Office de la Propriété Intellectuelle du Canada
Un organisme d'Industrie Canada

Canadian Intellectual Property Office
An agency of Industry Canada

CA 2296841 C 2012/02/21

(11)(21) 2 296 841

(12) BREVET CANADIEN
CANADIAN PATENT

(13) C

(86) Date de dépôt PCT/PCT Filing Date: 1998/07/17
(87) Date publication PCT/PCT Publication Date: 1999/01/28
(45) Date de délivrance/Issue Date: 2012/02/21
(85) Entrée phase nationale/National Entry: 2000/01/17
(86) N° demande PCT/PCT Application No.: AU 1998/000564
(87) N° publication PCT/PCT Publication No.: 1999/003886
(30) Priorité/Priority: 1997/07/18 (AU PO 8117)

(51) Cl.Int./Int.Cl. *C12N 15/12* (2006.01), *A61K 38/17* (2006.01), *C07H 21/04* (2006.01), *C07K 14/435* (2006.01), *C07K 14/78* (2006.01), *C12N 15/63* (2006.01), *C12P 21/02* (2006.01), *A61K 38/00* (2006.01)

(72) Inventeur/Inventor:
WEISS, ANTHONY STEVEN, AU

(73) Propriétaire/Owner:
THE UNIVERSITY OF SYDNEY, AU

(74) Agent: RICHES, MCKENZIE & HERBERT LLP

(54) Titre : DERIVES DE TROPOELASTINE
(54) Title: TROPOELASTIN DERIVATIVES

(57) Abrégé/Abstract:
The invention relates to derivatives of tropoelastin and variants of those derivatives. The invention further provides expression products and hybrid molecules of the derivatives and variants of the invention. The invention further provides methods for the production of the derivatives, variants, expression products and hybrid molecules. Further provided are formulations, cross-linked structures and implants comprising the derivatives, variants, expression products and hybrid molecules of the invention. Further provided are uses of the derivatives, variants, expression products and hybrid molecules of the invention.

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C07K 14/435, C07H 21/04, A61K 38/17, C12N 15/12, C12P 21/02

A1

(11) International Publication Number: WO 99/03886

(43) International Publication Date: 28 January 1999 (28.01.99)

(21) International Application Number: PCT/AU98/00564

(22) International Filing Date: 17 July 1998 (17.07.98)

(30) Priority Data:
PO 8117 18 July 1997 (18.07.97) AU

(71) Applicant *(for all designated States except US)*: THE UNIVERSITY OF SYDNEY [AU/AU]; Parramatta Road, Sydney, NSW 2006 (AU).

(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: WEISS, Anthony, Steven [AU/AU]; 235 Rainbow Street, Randwick, NSW 2031 (AU).

(74) Agent: GRIFFITH HACK; G.P.O. Box 4164, Sydney, NSW 2001 (AU).

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

Published
With international search report.

(54) Title: TROPOELASTIN DERIVATIVES

(57) Abstract

The invention relates to derivatives of tropoelastin and variants of those derivatives. The invention further provides expression products and hybrid molecules of the derivatives and variants of the invention. The invention further provides methods for the production of the derivatives, variants, expression products and hybrid molecules. Further provided are formulations, cross–linked structures and implants comprising the derivatives, variants, expression products and hybrid molecules of the invention. Further provided are uses of the derivatives, variants, expression products and hybrid molecules of the invention.

TROPOELASTIN DERIVATIVES

TECHNICAL FIELD

The present invention relates to derivatives of human tropoelastin and variants thereof, to genetic constructs encoding the amino acid sequences of the derivatives and variants and to uses of the derivatives and variants. In particular, the derivatives of the present invention have elastin-like properties or macro-molecular binding properties.

BACKGROUND ART

There are various forms of tropoelastin that typically appear to consist of two types of alternating domains: those rich in hydrophobic amino acids (responsible for the elastic properties) and those rich in lysine residues (responsible for cross-link formation). Hydrophobic and cross-linking domains are encoded in separate exons (Indik *et al* 1987).

The 26 A region of human tropoelastin is unique amongst tropoelastin domains in that, due to the absence of lysine, this region does not participate in elastin cross-link formation. Furthermore, this region is a serine-rich domain and lacks hydrophobic stretches, indicating that it is unlikely to contribute to the elasticity of tropoelastin. There is otherwise limited information on the structure and functional relationships of the 26 A region (Bedell-Hogan *et al.*, 1993).

The gene for tropoelastin is believed to be present as a single copy in the mammalian genome, and is expressed in the form of multiple transcripts, distinguished by alternative splicing of the pre-mRNA (Indik *et al*, 1990; Oliver *et al*, 1987). Modest expression of a natural human tropoelastin sequence has been achieved by Indik *et al* (1990) using cDNA, providing free polypeptide which unfortunately was unstable.

Expression of substantial amounts of human tropoelastin using synthetic polynucleotides is reported

in WO94/14958. In particular, a construct, SHEL, providing substantial amounts of full length human tropoelastin is described.

DESCRIPTION OF THE INVENTION

In the specification and claims, "derivatives of human tropoelastin" or "tropoelastin derivatives" means novel peptides, polypeptides or proteins which contain amino acid sequences derived from the native amino acid sequences of human tropoelastin molecules. The amino acid sequences of the derivatives of human tropoelastin may be derived from any of the amino acid sequences of the isoforms of human tropoelastin. Derivatives of human tropoelastin are distinguished from human tropoelastin molecules in that the amino acid sequences of derivatives are altered with respect to native tropoelastin sequences by substitution, addition or deletion of residues, or a combination of these alterations, in derivative amino acid sequences.

In a first aspect, the present invention provides derivatives of human tropoelastin which have elastin-like properties. Elastin-like properties are a combination of elastic properties, including the phenomenon of recoil following molecular distention under appropriate conditions, and the ability to be cross-linked to other elastin molecules and/or other elastin-like molecules.

In a second aspect, the present invention provides derivatives of human tropoelastin which have macro-molecular binding properties including the ability to bind glycosaminoglycans.

In a third aspect, the present invention provides derivatives of human tropoelastin which have elastin-like properties and macro-molecular binding properties.

The present invention further provides amino acid sequence variants of the derivatives of the invention. In the specification and claims "variants" means amino acid sequences which retain the properties of the corresponding derivative of human tropoelastin, for example, elastin-

like properties or macro-molecular binding properties, or a combination of elastin-like properties and macro-molecular binding properties, and have an amino acid sequence which is homologous with the amino acid sequence of the corresponding derivative. For the purposes of this description, "homology" between the amino acid sequence of a particular derivative of human tropoelastin and another amino acid sequence connotes a likeness short of identity, indicative of a derivation of one sequence from the other. In particular, an amino acid sequence is homologous to a derivative of human tropoelastin if the alignment of that amino acid sequence with the sequence of the derivative of human tropoelastin reveals a similarity of about 65% over any 20 amino acid stretch or over any repetitive element of the molecules shorter than 20 amino acids in length. Such a sequence comparison can be performed via known algorithims, such as that of Lipman and Pearson (1985). Similarity is observed between amino acids where those amino acids have a side chain which confers a similar chemical property in the same chemical environment. For example, threonine and serine are similar amino acids; aspartic acid and glutamic acid are similar amino acids; valine, leucine and isoleucine are similar amino acids etc. Thus, an amino acid sequence may be considered homologous with the amino acid sequence of a human tropoelastin derivative because the alignment of those sequences reveals a similarity of 65%, although at each amino acid position in the aligned sequences, none of the residues are identical.

Inasmuch as the present invention provides derivatives of human tropoelastin and amino acid sequence variants of those derivatives, the invention thus extends to amino acid sequence variants incorporating amino acid sequences of non-human tropoelastins. Amino acid sequence variants which are non-human tropoelastin derivatives, or are based all, or in part, on non-human tropoelastin derivatives retain properties of the corresponding derivative of non-human tropoelastin, for example,

elastin-like properties or macro-molecular binding properties, or a combination of elastin-like properties and macro-molecular binding properties, and have an amino acid sequence which is homologous with the amino acid sequence of the corresponding human derivative. The variants of the invention also include variants of the non-human tropoelastin derivatives, or of derivatives based on the non-human tropoelastin derivatives.

"Homology" between the amino acid sequence of a particular derivative of non-human tropoelastin and another amino acid sequence connotes a likeness short of identity, indicative of a derivation of one sequence from the other. In particular, an amino acid sequence is homologous to a derivative of non-human tropoelastin if the alignment of that amino acid sequence with the sequence of the derivative of non-human tropoelastin reveals a similarity of about 65% over any 20 amino acid stretch or over any repetitive element of the molecules shorter than 20 amino acids in length. The skilled addressee will understand that species that are substantially phylogenetically related to humans express tropoelastin molecules which consist of amino acid sequences with homology to human tropoelastin amino acid sequences. Indeed, amino acid sequences of non-human tropoelastins have been determined, including the amino acid sequences of chick tropoelastin, bovine tropoelastin and rat tropoelastin (Bressan *et al.* 1987, Raju *et al.* 1987, Pierce *et al.* 1992) and over multiple regions, these are homologous with the human tropoelastin amino acid sequences. The skilled addressee will recognise therefore, that derivatives of human tropoelastin and amino acid sequence variants of those derivatives will necessarily encompass corresponding tropoelastin amino acid sequences from these and other non-human species.

The present invention provides a tropoelastin derivative comprising the amino acid sequence of SHELδmodified (SEQ ID NO:5). The amino acid sequence of

SHELδmodified and the alignment of that amino acid sequence with the human tropoelastin sequence is shown in Figure 5.

The invention also provides an amino acid sequence variant of the derivative comprising the amino acid sequence of SHELδmodified.

The invention also provides a polynucleotide encoding a tropoelastin derivative comprising the amino acid sequence of SHELδmodified. The nucleotide sequence encoding SHELδmodified is shown in Figure 3 (SEQ ID NO: 4). Preferably the polynucleotide comprises the nucleotide sequence which corresponds to SHELδmodified shown in Figure 3.

The invention also provides a polynucleotide encoding an amino acid sequence variant of the derivative SHELδmodified.

The present invention further provides a synthetic polynucleotide encoding a tropoelastin derivative comprising the amino acid sequence of SHELδ26A (SEQ ID NO:3). A synthetic polynucleotide is a molecule which comprises a nucleotide sequence that contains silent mutations with respect to the corresponding native polynucleotide molecule. The silent mutations enhance the expression of the synthetic polynucleotide. The amino acid sequence of SHELδ26A and the alignment of that amino acid sequence with the human tropoelastin sequence is shown in Figure 2. The SHELδ26A derivative excludes the SHEL coding sequence corresponding to exon 26A.

Preferably the synthetic polynucleotide comprises the sequence shown in Figure 1 (SEQ ID NO:1) from nucleotide position 1 to 1676 contiguous with nucleotide position 1775 to 2210.

The invention also provides a polynucleotide encoding an amino acid sequence variant of the derivative SHELδ26A.

The invention also provides an amino acid sequence

variant of the derivative comprising the amino acid sequence of SHELδ26A.

The present inventor has, for the first time, shown that the region encoded by exon 26A (peptide 26A) of the tropoelastin gene binds glycosaminoglycans (GAGs) (Figure 6A and B). GAGs are macro-molecules particularly associated with the extracellular environment. These molecules play an important role in the architecture and mechanical properties of connective tissues and mediate interactions with and availability of other molecules.

Thus, the present invention provides a tropoelastin derivative comprising the amino acid sequence of peptide 26A. Peptide 26A has the amino acid sequence: GADEGVRRSLSPELREGDPSSSQHLPSTPSSPRV (SEQ ID NO: 12) or GADEGVRRSLSPELREGDPSSSQHLPSTPSSPRF (SEQ ID NO: 13).

The present invention also provides an amino acid sequence variant of the derivative comprising the amino acid sequence of peptide 26A.

The invention also provides a polynucleotide encoding a tropoelastin derivative comprising the amino acid sequence of peptide 26A. Preferably the polynucleotide comprises the nucleotide sequence shown in Figure 1 (SEQ ID NO: 1) from nucleotide position 1687 to 1778. Preferably the 3' terminal codon is GTT (which encodes V) or TTT (which encodes F).

The invention also provides a polynucleotide encoding an amino acid sequence variant of the derivative comprising the amino acid sequence of peptide 26A.

In appreciating the GAG binding property of peptide 26A, the present inventor envisages the generation of novel subsets of hybrid molecules, comprising biological polymers which are linked to peptide 26A, wherein the peptide 26A imparts GAG binding activity to the polymer. In particular, the present inventor has recognised that the deletion or insertion of the peptide 26A amino acid sequence, or a variant of that amino acid sequence will alter GAG binding activity. Thus, the present invention relates to tropoelastin derivatives in which full length

or partial length tropoelastin molecules have been modified by the addition of one or more exon 26A regions to enhance interactions with GAGs. Moreover, the invention relates to site directed modification of the amino acid sequence of peptide 26A so as to generate variants of the peptide 26A amino acid sequence which have altered affinity or altered specificity for GAGs. Tropoelastin derivatives or variants of the derivatives which contain altered GAG binding activity may be uncross-linked or cross-linked.

In another aspect, the invention provides a hybrid molecule. In the specification and claims, "hybrid molecule" means a molecule comprising a biological polymer which is linked to a tropoelastin derivative comprising the amino acid sequence of peptide 26A or an amino acid sequence variant of a derivative comprising the amino acid sequence of peptide 26A. Preferably the biological polymer is a protein. More preferably the protein is selected from the group consisting of growth factors, cytokines and antibodies. Alternatively the biological polymer is selected from the group consisting of lipids, sugars or nucleic acids.

In one embodiment, and where the biological polymer is a protein, the hybrid molecule is produced by recombinant DNA techniques, including for example the construction of a nucleotide sequence which encodes the biological polymer and the tropoelastin derivative comprising the amino acid sequence of peptide 26A, or the amino acid sequence variant of a derivative comprising the amino acid sequence of peptide 26 A, in a single open reading frame. Alternatively, the hybrid molecule may be produced synthetically by solid phase peptide synthesis, including, for example the methods of synthesis disclosed in Merrifield (1963) or Knorr *et al*. (1989). Examples of peptide synthesis also include the synthesis methods used by peptide synthesisers of Perkin Elmer/Applied Biosystems, CA, US.

In another aspect, the invention provides a

polynucleotide sequence encoding a hybrid molecule of the invention.

In another aspect, the invention provides a hybrid molecule which comprises a synthetic polymer which is linked in a tropoelastin derivative comprising the amino acid sequence of peptide 26A, or an amino acid sequence variant of the derivative comprising the amino acid sequence of peptide 26A.

The invention further provides a method of imparting or enhancing GAG binding activity to a biological polymer comprising the step of linking a tropoelastin derivative comprising the amino acid sequence of peptide 26A, or an amino acid sequence variant of peptide 26A with the biological polymer. Preferably the biological polymer is a protein.

The invention further provides a method of deleting or reducing GAG binding activity from a biological polymer comprising the step of deleting a tropoelastin derivative comprising the amino acid sequence of peptide 26A, or an amino acid sequence variant of peptide 26A from the biological polymer. Preferably the biological polymer is a protein.

The present invention also provides a tropoelastin derivative comprising the amino acid sequence of SHELgamma. SHELgamma has the amino acid sequence:
SAMGALVGLGVPGLGVGAGVPGFGAGADEGVRRSLSPELREGDPSSSQHLPSTPSSPR
VPGALAAAKAAKYGAAVPGVLGGLGALGGVGIPGGVVGAGPAAAAAAAKAAAKAAQFG
LVGAAGLGGLGVGGLGVPGVGGLGGIPPAAAAKAAKYGAAGLGGVLGGAGQFPLGGVA
ARPGFGLSPIFPGGACLGKACGRKRK (SEQ ID NO: 9).

The invention also provides an amino acid sequence variant of the derivative comprising the amino acid sequence of SHELgamma.

The invention also provides a polynucleotide encoding a tropoelastin derivative, the derivative comprising the amino acid sequence of SHELgamma. The nucleotide sequence of the polynucleotide SHELgamma (SEQ ID NO: 8) is shown in Figure 8. In this nucleotide sequence, the first 9 codons from nucleotide position 948 to 974 are derived

from the glutathione *S*-transferase (GST) fusion nucleotide sequence. Preferably the polynucleotide comprises the nucleotide sequence shown in Figure 8. More preferably the polynucleotide comprises the nucleotide sequence shown in Figure 8 from nucleotide sequence position 975 to 1547.

The invention also provides a polynucleotide encoding an amino acid sequence variant of the derivative comprising the amino acid sequence of SHELgamma.

The present invention also provides a polynucleotide encoding a tropoelastin derivative, the derivative comprising the amino acid sequence of SHELgamma excluding exon 26A. The nucleotide sequence of the polynucleotide SHELgamma excluding exon 26A (SEQ ID NO: 6) is shown in Figure 7. In this nucleotide sequence, the first 5 codons from nucleotide position 948 to 962 are derived from the GST nucleotide sequence. SHELgamma excluding exon 26A has the following amino acid sequence:

VPGALAAAKAAKYGAAVPGVLGGLGALGGVGIPGGVVGAGPAAAAAAAKAAAKAAQFG
LVGAAGLGGLGVGGLGVPGVGGLGGIPPAAAAKAAKYGAAGLGGVLGGAGQFPLGGVA
ARPGFGLSPIFPGGACLGKACGRKRK (SEQ ID NO: 7).

Preferably the polynucleotide comprises the nucleotide sequence shown in SEQ ID NO:6. More preferably the polynucleotide comprises the nucleotide sequence shown in SEQ ID NO: 6 from nucleotide sequence position 15 to 441.

The invention also provides a polynucleotide encoding an amino acid sequence variant of the derivative comprising the amino acid sequence of SHELgamma excluding exon 26A.

The invention also provides a tropoelastin derivative comprising the amino acid sequence of SHELgamma excluding exon 26A.

The invention also provides an amino acid sequence variant of the derivative comprising SHELgamma excluding exon 26A.

The derivatives of the invention based on SHELgamma can also be produced by *in vitro* biochemical cleavage of tropoelastin products such as SHEL, so as to release a carboxy- terminal fragment. The carboxy-terminal fragment

may be purified by reverse phase HPLC.

The present invention also provides a tropoelastin derivative comprising the amino acid sequence of SHEL31-36. SHEL31-36 has the following amino acid sequence: GIPPAAAAKAAKYGAAGLGGVLGGAGQFPLGGVAARPGFGLSPIFPGGACLGKACG-RKRK (SEQ ID NO: 10).

SHEL31-36 retains a crosslinking domain. As a consequence of its elastin-like properties, it is envisaged that this and related tropoelastin derivatives can be used to interfere with tropoelastin deposition and formation of unaltered elastic fibre.

The invention also provides an amino acid sequence variant of the derivative comprising the amino acid sequence of SHEL31-36.

The invention also provides a polynucleotide encoding a tropoelastin derivative, the derivative comprising the amino acid sequence of SHEL31-36. Preferably the polynucleotide comprises the nucleotide sequence shown in Figure 1 (SEQ ID NO:1) from nucleotide position 2022 to 2210.

The invention also provides a polynucleotide encoding an amino acid variant of the derivative comprising the amino acid sequence of SHEL31-36.

The present invention also provides a tropoelastin derivative, comprising the amino acid sequence of SHEL32-36. SHEL32-36 has the following amino acid sequence: GAAGLGGVLGGAGQFPLGGVAARPGFGLSPIFPGGACLGKACGRKRK (SEQ ID NO: 11).

The invention also provides an amino acid sequence variant of the derivative comprising the amino acid sequence of SHEL32-36.

The invention also provides a polynucleotide encoding a tropoelastin derivative, the derivative comprising the amino acid sequence of SHEL32-36. Preferably the polynucleotide comprises the nucleotide sequence shown in Figure 1 (SEQ ID NO: 1) from nucleotide position 2061 to 2210.

The present invention also provides a polynucleotide

encoding an amino acid sequence variant of the derivative comprising the amino acid sequence of SHEL32-36.

As a consequence of its elastin-like properties, it is envisaged that SHEL32-36 and related tropoelastin derivatives can be used to interfere with tropoelastin deposition and formation of an unaltered elastic fibre.

The present invention also provides a tropoelastin derivative, comprising the amino acid sequence of SHEL26-36. SHEL26-36 has the following amino acid sequence:

AAAGLGAGIPGLGVGVGVPGLGVGAGVPGLGVGAGVPGFGAGADEGVRRSLSPELREGD PSSSQHLPSTPSSPRVPGALAAAKAAKYGAAVPGVLGGLGALGGVGIPGGVVGAGPAAA AAAAKAAAKAAQFGLVGAAGLGGLGVGGLGVPGVGGLGGIPPAAAAKAAKYGAAGLGGV LGGAGQFPLGGVAARPGFGLSPIFPGGACLGKACGRKRK (SEQ ID NO: 14)

The invention also provides an amino acid sequence variant of the derivative comprising the amino acid sequence of SHEL26-36.

The invention also provides a polynucleotide encoding a tropoelastin derivative, the derivative comprising the amino acid sequence of SHEL26-36. Preferably the polynucleotide comprises the nucleotide sequence shown in Figure 1 from nucleotide position 1554-2210.

The present invention also provides a tropoelastin derivative, comprising the amino acid sequence of SHEL26-36 excluding exon 26A. SHEL26-36 excluding exon 26A has the following amino acid sequence:

AAAGLGAGIPGLGVGVGVPGLGVGAGVPGLGVGAGVPGFGAVPGALAAAKAAKYGAAVP GVLGGLGALGGVGIPGGVVGAGPAAAAAAAKAAAKAAQFGLVGAAGLGGLGVGGLGVPG VGGLGGIPPAAAAKAAKYGAAGLGGVLGGAGQFPLGGVAARPGFGLSPIFPGGACLGKA CGRKRK (SEQ ID NO: 15)

The invention also provides an amino acid sequence variant of the derivative comprising the amino acid sequence of SHEL26-36 excluding exon 26A.

The invention also provides a polynucleotide encoding a tropoelastin derivative, the derivative comprising the amino acid sequence of SHEL26-36 excluding exon 26A. Preferably the polynucleotide comprises the nucleotide sequence shown in Figure 1 from nucleotide position 1554

to 1676 contiguous with 1776 to 2210.

The present invention also provides a polynucleotide encoding an amino acid sequence variant of the derivative comprising the amino acid sequence of SHEL26-36.

In another aspect the present invention provides a formulation comprising a tropoelastin derivative, a variant of the derivative or a hybrid molecule of the invention, together with a carrier or diluent.

Formulations of the derivatives, variants or hybrid molecules of the invention can be prepared in accordance with standard techniques appropriate to the field in which they are to be used.

The polynucleotides and synthetic polynucleotides of the invention can be provided in association with other polynucleotide sequences including 5' and 3' untranslated sequences, and 5' upstream and 3' downstream transcriptional regulatory sequences. The polynucleotides and synthetic polynucleotides may be provided as a recombinant DNA molecule including plasmid DNA.

The polynucleotides and synthetic polynucleotides of the invention can be prepared using the techniques of chemical synthesis or recombinant DNA technology, or by a combination of both techniques.

In a further aspect the invention provides a vector comprising a polynucleotide or synthetic polynucleotide encoding a tropoelastin derivative, a variant of the derivative or a hybrid molecule of the invention.

Vectors useful in this invention include plasmids, phages and phagemids. The polynucleotides and synthetic polynucleotides of the present invention can also be used in integrative expression systems or lytic or comparable expression systems.

Suitable vectors will generally contain origins of replication and control sequences which are derived from species compatible with the intended expression host. Typically these vectors include a promoter located upstream from the polynucleotide, together with a ribosome binding site if intended for prokaryotic expression, and a

phenotypic selection gene such as one conferring antibiotic resistance or supplying an auxotrophic requirement. For production vectors, vectors which provide for enhanced stability through partitioning may be chosen. Where integrative vectors are used it is not necessary for the vector to have an origin of replication. Lytic and other comparable expression systems do not need to have those functions required for maintenance of vectors in hosts.

For *E. coli* typical vectors include pBR322, pBluescript II SK*, pGEX-2T, pTrc99A, pET series vectors, particularly pET3d, (Studier *et al.*, 1990) and derivatives of these vectors. Derivatives include those plasmids with a modified protease recognition sequence to facilitate purification of a protein domain.

In another aspect the invention provides a cell capable of expressing a polynucleotide or a synthetic polynucleotide which encodes a derivative or variant of the invention, or a polynucleotide which encodes a hybrid molecule of the invention.

A preferred expression system is an *E. coli* expression system. However, the invention includes within its scope the use of other hosts capable of expressing protein from the polynucleotides designed for use in *E. coli*. The invention also includes the use of polynucleotides and synthetic polynucleotides suitable for use in other expression systems such as other microbial expression systems. These other expression systems include yeast, and bacterial expression systems, insect cell expression systems, and expression systems involving other eukaryotic cell lines or whole organisms.

Examples of *E. coli* hosts include *E. coli* B strain derivatives (Studier *et al*, 1990), and K-strain derivatives such as NM522 (Gough and Murray, 1983) and XL1-Blue (Bullock *et al*, 1987).

In a further aspect the present invention provides an expression product. In the specification and claims, "expression product" means a derivative or variant of the

invention expressed by a cell containing a polynucleotide or a synthetic polynucleotide encoding a derivative or variant of the invention.

The expression products of the invention may be fused expression products which include all or part of a protein encoded by the vector in peptide linkage with the derivative or variant. They may also include, for example, an N-terminal methionine or other additional residues which do not permanently impair the elastin-like, or macro-molecular binding properties of the product.

Typically the fusion is to the N-terminus of the expression product. An example of a suitable protein is to the C-terminus of glutathione *S*-transferase. The fused protein sequence may be chosen in order to cause the expression product to be secreted or expressed as a cell surface protein to simplify purification or expressed as a cytoplasmic protein.

The expressed fusion products may subsequently be treated to remove the fused protein sequences to provide free tropoelastin derivative or variant. Treatment is typically through protease treatment or, in the case of secretion, removal is effected by endogenous host secretion machinery. An example of this is secretion by yeasts.

Non-fused systems include the introduction of or use of a pre-existing methionine codon. An example of this is the use of pET3a or pET3d in *E. coli*.

In another aspect the invention provides a polynucleotide encoding an expression product of the invention.

In another aspect the present invention provides a formulation comprising an expression product of the invention together with a carrier or diluent. The formulation of the expression product can be prepared in accordance with standard techniques appropriate to the field in which they are to be used.

According to a further aspect of the present invention there is provided a method for producing a

tropoelastin derivative or a variant of the derivative comprising providing a vector containing a polynucleotide or a synthetic polynucleotide encoding the derivative or variant; introducing the vector into a suitable host cell; maintaining the cell in conditions suitable for expression of the polynucleotide or synthetic polynucleotide and isolating the derivative or variant of the invention. The method can be applied to the production of the expression products and hybrid molecules (in which the hybrid molecules comprise the peptide 26A or a variant thereof and a further amino acid sequence) of the invention, by providing a vector containing a polynucleotide encoding an expression product or a hybrid molecule; introducing the vector into a suitable host cell; maintaining the cell in conditions suitable for expression of the polynucleotide and isolating the expression product or hybrid molecule.

In one embodiment, the polynucleotide or synthetic polynucleotide encoding the derivative, variant, expression product or hybrid molecule of the invention is expressed in a host cell which is maintained in culture *in vitro*.

Alternatively, the polynucleotide or synthetic polynucleotide encoding the derivative, variant, expression product or hybrid molecule of the invention is expressed in a host cell which is maintained *in vivo*. Thus, in another embodiment, the polynucleotide or synthetic polynucleotide encoding the derivative, variant, expression product or hybrid molecule of the invention is expressed in a transgenic animal. Methods for the generation of transgenic animals are known in the art. Exemplary methods are described in Slack et al. 1991 and Janne et al. 1992.

The tropoelastin derivatives, variants of the derivatives, and hybrid molecules (in which the hybrid molecules comprise the peptide 26A or a variant thereof and a further amino acid sequence) of the invention may be produced by solid phase peptide synthesis, including, for example, the methods of synthesis disclosed in Merrifield

(1963) or Knorr *et al (1989)*. Examples of peptide synthesis also include the synthesis methods used by peptide synthesisers of Perkin Elmer/Applied Biosystems, CA, US. As an alternative to cell synthesis from a polynucleotide or synthetic polynucleotide, the expression products of the invention may be produced by solid phase peptide synthesis.

In a further aspect the present invention provides an implant formed from at least one tropoelastin derivative and/or variant of the derivative of the invention. The implant may alternatively contain at least one expression product and/or at least one hybrid molecule of the invention.

The implants are formed into the required shape by cross-linking the tropoelastin derivative, variant of the derivative, expression product, or hybrid molecule of the invention, in a mould which conforms to the desired shape of the implant. Where the implant is required to be used in sheet form the tropoelastin derivative, variant of the derivative, expression product, or hybrid molecule of the invention can be cross-linked on a flat surface. Relevant methodologies are described in, for example, US Patent No. 4 474 851 and US Patent No. 5 250 516. The elastomeric materials may be exclusively prepared from one or more tropoelastin derivatives, variants of the derivative, expression products, or hybrid molecules of the invention or may be composites prepared from one or more of these constituents together with other materials.

The tropoelastin derivatives or variants of the derivatives can be cross-linked to form elastin or elastin-like material or can be cross-linked in conjunction with other biological or synthetic molecules to form a composite material.

Thus in another aspect the invention provides a cross-linked complex which comprises at least one tropoelastin derivative of the invention and/or at least one variant of a derivative of the invention. The cross-linked complexes may additionally contain at least one

expression product and/or at least one hybrid molecule of the invention, which may be cross-linked to the at least one tropoelastin derivative and/or variant of the derivative of the invention.

The cross-linking of the tropoelastin derivatives, variants of the derivatives, hybrid molecules and expression products of the invention can be achieved by chemical oxidation of lysine side chains using processes such as ruthenium tetroxide mediated oxidation and quinone mediated oxidation, or by using homobifunctional chemical cross-linking agents such as dithiobis (succinimidylpropionate), dimethyl adipimidate or dimethyl pimelimidate. Glutaraldehyde cross-linking is an important addition to this repetoire. Another alternative is the cross-linking of lysine and glutamic side chains.

The tropoelastin derivatives, variants of the derivatives, hybrid molecules and expression products of the invention may also be enzymatically cross-linked by methods including lysyl oxidase mediated oxidation or may be cross-linked using gamma irradiation.

Accordingly, in another aspect, the present invention resides in a synthetic polynucleotide encoding a tropoelastin derivative, wherein the polynucleotide has a nucleotide sequence of: nucleotide position 1 to 1676 of SEQ ID NO: 1 contiguous with 1775 to 2210 of SEQ ID NO: 1; or nucleotide position 1554 to 1676 of SEQ ID NO: 1 contiguous with 1776 to 2210 of SEQ ID NO: 1.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1: Nucleotide (SEQ ID NO: 1) and predicted amino acid (SEQ ID NO:2) sequences of synthetic human tropoelastin SHEL. The upper (numbered) nucleotide sequence represents the coding strand.

Figure 2: Alignment of SHEL (SEQ ID NO:2)(upper line) and SHELδ26A (SEQ ID NO: 3) amino acid sequences.

Figure 3: Nucleotide (SEQ ID NO: 4) and predicted amino acid (SEQ ID NO: 5) sequences of SHELδmodified.

Figure 4: Alignment of SHELδmodified (SEQ ID NO: 4) (upper line) and SHEL (SEQ ID NO:1) nucleotide sequences.

Figure 5: Alignment of SHELδmodified (SEQ ID NO: 5)(lower line) and SHEL (SEQ ID NO: 1) amino acid sequences.

Figure 6A: HPLC elution profile of GST-exon 26A fusion protein tropoelastin derivative loaded in from

heparin sepharose. 6B: Binding of peptide 26A (SEQ ID NO: 12 and SEQ ID NO: 13) to glycosaminoglycans.

Figure 7: Nucleotide (SEQ ID NO: 6) and predicted amino acid (SEQ ID NO: 7) sequences of SHELgamma excluding exon 26A.

Figure 8: Nucleotide (SEQ ID NO: 8) and predicted amino acid (SEQ ID NO: 9) sequences of SHELgamma.

BEST METHOD OF PERFORMING THE INVENTION

The recombinant and synthetic procedures used for the synthesis of the derivatives, variants, expression products and hybrid molecules of the invention are described in standard texts such as Sambrook *et al* (1989).

Tropoelastin nucleotide sequences may be modified so as to provide derivatives, variants, expression products or hybrid molecules, by conventional site-directed or random mutagenesis. The sequences may also be modified by oligonucleotide-directed mutagenesis, which comprises the following steps:

1. synthesis of an oligonucleotide with a sequence that contains the desired nucleotide substitution (mutation);
2. hybridising the oligonucleotide to a template comprising a structural sequence encoding tropoelastin; and
3. using a DNA polymerase to extend the oligonucleotide as a primer.

Another approach which is particularly suited to situations where a synthetic polynucleotide encoding the tropoelastin derivative is prepared from oligonucleotide blocks bounded by restriction sites, is cassette mutagenesis where entire restriction fragments are replaced.

Purification of the derivatives, variants, expression products or hybrid molecules of the invention is performed using standard techniques including HPLC. The actual sequence of steps in the purification of a particular derivative, variant, expression product or hybrid molecule

is limited by the environment from which the molecule is to be purified. By way of example, reference is made to the purification scheme disclosed in WO94/14958.

Formulations in accordance with the invention are formulated in accordance with standard techniques.

The amount of derivative, variant, expression product or hybrid molecule that may be combined with a carrier or diluent to produce a single dosage will vary depending on the situation in which the formulation is to be used and the particular mode of administration.

It will be understood also that specific doses for any particular host may be influenced by factors such as the age, sex, weight and general health of the host as well as the particular characteristics of the derivative, variant, expression product or hybrid molecule of the invention being used, and how it is administered.

Injectable preparations, for example, sterile injectable aqueous or oleagenous suspensions may be formulated according to the known art using suitable dispersing or wetting agents and suspending agents. The sterile injectable preparation may also be a sterile injectable solution or suspension in a non-toxic parenterally acceptable diluent or solvent. Among the acceptable vehicles or solvents that may be employed are water, Ringer's solution, alcohols and isotonic sodium chloride solution. In addition, sterile, fixed oils are conventionally employed as a solvent or suspending medium. For this purpose any bland fixed oil may be employed including synthetic mono- or diglycerides. In addition, fatty acids such as oleic acid and organic solvents find use in the preparation of injectables.

Routes of administration, dosages to be administered as well as frequency of administration are all factors which can be optimised using ordinary skill in the art.

In addition, the derivatives, variants, expression products and hybrid molecules of the invention may be prepared as topical preparations for instance as anti-wrinkle and hand lotions using standard techniques for the

preparation of such formulations. They may be prepared in aerosol form for, for instance, administration to a patient's lungs, or in the form of surgical implants, foods or industrial products by standard techniques.

SHEL

The preparation of SHEL is described in WO94/14958. It is directly expressed as a full length human protein with a calculated molecular weight of 64kDa. The full nucleotide sequence and corresponding amino acid sequence of SHEL is shown in Figure 1. The preparation of pSHELF is described in WO94/14958.

pSHELF differs from the natural coding sequence(s) in a number of ways. As described in WO94/14958, the untranslated regions present in the tropoelastin cDNA sequence were disregarded in designing the synthetic gene, and the nucleotides encoding the signal peptide were removed. Restriction endonuclease recognition sites were incorporated at regular intervals into the gene by typically altering only the third base of the relevant codons, thereby maintaining the primary sequence of the gene product. The facility for silent alteration of the coding sequence was also exploited to change the codon bias of the tropoelastin gene to that commonly found in highly expressed *E.coli* genes. [Genetics Computer Group (GCG) package version 7-UNIX using Codon Frequency and Gen Run Data: ecohigh-cod]. Two additional stop codons were added to the 3'-end, and an ATG start codon comprising a novel NcoI site was appended to the 5'-end. *Bam* HI cloning sites were engineered at both ends of the synthetic sequence. Since the gene contains no internal methionine residues, treatment of the newly-synthesized gene product (expressed directly or as a fusion with another gene) with cyanogen bromide would liberate a protein with the same or similar sequence as one form of natural tropoelastin comprising 731 amino acids. Other forms of processing are envisaged, which may generate tropoelastin species of the same or different lengths.

Two stop codons were added in order to allow the possible use of the construct in suppressor hosts, and also to avoid any potential depletion of termination (release) factors for translation.

As described in the following examples, the derivatives, pSHELFδ26A, pSHELFδ modified, pSHELgamma, pSHEL31-36, pSHEL32-36 and pSHELgammaδ26A were derived from the pSHELF nucleotide sequence. These particular derivatives, and indeed the derivaties, variants, expression products and hybrid molecules of the invention can equally be derived from a native human or non -human tropoelastin nucleotide sequence.

Example 1: Construction of pSHELFδ26A and pSHELFδ modified

Mutagenesis was used with pSHELF to remove DNA corresponding to exon 26A. The sequence of the mutagenic primer was:

5'CGG GTT TCG GTG CTG TTC CGG GCG CGC TGG 3'

This flanked either side of exon 26A by 15bp resulting in its precise deletion. A second selection primer, which mutates a unique restriction site to another restriction site is normally used in the protocol but was not in this case since deletion of exon 26A also resulted in the deletion of a unique restriction site, *Pml*I. The enzyme *Pml*I was used to treat the mutation reaction to linearise any unmutated parental plasmid and consequently to enrich for mutant plasmid. The reaction mixture was used to transform competent BMH17-18 *mutS E. coli*, defective in mismatch repair, by electroporation and the entire transformed culture was grown overnight in LB+ampicillin. Mixed plasmid DNA, containing both mutated and parental plasmids, was isolated from the culture and the plasmid DNA was digested with *Pml*I to linearise the parental plasmid. The plasmid DNA, now enriched for mutated plasmid, was used to transform *E. coli* HMS174 by electroporation and transformants selected on LB plates

containing 75μgml^{-1} ampicillin.

Colonies were grown overnight and plasmid mini-preparations performed. Constructs were screened using *Pml*I and those which were insensitive to digestion were further screened by *Kpn*I/*Pst*I double digestion. Candidate clones were sequenced to verify the sequence, named pSHELFδmodified.

Sequencing confirmed the region immediately surrounding the deletion was correct. *Pst*I and *Bss*HII restriction sites surrounding the correct region of pSHELFδmodified were used to remove the desired segment and re-insert it into the corresponding site of pSHELF. 6.5μg pSHELF and 7.5μg pSHELFδmodified were digested with *Bss*HII, precipitated and digested with *Pst*I. The appropriate three fragments were gel-purified and ligated. DNA was transformed into *E. coli* XL1-Blue and transformants selected on plates containing 75μgml^{-1} ampicillin.

Plasmids were isolated by mini-preparations and screened using *Bgl*I digestion. A candidate clone was further analysed by restriction enzyme digestion and sequenced, and named pSHELFδ26A.

Example 2: Synthesis of Exon 26A

The region of SHEL corresponding to exon 26A was amplified by PCR, with primers modified to introduce an in-frame BamH1 site upstream and a stop codon downstream at the 3′ end. Two forms were generated: one terminating in valine (26AV) and the other terminating in phenyalanine (26AF). These forms are as follows:

GADEGVRRSLSPELREGDPSSSQHLPSTPSSPRV with properties:

Molecular weight = 3588.80

Residues = 34

Average Residue Weight = 105.553

Charge = -1

Isoelectric point = 5.71

and

GADEGVRRSLSPELREGDPSSSQHLPSTPSSPRF

A 26A coding region was expressed as a glutathione S-transferase (GST) fusion protein.

Example 3: Glycosaminoglycan binding activity of Exon 26A

Ultrafiltration assay methodology was developed to examine and quantify interactions occurring *in vitro* between the 26A region and purified extracellular matrix glcosaminoglycans. GST26A fusion protein and tropoelastin were compared with GST and tropoelastin lacking exon 26A at physiologicaly relevant conditions of pH and ionic strength.

Experimental evidence supports the notion that peptide 26A (26AF and 26AV) binds GAGs. Immobilised heparin column binding shows that GST26A binds more tightly than does GST, and requires a higher sodium chloride concentration for elution (Figure 6B).

Furthermore, GST26A fusion protein binds radioactive heparin with greater efficiencies than GST, and these can be compared with GAGs including chondroitin sulphates and keratin sulphates. An implication of this is that GAGs binding to tropoelastin can be adjusted based upon the content of 26A. Cross-linked tropoelastin would be expected to show differential binding to GAGs based on the relative amounts of SHEL vs. SHELδ26A.

In summary, these studies reveal that the 26A region is a functional glycosaminoglycan binding domain, which functions in intact tropoelastin. It is also active when isolated as a fusion entity yet displays no detectable structure in the absence of bound GAG. Furthermore, panel competition studies indicate a preference for those GAGs found in close association with the elastic fibre in the extracellular matrix.

Example 4: Construction of pSHELgamma, pSHEL31-36, pSHEL32-36 and pSHELgammaδ26A

pSHELgamma is derived from the pSHELgamma construct disclosed in WO94/14958. pSHEL31-36, pSHEL32-36 and pSHELgammaδ26A were derived from pSHELgamma. pSHELgamma was modified by introducing an oligonucleotide linker at the *Kpn*I site. This encoded a factor Xa cleavage site which could be utilised in subsequent constructs. PCR and site directed mutagenesis was then used to generate further, shorter forms which provided fusions with GST. Constructs were DNA sequenced for verification. Induced protein was isolated as GST-fusion proteins, which were subsequently bound to glutathione agarose. Protease cleavage was optional where fusion proteins were desired; otherwise the cleaved proteins and peptides were further purified by reverse phase HPLC.

INDUSTRIAL APPLICATION

The derivatives and expression products of the invention are of use in *inter alia* the medical, pharmaceutical, veterinary and cosmetic fields.

REFERENCES

1. Indik Z, Yeh H, Ornstein-Goldstein N, Sheppard P, Anderson N, Rosenbloom JC, Peltonen L and Rosenbloom J (1987) PNAS (USA) **84** 5680-5684

2. Indik Z, Abrams W.R., Kucich U, Gibson C.W., Mecham R.P. and Rosenbloom J (1990) Arch. Biochem Biophys **280** 80-86

3. Oliver L, Luvalle PA, Davidson J.M., Rosenbloom J, Mathew C.G., Betser A.J. and Boyd C.D. (1987) Collagen Rel Res **7** 77-89

4. Sambrook J., Fritsch E.F., and Maniatis T. (1989) Molecular cloning: a laboratory manual, second edition Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York

5. Bressan G.M., Argos P. and Stanley K.K. (1987) Biochemistry **26** 1497-11503

6. Raju K. aand Anwar R.A. (1987) J. Biol Chem **262** 5755-5762

7. Pierce R.A., Alatawi A, Deak S.B. & Boyd C.D. (1992) Genomics **12** 651-658

8. Lipman and Pearson (1985) Science 227,1435.

9. Bedell-Hogan, D., Trackman, P., Abrams, W., Rosenbloom, J. and Kagan H. (1993) J. Biol. Chem. **268**, 10345-10350

10. Studier, F. W., Rosenberg, A. H., Dunn, J. J. and Dubendorff, J. W. (1990) Methods Enzymol. **185**, 60-89

11. Gough, J., and Murray, N. (1983) J. Mol. Biol. **166**,

1-19

12. Bullock, W. O., Fernandez, J. M. and Short, J. M. (1987) BioTechniques **5**, 376-379

13. Slack, J. L., Liska, D. J. and Bornstein P. (1991) Mol. Cell Biol. 11: 2066-2074

14. Janne, J., Hyttinen, J. M., Peura, T., Tolvanen, M., Alhonen, L. And Halmekyto M. (1992) Ann. Med. 24: 273-280.

15. Merrifield, R.B., (1963) J. Am. Chem. Soc. 85: 2149-2154.

16. Knorr R., Trzeciak, Bannarth W., Gillessen, D. (1989) Tetrahedron Letters 30: 1927-1930

SEQUENCE LISTING

GENERAL INFORMATION:

APPLICANT: THE UNIVERSITY OF SYDNEY

TITLE OF INVENTION: TROPOELASTIN DERIVATIVES

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:

ADDRESSEE: RICHES, McKENZIE & HERBERT

STREET: 2 BLOOR STREET EAST, SUITE 2900

CITY: TORONTO

PROVINCE: ONTARIO

COUNTRY: CANADA

POSTAL CODE: M4W 3J5

COMPUTER READABLE FORM:

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: ASCII TEXT

CURRENT APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE: 17 July 1998

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: AU PO8117

FILING DATE: 18-JUL-1997

CLASSIFICATION:

PRIOR APPLICATION DATA

APPLICATION NUMBER: PCT/AU98/00564

FILING DATE: 17 JULY 1998

CLASSIFICATION: CO7K 14/435, CO7h 21/04, 61K 38/17, C12N 15/12, C12P 21/02

PATENT AGENT INFORMATION

NAME: RICHES, McKENZIE & HERBERT

REFERENCE NUMBER: P195199

INFORMATION FOR SEQ ID NO:1:

SEQUENCE CHARACTERISTICS:

LENGTH: 2210 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: YES

ANTI-SENSE: NO

FRAGMENT TYPE:

ORIGINAL SOURCE:

IMMEDIATE SOURCE;

POSITION IN GENOME:

CHROMOSOME/SEGMENT:

MAP POSITION:

UNITS:

FEATURE

NAME/KEY:

LOCATION:

IDENTIFICATION METHOD:

OTHER INFORMATION:

PUBLICATION INFORMATION

AUTHORS:

TITLE:

JOURNAL:

VOLUME:

ISSUE:

PAGES:

DATA:

DOCUMENT NUMBER: WO 99/03886

FILING DATE: 17 JULY 1998

PUBLICATION DATE: 28 JANUARY 1999

RELEVANT RESIDUES IN SEQ ID NO.:

SEQUENCE DESCRIPTION: SEQ ID NO:1:

GATCCATGGG TGGCGTTCCG GGTGCTATCC CGGGTGGCGT TCCGGGTGGT GTATTCTACC 60

CAGGCGCGGG TCTGGGTGCA CTGGGCGGTG GTGCGCTGGG CCCGGGTGGT AAACCGCTGA 120

AACCGGTTCC AGGCGGTCTG GCAGGTGCTG GTCTGGGTGC AGGTCTGGGC GCGTTCCCGG 180

CGGTTACCTT CCCGGGTGCT CTGGTTCCGG GTGGCGTTGC AGACGCAGCT GCTGCGTACA 240

AAGCGGCAAA GGCAGGTGCG GGTCTGGGCG GGGTACCAGG TGTTGGCGGT CTGGGTGTAT 300

CTGCTGGCGC AGTTGTTCCG CAGCCGGGTG CAGGTGTAAA ACCGGGCAAA GTTCCAGGTG 360

TTGGTCTGCC GGGCGTATAC CCGGGTGGTG TTCTGCCGGG CGCGCGTTTC CCAGGTGTTG 420

GTGTACTGCC GGGCGTTCCG ACCGGTGCAG GTGTTAAACC GAAGGCACCA GGTGTAGGCG 480

GCGCGTTCGC GGGTATCCCG GGTGTTGGCC CGTTCGGTGG TCCGCAGCCA GGCGTTCCGC 540

TGGGTTACCC GATCAAAGCG CCGAAGCTTC CAGGTGGCTA CGGTCTGCCG TACACCACCG 600

GTAAACTGCC GTACGGCTAC GGTCCGGGTG GCGTAGCAGG TGCTGCGGGT AAAGCAGGCT 660

ACCCAACCGG TACTGGTGTT GGTCCGCAGG CTGCTGCGGC AGCTGCGGCG AAGGCAGCAG 720
CAAAATTCGG CGCGGGTGCA GCGGGTGTTC TGCCGGGCGT AGGTGGTGCT GGCGTTCCGG 780
GTGTTCCAGG TGCGATCCCG GGCATCGGTG GTATCGCAGG CGTAGGTACT CCGGCGGCCG 840
CTGCGGCTGC GGCAGCTGCG GCGAAAGCAG CTAAATACGG TGCGGCAGCA GGCCTGGTTC 900
CGGGTGGTCC AGGCTTCGGT CCGGGTGTTG TAGGCGTTCC GGGTGCTGGT GTTCCGGGCG 960
TAGGTGTTCC AGGTGCGGGC ATCCCGGTTG TACCGGGTGC AGGTATCCCG GGCGCTGCGG 1020
TTCCAGGTGT TGTATCCCCG GAAGCGGCAG CTAAGGCTGC TGCGAAAGCT GCGAAATACG 1080
GAGCTCGTCC GGGCGTTGGT GTTGGTGGCA TCCCGACCTA CGGTGTAGGT GCAGGCGGTT 1140
TCCCAGGTTT CGGCGTTGGT GTTGGTGGCA TCCCGGGTGT AGCTGGTGTT CCGTCTGTTG 1200
GTGGCGTACC GGGTGTTGGT GGCGTTCCAG GTGTAGGTAT CTCCCCGGAA GCGCAGGCAG 1260
CTGCGGCAGC TAAAGCAGCG AAGTACGGCG TTGGTACTCC GGCGGCAGCA GCTGCTAAAG 1320
CAGCGGCTAA AGCAGCGCAG TTCGGACTAG TTCCGGGCGT AGGTGTTGCG CCAGGTGTTG 1380
GCGTAGCACC GGGTGTTGGT GTTGCTCCGG GCGTAGGTCT GGCACCGGGT GTTGGCGTTG 1440
CACCAGGTGT AGGTGTTGCG CCGGGCGTTG GTGTAGCACC GGGTATCGGT CCGGGTGGCG 1500
TTGCGGCTGC TGCGAAATCT GCTGCGAAGG TTGCTGCGAA AGCGCAGCTG CGTGCAGCAG 1560
CTGGTCTGGG TGCGGGCATC CCAGGTCTGG GTGTAGGTGT TGGTGTTCCG GGCCTGGGTG 1620
TAGGTGCAGG GGTACCGGGC CTGGGTGTTG GTGCAGGCGT TCCGGGTTTC GGTGCTGGCG 1680
CGGACGAAGG TGTACGTCGT TCCCTGTCTC CAGAACTGCG TGAAGGTGAC CCGTCCTCTT 1740
CCCAGCACCT GCCGTCTACC CCGTCCTCTC CACGTGTTCC GGGCGCGCTG GCTGCTGCGA 1800
AAGCGGCGAA ATACGGTGCA GCGGTTCCGG GTGTACTGGG CGGTCTGGGT GCTCTGGGCG 1860
GTGTTGGTAT CCCGGGCGGT GTTGTAGGTG CAGGCCCAGC TGCAGCTGCT GCTGCGGCAA 1920
AGGCAGCGGC GAAAGCAGCT CAGTTCGGTC TGGTTGGTGC AGCAGGTCTG GGCGGTCTGG 1980
GTGTTGGCGG TCTGGGTGTA CCGGGCGTTG GTGGTCTGGG TGGCATCCCG CCGGCGGCGG 2040
CAGCTAAAGC GGCTAAATAC GGTGCAGCAG GTCTGGGTGG CGTTCTGGGT GGTGCTGGTC 2100
AGTTCCCACT GGGCGGTGTA GCGGCACGTC CGGGTTTCGG TCTGTCCCCG ATCTTCCCAG 2160
GCGGTGCATG CCTGGGTAAA GCTTGCGGCC GTAAACGTAA ATAATGATAG 2210

INFORMATION FOR SEQ ID NO:2:

SEQUENCE CHARACTERISTICS:

LENGTH: 733 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL:

ANTI-SENSE:

FRAGMENT TYPE:

ORIGINAL SOURCE:

IMMEDIATE SOURCE;

POSITION IN GENOME:

CHROMOSOME/SEGMENT:

MAP POSITION:

UNITS:

FEATURE

NAME/KEY:

LOCATION:

IDENTIFICATION METHOD:

OTHER INFORMATION:

PUBLICATION INFORMATION

AUTHORS:

TITLE:

JOURNAL:

VOLUME:

ISSUE:

PAGES:

DATA:

DOCUMENT NUMBER: WO 99/03886

FILING DATE: 17 JULY 1998

PUBLICATION DATE: 28 JANUARY 1999

RELEVANT RESIDUES IN SEQ ID NO.:

SEQUENCE DESCRIPTION: SEQ ID NO:2:

Ser Met Gly Gly Val Pro Gly Ala Ile Pro Gly Gly Val Pro Gly Gly
1 5 10 15
Val Phe Tyr Pro Gly Ala Gly Leu Gly Ala Leu Gly Gly Gly Ala Leu
20 25 30
Gly Pro Gly Gly Lys Pro Leu Lys Pro Val Pro Gly Gly Leu Ala Gly
35 40 45
Ala Gly Leu Gly Ala Gly Leu Gly Ala Phe Pro Ala Val Thr Phe Pro
50 55 60
Gly Ala Leu Val Pro Gly Gly Val Ala Asp Ala Ala Ala Ala Tyr Lys
65 70 75 80
Ala Ala Lys Ala Gly Ala Gly Leu Gly Gly Val Pro Gly Val Gly Gly
85 90 95
Leu Gly Val Ser Ala Gly Ala Val Val Pro Gln Pro Gly Ala Gly Val
100 105 110
Lys Pro Gly Lys Val Pro Gly Val Gly Leu Pro Gly Val Tyr Pro Gly
115 120 125
Gly Val Leu Pro Gly Ala Arg Phe Pro Gly Val Gly Val Leu Pro Gly
130 135 140
Val Pro Thr Gly Ala Gly Val Lys Pro Lys Ala Pro Gly Val Gly Gly
145 150 155 160
Ala Phe Ala Gly Ile Pro Gly Val Gly Pro Phe Gly Gly Pro Gln Pro
165 170 175
Gly Val Pro Leu Gly Tyr Pro Ile Lys Ala Pro Lys Leu Pro Gly Gly
180 185 190
Tyr Gly Leu Pro Tyr Thr Thr Gly Lys Leu Pro Tyr Gly Tyr Gly Pro
195 200 205

Gly Gly Val Ala Gly Ala Ala Gly Lys Ala Gly Tyr Pro Thr Gly Thr
210 215 220
Gly Val Gly Pro Gln Ala Ala Ala Ala Ala Ala Ala Lys Ala Ala Ala
225 230 235 240
Lys Phe Gly Ala Gly Ala Ala Gly Val Leu Pro Gly Val Gly Gly Ala
245 250 255
Gly Val Pro Gly Val Pro Gly Ala Ile Pro Gly Ile Gly Gly Ile Ala
260 265 270
Gly Val Gly Thr Pro Ala Ala Ala Ala Ala Ala Ala Ala Ala Ala Lys
275 280 285
Ala Ala Lys Tyr Gly Ala Ala Ala Gly Leu Val Pro Gly Gly Pro Gly
290 295 300
Phe Gly Pro Gly Val Val Gly Val Pro Gly Ala Gly Val Pro Gly Val
305 310 315 320
Gly Val Pro Gly Ala Gly Ile Pro Val Val Pro Gly Ala Gly Ile Pro
325 330 335
Gly Ala Ala Val Pro Gly Val Val Ser Pro Glu Ala Ala Ala Lys Ala
340 345 350
Ala Ala Lys Ala Ala Lys Tyr Gly Ala Arg Pro Gly Val Gly Val Gly
355 360 365
Gly Ile Pro Thr Tyr Gly Val Gly Ala Gly Gly Phe Pro Gly Phe Gly
370 375 380
Val Gly Val Gly Gly Ile Pro Gly Val Ala Gly Val Pro Ser Val Gly
385 390 395 400
Gly Val Pro Gly Val Gly Gly Val Pro Gly Val Gly Ile Ser Pro Glu
405 410 415
Ala Gln Ala Ala Ala Ala Ala Lys Ala Ala Lys Tyr Gly Val Gly Thr
420 425 430
Pro Ala Ala Ala Ala Ala Lys Ala Ala Ala Lys Ala Ala Gln Phe Gly
435 440 445
Leu Val Pro Gly Val Gly Val Ala Pro Gly Val Gly Val Ala Pro Gly

450 455 460

Val Gly Val Ala Pro Gly Val Gly Leu Ala Pro Gly Val Gly Val Ala

465 470 475 480

Pro Gly Val Gly Val Ala Pro Gly Val Gly Val Ala Pro Gly Ile Gly

485 490 495

Pro Gly Gly Val Ala Ala Ala Ala Lys Ser Ala Ala Lys Val Ala Ala

500 505 510

Lys Ala Gln Leu Arg Ala Ala Ala Gly Leu Gly Ala Gly Ile Pro Gly

515 520 525

Leu Gly Val Gly Val Gly Val Pro Gly Leu Gly Val Gly Ala Gly Val

530 535 540

Pro Gly Leu Gly Val Gly Ala Gly Val Pro Gly Phe Gly Ala Gly Ala

545 550 555 560

Asp Glu Gly Val Arg Arg Ser Leu Ser Pro Glu Leu Arg Glu Gly Asp

565 570 575

Pro Ser Ser Ser Gln His Leu Pro Ser Thr Pro Ser Ser Pro Arg Val

580 585 590

Pro Gly Ala Leu Ala Ala Ala Lys Ala Ala Lys Tyr Gly Ala Ala Val

595 600 605

Pro Gly Val Leu Gly Gly Leu Gly Ala Leu Gly Gly Val Gly Ile Pro

610 615 620

Gly Gly Val Val Gly Ala Gly Pro Ala Ala Ala Ala Ala Ala Ala Lys

625 630 635 640

Ala Ala Ala Lys Ala Ala Gln Phe Gly Leu Val Gly Ala Ala Gly Leu

645 650 655

Gly Gly Leu Gly Val Gly Gly Leu Gly Val Pro Gly Val Gly Gly Leu

660 665 670

Gly Gly Ile Pro Pro Ala Ala Ala Ala Lys Ala Ala Lys Tyr Gly Ala

675 680 685

Ala Gly Leu Gly Gly Val Leu Gly Gly Ala Gly Gln Phe Pro Leu Gly

690 695 700

Gly Val Ala Ala Arg Pro Gly Phe Gly Leu Ser Pro Ile Phe Pro Gly
705 710 715 720
Gly Ala Cys Leu Gly Lys Ala Cys Gly Arg Lys Arg Lys
725 730

INFORMATION FOR SEQ ID NO:3:

SEQUENCE CHARACTERISTICS:

LENGTH: 698 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL:

ANTI-SENSE:

FRAGMENT TYPE:

ORIGINAL SOURCE:

IMMEDIATE SOURCE;

POSITION IN GENOME:

CHROMOSOME/SEGMENT:

MAP POSITION:

UNITS:

FEATURE

NAME/KEY:

LOCATION:

IDENTIFICATION METHOD:

OTHER INFORMATION:

PUBLICATION INFORMATION

AUTHORS:

TITLE:

JOURNAL:

VOLUME:

ISSUE:

PAGES:

DATA:

DOCUMENT NUMBER: WO 99/03886

FILING DATE: 17 JULY 1998

PUBLICATION DATE: 28 JANUARY 1999

RELEVANT RESIDUES IN SEQ ID NO.:

SEQUENCE DESCRIPTION: SEQ ID NO:3:

Gly Gly Val Pro Gly Ala Ile Pro Gly Gly Val Pro Gly Gly Val Phe
1 5 10 15
Tyr Pro Gly Ala Gly Leu Gly Ala Leu Gly Gly Gly Ala Leu Gly Pro
20 25 30
Gly Gly Lys Pro Leu Lys Pro Val Pro Gly Gly Leu Ala Gly Ala Gly
35 40 45
Leu Gly Ala Gly Leu Gly Ala Phe Pro Ala Val Thr Phe Pro Gly Ala
50 55 60
Leu Val Pro Gly Gly Val Ala Asp Ala Ala Ala Ala Tyr Lys Ala Ala
65 70 75 80
Lys Ala Gly Ala Gly Leu Gly Gly Val Pro Gly Val Gly Gly Leu Gly
85 90 95
Val Ser Ala Gly Ala Val Val Pro Gln Pro Gly Ala Gly Val Lys Pro

100 105 110

Gly Lys Val Pro Gly Val Gly Leu Pro Gly Val Tyr Pro Gly Gly Val

115 120 125

Leu Pro Gly Ala Arg Phe Pro Gly Val Gly Val Leu Pro Gly Val Pro

130 135 140

Thr Gly Ala Gly Val Lys Pro Lys Ala Pro Gly Val Gly Gly Ala Phe

145 150 155 160

Ala Gly Ile Pro Gly Val Gly Pro Phe Gly Gly Pro Gln Pro Gly Val

165 170 175

Pro Leu Gly Tyr Pro Ile Lys Ala Pro Lys Leu Pro Gly Gly Tyr Gly

180 185 190

Leu Pro Tyr Thr Thr Gly Lys Leu Pro Tyr Gly Tyr Gly Pro Gly Gly

195 200 205

Val Ala Gly Ala Ala Gly Lys Ala Gly Tyr Pro Thr Gly Thr Gly Val

210 215 220

Gly Pro Gln Ala Ala Ala Ala Ala Ala Ala Lys Ala Ala Ala Lys Phe

225 230 235 240

Gly Ala Gly Ala Ala Gly Val Leu Pro Gly Val Gly Gly Ala Gly Val

245 250 255

Pro Gly Val Pro Gly Ala Ile Pro Gly Ile Gly Gly Ile Ala Gly Val

260 265 270

Gly Thr Pro Ala Ala Ala Ala Ala Ala Ala Ala Ala Ala Lys Ala Ala

275 280 285

Lys Tyr Gly Ala Ala Ala Gly Leu Val Pro Gly Gly Pro Gly Phe Gly

290 295 300

Pro Gly Val Val Gly Val Pro Gly Ala Gly Val Pro Gly Val Gly Val

305 310 315 320

Pro Gly Ala Gly Ile Pro Val Val Pro Gly Ala Gly Ile Pro Gly Ala

325 330 335

Ala Val Pro Gly Val Val Ser Pro Glu Ala Ala Ala Lys Ala Ala Ala

340 345 350

Lys Ala Ala Lys Tyr Gly Ala Arg Pro Gly Val Gly Val Gly Gly Ile
355 360 365
Pro Thr Tyr Gly Val Gly Ala Gly Gly Phe Pro Gly Phe Gly Val Gly
370 375 380
Val Gly Gly Ile Pro Gly Val Ala Gly Val Pro Ser Val Gly Gly Val
385 390 395 400
Pro Gly Val Gly Gly Val Pro Gly Val Gly Ile Ser Pro Glu Ala Gln
405 410 415
Ala Ala Ala Ala Ala Lys Ala Ala Lys Tyr Gly Val Gly Thr Pro Ala
420 425 430
Ala Ala Ala Ala Lys Ala Ala Ala Lys Ala Ala Gln Phe Gly Leu Val
435 440 445
Pro Gly Val Gly Val Ala Pro Gly Val Gly Val Ala Pro Gly Val Gly
450 455 460
Val Ala Pro Gly Val Gly Leu Ala Pro Gly Val Gly Val Ala Pro Gly
465 470 475 480
Val Gly Val Ala Pro Gly Val Gly Val Ala Pro Gly Ile Gly Pro Gly
485 490 495
Gly Val Ala Ala Ala Ala Lys Ser Ala Ala Lys Val Ala Ala Lys Ala
500 505 510
Gln Leu Arg Ala Ala Ala Gly Leu Gly Ala Gly Ile Pro Gly Leu Gly
515 520 525
Val Gly Val Gly Val Pro Gly Leu Gly Val Gly Ala Gly Val Pro Gly
530 535 540
Leu Gly Val Gly Ala Gly Val Pro Gly Phe Gly Ala Val Pro Gly Ala
545 550 555 560
Leu Ala Ala Ala Lys Ala Ala Lys Tyr Gly Ala Ala Val Pro Gly Val
565 570 575
Leu Gly Gly Leu Gly Ala Leu Gly Gly Val Gly Ile Pro Gly Gly Val
580 585 590
Val Gly Ala Gly Pro Ala Ala Ala Ala Ala Ala Ala Ala Lys Ala Ala Ala

595 600 605

Lys Ala Ala Gln Phe Gly Leu Val Gly Ala Ala Gly Leu Gly Gly Leu

610 615 620

Gly Val Gly Gly Leu Gly Val Pro Gly Val Gly Gly Leu Gly Gly Ile

625 630 635 640

Pro Pro Ala Ala Ala Ala Lys Ala Ala Lys Tyr Gly Ala Ala Gly Leu

645 650 655

Gly Gly Val Leu Gly Gly Ala Gly Gln Phe Pro Leu Gly Gly Val Ala

660 665 670

Ala Arg Pro Gly Phe Gly Leu Ser Pro Ile Phe Pro Gly Gly Ala Cys

675 680 685

Leu Gly Lys Ala Cys Gly Arg Lys Arg Lys

690 695

INFORMATION FOR SEQ ID NO:4:

SEQUENCE CHARACTERISTICS:

LENGTH: 1983 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: YES

ANTI-SENSE: NO

FRAGMENT TYPE:

ORIGINAL SOURCE:

IMMEDIATE SOURCE;

POSITION IN GENOME:

CHROMOSOME/SEGMENT:

MAP POSITION:

UNITS:

FEATURE

NAME/KEY:

LOCATION:

IDENTIFICATION METHOD:

OTHER INFORMATION:

PUBLICATION INFORMATION

AUTHORS:

TITLE:

JOURNAL:

VOLUME:

ISSUE:

PAGES:

DATA:

DOCUMENT NUMBER: WO 99/03886

FILING DATE: 17 JULY 1998

PUBLICATION DATE: 28 JANUARY 1999

RELEVANT RESIDUES IN SEQ ID NO.:

SEQUENCE DESCRIPTION: SEQ ID NO:4:

ATGGGTGGCG TTCCGGGTGC TGTTCCGGGT GGCGTTCCGG GTGGTGTATT CTACCCAGGC 60

GCGGGTTTCG GTGCTGTTCC GGGTGGCGTT GCAGACGCAG CTGCTGCGTA CAAAGCGGCA 120

AAGGCAGGTG CGGGTCTGGG CGGGGTACCA GGTGTTGGCG GTCTGGGTGT ATCTGCTGGC 180

GCAGTTGTTC CGCAGCCGGG TGCAGGTGTA AAACCGGGCA AAGTTCCAGG TGTTGGTCTG 240

CCGGGCGTAT ACCCGGGTTT CGGTGCTGTT CCGGGCGCGC GTTTCCCAGG TGTTGGTGTA 300

CTGCCGGGCG TTCCGACCGG TGCAGGTGTT AAACCGAAGG CACCAGGTGT AGGCGGCGCG 360

TTCGCGGGTA TCCCGGGTGT TGGCCCGTTC GGTGGTCCGC AGCCAGGCGT TCCGCTGGGT 420

TACCCGATCA AAGCGCCGAA GCTTCCAGGT GGCTACGGTC TGCCGTACAC CACCGGTAAA 480
CTGCCGTACG GCTACGGTCC GGGTGGCGTA GCAGGTGCTG CGGGTAAAGC AGGCTACCCA 540
ACCGGTACTG GTGTTGGTCC GCAGGCTGCT GCGGCAGCTG CGGCGAAGGC AGCAGCAAAA 600
TTCGGCGCGG GTGCAGCGGG TTTCGGTGCT GTTCCGGGCG TAGGTGGTGC TGGCGTTCCG 660
GGTGTTCCAG GTGCGATCCC GGGCATCGGT GGTATCGCAG GCGTAGGTAC TCCGGCGGCC 720
GCTGCGGCTG CGGCAGCTGC GGCGAAAGCA GCTAAATACG GTGCGGCAGC AGGCCTGGTT 780
CCGGGTGGTC CAGGCTTCGG TCCGGGTGTT GTAGGCGTTC CGGGTTTCGG TGCTGTTCCG 840
GGCGTAGGTG TTCCAGGTGC GGGCATCCCG GTTGTACCGG GTGCAGGTAT CCCGGGCGCT 900
GCGGGTTTCG GTGCTGTATC CCCGGAAGCG GCAGCTAAGG CTGCTGCGAA AGCTGCGAAA 960
TACGGAGCTC GTCCGGGCGT TGGTGTTGGT GGCATCCCGA CCTACGGTGT AGGTGCAGGC 1020
GGTTTCCCAG GTTTCGGCGT TGGTGTTGGT GGCATCCCGG GTGTAGCTGG TGTTCCGTCT 1080
GTTGGTGGCG TACCGGGTGT TGGTGGCGTT CCAGGTGTAG GTATCTCCCC GGAAGCGCAG 1140
GCAGCTGCGG CAGCTAAAGC AGCGAAGTAC GGCGTTGGTA CTCCGGCGGC AGCAGCTGCT 1200
AAAGCAGCGG CTAAAGCAGC GCAGTTCGGA CTAGTTCCGG GCGTAGGTGT TGCGCCAGGT 1260
GTTGGCGTAG CACCGGGTGT TGGTGTTGCT CCGGGCGTAG GTCTGGCACC GGGTGTTGGC 1320
GTTGCACCAG GTGTAGGTGT TGCGCCGGGC GTTGGTGTAG CACCGGGTAT CGGTCCGGGT 1380
GGCGTTGCGG CTGCTGCGAA ATCTGCTGCG AAGGTTGCTG CGAAAGCGCA GCTGCGTGCA 1440
GCAGCTGGTC TGGGTGCGGG CATCCCAGGT CTGGGTGTAG GTGTTGGTGT TCCGGGCCTG 1500
GGTGTAGGTG CAGGGGTACC GGGCCTGGGT GTTGGTGCAG GCGTTCCGGG TTTCGGTGCT 1560
GTTCCGGGCG CGCTGGCTGC TGCGAAAGCG GCGAAATACG GTGCTGTTCC GGGTGTACTG 1620
GGCGGTCTGG GTGCTCTGGG CGGTGTTGGT ATCCCGGGCG GTGTTGTAGG TGCAGGCCCA 1680
GCTGCAGCTG CTGCTGCGGC AAAGGCAGCG GCGAAAGCAG CTCAGTTCGG TCTGGTTGGT 1740
GCAGCAGGTC TGGGCGGTCT GGGTGTTGGC GGTCTGGGTG TACCGGGCGT TGGTGGTCTG 1800
GGTGGCATCC CGCCGGCGGC GGCAGCTAAA GCGGCTAAAT ACGGTGCAGC AGGTCTGGGT 1860
GGCGTTCTGG GTGGTGCTGG TCAGTTCCCA CTGGGCGGTG TAGCGGCACG TCCGGGTTTC 1920
GGTCTGTCCC CGATCTTCCC AGGCGGTGCA TGCCTGGGTA AAGCTTGCGG CCGTAAACGT 1980
AAA 1983

INFORMATION FOR SEQ ID NO:5:

SEQUENCE CHARACTERISTICS:

LENGTH: 660 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL:

ANTI-SENSE:

FRAGMENT TYPE:

ORIGINAL SOURCE:

IMMEDIATE SOURCE;

POSITION IN GENOME:

CHROMOSOME/SEGMENT:

MAP POSITION:

UNITS:

FEATURE

NAME/KEY:

LOCATION:

IDENTIFICATION METHOD:

OTHER INFORMATION:

PUBLICATION INFORMATION

AUTHORS:

TITLE:

JOURNAL:

VOLUME:

ISSUE:

PAGES:

DATA:

DOCUMENT NUMBER: WO 99/03886

FILING DATE: 17 JULY 1998

PUBLICATION DATE: 28 JANUARY 1999

RELEVANT RESIDUES IN SEQ ID NO.:

SEQUENCE DESCRIPTION: SEQ ID NO:5:

Met Gly Gly Val Pro Gly Ala Val Pro Gly Gly Val Pro Gly Gly Val
1 5 10 15
Phe Tyr Pro Gly Ala Gly Phe Gly Ala Val Pro Gly Gly Val Ala Asp
20 25 30
Ala Ala Ala Ala Tyr Lys Ala Ala Lys Ala Gly Ala Gly Leu Gly Gly
35 40 45
Val Pro Gly Val Gly Gly Leu Gly Val Ser Ala Gly Ala Val Val Pro
50 55 60
Gln Pro Gly Ala Gly Val Lys Pro Gly Lys Val Pro Gly Val Gly Leu
65 70 75 80
Pro Gly Val Tyr Pro Gly Phe Gly Ala Val Pro Gly Ala Arg Phe Pro
85 90 95
Gly Val Gly Val Leu Pro Gly Val Pro Thr Gly Ala Gly Val Lys Pro
100 105 110
Lys Ala Pro Gly Val Gly Gly Ala Phe Ala Gly Ile Pro Gly Val Gly
115 120 125
Pro Phe Gly Gly Pro Gln Pro Gly Val Pro Leu Gly Tyr Pro Ile Lys
130 135 140
Ala Pro Lys Leu Pro Gly Gly Tyr Gly Leu Pro Tyr Thr Thr Gly Lys
145 150 155 160
Leu Pro Tyr Gly Tyr Gly Pro Gly Gly Val Ala Ala Ala Gly Lys Ala
165 170 175
Gly Tyr Pro Thr Gly Thr Gly Val Gly Pro Gln Ala Ala Ala Ala Ala
180 185 190
Ala Ala Lys Ala Ala Ala Lys Phe Gly Ala Gly Ala Ala Gly Phe Gly
195 200 205

Ala Val Pro Gly Val Gly Gly Ala Gly Val Pro Gly Val Pro Gly Ala
210 215 220
Ile Pro Gly Ile Gly Gly Ile Ala Gly Val Gly Thr Pro Ala Ala Ala
225 230 235 240
Ala Ala Ala Ala Ala Ala Ala Lys Ala Ala Lys Tyr Gly Ala Ala Ala
245 250 255
Gly Leu Val Pro Gly Gly Pro Gly Phe Gly Pro Gly Val Val Gly Val
260 265 270
Pro Gly Phe Gly Ala Val Pro Gly Val Gly Val Pro Gly Ala Gly Ile
275 280 285
Pro Val Val Pro Gly Ala Gly Ile Pro Gly Ala Ala Gly Phe Gly Ala
290 295 300
Val Ser Pro Glu Ala Ala Ala Lys Ala Ala Ala Lys Ala Ala Lys Tyr
305 310 315 320
Gly Ala Arg Pro Gly Val Gly Val Gly Gly Ile Pro Thr Tyr Gly Val
325 330 335
Gly Ala Gly Phe Phe Pro Gly Phe Gly Val Gly Val Gly Gly Ile Pro
340 345 350
Gly Val Ala Gly Val Pro Ser Val Gly Gly Val Pro Gly Val Gly Gly
355 360 365
Val Pro Gly Val Gly Ile Ser Pro Glu Ala Gln Ala Ala Ala Ala Ala
370 375 380
Lys Ala Ala Lys Tyr Gly Val Gly Thr Pro Ala Ala Ala Ala Ala Lys
385 390 395 400
Ala Ala Ala Lys Ala Ala Gln Phe Gly Leu Val Pro Gly Val Gly Val
405 410 415
Ala Pro Gly Val Gly Val Ala Pro Gly Val Gly Val Ala Pro Gly Val
420 425 430
Gly Leu Ala Pro Gly Val Gly Val Ala Pro Gly Val Gly Val Ala Pro
435 440 445
Gly Val Gly Val Ala Pro Gly Ile Gly Pro Gly Gly Val Ala Ala Ala

450 455 460

Ala Lys Ser Ala Ala Lys Val Ala Ala Lys Ala Gln Leu Arg Ala Ala

465 470 475 480

Ala Gly Leu Gly Ala Gly Ile Pro Gly Leu Gly Val Gly Val Gly Val

485 490 495

Pro Gly Leu Gly Val Gly Ala Gly Val Pro Gly Leu Gly Val Gly Ala

500 505 510

Gly Val Pro Gly Phe Gly Ala Val Pro Gly Ala Leu Ala Ala Ala Lys

515 520 525

Ala Ala Lys Tyr Gly Ala Val Pro Gly Val Leu Gly Gly Leu Gly Ala

530 535 540

Leu Gly Gly Val Gly Ile Pro Gly Gly Val Val Gly Ala Gly Pro Ala

545 550 555 560

Ala Ala Ala Ala Ala Ala Lys Ala Ala Ala Lys Ala Ala Gln Phe Gly

565 570 575

Leu Val Gly Ala Ala Gly Leu Gly Gly Leu Gly Val Gly Gly Leu Gly

580 585 590

Val Pro Gly Val Gly Gly Leu Gly Gly Ile Pro Pro Ala Ala Ala Ala

595 600 605

Lys Ala Ala Lys Tyr Gly Ala Ala Gly Leu Gly Gly Val Leu Gly Gly

610 615 620

Ala Gly Gln Phe Pro Leu Gly Gly Val Ala Ala Arg Pro Gly Phe Gly

625 630 635 640

Leu Ser Pro Ile Phe Pro Gly Gly Ala Cys Leu Gly Lys Ala Cys Gly

645 650 655

Arg Lys Arg Lys

660

INFORMATION FOR SEQ ID NO:6:

SEQUENCE CHARACTERISTICS:

LENGTH: 441 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: YES

ANTI-SENSE: NO

FRAGMENT TYPE:

ORIGINAL SOURCE:

IMMEDIATE SOURCE;

POSITION IN GENOME:

CHROMOSOME/SEGMENT:

MAP POSITION:

UNITS:

FEATURE

NAME/KEY:

LOCATION:

IDENTIFICATION METHOD:

OTHER INFORMATION:

PUBLICATION INFORMATION

AUTHORS:

TITLE:

JOURNAL:

VOLUME:

ISSUE:

PAGES:

DATA:

DOCUMENT NUMBER: WO 99/03886

FILING DATE: 17 JULY 1998

PUBLICATION DATE: 28 JANUARY 1999

RELEVANT RESIDUES IN SEQ ID NO.:

SEQUENCE DESCRIPTION: SEQ ID NO:6:

TCCGCCATGG GAGGTGTTCC GGGCGCGCTG GCTGCTGCGA AAGCGGCGAA ATACGGTGCA 60

GCGGTTCCGG GTGTACTGGG CGGTCTGGGT GCTCTGGGCG GTGTTGGTAT CCCGGGCGGT 120

GTTGTAGGTG CAGGCCCAGC TGCAGCTGCT GCTGCGGCAA AGGCAGCGGC GAAAGCAGCT 180

CAGTTCGGTC TGGTTGGTGC AGCAGGTGTG GGCGGTCTGG GTGTTGGCGG TCTGGGTGTA 240

CCGGGCGTTG GTGGTCTGGG TGGCATCCCG CCGGCGGCGG CAGCTAAAGC GGCTAAATAC 300

GGTGCAGCAG GTCTGGGTGG CGTTCTGGGT GGTGCTGGTC AGTTCCCACT GGGCGGTGTA 360

GCGGCACGTC CGGGTTTCGG TCTGTCCCCG ATCTTCCCAG GCGGTGCATG CCTGGGTAAA 420

GCTTGCGGCC GTAAACGTAA A 441

INFORMATION FOR SEQ ID NO:7:

SEQUENCE CHARACTERISTICS:

LENGTH: 147 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL:

ANTI-SENSE:

FRAGMENT TYPE:

ORIGINAL SOURCE:

IMMEDIATE SOURCE;

POSITION IN GENOME:

CHROMOSOME/SEGMENT:

MAP POSITION:

UNITS:

FEATURE

NAME/KEY:

LOCATION:

IDENTIFICATION METHOD:

OTHER INFORMATION:

PUBLICATION INFORMATION

AUTHORS:

TITLE:

JOURNAL:

VOLUME:

ISSUE:

PAGES:

DATA:

DOCUMENT NUMBER: WO 99/03886

FILING DATE: 17 JULY 1998

PUBLICATION DATE: 28 JANUARY 1999

RELEVANT RESIDUES IN SEQ ID NO.:

SEQUENCE DESCRIPTION: SEQ ID NO:7:

Ser	Ala	Met	Gly	Gly	Val	Pro	Gly	Ala	Leu	Ala	Ala	Ala	Lys	Ala	Ala
1				5					10					15	
Lys	Tyr	Gly	Ala	Ala	Val	Pro	Gly	Val	Leu	Gly	Gly	Leu	Gly	Ala	Leu
			20					25					30		
Gly	Gly	Val	Gly	Ile	Pro	Gly	Gly	Val	Val	Gly	Ala	Gly	Pro	Ala	Ala
		35					40					45			

Ala Ala Ala Ala Ala Lys Ala Ala Ala Lys Ala Ala Gln Phe Gly Leu
50 55 60
Val Gly Ala Ala Gly Leu Gly Gly Leu Gly Val Gly Gly Leu Gly Val
65 70 75 80
Pro Gly Val Gly Gly Leu Gly Gly Ile Pro Pro Ala Ala Ala Ala Lys
85 90 95
Ala Ala Lys Tyr Gly Ala Ala Gly Leu Gly Gly Val Leu Gly Gly Ala
100 105 110
Gly Gln Phe Pro Leu Gly Gly Val Ala Ala Arg Pro Gly Phe Gly Leu
115 120 125
Ser Pro Ile Phe Pro Gly Gly Ala Cys Leu Gly Lys Ala Cys Gly Arg
130 135 140
Lys Arg Lys
145

INFORMATION FOR SEQ ID NO:8:

SEQUENCE CHARACTERISTICS:

LENGTH: 600 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: YES

ANTI-SENSE: NO

FRAGMENT TYPE:

ORIGINAL SOURCE:

IMMEDIATE SOURCE;

POSITION IN GENOME:

CHROMOSOME/SEGMENT:

MAP POSITION:

UNITS:

FEATURE

NAME/KEY:

LOCATION:

IDENTIFICATION METHOD:

OTHER INFORMATION:

PUBLICATION INFORMATION

AUTHORS:

TITLE:

JOURNAL:

VOLUME:

ISSUE:

PAGES:

DATA:

DOCUMENT NUMBER: WO 99/03886

FILING DATE: 17 JULY 1998

PUBLICATION DATE: 28 JANUARY 1999

RELEVANT RESIDUES IN SEQ ID NO.:

SEQUENCE DESCRIPTION: SEQ ID NO:8:

TCCGCCATGG GAGCTCTGGT AGGCCTGGGC GTACCGGGCC TGGGTGTTGG TGCAGGCGTT 60

CCGGGTTTCG GTGCTGGCGC GGACGAAGGT GTACGTCGTT CCCTGTCTCC AGAACTGCGT 120

GAAGGTGACC CGTCCTCTTC CCAGCACCTG CCGTCTACCC CGTCCTCTCC ACGTGTTCCG 180

GGCGCGCTGG CTGCTGCGAA AGCGGCGAAA TACGGTGCAG CGGTTCCGGG TGTACTGGGC 240

GGTCTGGGTG CTCTGGGCGG TGTTGGTATC CCGGGCGGTG TTGTAGGTGC AGGCCCAGCT 300

GCAGCTGCTG CTGCGGCAAA GGCAGCGGCG AAAGCAGCTC AGTTCGGTCT GGTTGGTGCA 360

GCAGGTCTGG GCGGTCTGGG TGTTGGCGGT CTGGGTGTAC CGGGCGTTGG TGGTCTGGGT 420

GGCATCCCGC CGGCGGCGGC AGCTAAAGCG GCTAAATACG GTGCAGCAGG TCTGGGTGGC 480

GTTCTGGGTG GTGCTGGTCA GTTCCCACTG GGCGGTGTAG CGGCACGTCC GGGTTTCGGT 540

CTGTCCCCGA TCTTCCCAGG CGGTGCATGC CTGGGTAAAG CTTGCGGCCG TAAACGTAAA 600

INFORMATION FOR SEQ ID NO:9:

SEQUENCE CHARACTERISTICS:

LENGTH: 200 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL:

ANTI-SENSE:

FRAGMENT TYPE:

ORIGINAL SOURCE:

IMMEDIATE SOURCE;

POSITION IN GENOME:

CHROMOSOME/SEGMENT:

MAP POSITION:

UNITS:

FEATURE

NAME/KEY:

LOCATION:

IDENTIFICATION METHOD:

OTHER INFORMATION:

PUBLICATION INFORMATION

AUTHORS:

TITLE:

JOURNAL:

VOLUME:

ISSUE:

PAGES:

DATA:

DOCUMENT NUMBER: WO 99/03886

FILING DATE: 17 JULY 1998

PUBLICATION DATE: 28 JANUARY 1999

RELEVANT RESIDUES IN SEQ ID NO.:

SEQUENCE DESCRIPTION: SEQ ID NO:9:

Ser	Ala	Met	Gly	Ala	Leu	Val	Gly	Leu	Gly	Val	Pro	Gly	Leu	Gly	Val
1				5				10						15	
Gly	Ala	Gly	Val	Pro	Gly	Phe	Gly	Ala	Gly	Ala	Asp	Glu	Gly	Val	Arg
			20					25					30		
Arg	Ser	Leu	Ser	Pro	Glu	Leu	Arg	Glu	Gly	Asp	Pro	Ser	Ser	Ser	Gln
		35					40					45			
His	Leu	Pro	Ser	Thr	Pro	Ser	Ser	Pro	Arg	Val	Pro	Gly	Ala	Leu	Ala
	50					55					60				
Ala	Ala	Lys	Ala	Ala	Lys	Tyr	Gly	Ala	Ala	Val	Pro	Gly	Val	Leu	Gly
65					70					75					80

Gly Leu Gly Ala Leu Gly Gly Val Gly Ile Pro Gly Gly Val Val Gly
85 90 95
Ala Gly Pro Ala Ala Ala Ala Ala Ala Ala Lys Ala Ala Ala Lys Ala
100 105 110
Ala Gln Phe Gly Leu Val Gly Ala Ala Gly Leu Gly Gly Leu Gly Val
115 120 125
Gly Gly Leu Gly Val Pro Gly Val Gly Gly Leu Gly Gly Ile Pro Pro
130 135 140
Ala Ala Ala Ala Lys Ala Ala Lys Tyr Gly Ala Ala Gly Leu Gly Gly
145 150 155 160
Val Leu Gly Gly Ala Gly Gln Phe Pro Leu Gly Gly Val Ala Ala Arg
165 170 175
Pro Gly Phe Gly Leu Ser Pro Ile Phe Pro Gly Gly Ala Cys Leu Gly
180 185 190
Lys Ala Cys Gly Arg Lys Arg Lys
195 200

INFORMATION FOR SEQ ID NO:10:

SEQUENCE CHARACTERISTICS:

LENGTH: 60 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL:

ANTI-SENSE:

FRAGMENT TYPE:

ORIGINAL SOURCE:

IMMEDIATE SOURCE;

POSITION IN GENOME:

CHROMOSOME/SEGMENT:

MAP POSITION:

UNITS:

FEATURE

NAME/KEY:

LOCATION:

IDENTIFICATION METHOD:

OTHER INFORMATION:

PUBLICATION INFORMATION

AUTHORS:

TITLE:

JOURNAL:

VOLUME:

ISSUE:

PAGES:

DATA:

DOCUMENT NUMBER: WO 99/03886

FILING DATE: 17 JULY 1998

PUBLICATION DATE: 28 JANUARY 1999

RELEVANT RESIDUES IN SEQ ID NO.:

SEQUENCE DESCRIPTION: SEQ ID NO:10:

Gly	Ile	Pro	Pro	Ala	Ala	Ala	Ala	Lys	Ala	Ala	Lys	Tyr	Gly	Ala	Ala
1				5					10					15	
Gly	Leu	Gly	Gly	Val	Leu	Gly	Gly	Ala	Gly	Gln	Phe	Pro	Leu	Gly	Gly

20 25 30

Val Ala Ala Arg Pro Gly Phe Gly Leu Ser Pro Ile Phe Pro Gly Gly

35 40 45

Ala Cys Leu Gly Lys Ala Cys Gly Arg Lys Arg Lys

50 55 60

INFORMATION FOR SEQ ID NO:11:

SEQUENCE CHARACTERISTICS:

LENGTH: 47 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL:

ANTI-SENSE:

FRAGMENT TYPE:

ORIGINAL SOURCE:

IMMEDIATE SOURCE;

POSITION IN GENOME:

CHROMOSOME/SEGMENT:

MAP POSITION:

UNITS:

FEATURE

NAME/KEY:

LOCATION:

IDENTIFICATION METHOD:

OTHER INFORMATION:

PUBLICATION INFORMATION

AUTHORS:

TITLE:

JOURNAL:

VOLUME:

ISSUE:

PAGES:

DATA:

DOCUMENT NUMBER: WO 99/03886

FILING DATE: 17 JULY 1998

PUBLICATION DATE: 28 JANUARY 1999

RELEVANT RESIDUES IN SEQ ID NO.:

SEQUENCE DESCRIPTION: SEQ ID NO:11:

Gly	Ala	Ala	Gly	Leu	Gly	Gly	Val	Leu	Gly	Gly	Ala	Gly	Gln	Phe	Pro
1				5					10					15	
Leu	Gly	Gly	Val	Ala	Ala	Arg	Pro	Gly	Phe	Gly	Leu	Ser	Pro	Ile	Phe
			20					25					30		
Pro	Gly	Gly	Ala	Cys	Leu	Gly	Lys	Ala	Cys	Gly	Arg	Lys	Arg	Lys	
		35					40					45			

INFORMATION FOR SEQ ID NO:12:

SEQUENCE CHARACTERISTICS:

LENGTH: 34 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL:

ANTI-SENSE:

FRAGMENT TYPE:

ORIGINAL SOURCE:

IMMEDIATE SOURCE;

POSITION IN GENOME:

CHROMOSOME/SEGMENT:

MAP POSITION:

UNITS:

FEATURE

NAME/KEY:

LOCATION:

IDENTIFICATION METHOD:

OTHER INFORMATION:

PUBLICATION INFORMATION

AUTHORS:

TITLE:

JOURNAL:

VOLUME:

ISSUE:

PAGES:

DATA:

DOCUMENT NUMBER: WO 99/03886

FILING DATE: 17 JULY 1998

PUBLICATION DATE: 28 JANUARY 1999

RELEVANT RESIDUES IN SEQ ID NO.:

SEQUENCE DESCRIPTION: SEQ ID NO:12:

Gly Ala Asp Glu Gly Val Arg Arg Ser Leu Ser Pro Glu Leu Arg Glu
1 5 10 15

Gly Asp Pro Ser Ser Ser Gln His Leu Pro Ser Thr Pro Ser Ser Pro
20 25 30

Arg Val

INFORMATION FOR SEQ ID NO:13:

SEQUENCE CHARACTERISTICS:

LENGTH: 34 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL:

ANTI-SENSE:

FRAGMENT TYPE:

ORIGINAL SOURCE:

IMMEDIATE SOURCE;

POSITION IN GENOME:

CHROMOSOME/SEGMENT:

MAP POSITION:

UNITS:

FEATURE

NAME/KEY:

LOCATION:

IDENTIFICATION METHOD:

OTHER INFORMATION:

PUBLICATION INFORMATION

AUTHORS:

TITLE:

JOURNAL:

VOLUME:

ISSUE:

PAGES:

DATA:

DOCUMENT NUMBER: WO 99/03886

FILING DATE: 17 JULY 1998

PUBLICATION DATE: 28 JANUARY 1999

RELEVANT RESIDUES IN SEQ ID NO.:

SEQUENCE DESCRIPTION: SEQ ID NO:13:

Gly Ala Asp Glu Gly Val Arg Arg Ser Leu Ser Pro Glu Leu Arg Glu
1 5 10 15
Gly Asp Pro Ser Ser Ser Gln His Leu Pro Ser Thr Pro Ser Ser Pro
20 25 30
Arg Phe

INFORMATION FOR SEQ ID NO:14:

SEQUENCE CHARACTERISTICS:

LENGTH: 216 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL:

ANTI-SENSE:

FRAGMENT TYPE:

ORIGINAL SOURCE:

IMMEDIATE SOURCE;

POSITION IN GENOME:

CHROMOSOME/SEGMENT:

MAP POSITION:

UNITS:

FEATURE

NAME/KEY:

LOCATION:

IDENTIFICATION METHOD:

OTHER INFORMATION:

PUBLICATION INFORMATION

AUTHORS:

TITLE:

JOURNAL:

VOLUME:

ISSUE:

PAGES:

DATA:

DOCUMENT NUMBER: WO 99/03886

FILING DATE: 17 JULY 1998

PUBLICATION DATE: 28 JANUARY 1999

RELEVANT RESIDUES IN SEQ ID NO.:

SEQUENCE DESCRIPTION: SEQ ID NO:14:

Ala Ala Ala Gly Leu Gly Ala Gly Ile Pro Gly Leu Gly Val Gly Val
1 5 10 15
Gly Val Pro Gly Leu Gly Val Gly Ala Gly Val Pro Gly Leu Gly Val
20 25 30
Gly Ala Gly Val Pro Gly Phe Gly Ala Gly Ala Asp Glu Gly Val Arg
35 40 45
Arg Ser Leu Ser Pro Glu Leu Arg Glu Gly Asp Pro Ser Ser Ser Gln
50 55 60
His Leu Pro Ser Thr Pro Ser Ser Pro Arg Val Pro Gly Ala Leu Ala
65 70 75 80
Ala Ala Lys Ala Ala Lys Tyr Gly Ala Ala Val Pro Gly Val Leu Gly
85 90 95
Gly Leu Gly Ala Leu Gly Gly Val Gly Ile Pro Gly Gly Val Val Gly
100 105 110
Ala Gly Pro Ala Ala Ala Ala Ala Ala Ala Lys Ala Ala Ala Lys Ala
115 120 125
Ala Gln Phe Gly Leu Val Gly Ala Ala Gly Leu Gly Gly Leu Gly Val
130 135 140
Gly Gly Leu Gly Val Pro Gly Val Gly Gly Leu Gly Gly Ile Pro Pro
145 150 155 160
Ala Ala Ala Ala Lys Ala Ala Lys Tyr Gly Ala Ala Gly Leu Gly Gly

```
                    165                 170                 175
Val Leu Gly Gly Ala Gly Gln Phe Pro Leu Gly Gly Val Ala Ala Arg
                180                 185                 190
Pro Gly Phe Gly Leu Ser Pro Ile Phe Pro Gly Gly Ala Cys Leu Gly
        195                 200                 205
Lys Ala Cys Gly Arg Lys Arg Lys
  210                 215
```

INFORMATION FOR SEQ ID NO:15:

SEQUENCE CHARACTERISTICS:

LENGTH: 183 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL:

ANTI-SENSE:

FRAGMENT TYPE:

ORIGINAL SOURCE:

IMMEDIATE SOURCE;

POSITION IN GENOME:

CHROMOSOME/SEGMENT:

MAP POSITION:

UNITS:

FEATURE

NAME/KEY:

LOCATION:

IDENTIFICATION METHOD:

OTHER INFORMATION:

PUBLICATION INFORMATION

AUTHORS:

TITLE:

JOURNAL:

VOLUME:

ISSUE:

PAGES:

DATA:

DOCUMENT NUMBER: WO 99/03886

FILING DATE: 17 JULY 1998

PUBLICATION DATE: 28 JANUARY 1999

RELEVANT RESIDUES IN SEQ ID NO.:

SEQUENCE DESCRIPTION: SEQ ID NO:15:

Ala Ala Ala Gly Leu Gly Ala Gly Ile Pro Gly Leu Gly Val Gly Val
1 5 10 15

Gly Val Pro Gly Leu Gly Val Gly Ala Gly Val Pro Gly Leu Gly Val
20 25 30

Gly Ala Gly Val Pro Gly Phe Gly Ala Val Pro Gly Ala Leu Ala Ala
35 40 45

Ala Lys Ala Ala Lys Tyr Gly Ala Ala Val Pro Gly Val Leu Gly Gly
50 55 60

Leu Gly Ala Leu Gly Gly Val Gly Ile Pro Gly Gly Val Val Gly Ala
65 70 75 80

Gly Pro Ala Ala Ala Ala Ala Ala Ala Lys Ala Ala Ala Lys Ala Ala
85 90 95
Gln Phe Gly Leu Val Gly Ala Ala Gly Leu Gly Gly Leu Gly Val Gly
100 105 110
Gly Leu Gly Val Pro Gly Val Gly Gly Leu Gly Gly Ile Pro Pro Ala
115 120 125
Ala Ala Ala Lys Ala Ala Lys Tyr Gly Ala Ala Gly Leu Gly Gly Val
130 135 140
Leu Gly Gly Ala Gly Gln Phe Pro Leu Gly Gly Val Ala Ala Arg Pro
145 150 155 160
Gly Phe Gly Leu Ser Pro Ile Phe Pro Gly Gly Ala Cys Leu Gly Lys
165 170 175
Ala Cys Gly Arg Lys Arg Lys
180

We Claim:

1. A synthetic polynucleotide encoding a tropoelastin derivative, wherein the polynucleotide has a nucleotide sequence of:

nucleotide position 1 to 1676 of SEQ ID NO: 1 contiguous with 1775 to 2210 of SEQ ID NO: 1; or

nucleotide position 1554 to 1676 of SEQ ID NO: 1 contiguous with 1776 to 2210 of SEQ ID NO: 1.

2. A vector comprising a polynucleotide according to claim 1.

3. The vector according to claim 2 wherein the polynucleotide is operatively linked to a promoter or enhancer regulatory sequence.

4. A cell containing the vector according to claim 2 or 3.

5. A method for producing a derivative of tropoelastin, the method comprising:

a) providing a vector comprising a polynucleotide according to claim 1;

b) introducing the vector into a cell;

c) maintaining the cell in conditions suitable for expression of the tropoelastin; and

d) isolating the tropoelastin derivative.

6. A tropoelastin derivative produced by the method of claim 5.

7. A formulation comprising the tropoelastin derivative according to claim 6, together with a pharmaceutically acceptable carrier or diluent.

```
1   GATCCATGGGTGGCGTTCCGGGTGCTATCCCGGGTGGCGTTCCGGGTGGTGTATTCTACC   60
    GTACCCACCGCAAGGCCCACGATAGGGCCCACCGCAAGGCCCACCACATAAGATGG

      S  M  G  G  V  P  G  A  I  P  G  G  V  P  G  G  V  F  Y  P

61  CAGGCGCGGGTCTGGGTGCACTGGGCGGTGGTGCGCTGGGCCCGGGTGGTAAACCGCTGA   120
    GTCCGCGCCCAGACCCACGTGACCCGCCACCACGCGACCCGGGCCCACCATTTGGCGACT

      G  A  G  L  G  A  L  G  G  G  A  L  G  P  G  G  K  P  L  K

121 AACCGGTTCCAGGCGGTCTGGCAGGTGCTGGTCTGGGTGCAGGTCTGGGCGCGTTCCCGG   180
    TTGGCCAAGGTCCGCCAGACCGTCCACGACCAGACCCACGTCCAGACCCGCGCAAGGGCC

      P  V  P  G  G  L  A  G  A  G  L  G  A  G  L  G  A  F  P  A

181 CGGTTACCTTCCCGGGTGCTCTGGTTCCGGGTGGCGTTGCAGACGCAGCTGCTGCGTACA   240
    GCCAATGGAAGGGCCCACGAGACCAAGGCCCACCGCAACGTCTGCGTCGACGACGCATGT

      V  T  F  P  G  A  L  V  P  G  G  V  A  D  A  A  A  A  Y  K

241 AAGCGGCAAAGGCAGGTGCGGGTCTGGGCGGGGTACCAGGTGTTGGCGGTCTGGGTGTAT   300
    TTCGCCGTTTCCGTCCACGCCCAGACCCGCCCCATGGTCCACAACCGCCAGACCCACATA

      A  A  K  A  G  A  G  L  G  G  V  P  G  V  G  G  L  G  V  S

301 CTGCTGGCGCAGTTGTTCCGCAGCCGGGTGCAGGTGTAAAACCGGGCAAAGTTCCAGGTG   360
    GACGACCGCGTCAACAAGGCGTCGGCCCACGTCCACATTTTGGCCCGTTTCAAGGTCCAC

      A  G  A  V  V  P  Q  P  G  A  G  V  K  P  G  K  V  P  G  V

361 TTGGTCTGCCGGGCGTATACCCGGGTGGTGTTCTGCCGGGCGCGCGTTTCCCAGGTGTTG   420
    AACCAGACGGCCCGCATATGGGCCCACCACAAGACGGCCCGCGCGCAAAGGGTCCACAAC

      G  L  P  G  V  Y  P  G  G  V  L  P  G  A  R  F  P  G  V  G
```

Figure 1(1)

```
421  GTGTACTGCCGGGCGTTCCGACCGGTGCAGGTGTTAAACCGAAGGCACCAGGTGTAGGCG   480
     CACATGACGGCCCGCAAGGCTGGCCACGTCCACAATTTGGCTTCCGTGGTCCACATCCGC
       V  L  P  G  V  P  T  G  A  G  V  K  P  K  A  P  G  V  G  G

481  GCGCGTTCGCGGGTATCCCGGGTGTTGGCCCGTTCGGTGGTCCGCAGCCAGGCGTTCCGC   540
     CGCGCAAGCGCCCATAGGGCCCACAACCGGGCAAGCCACCAGGCGTCGGTCCGCAAGGCG
       A  F  A  G  I  P  G  V  G  P  F  G  G  P  Q  P  G  V  P  L

541  TGGGTTACCCGATCAAAGCGCCGAAGCTTCCAGGTGGCTACGGTCTGCCGTACACCACCG   600
     ACCCAATGGGCTAGTTTCGCGGCTTCGAAGGTCCACCGATGCCAGACGGCATGTGGTGGC
       G  Y  P  I  K  A  P  K  L  P  G  G  Y  G  L  P  Y  T  T  G

601  GTAAACTGCCGTACGGCTACGGTCCGGGTGGCGTAGCAGGTGCTGCGGGTAAAGCAGGCT   660
     CATTTGACGGCATGCCGATGCCAGGCCCACCGCATCGTCCACGACGCCCATTTCGTCCGA
       K  L  P  Y  G  Y  G  P  G  G  V  A  G  A  A  G  K  A  G  Y

661  ACCCAACCGGTACTGGTGTTGGTCCGCAGGCTGCTGCGGCAGCTGCGGCGAAGGCAGCAG   720
     TGGGTTGGCCATGACCACAACCAGGCGTCCGACGACGCCGTCGACGCCGCTTCCGTCGTC
       P  T  G  T  G  V  G  P  Q  A  A  A  A  A  A  A  K  A  A  A

721  CAAAATTCGGCGCGGGTGCAGCGGGTGTTCTGCCGGGCGTAGGTGGTGCTGGCGTTCCGG   780
     GTTTTAAGCCGCGCCCACGTCGCCCACAAGACGGCCCGCATCCACCACGACCGCAAGGCC
       K  F  G  A  G  A  A  G  V  L  P  G  V  G  G  A  G  V  P  G

781  GTGTTCCAGGTGCGATCCCGGGCATCGGTGGTATCGCAGGCGTAGGTACTCCGGCGGCCG   840
     CACAAGGTCCACGCTAGGGCCCGTAGCCACCATAGCGTCCGCATCCATGAGGCCGCCGGC
       V  P  G  A  I  P  G  I  G  G  I  A  G  V  G  T  P  A  A  A

841  CTGCGGCTGCGGCAGCTGCGGCGAAAGCAGCTAAATACGGTGCGGCAGCAGGCCTGGTTC   900
     GACGCCGACGCCGTCGACGCCGCTTTCGTCGATTTATGCCACGCCGTCGTCCGGACCAAG
       A  A  A  A  A  A  A  K  A  A  K  Y  G  A  A  A  G  L  V  P
```

Figure 1(2)

```
901   CGGGTGGTCCAGGCTTCGGTCCGGGTGTTGTAGGCGTTCCGGGTGCTGGTGTTCCGGGCG   960
      GCCCACCAGGTCCGAAGCCAGGCCCACAACATCCGCAAGGCCCACGACCACAAGGCCCGC

        G   G   P   G   F   G   P   G   V   V   G   V   P   G   A   G   V   P   G   V

961   TAGGTGTTCCAGGTGCGGGCATCCCGGTTGTACCGGGTGCAGGTATCCCGGGCGCTGCGG   1020
      ATCCACAAGGTCCACGCCCGTAGGGCCAACATGGCCCACGTCCATAGGGCCCGCGACGCC

        G   V   P   G   A   G   I   P   V   V   P   G   A   G   I   P   G   A   A   V

1021  TTCCAGGTGTTGTATCCCCGGAAGCGGCAGCTAAGGCTGCTGCGAAAGCTGCGAAATACG   1080
      AAGGTCCACAACATAGGGGCCTTCGCCGTCGATTCCGACGACGCTTTCGACGCTTTATGC

        P   G   V   V   S   P   E   A   A   A   K   A   A   A   K   A   A   K   Y   G

1081  GAGCTCGTCCGGGCGTTGGTGTTGGTGGCATCCCGACCTACGGTGTAGGTGCAGGCGGTT   1140
      CTCGAGCAGGCCCGCAACCACAACCACCGTAGGGCTGGATGCCACATCCACGTCCGCCAA

        A   R   P   G   V   G   V   G   G   I   P   T   Y   G   V   G   A   G   G   F

1141  TCCCAGGTTTCGGCGTTGGTGTTGGTGGCATCCCGGGTGTAGCTGGTGTTCCGTCTGTTG   1200
      AGGGTCCAAAGCCGCAACCACAACCACCGTAGGGCCCACATCGACCACAAGGCAGACAAC

        P   G   F   G   V   G   V   G   G   I   P   G   V   A   G   V   P   S   V   G

1201  GTGGCGTACCGGGTGTTGGTGGCGTTCCAGGTGTAGGTATCTCCCCGGAAGCGCAGGCAG   1260
      CACCGCATGGCCCACAACCACCGCAAGGTCCACATCCATAGAGGGGCCTTCGCGTCCGTC

        G   V   P   G   V   G   G   V   P   G   V   G   I   S   P   E   A   Q   A   A

1261  CTGCGGCAGCTAAAGCAGCGAAGTACGGCGTTGGTACTCCGGCGGCAGCAGCTGCTAAAG   1320
      GACGCCGTCGATTTCGTCGCTTCATGCCGCAACCATGAGGCCGCCGTCGTCGACGATTTC

        A   A   A   K   A   A   K   Y   G   V   G   T   P   A   A   A   A   A   K   A

1321  CAGCGGCTAAAGCAGCGCAGTTCGGACTAGTTCCGGGCGTAGGTGTTGCGCCAGGTGTTG   1380
      GTCGCCGATTTCGTCGCGTCAAGCCTGATCAAGGCCCGCATCCACAACGCGGTCCACAAC

        A   A   K   A   A   Q   F   G   L   V   P   G   V   G   V   A   P   G   V   G
```

Figure 1(3)

```
1381  GCGTAGCACCGGGTGTTGGTGTTGCTCCGGGCGTAGGTCTGGCACCGGGTGTTGGCGTTG   1440
      CGCATCGTGGCCCACAACCACAACGAGGCCCGCATCCAGACCGTGGCCCACAACCGCAAC

        V   A   P   G   V   G   V   A   P   G   V   G   L   A   P   G   V   G   V   A

1441  CACCAGGTGTAGGTGTTGCGCCGGGCGTTGGTGTAGCACCGGGTATCGGTCCGGGTGGCG   1500
      GTGGTCCACATCCACAACGCGGCCCGCAACCACATCGTGGCCCATAGCCAGGCCCACCGC

        P   G   V   G   V   A   P   G   V   G   V   A   P   G   I   G   P   G   G   V

1501  TTGCGGCTGCTGCGAAATCTGCTGCGAAGGTTGCTGCGAAAGCGCAGCTGCGTGCAGCAG   1560
      AACGCCGACGACGCTTTAGACGACGCTTCCAACGACGCTTTCGCGTCGACGCACGTCGTC

        A   A   A   A   K   S   A   A   K   V   A   A   K   A   Q   L   R   A   A   A

1561  CTGGTCTGGGTGCGGGCATCCCAGGTCTGGGTGTAGGTGTTGGTGTTCCGGGCCTGGGTG   1620
      GACCAGACCCACGCCCGTAGGGTCCAGACCCACATCCACAACCACAAGGCCCGGACCCAC

        G   L   G   A   G   I   P   G   L   G   V   G   V   G   V   P   G   L   G   V

1621  TAGGTGCAGGGGTACCGGGCCTGGGTGTTGGTGCAGGCGTTCCGGGTTTCGGTGCTGGCG   1680
      ATCCACGTCCCCATGGCCCGGACCCACAACCACGTCCGCAAGGCCCAAAGCCACGACCGC

        G   A   G   V   P   G   L   G   V   G   A   G   V   P   G   F   G   A   G   A

1681  CGGACGAAGGTGTACGTCGTTCCCTGTCTCCAGAACTGCGTGAAGGTGACCCGTCCTCTT   1740
      GCCTGCTTCCACATGCAGCAAGGGACAGAGGTCTTGACGCACTTCCACTGGGCAGGAGAA

        D   E   G   V   R   R   S   L   S   P   E   L   R   E   G   D   P   S   S   S

1741  CCCAGCACCTGCCGTCTACCCCGTCCTCTCCACGTGTTCCGGGCGCGCTGGCTGCTGCGA   1800
      GGGTCGTGGACGGCAGATGGGGCAGGAGAGGTGCACAAGGCCCGCGCGACCGACGACGCT

        Q   H   L   P   S   T   P   S   S   P   R   V   P   G   A   L   A   A   A   K

1801  AAGCGGCGAAATACGGTGCAGCGGTTCCGGGTGTACTGGGCGGTCTGGGTGCTCTGGGCG   1860
      TTCGCCGCTTTATGCCACGTCGCCAAGGCCCACATGACCCGCCAGACCCACGAGACCCGC

        A   A   K   Y   G   A   A   V   P   G   V   L   G   G   L   G   A   L   G   G
```

Figure 1(4)

```
1861  GTGTTGGTATCCCGGGCGGTGTTGTAGGTGCAGGCCCAGCTGCAGCTGCTGCTGCGGCAA   1920
      CACAACCATAGGGCCCGCCACAACATCCACGTCCGGGTCGACGTCGACGACGACGCCGTT
        V  G  I  P  G  G  V  V  G  A  G  P  A  A  A  A  A  A  A  K

1921  AGGCAGCGGCGAAAGCAGCTCAGTTCGGTCTGGTTGGTGCAGCAGGTCTGGGCGGTCTGG   1980
      TCCGTCGCCGCTTTCGTCGAGTCAAGCCAGACCAACCACGTCGTCCAGACCCGCCAGACC
        A  A  A  K  A  A  Q  F  G  L  V  G  A  A  G  L  G  G  L  G

1981  GTGTTGGCGGTCTGGGTGTACCGGGCGTTGGTGGTCTGGGTGGCATCCCGCCGGCGGCGG   2040
      CACAACCGCCAGACCCACATGGCCCGCAACCACCAGACCCACCGTAGGGCGGCCGCCGCC
        V  G  G  L  G  V  P  G  V  G  G  L  G  G  I  P  P  A  A  A

2041  CAGCTAAAGCGGCTAAATACGGTGCAGCAGGTCTGGGTGGCGTTCTGGGTGGTGCTGGTC   2100
      GTCGATTTCGCCGATTTATGCCACGTCGTCCAGACCCACCGCAAGACCCACCACGACCAG
        A  K  A  A  K  Y  G  A  A  G  L  G  G  V  L  G  G  A  G  Q

2101  AGTTCCCACTGGGCGGTGTAGCGGCACGTCCGGGTTTCGGTCTGTCCCCGATCTTCCCAG   2160
      TCAAGGGTGACCCGCCACATCGCCGTGCAGGCCCAAAGCCAGACAGGGGCTAGAAGGGTC
        F  P  L  G  G  V  A  A  R  P  G  F  G  L  S  P  I  F  P  G

2161  GCGGTGCATGCCTGGGTAAAGCTTGCGGCCGTAAACGTAAATAATGATAG             2210
      CGCCACGTACGGACCCATTTCGAACGCCGGCATTTGCATTTATTACTATCCTAG
        G  A  C  L  G  K  A  C  G  R  K  R  K  *  *  *
```

Figure 1(5)

```
1 GGVPGAIPGGVPGGVFYPGAGLGALGGGALGPGGKPLKPVPGGLAGAGLG 50
  ||||||||||||||||||||||||||||||||||||||||||||||||||
1 GGVPGAIPGGVPGGVFYPGAGLGALGGGALGPGGKPLKPVPGGLAGAGLG 50

51 AGLGAFPAVTFPGALVPGGVADAAAAYKAAKAGAGLGGVPGVGGLGVSAG 100
   ||||||||||||||||||||||||||||||||||||||||||||||||||
51 AGLGAFPAVTFPGALVPGGVADAAAAYKAAKAGAGLGGVPGVGGLGVSAG 100

101 AVVPQPGAGVKPGKVPGVGLPGVYPGGVLPGARFPGVGVLPGVPTGAGVK 150
    ||||||||||||||||||||||||||||||||||||||||||||||||||
101 AVVPQPGAGVKPGKVPGVGLPGVYPGGVLPGARFPGVGVLPGVPTGAGVK 150

151 PKAPGVGGAFAGIPGVGPFGGPQPGVPLGYPIKAPKLPGGYGLPYTTGKL 200
    ||||||||||||||||||||||||||||||||||||||||||||||||||
151 PKAPGVGGAFAGIPGVGPFGGPQPGVPLGYPIKAPKLPGGYGLPYTTGKL 200

201 PYGYGPGGVAGAAGKAGYPTGTGVGPQAAAAAAAKAAAKFGAGAAGVLPG 250
    ||||||||||||||||||||||||||||||||||||||||||||||||||
201 PYGYGPGGVAGAAGKAGYPTGTGVGPQAAAAAAAKAAAKFGAGAAGVLPG 250

251 VGGAGVPGVPGAIPGIGGIAGVGTPAAAAAAAAAAKAAKYGAAAGLVPGG 300
    ||||||||||||||||||||||||||||||||||||||||||||||||||
251 VGGAGVPGVPGAIPGIGGIAGVGTPAAAAAAAAAAKAAKYGAAAGLVPGG 300

301 PGFGPGVVGVPGAGVPGVGVPGAGIPVVPGAGIPGAAVPGVVSPEAAAKA 350
    ||||||||||||||||||||||||||||||||||||||||||||||||||
301 PGFGPGVVGVPGAGVPGVGVPGAGIPVVPGAGIPGAAVPGVVSPEAAAKA 350

351 AAKAAKYGARPGVGVGGIPTYGVGAGGFPGFGVGVGGIPGVAGVPSVGGV 400
    ||||||||||||||||||||||||||||||||||||||||||||||||||
351 AAKAAKYGARPGVGVGGIPTYGVGAGGFPGFGVGVGGIPGVAGVPSVGGV 400

401 PGVGGVPGVGISPEAQAAAAAKAAKYGVGTPAAAAAKAAAKAAQFGLVPG 450
    ||||||||||||||||||||||||||||||||||||||||||||||||||
401 PGVGGVPGVGISPEAQAAAAAKAAKYGVGTPAAAAAKAAAKAAQFGLVPG 450

451 VGVAPGVGVAPGVGVAPGVGLAPGVGVAPGVGVAPGVGVAPGIGPGGVAA 500
    ||||||||||||||||||||||||||||||||||||||||||||||||||
451 VGVAPGVGVAPGVGVAPGVGLAPGVGVAPGVGVAPGVGVAPGIGPGGVAA 500

501 AAKSAAKVAAKAQLRAAAGLGAGIPGLGVGVGVPGLGVGAGVPGLGVGAG 550
    ||||||||||||||||||||||||||||||||||||||||||||||||||
501 AAKSAAKVAAKAQLRAAAGLGAGIPGLGVGVGVPGLGVGAGVPGLGVGAG 550

551 VPGFGAGADEGVRRSLSPELREGDPSSSQHLPSTPSSPRVPGALAAAKAA 600
    ||||||                                ||||||||||
551 VPGFGA..............................VPGALAAAKAA 567

601 KYGAAVPGVLGGLGALGGVGIPGGVVGAGPAAAAAAAKAAAKAAQFGLVG 650
    ||||||||||||||||||||||||||||||||||||||||||||||||||
568 KYGAAVPGVLGGLGALGGVGIPGGVVGAGPAAAAAAAKAAAKAAQFGLVG 617

651 AAGLGGLGVGGLGVPGVGGLGGIPPAAAAKAAKYGAAGLGGVLGGAGQFP 700
    ||||||||||||||||||||||||||||||||||||||||||||||||||
618 AAGLGGLGVGGLGVPGVGGLGGIPPAAAAKAAKYGAAGLGGVLGGAGQFP 667

701 LGGVAARPGFGLSPIFPGGACLGKACGRKRK 731
    ||||||||||||||||||||||||||||||
668 LGGVAARPGFGLSPIFPGGACLGKACGRKRK 698
```

Figure 2(1)

```
1   ATGGGTGGCGTTCCGGGTGCTGTTCCGGGTGGCGTTCCGGGTGGTGTATT  50
    ||||||||||||||||||||||||||||||||||||||||||||||||||
1   MetGlyGlyValProGlyAlaValProGlyGlyValProGlyGlyValPh  17

51  CTACCCAGGCGCGGGTTTCGGTGCTGTTCCGGGTGGCGTTGCAGACGCAG 100
    ||||||||||||||||||||||||||||||||||||||||||||||||||
18  eTyrProGlyAlaGlyPheGlyAlaValProGlyGlyValAlaAspAlaA  34

101 CTGCTGCGTACAAAGCGGCAAAGGCAGGTGCGGGTCTGGGCGGGGTACCA 150
    ||||||||||||||||||||||||||||||||||||||||||||||||||
35  laAlaAlaTyrLysAlaAlaLysAlaGlyAlaGlyLeuGlyGlyValPro  50

151 GGTGTTGGCGGTCTGGGTGTATCTGCTGGCGCAGTTGTTCCGCAGCCGGG 200
    ||||||||||||||||||||||||||||||||||||||||||||||||||
51  GlyValGlyGlyLeuGlyValSerAlaGlyAlaValValProGlnProGl  67

201 TGCAGGTGTAAAACCGGGCAAAGTTCCAGGTGTTGGTCTGCCGGGCGTAT 250
    ||||||||||||||||||||||||||||||||||||||||||||||||||
68  yAlaGlyValLysProGlyLysValProGlyValGlyLeuProGlyValT  84

251 ACCCGGGTTTCGGTGCTGTTCCGGGCGCGCGTTTCCCAGGTGTTGGTGTA 300
    ||||||||||||||||||||||||||||||||||||||||||||||||||
85  yrProGlyPheGlyAlaValProGlyAlaArgPheProGlyValGlyVal 100

301 CTGCCGGGCGTTCCGACCGGTGCAGGTGTTAAACCGAAGGCACCAGGTGT 350
    ||||||||||||||||||||||||||||||||||||||||||||||||||
101 LeuProGlyValProThrGlyAlaGlyValLysProLysAlaProGlyVa 117

351 AGGCGGCGCGTTCGCGGGTATCCCGGGTGTTGGCCCGTTCGGTGGTCCGC 400
    ||||||||||||||||||||||||||||||||||||||||||||||||||
118 lGlyGlyAlaPheAlaGlyIleProGlyValGlyProPheGlyGlyProG 134

401 AGCCAGGCGTTCCGCTGGGTTACCCGATCAAAGCGCCGAAGCTTCCAGGT 450
    ||||||||||||||||||||||||||||||||||||||||||||||||||
135 lnProGlyValProLeuGlyTyrProIleLysAlaProLysLeuProGly 150

451 GGCTACGGTCTGCCGTACACCACCGGTAAACTGCCGTACGGCTACGGTCC 500
    ||||||||||||||||||||||||||||||||||||||||||||||||||
151 GlyTyrGlyLeuProTyrThrThrGlyLysLeuProTyrGlyTyrGlyPr 167

501 GGGTGGCGTAGCAGGTGCTGCGGGTAAAGCAGGCTACCCAACCGGTACTG 550
    ||||||||||||||||||||||||||||||||||||||||||||||||||
168 oGlyGlyValAlaGlyAlaAlaGlyLysAlaGlyTyrProThrGlyThrG 184

551 GTGTTGGTCCGCAGGCTGCTGCGGCAGCTGCGGCGAAGGCAGCAGCAAAA 600
    ||||||||||||||||||||||||||||||||||||||||||||||||||
185 lyValGlyProGlnAlaAlaAlaAlaAlaAlaAlaLysAlaAlaAlaLys 200

601 TTCGGCGCGGGTGCAGCGGGTTTCGGTGCTGTTCCGGGGCGTAGGTGGTGC 650
    ||||||||||||||||||||||||||||||||||||||||||||||||||
201 PheGlyAlaGlyAlaAlaGlyPheGlyAlaValProGlyValGlyGlyAl 217

651 TGGCGTTCCGGGTGTTCCAGGTGCGATCCCGGGGCATCGGTGGTATCGCAG 700
    ||||||||||||||||||||||||||||||||||||||||||||||||||
218 aGlyValProGlyValProGlyAlaIleProGlyIleGlyGlyIleAlaG 234

701 GCGTAGGTACTCCGGCGGCCGCTGCGGCTGCGGCAGCTGCGGCGAAAGCA 750
    ||||||||||||||||||||||||||||||||||||||||||||||||||
235 lyValGlyThrProAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaLys 250
```

Figure 3(1)

```
751 GCTAAATACGGTGCGGCAGCAGGCCTGGTTCCGGGTGGTCCAGGCTTCGG 800
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
251 AlaLysTyrGlyAlaAlaAlaGlyLeuValProGlyGlyProGlyPheGl 267

801 TCCGGGTGTTGTAGGCGTTCCGGGTTTCGGTGCTGTTCCGGGCGTAGGTG 850
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
268 yProGlyValValGlyValProGlyPheGlyAlaValProGlyValGlyV 284

851 TTCCAGGTGCGGGCATCCCGGTTGTACCGGGTGCAGGTATCCCGGGCGCT 900
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
285 alProGlyAlaGlyIleProValValProGlyAlaGlyIleProGlyAla 300

901 GCGGGTTTCGGTGCTGTATCCCCGGAAGCGGCAGCTAAGGCTGCTGCGAA 950
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
301 AlaGlyPheGlyAlaValSerProGluAlaAlaAlaLysAlaAlaAlaLy 317

951 AGCTGCGAAATACGGAGCTCGTCCGGGCGTTGGTGTTGGTGGCATCCCGA 1000
    ||||||||||||||||||||||||||||||||||||||||||||||||||||
318 sAlaAlaLysTyrGlyAlaArgProGlyValGlyValGlyGlyIlePro T 334

1001 CCTACGGTGTAGGTGCAGGCGGTTTCCCAGGTTTCGGCGTTGGTGTTGGT 1050
     ||||||||||||||||||||||||||||||||||||||||||||||||||||
335 hrTyrGlyValGlyAlaGlyGlyPheProGlyPheGlyValGlyValGly 350

1051 GGCATCCCGGGTGTAGCTGGTGTTCCGTCTGTTGGTGGCGTACCGGGTGT 1100
     ||||||||||||||||||||||||||||||||||||||||||||||||||||
351 GlyIleProGlyValAlaGlyValProSerValGlyGlyValProGlyVa 367

1101 TGGTGGCGTTCCAGGTGTAGGTATCTCCCCGGAAGCGCAGGCAGCTGCGG 1150
     ||||||||||||||||||||||||||||||||||||||||||||||||||||
368 lGlyGlyValProGlyValGlyIleSerProGluAlaGlnAlaAlaAlaA 384

1151 CAGCTAAAGCAGCGAAGTACGGCGTTGGTACTCCGGCGGCAGCAGCTGCT 1200
     ||||||||||||||||||||||||||||||||||||||||||||||||||||
385 laAlaLysAlaAlaLysTyrGlyValGlyThrProAlaAlaAlaAlaAla 400

1201 AAAGCAGCGGCTAAAGCAGCGCAGTTCGGACTAGTTCCGGGCGTAGGTGT 1250
     ||||||||||||||||||||||||||||||||||||||||||||||||||||
401 LysAlaAlaAlaLysAlaAlaGlnPheGlyLeuValProGlyValGlyVa 417

1251 TGCGCCAGGTGTTGGCGTAGCACCGGGTGTTGGTGTTGCTCCGGGCGTAG 1300
     ||||||||||||||||||||||||||||||||||||||||||||||||||||
418 lAlaProGlyValGlyValAlaProGlyValGlyValAlaProGlyValG 434

1301 GTCTGGCACCGGGTGTTGGCGTTGCACCAGGTGTAGGTGTTGCGCCGGGC 1350
     ||||||||||||||||||||||||||||||||||||||||||||||||||||
435 lyLeuAlaProGlyValGlyValAlaProGlyValGlyValAlaProGly 450

1351 GTTGGTGTAGCACCGGGTATCGGTCCGGGTGGCGTTGCGGCTGCTGCGAA 1400
     ||||||||||||||||||||||||||||||||||||||||||||||||||||
451 ValGlyValAlaProGlyIleGlyProGlyGlyValAlaAlaAlaAlaLy 467

1401 ATCTGCTGCGAAGGTTGCTGCGAAAGCGGCAGCTGCGTGCAGCAGCTGGTC 1450
     ||||||||||||||||||||||||||||||||||||||||||||||||||||
468 sSerAlaAlaLysValAlaAlaLysAlaGlnLeuArgAlaAlaAlaGlyL 484

1451 TGGGTGCGGGCATCCCAGGTCTGGGTGTAGGTGTTGGTGTTCCGGGGCCTG 1500
     ||||||||||||||||||||||||||||||||||||||||||||||||||||
485 euGlyAlaGlyIleProGlyLeuGlyValGlyValGlyValProGlyLeu 500
```

Figure 3(2)

```
1501 GGTGTAGGTGCAGGGGTACCGGGCCTGGGTGTTGGTGCAGGCGTTCCGGG 1550
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 501 GlyValGlyAlaGlyValProGlyLeuGlyValGlyAlaGlyValProGl 517

1551 TTTCGGTGCTGTTCCGGGCGCGCTGGCTGCTGCGAAAGCGGCGAAATACG 1600
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 518 yPheGlyAlaValProGlyAlaLeuAlaAlaAlaLysAlaAlaLysTyrG 534

1601 GTGCTGTTCCGGGTGTACTGGGCGGTCTGGGTGCTCTGGGCGGTGTTGGT 1650
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 535 lyAlaValProGlyValLeuGlyGlyLeuGlyAlaLeuGlyGlyValGly 550

1651 ATCCCGGGCGGTGTTGTAGGTGCAGGCCCAGCTGCAGCTGCTGCTGCGGC 1700
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 551 IleProGlyGlyValValGlyAlaGlyProAlaAlaAlaAlaAlaAlaAl 567

1701 AAAGGCAGCGGCGAAAGCAGCTCAGTTCGGTCTGGTTGGTGCAGCAGGTC 1750
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 568 aLysAlaAlaAlaLysAlaAlaGlnPheGlyLeuValGlyAlaAlaGlyL 584

1751 TGGGCGGTCTGGGTGTTGGCGGTCTGGGTGTACCGGGCGTTGGTGGTCTG 1800
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 585 euGlyGlyLeuGlyValGlyGlyLeuGlyValProGlyValGlyGlyLeu 600

1801 GGTGGCATCCCGCCGGCGGCGGCAGCTAAAGCGGCTAAATACGGTGCAGC 1850
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 601 GlyGlyIleProProAlaAlaAlaAlaLysAlaAlaLysTyrGlyAlaAl 617

1851 AGGTCTGGGTGGCGTTCTGGGTGGTGCTGGTCAGTTCCCACTGGGCGGTG 1900
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 618 aGlyLeuGlyGlyValLeuGlyGlyAlaGlyGlnPheProLeuGlyGlyV 634

1901 TAGCGGCACGTCCGGGTTTCGGTCTGTCCCCGATCTTCCCAGGCGGTGCA 1950
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 635 alAlaAlaArgProGlyPheGlyLeuSerProIlePheProGlyGlyAla 650

1951 TGCCTGGGTAAAGCTTGCGGCCGTAAACGTAAA 1983
     |||||||||||||||||||||||||||||||||
 651 CysLeuGlyLysAlaCysGlyArgLysArgLys 661
```

Figure 3(3)

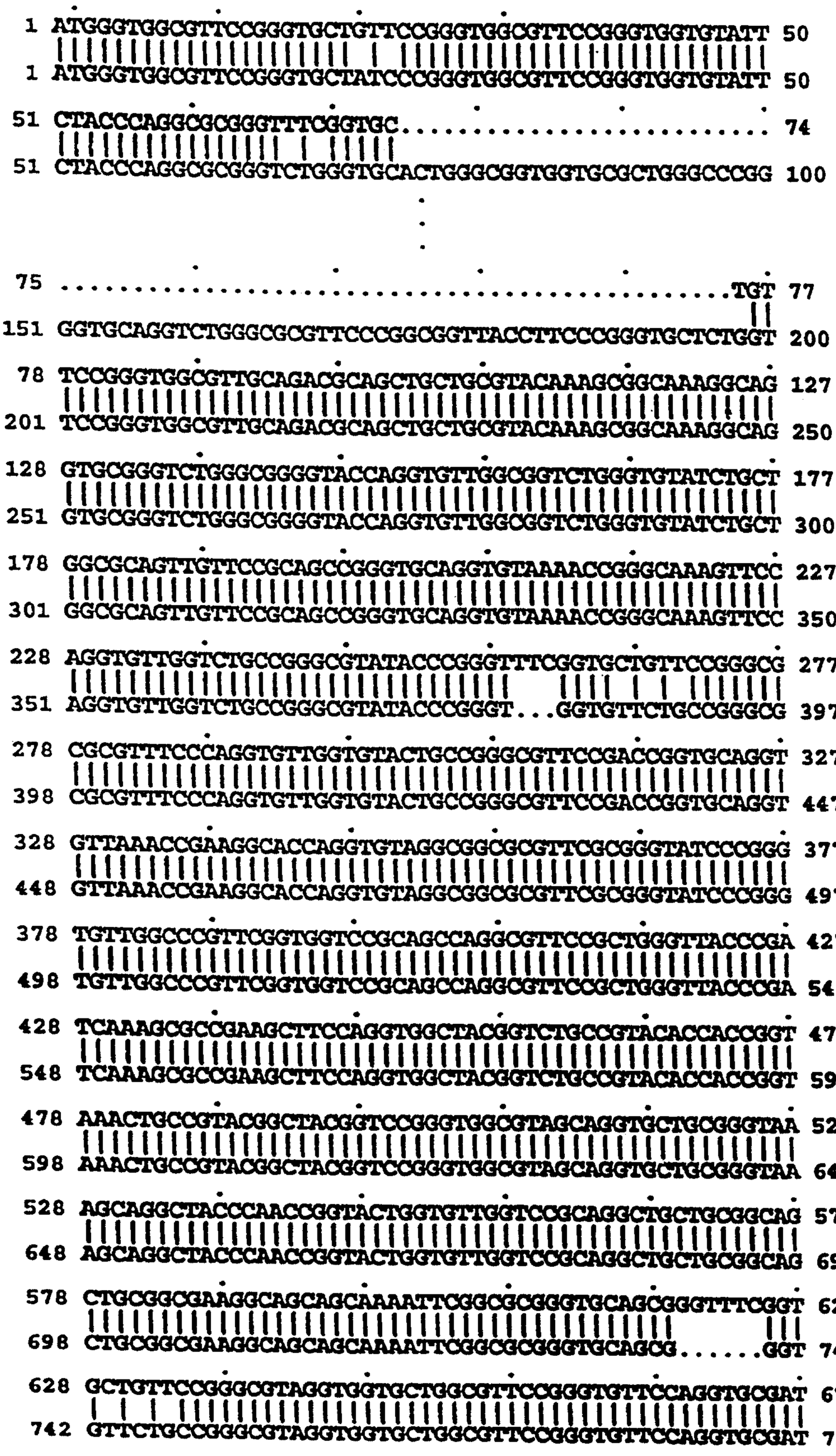
1 ATGGGTGGCGTTCCGGGTGCTGTTCCGGGTGGCGTTCCGGGTGGTGTATT 50
1 ATGGGTGGCGTTCCGGGTGCTATCCCGGGTGGCGTTCCGGGTGGTGTATT 50
51 CTACCCAGGCGCGGGTTTCGGTGC 74
51 CTACCCAGGCGCGGGTCTGGGTGCACTGGGCGGTGGTGCGCTGGGCCCGG 100
75 TGT 77
151 GGTGCAGGTCTGGGCGCGTTCCCGGCGGTTACCTTCCCGGGTGCTCTGGT 200
78 TCCGGGTGGCGTTGCAGACGCAGCTGCTGCGTACAAAGCGGCAAAGGCAG 127
201 TCCGGGTGGCGTTGCAGACGCAGCTGCTGCGTACAAAGCGGCAAAGGCAG 250
128 GTGCGGGTCTGGGCGGGGTACCAGGTGTTGGCGGTCTGGGTGTATCTGCT 177
251 GTGCGGGTCTGGGCGGGGTACCAGGTGTTGGCGGTCTGGGTGTATCTGCT 300
178 GGCGCAGTTGTTCCGCAGCCGGGTGCAGGTGTAAAACCGGGCAAAGTTCC 227
301 GGCGCAGTTGTTCCGCAGCCGGGTGCAGGTGTAAAACCGGGCAAAGTTCC 350
228 AGGTGTTGGTCTGCCGGGCGTATACCCGGGTTTCGGTGCTGTTCCGGGCG 277
351 AGGTGTTGGTCTGCCGGGCGTATACCCGGGT...GGTGTTCTGCCGGGCG 397
278 CGCGTTTCCCAGGTGTTGGTGTACTGCCGGGCGTTCCGACCGGTGCAGGT 327
398 CGCGTTTCCCAGGTGTTGGTGTACTGCCGGGCGTTCCGACCGGTGCAGGT 447
328 GTTAAACCGAAGGCACCAGGTGTAGGCGGCGCGTTCGCGGGTATCCCGGG 377
448 GTTAAACCGAAGGCACCAGGTGTAGGCGGCGCGTTCGCGGGTATCCCGGG 497
378 TGTTGGCCCGTTCGGTGGTCCGCAGCCAGGCGTTCCGCTGGGTTACCCGA 427
498 TGTTGGCCCGTTCGGTGGTCCGCAGCCAGGCGTTCCGCTGGGTTACCCGA 547
428 TCAAAGCGCCGAAGCTTCCAGGTGGCTACGGTCTGCCGTACACCACCGGT 477
548 TCAAAGCGCCGAAGCTTCCAGGTGGCTACGGTCTGCCGTACACCACCGGT 597
478 AAACTGCCGTACGGCTACGGTCCGGGTGGCGTAGCAGGTGCTGCGGGTAA 527
598 AAACTGCCGTACGGCTACGGTCCGGGTGGCGTAGCAGGTGCTGCGGGTAA 647
528 AGCAGGCTACCCAACCGGTACTGGTGTTGGTCCGCAGGCTGCTGCGGCAG 577
648 AGCAGGCTACCCAACCGGTACTGGTGTTGGTCCGCAGGCTGCTGCGGCAG 697
578 CTGCGGCGAAGGCAGCAGCAAAATTCGGCGCGGGTGCAGCGGGTTTCGGT 627
698 CTGCGGCGAAGGCAGCAGCAAAATTCGGCGCGGGTGCAGCG.....GGT 741
628 GCTGTTCCGGGCGTAGGTGGTGCTGGCGTTCCGGGTGTTCCAGGTGCGAT 677
742 GTTCTGCCGGGCGTAGGTGGTGCTGGCGTTCCGGGTGTTCCAGGTGCGAT 791

Figure 4(1)

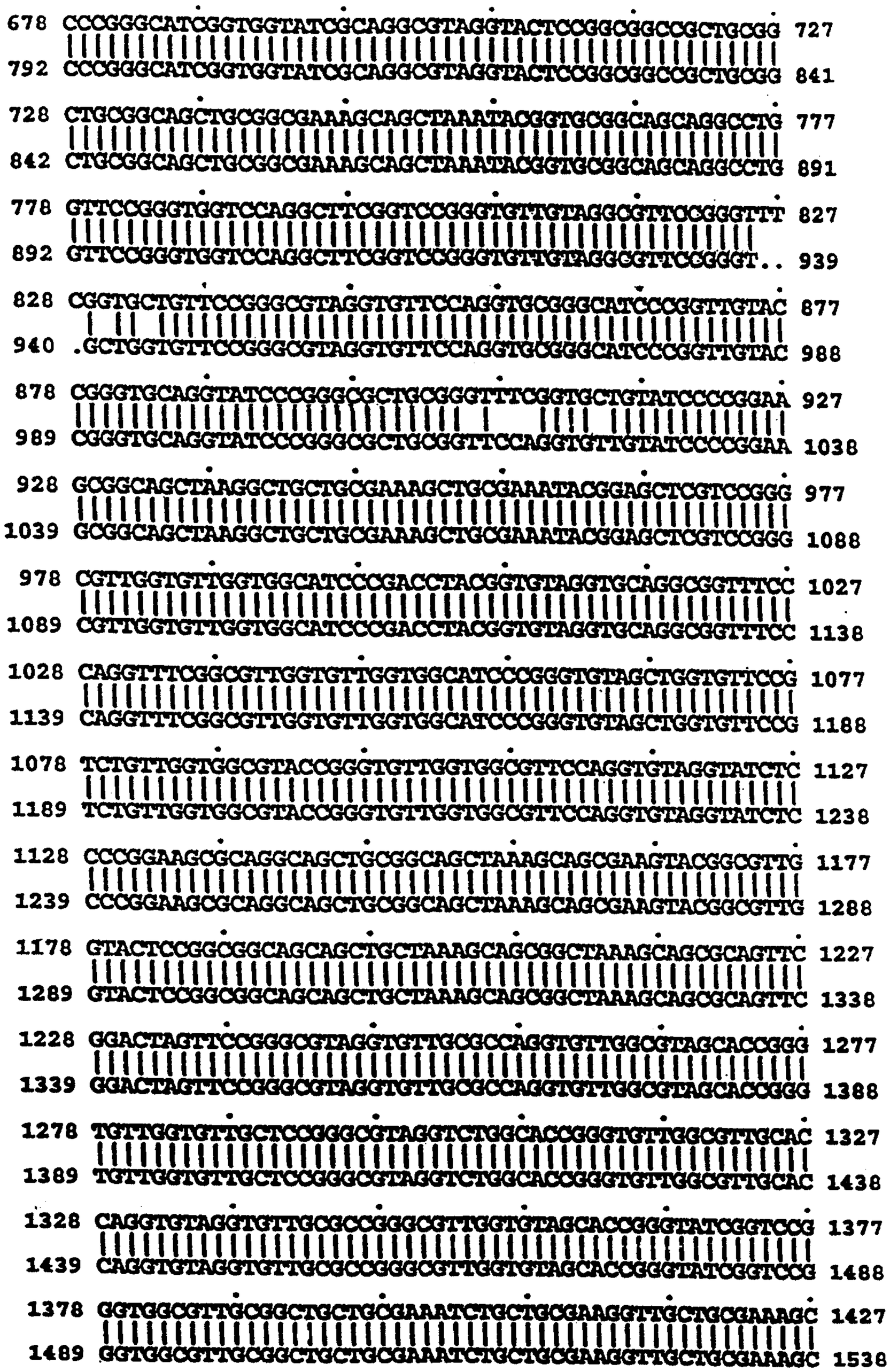
678 CCCGGGCATCGGTGGTATCGCAGGCGTAGGTACTCCGGCGGCCGCTGCGG 727
792 CCCGGGCATCGGTGGTATCGCAGGCGTAGGTACTCCGGCGGCCGCTGCGG 841
728 CTGCGGCAGCTGCGGCGAAAGCAGCTAAATACGGTGCGGCAGCAGGCCTG 777
842 CTGCGGCAGCTGCGGCGAAAGCAGCTAAATACGGTGCGGCAGCAGGCCTG 891
778 GTTCCGGGTGGTCCAGGCTTCGGTCCGGGTGTTGTAGGCGTTCCGGGTTT 827
892 GTTCCGGGTGGTCCAGGCTTCGGTCCGGGTGTTGTAGGGGTTCGGGGT.. 939
828 CGGTGCTGTTCCGGGCGTAGGTGTTCCAGGTGCGGGCATCCCGGTTGTAC 877
940 .GCTGGTGTTCCGGGCGTAGGTGTTCCAGGTGCGGGCATCCCGGTTGTAC 988
878 CGGGTGCAGGTATCCCGGGCGCTGCGGGTTTCGGTGCTGTATCCCCGGAA 927
989 CGGGTGCAGGTATCCCGGGCGCTGCGGTTCCAGGTGTTGTATCCCCGGAA 1038
928 GCGGCAGCTAAGGCTGCTGCGAAAGCTGCGAAATACGGAGCTCGTCCGGG 977
1039 GCGGCAGCTAAGGCTGCTGCGAAAGCTGCGAAATACGGAGCTCGTCCGGG 1088
978 CGTTGGTGTTGGTGGCATCCCGACCTACGGTGTAGGTGCAGGCGGTTTCC 1027
1089 CGTTGGTGTTGGTGGCATCCCGACCTACGGTGTAGGTGCAGGCGGTTTCC 1138
1028 CAGGTTTCGGCGTTGGTGTTGGTGGCATCCCGGGTGTAGCTGGTGTTCCG 1077
1139 CAGGTTTCGGCGTTGGTGTTGGTGGCATCCCGGGTGTAGCTGGTGTTCCG 1188
1078 TCTGTTGGTGGCGTACCGGGTGTTGGTGGCGTTCCAGGTGTAGGTATCTC 1127
1189 TCTGTTGGTGGCGTACCGGGTGTTGGTGGCGTTCCAGGTGTAGGTATCTC 1238
1128 CCCGGAAGCGCAGGCAGCTGCGGCAGCTAAAGCAGCGAAGTACGGCGTTG 1177
1239 CCCGGAAGCGCAGGCAGCTGCGGCAGCTAAAGCAGCGAAGTACGGCGTTG 1288
1178 GTACTCCGGCGGCAGCAGCTGCTAAAGCAGCGGCTAAAGCAGCGCAGTTC 1227
1289 GTACTCCGGCGGCAGCAGCTGCTAAAGCAGCGGCTAAAGCAGCGCAGTTC 1338
1228 GGACTAGTTCCGGGCGTAGGTGTTGCGCCAGGTGTTGGCGTAGCACCGGG 1277
1339 GGACTAGTTCCGGGCGTAGGTGTTGCGCCAGGTGTTGGCGTAGCACCGGG 1388
1278 TGTTGGTGTTGCTCCGGGCGTAGGTCTGGCACCGGGTGTTGGCGTTGCAC 1327
1389 TGTTGGTGTTGCTCCGGGCGTAGGTCTGGCACCGGGTGTTGGCGTTGCAC 1438
1328 CAGGTGTAGGTGTTGCGCCGGGCGTTGGTGTAGCACCGGGTATCGGTCCG 1377
1439 CAGGTGTAGGTGTTGCGCCGGGCGTTGGTGTAGCACCGGGTATCGGTCCG 1488
1378 GGTGGCGTTGCGGCTGCTGCGAAATCTGCTGCGAAGGTTGCTGCGAAAGC 1427
1489 GGTGGCGTTGCGGCTGCTGCGAAATCTGCTGCGAAGGTTGCTGCGAAAGC 1538

Figure 4(2)

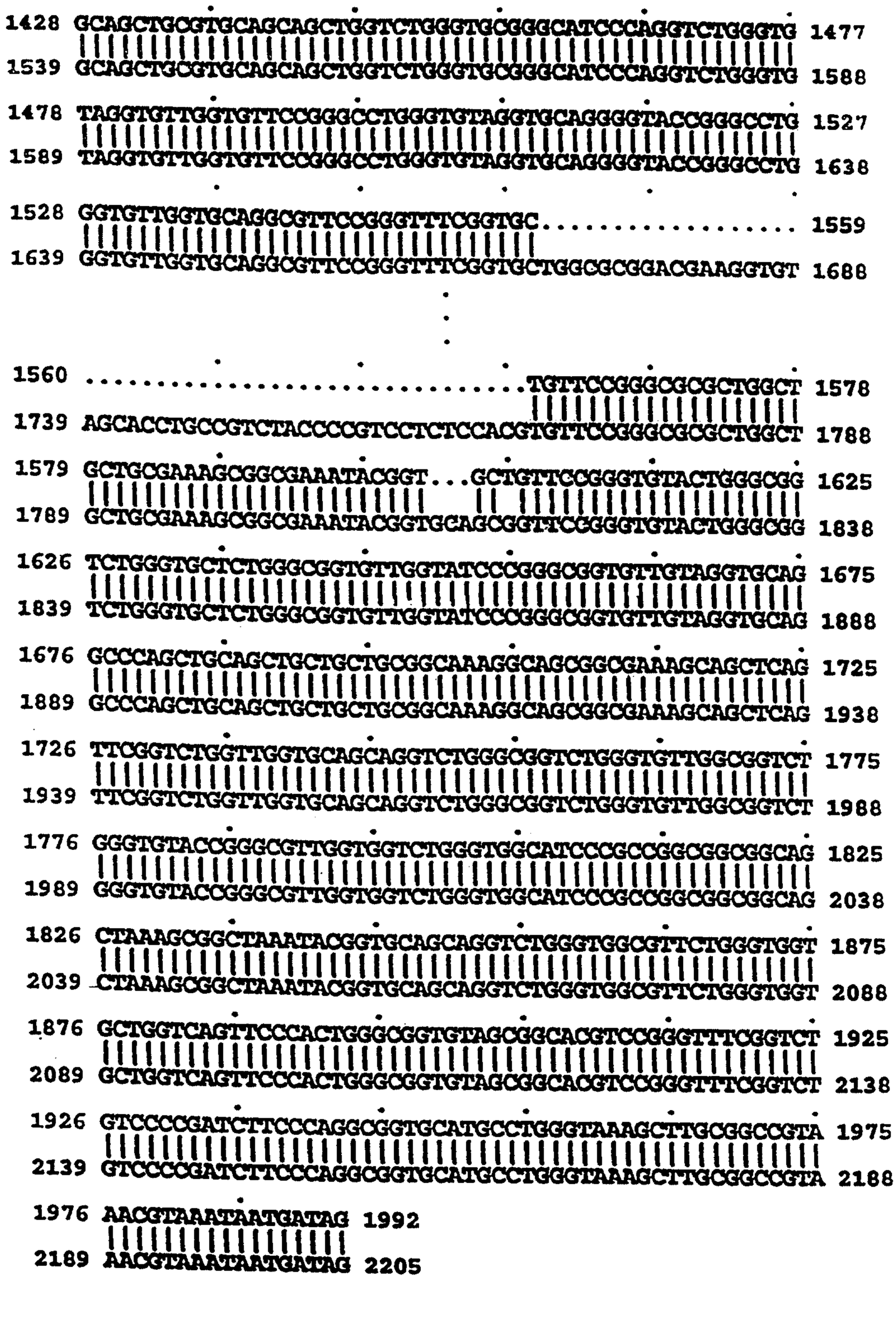
1428 GCAGCTGCGTGCAGCAGCTGGTCTGGGTGCGGGCATCCCAGGTCTGGGTG 1477
1539 GCAGCTGCGTGCAGCAGCTGGTCTGGGTGCGGGCATCCCAGGTCTGGGTG 1588
1478 TAGGTGTTGGTGTTCCGGGCCTGGGTGTAGGTGCAGGGGTACCGGGGCCTG 1527
1589 TAGGTGTTGGTGTTCCGGGCCTGGGTGTAGGTGCAGGGGTACCGGGGCCTG 1638
1528 GGTGTTGGTGCAGGCGTTCCGGGTTTCGGTGC..... 1559
1639 GGTGTTGGTGCAGGCGTTCCGGGTTTCGGTGCTGGCGCCGGACGAAGGTGT 1688
1560 .....TGTTCCGGGCGCGCTGGCT 1578
1739 AGCACCTGCCGTCTACCCCGTCCTCTCCACGTGTTCCGGGCGCGCTGGCT 1788
1579 GCTGCGAAAGCGGCGAAATACGGT...GCTGTTCCGGGTGTACTGGGCGG 1625
1789 GCTGCGAAAGCGGCGAAATACGGTGCAGCGGTTCCGGGGTGTACTGGGCGG 1838
1626 TCTGGGTGCTCTGGGCGGTGTTGGTATCCCGGGCGGTGTTGTAGGTGCAG 1675
1839 TCTGGGTGCTCTGGGCGGTGTTGGTATCCCGGGCGGTGTTGTAGGTGCAG 1888
1676 GCCCAGCTGCAGCTGCTGCTGCGGCAAAGGCAGCGGCGAAAGCAGCTCAG 1725
1889 GCCCAGCTGCAGCTGCTGCTGCGGCAAAGGCAGCGGCGAAAGCAGCTCAG 1938
1726 TTCGGTCTGGTTGGTGCAGCAGGTCTGGGCGGTCTGGGTGTTGGCGGTCT 1775
1939 TTCGGTCTGGTTGGTGCAGCAGGTCTGGGCGGTCTGGGTGTTGGCGGTCT 1988
1776 GGGTGTACCGGGCGTTGGTGGTCTGGGTGGCATCCCGCCGGCGGCGGCAG 1825
1989 GGGTGTACCGGGCGTTGGTGGTCTGGGTGGCATCCCGCCGGCGGCGGCAG 2038
1826 CTAAAGCGGCTAAATACGGTGCAGCAGGTCTGGGTGGCGTTCTGGGTGGT 1875
2039 CTAAAGCGGCTAAATACGGTGCAGCAGGTCTGGGTGGCGTTCTGGGTGGT 2088
1876 GCTGGTCAGTTCCCACTGGGCGGTGTAGCGGCACGTCCGGGTTTCGGTCT 1925
2089 GCTGGTCAGTTCCCACTGGGCGGTGTAGCGGCACGTCCGGGTTTCGGTCT 2138
1926 GTCCCCGATCTTCCCAGGCGGTGCATGCCTGGGTAAAGCTTGCGGCCGTA 1975
2139 GTCCCCGATCTTCCCAGGCGGTGCATGCCTGGGTAAAGCTTGCGGCCGTA 2188
1976 AACGTAAATAATGATAG 1992
2189 AACGTAAATAATGATAG 2205

Figure 4(3)

```
        10         20         30         40         50         60         70         80
MGGVPGAIPGGVPGGVFYPGAGLGALGGGALGPGGKPLKPVPGGLAGAGLGAGLGAFPAVTFPGALVPGGVADAAAAYKA
:::::::  :::::::::: :                        :         ::   :       : ::  ::::::::::::::
MGGVPGAVPGGVPGGVFY----------------------p--------GA--G------F-GA-VPGGVADAAAAYKA
        10                                                  20                  30

        90        100        110        120        130        140        150
AKAGAGLGGVPGVGGLGVSAGAVVPQPGAGVKPGKVPGVGLPGVYPG-GVLPGARFPGVGVLPGVPTGAGVKPKAPGVGG
:::::::::::::::::::::::::::::::::::::::::::::::: :  ::::::::::::::::::::::::::::::
AKAGAGLGGVPGVGGLGVSAGAVVPQPGAGVKPGKVPGVGLPGVYPGFGAVPGARFPGVGVLPGVPTGAGVKPKAPGVGG
        50         60         70         80         90        100        110

        170        180        190        200        210        220        230
AFAGIPGVGPFGGPQPGVPLGYPIKAPKLPGGYGLPYTTGKLPYGYGPGGVAGAAGKAGYPTGTGVGPQAAAAAAAKAAA
:::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
AFAGIPGVGPFGGPQPGVPLGYPIKAPKLPGGYGLPYTTGKLPYGYGPGGVAGAAGKAGYPTGTGVGPQAAAAAAAKAAA
        130        140        150        160        170        180        190

         250        260        270        280        290        300        310
KFGAGAA--GVLPGVGGAGVPGVPGAIPGIGGIAGVGTPAAAAAAAAAAKAAKYGAAAGLVPGGPGFGPGVVGVPGAG-V
:::::::  :  ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::: : :
KFGAGAAGFGAVPGVGGAGVPGVPGAIPGIGGIAGVGTPAAAAAAAAAAKAAKYGAAAGLVPGGPGFGPGVVGVPGFGAV
        210        220        230        240        250        260        270

 320        330        340        350        360        370        380        390
PGVGVPGAGIPVVPGAGIPGAAVPGVVSPEAAAAKAAAKAAKYGARPGVGVGGIPTYGVGAGGFPGFGVGVGGIPGVAGVP
:::::::::::::::::::::  : ::::::::::::::::::::::::::::::::::::::::::::::::::::::::
PGVGVPGAGIPVVPGAGIPGAAGFGAVSPEAAAAKAAAKAAKYGARPGVGVGGIPTYGVGAGGFPGFGVGVGGIPGVAGVP
        290        300        310        320        330        340        350
```

Figure 5(1)

```
 400        410        420        430        440        450        460        470
SVGGVPGVGGVPGVGISPEAQAAAAAKAAKYGVGTPAAAAAKAAAKAAQFGLVPGVGVAPGVGVAPGVGVAPGVGLAPGV
:::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
SVGGVPGVGGVPGVGISPEAQAAAAAKAAKYGVGTPAAAAAKAAAKAAQFGLVPGVGVAPGVGVAPGVGVAPGVGLAPGV
         370        380        390        400        410        420        430

 480        490        500        510        520        530        540        550
GVAPGVGVAPGVGVAPGIGPGGVAAAAKSAAKVAAKAQLRAAAGLGAGIPGLGVGVGVPGLGVGAGVPGLGVGAGVPGFG
:::::::::::::::  ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
GVAPGVGVAPGVGVAPGIGPGGVAAAAKSAAKVAAKAQLRAAAGLGAGIPGLGVGVGVPGLGVGAGVPGLGVGAGVPGF-
         450        460        470        480        490        500        510

 560        570        580        590        600        610        620        630
AGADEGVRRSLSPELREGDPSSSQHLPSTPSSPRVPGALAAAKAAKYGAAVPGVLGGLGALGGVGIPGGVVGAGPAAAAA
::                                 :::::::::::::: ::::::::::::::::::::::::::::::::
·GA--------------------------------VPGALAAAKAAKYG-AVPGVLGGLGALGGVGIPGGVVGAGPAAAAA
                                                  530        540        550        560

 640        650        660        670        680        690        700        710
AAKAAAKAAQFGLVGAAGLGGLGVGGLGVPGVGGLGGIPPAAAAKAAKYGAAGLGGVLGGAGQFPLGGVAARPGFGLSPI
::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
AAKAAAKAAQFGLVGAAGLGGLGVGGLGVPGVGGLGGIPPAAAAKAAKYGAAGLGGVLGGAGQFPLGGVAARPGFGLSPI
  570        580        590        600        610        620        630        640

 720        730
FPGGACLGKACGRKRK
::::::::::::::::
FPGGACLGKACGRKRK
 650        660
```

Figure 5(2)

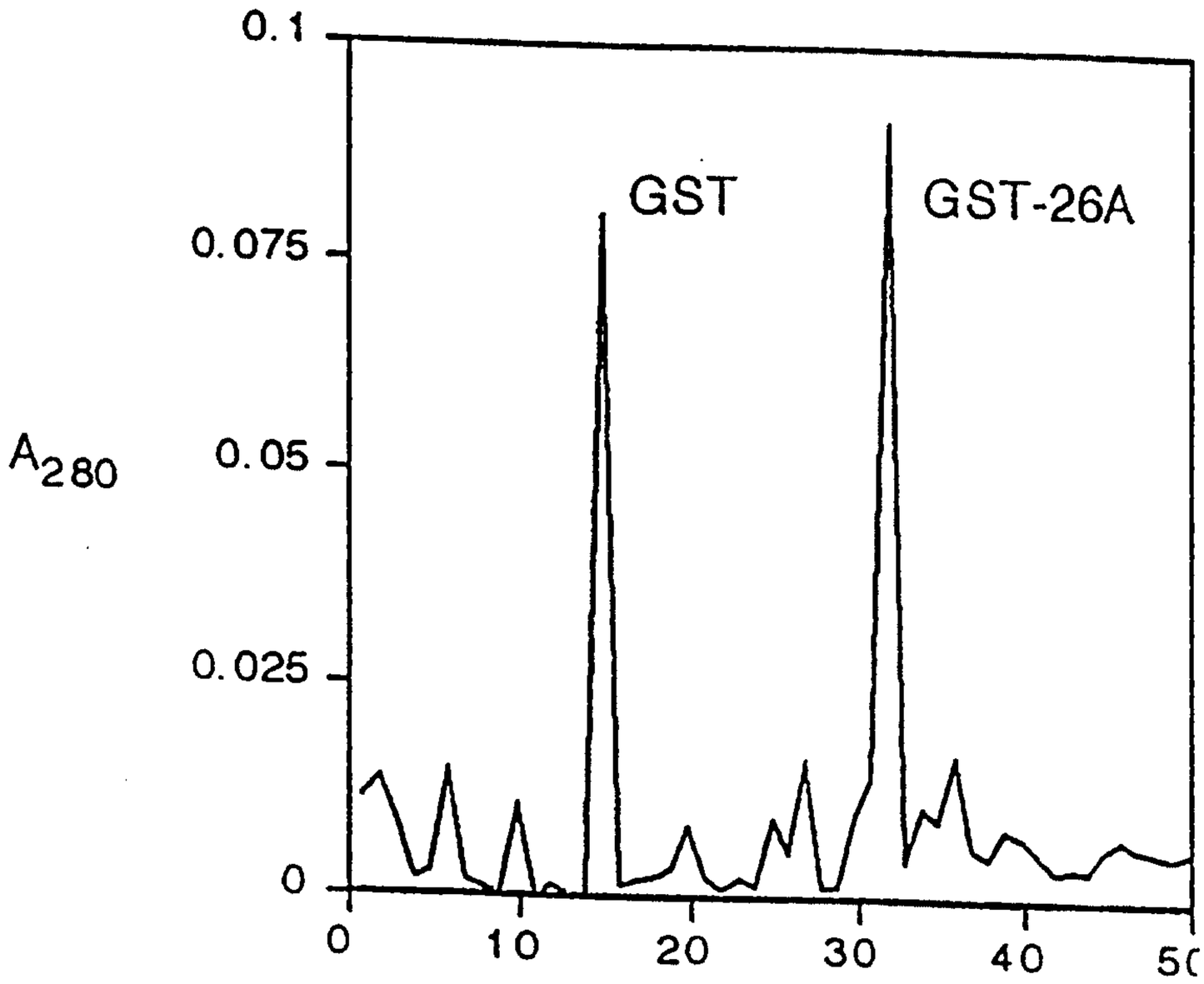
0.1
0.075
0.05
0.025
0
A_{280}
GST
GST-26A
0
10
20
30
40
50
Fraction number

Fig. 6(a)

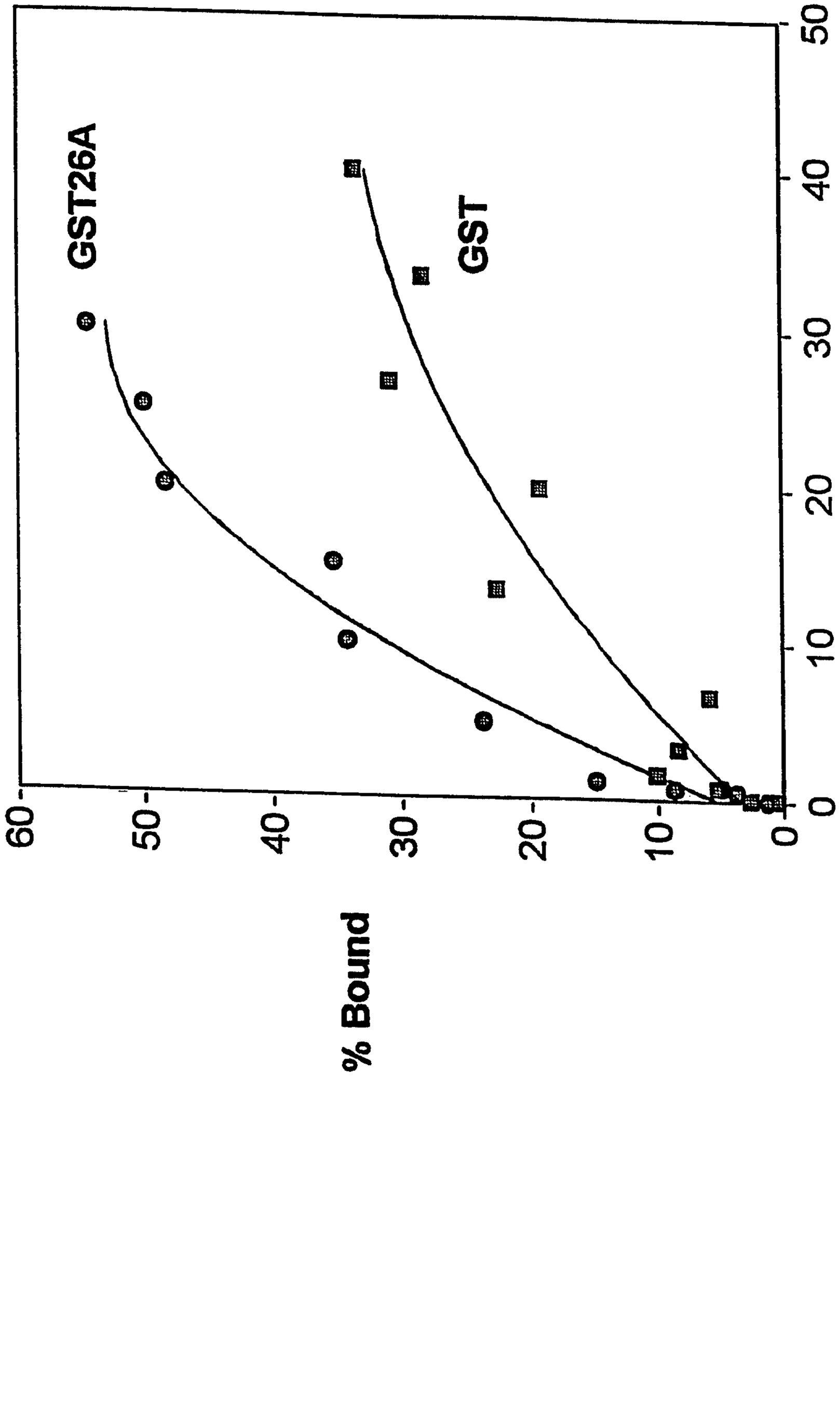

Fig. 6(b)

```
                .         .         .         .         .
948 TCCGCCATGGGAGGTGTTCCGGGCGCGCTGGCTGCTGCGAAAGCGGCGAA 997
    |||||||||||||||||||||||||||||||||||||||||||||||||||
  1 SerAlaMetGlyGlyValProGlyAlaLeuAlaAlaAlaLysAlaAlaLy 17
                .         .         .         .         .
998 ATACGGTGCAGCGGTTCCGGGTGTACTGGGCGGTCTGGGTGCTCTGGGCG 1047
    |||||||||||||||||||||||||||||||||||||||||||||||||||
 18 sTyrGlyAlaAlaValProGlyValLeuGlyGlyLeuGlyAlaLeuGlyG 34
                .         .         .         .         .
1048 GTGTTGGTATCCCGGGCGGTGTTGTAGGTGCAGGCCCAGCTGCAGCTGCT 1097
    |||||||||||||||||||||||||||||||||||||||||||||||||||
 35 lyValGlyIleProGlyGlyValValGlyAlaGlyProAlaAlaAlaAla 50
                .         .         .         .         .
1098 GCTGCGGCAAAGGCAGCGGCGAAAGCAGCTCAGTTCGGTCTGGTTGGTGC 1147
    |||||||||||||||||||||||||||||||||||||||||||||||||||
 51 AlaAlaAlaLysAlaAlaAlaLysAlaAlaGlnPheGlyLeuValGlyAl 67
                .         .         .         .         .
1148 AGCAGGTCTGGGCGGTCTGGGTGTTGGCGGTCTGGGTGTACCGGGCGTTG 1197
    |||||||||||||||||||||||||||||||||||||||||||||||||||
 68 aAlaGlyLeuGlyGlyLeuGlyValGlyGlyLeuGlyValProGlyValG 84
                .         .         .         .         .
1198 GTGGTCTGGGTGGCATCCCGCCGGCGGCGGCAGCTAAAGCGGCTAAATAC 1247
    |||||||||||||||||||||||||||||||||||||||||||||||||||
 85 lyGlyLeuGlyGlyIleProProAlaAlaAlaAlaLysAlaAlaLysTyr 100
                .         .         .         .         .
1248 GGTGCAGCAGGTCTGGGTGGCGTTCTGGGTGGTGCTGGTCAGTTCCCACT 1297
    |||||||||||||||||||||||||||||||||||||||||||||||||||
101 GlyAlaAlaGlyLeuGlyGlyValLeuGlyGlyAlaGlyGlnPheProLe 117
                .         .         .         .         .
1298 GGGCGGTGTAGCGGCACGTCCGGGTTTCGGTCTGTCCCCGATCTTCCCAG 1347
    |||||||||||||||||||||||||||||||||||||||||||||||||||
118 uGlyGlyValAlaAlaArgProGlyPheGlyLeuSerProIlePheProG 134
                .         .         .         .
1348 GCGGTGCATGCCTGGGTAAAGCTTGCGGCCGTAAACGTAAA 1388
    |||||||||||||||||||||||||||||||||||||||||
135 lyGlyAlaCysLeuGlyLysAlaCysGlyArgLysArgLys 147
```

Figure 7

```
948 TCCGCCATGGGAGCTCTGGTAGGCCTGGGCGTACCGGGCCTGGGTGTTGG 997
    ||||||||||||||||||||||||||||||||||||||||||||||||||
  1 SerAlaMetGlyAlaLeuValGlyLeuGlyValProGlyLeuGlyValGl 17

998 TGCAGGCGTTCCGGGTTTCGGTGCTGGCGCGGACGAAGGTGTACGTCGTT 1047
    ||||||||||||||||||||||||||||||||||||||||||||||||||
 18 yAlaGlyValProGlyPheGlyAlaGlyAlaAspGluGlyValArgArgS 34

1048 CCCTGTCTCCAGAACTGCGTGAAGGTGACCCGTCCTCTTCCCAGCACCTG 1097
     ||||||||||||||||||||||||||||||||||||||||||||||||||
  35 erLeuSerProGluLeuArgGluGlyAspProSerSerSerGlnHisLeu 50

1098 CCGTCTACCCCGTCCTCTCCACGTGTTCCGGGCGCGCTGGCTGCTGCGAA 1147
     ||||||||||||||||||||||||||||||||||||||||||||||||||
  51 ProSerThrProSerSerProArgValProGlyAlaLeuAlaAlaAlaLy 67

1148 AGCGGCGAAATACGGTGCAGCGGTTCCGGGTGTACTGGGCGGTCTGGGTG 1197
     ||||||||||||||||||||||||||||||||||||||||||||||||||
  68 sAlaAlaLysTyrGlyAlaAlaValProGlyValLeuGlyGlyLeuGlyA 84

1198 CTCTGGGCGGTGTTGGTATCCCGGGCGGTGTTGTAGGTGCAGGCCCAGCT 1247
     ||||||||||||||||||||||||||||||||||||||||||||||||||
  85 laLeuGlyGlyValGlyIleProGlyGlyValValGlyAlaGlyProAla 100
```

Figure 8(1)

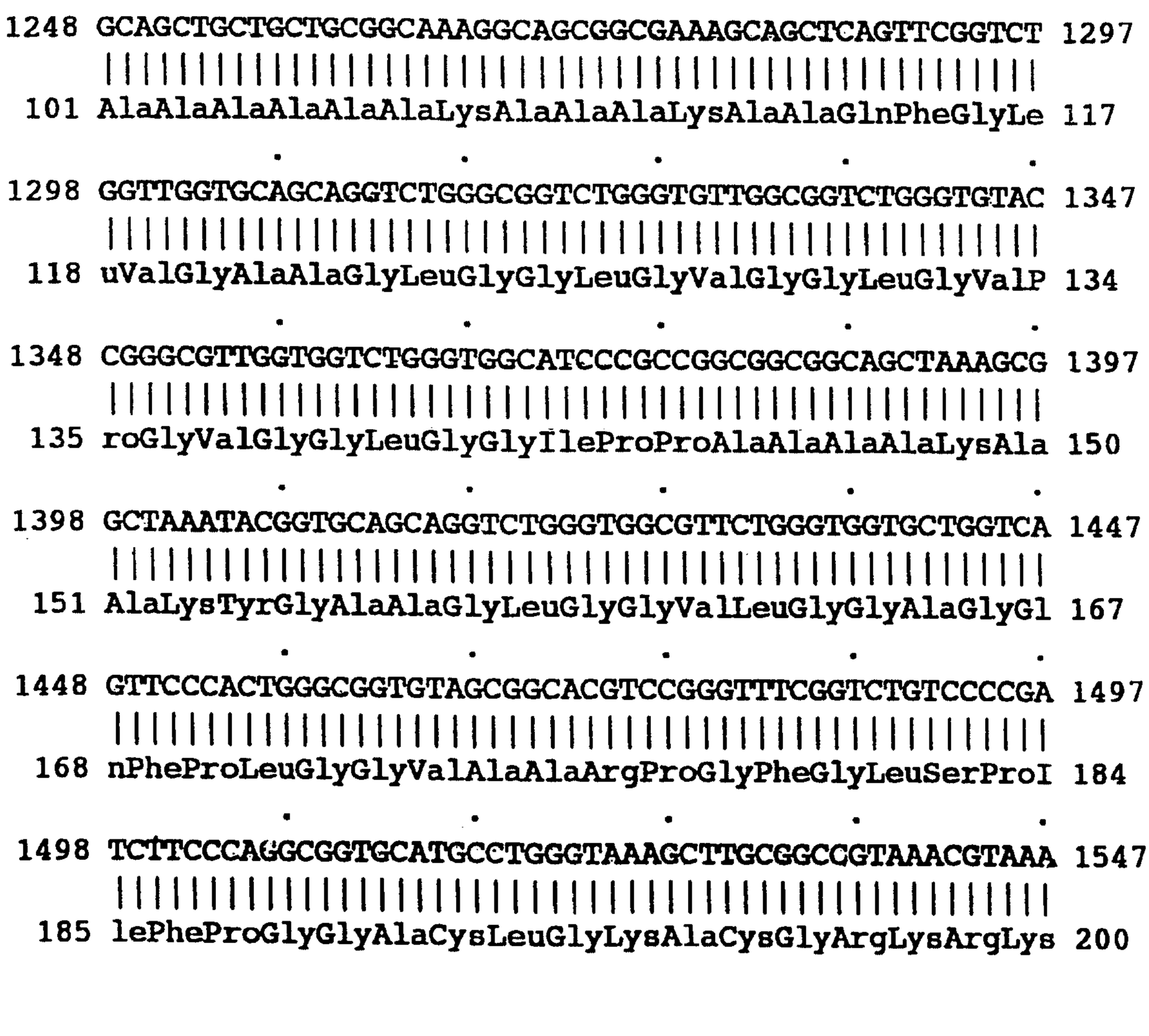
1248 GCAGCTGCTGCTGCGGCAAAGGCAGCGGCGAAAGCAGCTCAGTTCGGTCT 1297
101 AlaAlaAlaAlaAlaAlaLysAlaAlaAlaLysAlaAlaGlnPheGlyLe 117
1298 GGTTGGTGCAGCAGGTCTGGGCGGTCTGGGTGTTGGCGGTCTGGGTGTAC 1347
118 uValGlyAlaAlaGlyLeuGlyGlyLeuGlyValGlyGlyLeuGlyValP 134
1348 CGGGCGTTGGTGGTCTGGGTGGCATCCCGCCGGCGGCGGCAGCTAAAGCG 1397
135 roGlyValGlyGlyLeuGlyGlyIleProProAlaAlaAlaAlaLysAla 150
1398 GCTAAATACGGTGCAGCAGGTCTGGGTGGCGTTCTGGGTGGTGCTGGTCA 1447
151 AlaLysTyrGlyAlaAlaGlyLeuGlyGlyValLeuGlyGlyAlaGlyGl 167
1448 GTTCCCACTGGGCGGTGTAGCGGCACGTCCGGGTTTCGGTCTGTCCCCGA 1497
168 nPheProLeuGlyGlyValAlaAlaArgProGlyPheGlyLeuSerProI 184
1498 TCTTCCCAGGCGGTGCATGCCTGGGTAAAGCTTGCGGCCGTAAACGTAAA 1547
185 lePheProGlyGlyAlaCysLeuGlyLysAlaCysGlyArgLysArgLys 200

Figure 8(2)