gly Asp Leu Leu Val Gly Arg Asn Thr Gln Leu Met Ala Ser Ser Thr
35 40 45

50 Cys Gly Leu Ser Arg Ala Gln Lys Tyr Cys Ile Leu Ser Tyr Leu Glu
50 55 60

Gly Glu Gln Lys Cys Ser Ile Cys Asp Ser Arg Phe Pro Tyr Asp Pro
65 70 75 80

55

Tyr Asp Gln Pro Asn Ser His Thr Ile Glu Asn Val Thr Val Ser Phe
85 90 95

60 Glu Pro Asp Arg Glu Lys Lys Trp Trp Gln Ser Glu Asn Gly Leu Asp
100 105 110

His Val Ser Ile Arg Leu Asp Leu Glu Ala Leu Phe Arg Phe Ser His
115 120 125

5

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

10

Arg Ser Thr Asp Tyr Gly His Asn Trp Lys Val Phe Lys Tyr Phe Ala
145 150 155 160

Lys Asp Cys Ala Thr Ser Phe Pro Asn Ile Thr Ser Gly Gln Ala Gln
165 170 175

15

Gly Val Gly Asp Ile Val Cys Asp Ser Lys Tyr Ser Asp Ile Glu Pro
180 185 190

Ser Thr Gly Gly Glu Val Val Leu Lys Val Leu Asp Pro Ser Phe Glu
195 200 205

20

Ile Glu Asn Pro Tyr Ser Pro Tyr Ile Gln Asp Leu Val Thr Leu Thr
210 215 220

25

Asn Leu Arg Ile Asn Phe Thr Lys Leu His Thr Leu Gly Asp Ala Leu
225 230 235 240

Leu Gly Arg Arg Gln Asn Asp Ser Leu Asp Lys Tyr Tyr Tyr Ala Leu
245 250 255

30

Tyr Glu Met Ile Val Arg Gly Ser Cys Phe Cys Asn Gly His Ala Ser
260 265 270

Glu Cys Arg Pro Met Gln Lys Met Arg Gly Asp Val Phe Ser Pro Pro
275 280 285

35

Gly Met Val His Gly Gln Cys Val Cys Gln His Asn Thr Asp Gly Pro
290 295 300

40

Asn Cys Glu Arg Cys Lys Asp Phe Phe Gln Asp Ala Pro Trp Arg Pro
305 310 315 320

Ala Ala Asp Leu Gln Asp Asn Ala Cys Arg Ser Cys Ser Cys Asn Ser
325 330 335

45

His Ser Ser Arg Cys His Phe Asp Met Thr Thr Tyr Leu Ala Ser Gly
340 345 350

Gly Leu Ser Gly Gly Val Cys Glu Asp Cys Gln His Asn Thr Glu Gly
355 360 365

50

Gln His Cys Asp Arg Cys Arg Pro Leu Phe Tyr Arg Asp Pro Leu Lys
370 375 380

55

Thr Ile Ser Asp Pro Tyr Ala Cys Ile Pro Cys Glu Cys Asp Pro Asp
385 390 395 400

Gly Thr Ile Ser Gly Gly Ile Cys Val Ser His Ser Asp Pro Ala Leu
405 410 415

60

Gly Ser Val Ala Gly Gln Cys Leu Cys Lys Glu Asn Val Glu Gly Ala
420 425 430

Lys Cys Asp Gln Cys Lys Pro Asn His Tyr Gly Leu Ser Ala Thr Asp

5 435 440 445

Pro Leu Gly Cys Gln Pro Cys Asp Cys Asn Pro Leu Gly Ser Leu Pro
450 455 460

10 Phe Leu Thr Cys Asp Val Asp Thr Gly Gln Cys Leu Cys Leu Xaa Tyr
465 470 475 480

Val Thr Gly Ala His Cys Glu Glu Cys Thr Val Gly Tyr Trp Gly Leu
485 490 495

15

Gly Asn His Leu His Gly Cys Ser Pro Cys Asp Cys Asp Ile Gly Gly
500 505 510

Ala Tyr Ser Asn Val Cys Ser Pro Lys Asn Gly Gln Cys Glu Cys Arg
20 515 520 525

Pro His Val Thr Gly Arg Ser Cys Ser Glu Pro Ala Pro Gly Tyr Phe
530 535 540

25 Phe Ala Pro Leu Asn Phe Tyr Leu Tyr Glu Ala Glu Glu Ala Thr Thr
545 550 555 560

Leu Gln Gly Leu Ala Pro Leu Gly Ser Glu Thr Phe Gly Gln Ser Pro
565 570 575

30

Ala Val His Val Val Leu Gly Glu Pro Val Pro Gly Asn Pro Val Thr
580 585 590

Trp Thr Gly Pro Gly Phe Ala Arg Val Leu Pro Gly Ala Gly Leu Arg
35 595 600 605

Phe Ala Val Asn Asn Ile Pro Phe Pro Val Asp Phe Thr Ile Ala Ile
610 615 620

40 His Tyr Glu Thr Gln Ser Ala Ala Asp Trp Thr Val Gln Ile Val Val
625 630 635 640

Asn Pro Pro Gly Gly Ser Glu His Cys Ile Pro Lys Thr Leu Gln Ser
645 650 655

45

Lys Pro Gln Ser Phe Ala Leu Pro Ala Ala Thr Arg Ile Met Leu Leu
660 665 670

Pro Thr Pro Ile Cys Leu Glu Pro Asp Val Gln Tyr Ser Ile Asp Val
50 675 680 685

Tyr Phe Ser Gln Pro Leu Gln Gly Glu Ser His Ala His Ser His Val
690 695 700

55 Leu Val Asp Ser Leu Gly Leu Ile Pro Gln Ile Asn Ser Leu Glu Asn
705 710 715 720

Phe Cys Ser Lys Gln Asp Leu Asp Glu Tyr Gln Leu His Asn Cys Val
725 730 735

60

Glu Ile Ala Ser Ala Met Gly Pro Gln Val Leu Pro Gly Ala Cys Glu
740 745 750

5 Arg Leu Ile Ile Ser Met Ser Ala Lys Leu His Asp Gly Ala Val Ala
755 760 765

Cys Lys Cys His Pro Gln Gly Ser Val Gly Ser Ser Cys Ser Arg Leu
770 775 780

10

Gly Gly Gln Cys Gln Cys Lys Pro Leu Val Val Gly Arg Cys Cys Asp
785 790 795 800

Arg Cys Ser Thr Gly Ser Tyr Asp Leu Gly His His Gly Cys His Pro
15 805 810 815

Cys His Cys His Pro Gln Gly Ser Lys Asp Thr Val Cys Asp Gln Val
820 825 830

20 Thr Gly Gln Cys Pro Cys His Gly Glu Val Ser Gly Arg Arg Cys Asp
835 840 845

Arg Cys Leu Ala Gly Tyr Phe Gly Phe Pro Ser Cys His Pro Cys Pro
850 855 860

25

Cys Xaa Arg Phe Xaa Ala Xaa Asp Xaa Leu Xaa Xaa Asp Pro Glu Thr
865 870 875 880

Gly Ser Cys Phe Asn Cys Gly Gly Phe Thr Thr Gly Arg Asn Cys Glu
30 885 890 895

Arg Cys Ile Asp Gly Tyr Tyr Gly Asn Pro Ser Ser Gly Gln Pro Cys
900 905 910

35 Arg Pro Cys Leu Cys Pro Asp Asp Pro Ser Ser Asn Gln Tyr Phe Ala
915 920 925

His Ser Cys Tyr Gln Asn Leu Trp Ser Ser Asp Val Ile Cys Asn Cys
930 935 940

40

Leu Gln Gly Tyr Thr Gly Thr Gln Cys Gly Glu Cys Ser Thr Gly Phe
945 950 955 960

Tyr Gly Asn Pro Arg Ile Ser Gly Ala Pro Cys Gln Pro Cys Ala Cys
45 965 970 975

Asn Asn Asn Ile Asp Val Thr Asp Pro Glu Ser Cys Ser Arg Val Thr
980 985 990

50 Gly Glu Cys Leu Arg Cys Leu His Asn Thr Gln Gly Ala Asn Cys Gln
995 1000 1005

Leu Cys Lys Pro Gly His Tyr Gly Ser Ala Leu Asn Gln Thr Cys Arg
1010 1015 1020

55

Arg Cys Ser Cys His Ala Ser Gly Val Ser Pro Met Glu Cys Pro Pro
1025 1030 1035 1040

Gly Gly Gly Ala Cys Leu Cys Asp Pro Val Thr Gly Ala Cys Pro Cys
60 1045 1050 1055

Leu Pro Asn Val Thr Gly Leu Ala Cys Asp Arg Cys Ala Asp Gly Tyr
1060 1065 1070

5

Trp Asn Leu Val Pro Gly Arg Gly Cys Gln Ser Cys Asp Cys Asp Pro
1075 1080 1085

10 Arg Xaa Ser Gln Ser Ser His Cys Asp Gln Ala Arg Tyr Phe Lys Ala
1090 1095 1100

Tyr
1105

15 (2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 3754 base pairs
(B) TYPE: nucleic acid
20 (C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

25

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

ACATGCCCCG TTTGCTGCCT GAACCTCTCC ACAAAGACTC CCAGATCCTG AATTGAATTT 60
30
AATCATCTCC TGACAAAAGA ATGCAATTTC AACTGACCCT TTTTTTGCAC CTTGGGTGGC 120

TCAGTTACTC AAAAGCTCAA GATGACTGCA ACAGGGGTGC CTGTCATCCC ACCACTGGTG 180

35 ATCTCCTGGT GGGCAGGAAC ACGCAGCTTA TGGCTTCTTC TACCTGTGGG CTGAGCAGAG 240

CCCAGAAATA CTGCATCCTC AGTTACCTGG AGGGGGAACA AAAATGCTCC ATCTGTGACT 300

CTAGATTTCC ATATGATCCG TATGACCAAC CCAACAGCCA CACCATTGAG AATGTCACTG 360
40
TAAGTTTTGA ACCAGACAGA GAAAAGAAAT GGTGGCAATC TGAAAATGGT CTTGATCATG 420

TCAGCATCAG ACTGGACTTA GAGGCATTAT TTCGGTTCAG CCACCTTATC CTGACCTTTA 480

45 AGACTTTTCG GCCTGCTGCA ATGTTAGTTG AACGTTCCAC AGACTATGGA CACAACTGGA 540

AAGTGTTCAA ATATTTTGCA AAAGACTGTG CCACTTCCTT TCCTAACATC ACATCTGGCC 600

AGGCCCAGGG AGTGGGAGAC ATTGTTTGTG ACTCCAAATA CTCGGATATT GAACCCTCAA 660
50
CAGGTGGAGA GGTTGTTTTA AAAGTTTTGG ATCCCAGTTT TGAAATTGAA AACCCTTATA 720

GCCCCTACAT CCAAGACCTT GTGACATTGA CAAACCTGAG GATAAACTTT ACCAAGCTCC 780

55 ACACCCTTGG GGATGCTTTG CTTGGAAGGA GGCAAAATGA TTCCCTTGAT AAATACTACT 840

ATGCTCTGTA CGAGATGATT GTTCGGGGAA GCTGCTTTTG CAATGGCCAT GCTAGCGAAT 900

GTCGCCCTAT GCAGAAGATG CGGGGAGATG TTTTCAGCCC TCCTGGAATG GTTCACGGTC 960
60
AGTGTGTGTG TCAGCACAAT ACAGATGGTC CGAACTGTGA GAGATGCAAG GACTTCTTCC 1020

AGGATGCTCC TTGGAGGCCA GCTGCAGACC TCCAGGACAA CGCTTGCAGA TCGTGCAGCT 1080

```
5
GTAATAGCCA CTCCAGCCGC TGTCACTTTG ACATGACTAC GTACCTGGCA AGCGGTGGCC    1140
TCAGCGGGGG CGTGTGTGAA GACTGCCAGC ACAACACTGA GGGGCAGCAC TGCGACCGCT    1200
10  GCAGACCCCT CTTCTACAGG GACCCGCTCA AGACCATCTC AGATCCCTAC GCGTGCATTC    1260
CTTGTGAATG TGACCCCGAT GGGACCATAT CTGGTGGCAT TTGTGTGAGC CACTCTGATC    1320
CTGCCTTAGG GTCTGTGGCC GGCCAGTGCC TTTGTAAAGA GAACGTGGAA GGAGCCAAAT    1380
15
GCGACCAGTG CAAACCCAAC CACTACGGAC TAAGCGCCAC CGACCCCCTG GGCTGCCAGC    1440
CCTGCGACTG TAACCCCCTT GGGAGTCTGC CATTCTTGAC CTGTGATGTG GATACAGGCC    1500
20  AATGCTTGTG CCTGTATATG TCACCGGAGC ACACTGCGAA GAATGCACTG TTGGATACTG    1560
GGGCCTGGGA AATCATCTCC ATGGGTGTTC TCCCTGTGAC TGTGATATTG GAGGTGCTTA    1620
TTCTAACGTG TGCTCACCCA AGAATGGGCA GTGTGAATGC CGCCCACATG TCACTGGCCG    1680
25
TAGCTGCTCT GAACCAGCCC CTGGCTACTT CTTTGCTCCT TTGAATTTCT ATCTCTACGA    1740
GGCAGAGGAA GCCACAACAC TCCAAGGACT GGCGCCTTTG GGCTCGGAGA CGTTTGGCCA    1800
30  GAGTCCTGCT GTTCACGTTG TTTTAGGAGA GCCAGTTCCT GGGAACCCTG TTACATGGAC    1860
TGGACCTGGA TTTGCCAGGG TTCTCCCTGG GGCTGGCTTG AGATTTGCTG TCAACAACAT    1920
TCCCTTTCCT GTGGACTTCA CCATTGCCAT TCACTATGAA ACCCAGTCTG CAGCTGACTG    1980
35
GACTGTCCAG ATTGTGGTGA ACCCCCCTGG AGGGAGTGAG CACTGCATAC CCAAGACTCT    2040
ACAGTCAAAG CCTCAGTCTT TTGCCTTACC AGCGGCTACG AGAATCATGC TGCTTCCCAC    2100
40  ACCCATCTGT TTAGAACCAG ATGTACAATA TTCCATAGAT GTCTATTTTT CTCAGCCTTT    2160
GCAAGGAGAG TCCCACGCTC ATTCACATGT CCTGGTGGAC TCTCTTGGCC TTATTCCCCA    2220
AATCAATTCA TTGGAGAATT TCTGCAGCAA GCAGGACTTA GATGAGTATC AGCTTCACAA    2280
45
CTGTGTTGAA ATTGCCTCAG CAATGGGACC TCAAGTGCTC CCGGGTGCCT GTGAAAGGCT    2340
GATCATCAGC ATGTCTGCCA AGCTGCATGA TGGGGCTGTG GCCTGCAAGT GTCACCCCCA    2400
50  GGGCTCAGTC GGATCCAGCT GCAGCCGACT TGGAGGCCAG TGCCAGTGTA AACCTCTTGT    2460
GGTCGGGCGC TGCTGTGACA GGTGCTCAAC TGGAAGCTAT GATTTGGGGC ATCACGGCTG    2520
TCACCCATGT CACTGCCATC CTCAAGGATC AAAGGACACT GTATGTGACC AAGTAACAGG    2580
55
ACAGTGCCCC TGCCATGGAG AGGTGTCTGG CCGCCGCTGT GATCGCTGCC TGGCAGGCTA    2640
CTTTGGATTT CCCAGCTGCC ACCCTTGCCC TTGTAAAGGT TTCGCTAGAC ACTTTTGTGA    2700
60  TCCTGAGACA GGGTCATGCT TCAATTGTGG AGGCTTTACA ACTGGCAGAA ACTGTGAAAG    2760
GTGTATTGAT GGTTACTATG GAAATCCTTC TTCAGGACAG CCCTGTCGTC CTTGCCTGTG    2820
```

5 TCCAGATGAT CCCTCAAGCA ATCAGTATTT TGCCCATTCC TGTTATCAGA ATCTGTGGAG 2880

CTCAGATGTA ATCTGCAATT GTCTTCAAGG TTATACGGGT ACTCAGTGTG GAGAATGCTC 2940

TACTGGTTTC TATGGAAATC CAAGAATTTC AGGAGCACCT TGCCAACCAT GTGCCTGCAA 3000

10

CAACAACATA GATGTAACCG ATCCAGAGTC CTGCAGCCGG GTAACAGGGG AGTGCCTTCG 3060

ATGTTTGCAC AACACTCAGG GCGCAAACTG CCAGCTCTGC AAACCAGGTC ACTATGGATC 3120

15 AGCCCTCAAT CAGACCTGCA GAAGATGCTC CTGCCATGCT TCCGGCGTGA GTCCCATGGA 3180

GTGTCCCCCT GGTGGGGGAG CTTGCCTCTG TGACCCTGTC ACTGGTGCAT GTCCTTGTCT 3240

GCCGAATGTC ACAGGCCTGG CCTGTGACCG TTGTGCTGAT GGATACTGGA ATCTGGTCCC 3300

20

TGGCAGAGGA TGTCAGTCAT GTGACTGTGA CCCTAGCCTC TCAAAGTAGC CACTGTGACC 3360

AGGCAAGATA CTTTAAAGCT TACTAGTGCA TCAAAGTGAG CATGATAGTG AGACATGGTT 3420

25 TCTAATGTGT AAAGAAAGTT TCTTTTATGT ACTGTTGTTA ATTAGTGCAT TGAAACAGGA 3480

TGCCTTACAG GGATGGAGTC AGCCTCTATC AAGGAATGAA ACCAAAAAAG AGAATGAGCA 3540

TCTCAAGTTC AGCTTCGCCT ACTTCAGTTT CCCCTCTGTG ACTGAGGAAG TCAGAATTCA 3600

30

TACACAGTGA AACACAGACA TCAGCCTCAC CTTTCACTAT TTCATACATG TAACCATAGG 3660

GAAGACCTAA GAAATAGTTA ATCAGAAGAG ATTATGAATC AGAATGAAAA TAAACAGATA 3720

35 CCTTCAAAAC CTAAAAAAAA AAAAAAAAAA AAAA 3754

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/21391

A. CLASSIFICATION OF SUBJECT MATTER
IPC(6) :C07K 14/00
US CL :530/350
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)
U.S. : 530/350

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 practicable, search terms used)
APS and Dialog (Biotech)
search terms: laminin, alpha2, beta1, beta4, gamma3

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X,P	Database Genbank PID g4003505, Cloning and characterization of the human laminin beta-4 chain. Laminin beta-4 chain precusor. OLSEN et al. Publicly Available on 02 December 1997	3
A	US 5,660,982 A (TRYGGVASON et al.) 26 August 1997, see entire document.	1-3
A	AUMAILLEY et al. Laminins: A Family of Diverse Multifunctional Molecules of Basement Membranes. The Journal of Investigative Dermatology. February 1996, Vol. 106, pages 209-214, see entire document.	1-3

[X] Further documents are listed in the continuation of Box C. [ ] See patent family 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 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
01 MARCH 1999	11 MAR 1999

Name and mailing address of the ISA/US	Authorized officer
Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231	BRADLEY S. MAYHEW
Facsimile No. (703) 305-3230	Telephone No. (703) 308-0196

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/21391

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	ENGVALL, Eva Laminin Variants: Why, Where and When? Kidney International. September 1993, Vol. 43, pages 2-6, see entire document.	1-3