

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
29 July 2004 (29.07.2004)

PCT

(10) International Publication Number
WO 2004/063226 A2

(51) International Patent Classification⁷: **C07K 14/78**

(21) International Application Number:
PCT/EP2003/051089

(22) International Filing Date:
22 December 2003 (22.12.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/436,835 27 December 2002 (27.12.2002) US

(71) Applicant (*for all designated States except US*): **APPLIED RESEARCH SYSTEMS ARS HOLDING N.V.**
[NL/NL]; Pietermaai 15, Curacao (AN).

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): **MCALLISTER, Gregg** [US/US]; 65 Chelsea St. Apt 407, Charlestown, MA 02129 (US). **BIENKOWSKA, Jadwiga** [PL/US]; 45 Fayette St. Apt 3, Cambridge, MA 02139 (US).

(74) Agent: **HASSA, Jürgen**; Serono International SA, Intellectual Property, 12, chemin des Aulx, CH-1228 Plan-les-Ouates (CH).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (*regional*): ARIPO patent (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: NOVEL FIBRILLIN-LIKE POLYPEPTIDES

(57) Abstract: The present invention discloses open reading frames (ORFs) in human genome encoding for novel fibrillin-like polypeptides, and reagents related thereto including variants, mutants and fragments of said polypeptides, as well as ligands and antagonists directed against them. The invention provides methods for identifying and making these molecules, for preparing pharmaceutical compositions containing them, and for using them in the diagnosis, prevention and treatment of diseases.



- 1 -

NOVEL FIBRILLIN-LIKE POLYPEPTIDES

FIELD OF THE INVENTION

5 The present invention relates to nucleic acid sequences identified in the human genome as encoding for novel polypeptides, more specifically for fibrillin-like polypeptides.

All publications, patents and patent applications cited herein are incorporated in full by reference.

BACKGROUND OF THE INVENTION

10 Many novel polypeptides have been already identified by applying strict homology criteria to known polypeptides of the same family. However, since the actual content in polypeptide-encoding sequences in the human genome for fibrillin-like polypeptides (and for any other protein family) is still unknown, the possibility still exists to identify DNA sequence encoding polypeptide having fibrillin-like polypeptide activities by
15 applying alternative and less strict homology/structural criteria to the totality of Open Reading Frames (ORFs, that is, genomic sequences containing consecutive triplets of nucleotides coding for amino acids, not interrupted by a termination codon and potentially translatable in a polypeptide) present in the human genome.

The ability for cells to make and secrete extracellular proteins is central to many
20 biological processes. Enzymes, growth factors, extracellular matrix proteins and signalling molecules are all secreted by cells. This is through fusion of a secretory vesicle with the plasma membrane. In most cases, but not all, proteins are directed to the endoplasmic reticulum and into secretory vesicles by a signal peptide. Signal peptides are cis-acting sequences that affect the transport of polypeptide chains from
25 the cytoplasm to a membrane bound compartment such as a secretory vesicle. Polypeptides that are targeted to the secretory vesicles are either secreted into the extracellular matrix or are retained in the plasma membrane. The polypeptides that are retained in the plasma membrane will have one or more transmembrane domains. Examples of secreted proteins that play a central role in the functioning of a cell are
30 cytokines, hormones, extracellular matrix proteins (adhesion molecules), proteases, and growth and differentiation factors.

- 2 -

Microfibril bundles and proteins found in association with these bundles, particularly attachment molecules, are of interest in the field of dermatology, particularly in the study of skin which has been damaged from aging, injuries or the sun. Fibrillin microfibrils define the continuous elastic network of skin, and are present in dermis as microfibril bundles devoid of measurable elastin extending from the dermal-epithelial junction and as components of the thick elastic fibres present in the deep reticular dermis.

Fibrillin-1 is a modular glycoprotein with amino-terminal anaphylatoxin-like modules followed by nine epidermal growth factor (EGF)-like modules and, depending on alternative splicing, four possible carboxyl termini. Fibrillin-2 is a novel extracellular matrix protein frequently found in close association with microfibrils containing either fibronectin or fibulin. Thus, fibrillin, and molecules related thereto are of interest, particularly for the use of preventing skin from being damaged from aging, injuries or the sun, or for restoring skin damaged from same. Moreover, these molecules are generally of interest in the study of connective tissue and attachment molecules and related mechanisms. Fibrillin, and related molecules are further described in Adams, *et al.*, J. Mol. Biol., 272(2):226-36 (1997); Kielty and Shuttleworth, Microsc. Res. Tech., 38(4):413-27 (1997); and Child, J. Card. Surg., 12(2Supp.):131-3 (1997).

Thus, fibrillin, and molecules related thereto are of significant interest, particularly for the use of preventing skin from being damaged from aging, injuries or the sun, or for restoring skin damaged from same. Moreover, these molecules are generally of interest in the study of connective tissue and attachment molecules and related mechanisms.

The exact roles that fibrillins have in the formation and/or stabilization of extracellular matrix structures as well as their effects on cellular behavior are currently under investigation. However, it is clear that the identification of novel fibrillin-like proteins is of significant importance in increasing understanding of the underlying pathways that lead to certain disease states in which these proteins are implicated, and in developing more effective gene or drug therapies to treat these disorders.

SUMMARY OF THE INVENTION

The invention is based upon the identification of an Open Reading Frame (ORF) in the human genome encoding a novel fibrillin-like polypeptide. This polypeptide will be

- 3 -

referred to herein as the SCS0008 polypeptide. Based on the full coding sequence of the SCS0008 prediction, two splice variants of SCS0008, SCS0008-SV1 and SCS0008-SV2 were additionally discovered. SCS0008, SCS0008-SV1 and SCS0008-SV2 will be referred herein as SCS0008.

- 5 Accordingly, the invention provides isolated SCS0008 polypeptides having the amino acid sequence given by SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, and their mature forms, histidine forms, variants, and fragments, as
- 10 polypeptides having the activity of fibrillin-like polypeptides. The invention includes also the nucleic acids encoding them, vectors containing such nucleic acids, and cell containing these vectors or nucleic acids, as well as other related reagents such as fusion proteins, ligands, and antagonists.

- The invention provides methods for identifying and making these molecules, for
- 15 preparing pharmaceutical compositions containing them, and for using them in the diagnosis, prevention and treatment of diseases.

BRIEF DESCRIPTION OF THE FIGURES

- Figure 1:** alignment of the SCS0008 ORF with known related polypeptide sequences.
- 20 **Figure 2:** Clustal W alignment of SCS0008 predicted amino acid sequence with deduced amino acid sequences for cloned splice variants SCS0008SV1 and SCS0008SV2.
- Figure 3:** Nucleotide sequence of SCS0008 prediction with translation.
- Figure 4:** Nucleotide sequence with translation of SCS0008 product cloned using primers
- 25 SCS0008-CP1nest and SCS0008-CP2nest.
- Figure 5:** Map of pCR4-TOPO-SCS0008SV1.
- Figure 6:** Map of expression vector pEAK12d.
- Figure 7:** Map of pDONR 221.
- Figure 8:** Map of Expression vector pDEST12.2.
- 30 **Figure 9:** Map of pENTR-SCS0008SV1-6HIS.
- Figure 10:** Map of pEAK12d-SCS0008SV1-6HIS.
- Figure 11:** Map of pDEST12.2-SCS0008SV1-6HIS.

- 4 -

Figure 12: Nucleotide sequence of SCS0008 prediction with translation.

Figure 13: Nucleotide sequence with translation of SCS0008 product cloned using primers SCS0008-CP1nest and SCS0008-CP2nest.

Figure 14: Map of pCR4-TOPO-SCS0008-SV2.

5 **Figure 15:** Map of expression vector pEAK12d.

Figure 16: Map of Expression vector pDEST12.2.

Figure 17: Map of pDONR 221.

Figure 18: Map of pENTR-SCS0008SV2-6HIS.

Figure 19: Map of pEAK12d-SCS0008SV2-6HIS.

10 **Figure 20:** Map of pDEST12.2-SCS0008SV2-6HIS.

Figure 21: SMART Domains alignment of SCS0008, SCS0008-SV1, SCS0008-SV2, mouse nephronectin, and other related sequences.

DETAILED DESCRIPTION OF THE INVENTION

15 According to a first aspect of the present invention there is provided an isolated polypeptide having fibrillin-like activity selected from the group consisting of:

- a) the amino acid sequences as recited in SEQ ID NO: 2; SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 13, SEQ ID NO: 14;
- 20 b) the mature form of the polypeptides whose sequences are recited in SEQ ID NO: 2 (SEQ ID NO:3), SEQ ID NO: 6 (SEQ ID NO:8), SEQ ID NO: 7 (SEQ ID NO: 9), SEQ ID NO: 13 (SEQ ID NO: 15), SEQ ID NO: 14 (SEQ ID NO: 16);
- c) the histidine tag form of the polypeptides whose sequences are recited in SEQ ID NO: 2 (SEQ ID NO:4), SEQ ID NO: 6 (SEQ ID NO:10), SEQ ID NO: 7 (SEQ ID NO: 11), SEQ ID NO: 13 (SEQ ID NO: 17), SEQ ID NO: 14 (SEQ ID NO: 18);
- 25 d) a variant of the amino acid sequence recited in SEQ ID NO: 2, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 13, SEQ ID NO: 14, wherein any amino acid specified in the chosen sequence is non-conservatively substituted, provided that no more than 15% of the amino acid residues in the sequence are so
- 30 changed;

- 5 -

- e) an active fragment, precursor, salt, or derivative of the amino acid sequences given in a) to d).

The novel SCS0008 polypeptide described herein was identified on the basis of proprietary bioinformatics techniques. EGF domains were used as query sequences for searches of databases and the final annotation was attributed on the basis of amino acid sequence homology.

The totality of amino acid sequences obtained by translating the known ORFs in the human genome were challenged using this consensus sequence, and the positive hits were further screened for the presence of predicted specific structural and functional "signatures" that are distinctive of a polypeptide of this nature, and finally selected by comparing sequence features with known fibrillin-like polypeptides. Therefore, the novel polypeptides of the invention can be predicted to have fibrillin-like activities.

The terms "active" and "activity" refer to the fibrillin-like properties predicted for the fibrillin-like polypeptide whose amino acid sequence is presented in SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, in the present application. These properties include the ability to form microfibril bundles in connective tissue.

In a second aspect, the invention provides a purified nucleic acid molecule which encodes a polypeptide of the first aspect of the invention. Preferably, the purified nucleic acid molecule has the nucleic acid sequence as recited in SEQ ID NO:1 (encoding the fibrillin-like polypeptide whose amino acid sequence is recited in SEQ ID NO:2), SEQ ID NO:5 (encoding the amino acid sequence recited in SEQ ID NO:6), or SEQ ID NO:12 (encoding the amino acid sequence recited in SEQ ID NO:13).

In a third aspect, the invention provides a purified nucleic acid molecule which hybridizes under high stringency conditions with a nucleic acid molecule of the second aspect of the invention.

In a fourth aspect, the invention provides a vector, such as an expression vector, that contains a nucleic acid molecule of the second or third aspect of the invention.

In a fifth aspect, the invention provides a host cell transformed with a vector of the fourth aspect of the invention.

- 6 -

In a sixth aspect, the invention provides a ligand which binds specifically to, and which preferably inhibits the fibrillin-like activity of a polypeptide of the first aspect of the invention. Ligands to a polypeptide according to the invention may come in various forms, including natural or modified substrates, enzymes, receptors, small organic
5 molecules such as small natural or synthetic organic molecules of up to 2000Da, preferably 800Da or less, peptidomimetics, inorganic molecules, peptides, polypeptides, antibodies, structural or functional mimetics of the aforementioned.

In a seventh aspect, the invention provides a compound that is effective to alter the expression of a natural gene which encodes a polypeptide of the first aspect of the
10 invention or to regulate the activity of a polypeptide of the first aspect of the invention.

A compound of the seventh aspect of the invention may either increase (agonise) or decrease (antagonise) the level of expression of the gene or the activity of the polypeptide. Importantly, the identification of the function of the fibrillin-like polypeptide of the invention allows for the design of screening methods capable of identifying
15 compounds that are effective in the treatment and/or diagnosis of disease.

In an eighth aspect, the invention provides a polypeptide of the first aspect of the invention, or a nucleic acid molecule of the second or third aspect of the invention, or a vector of the fourth aspect of the invention, or a host cell of the fifth aspect of the invention, or a ligand of the sixth aspect of the invention, or a compound of the seventh
20 aspect of the invention, for use in therapy or diagnosis. These molecules may also be used in the manufacture of a medicament for the treatment of diseases such as those in which skin is damaged from aging, injuries or the sun, or for restoring skin damaged from the same, also multiple sclerosis, cancer, bone, joint or ligament reconstruction after fractures or lesions, osteoarthritis, rheumatoid arthritis, osteoporosis,
25 cardiovascular diseases and fibrosis (including liver fibrosis, kidney fibrosis, hepatitis and renal disorders) as well as those conditions and disorders mentioned in the therapeutic uses below.

In a ninth aspect, the invention provides a method of diagnosing a disease in a patient, comprising assessing the level of expression of a natural gene encoding a polypeptide
30 of the first aspect of the invention or the activity of a polypeptide of the first aspect of the invention in tissue from said patient and comparing said level of expression or activity to a control level, wherein a level that is different to said control level is

- 7 -

indicative of disease. Such a method will preferably be carried out *in vitro*. Similar methods may be used for monitoring the therapeutic treatment of disease in a patient, wherein altering the level of expression or activity of a polypeptide or nucleic acid molecule over the period of time towards a control level is indicative of regression of disease.

A preferred method for detecting polypeptides of the first aspect of the invention comprises the steps of: (a) contacting a ligand, such as an antibody, of the sixth aspect of the invention with a biological sample under conditions suitable for the formation of a ligand-polypeptide complex; and (b) detecting said complex.

A number of different such methods according to the ninth aspect of the invention exist, as the skilled reader will be aware, such as methods of nucleic acid hybridization with short probes, point mutation analysis, polymerase chain reaction (PCR) amplification and methods using antibodies to detect aberrant protein levels. Similar methods may be used on a short or long term basis to allow therapeutic treatment of a disease to be monitored in a patient. The invention also provides kits that are useful in these methods for diagnosing disease.

In a tenth aspect, the invention provides for the use of a polypeptide of the first aspect of the invention as a fibrillin-like protein. Suitable uses include use as a secreted glycoprotein, in particular in the context of tissue repair and remodeling, as a result of the ability of the protein to bind to extracellular matrix structures such as connective tissue fibers, basement membranes and blood clots through interacting with other extracellular matrix proteins such as fibronectin, laminins, nidogen, perlecan, fibulin and fibrinogen.

In an eleventh aspect, the invention provides a pharmaceutical composition comprising a polypeptide of the first aspect of the invention, or a nucleic acid molecule of the second or third aspect of the invention, or a vector of the fourth aspect of the invention, or a host cell of the fifth aspect of the invention, or a ligand of the sixth aspect of the invention, or a compound of the seventh aspect of the invention, in conjunction with a pharmaceutically-acceptable carrier.

In a twelfth aspect, the present invention provides a polypeptide of the first aspect of the invention, or a nucleic acid molecule of the second or third aspect of the invention, or a vector of the fourth aspect of the invention, or a host cell of the fifth aspect of the

- 8 -

invention, or a ligand of the sixth aspect of the invention, or a compound of the seventh aspect of the invention, for use in the manufacture of a medicament for the diagnosis or treatment of a disease, such as those in which skin is damaged from aging, injuries or the sun, or for restoring skin damaged from the same, also multiple sclerosis, cancer, bone, joint or ligament reconstruction after fractures or lesions, osteoarthritis, rheumatoid arthritis, osteoporosis, cardiovascular diseases and fibrosis (including liver fibrosis, kidney fibrosis, hepatitis and renal disorders) as well as those conditions and disorders mentioned in the therapeutic uses below.

10 In a thirteenth aspect, the invention provides a method of treating a disease in a patient comprising administering to the patient a polypeptide of the first aspect of the invention, or a nucleic acid molecule of the second or third aspect of the invention, or a vector of the fourth aspect of the invention, or a host cell of the fifth aspect of the invention, or a ligand of the sixth aspect of the invention, or a compound of the seventh aspect of the invention.

15 For diseases in which the expression of a natural gene encoding a polypeptide of the first aspect of the invention, or in which the activity of a polypeptide of the first aspect of the invention, is lower in a diseased patient when compared to the level of expression or activity in a healthy patient, the polypeptide, nucleic acid molecule, ligand or compound administered to the patient should be an agonist. Conversely, for diseases
20 in which the expression of the natural gene or activity of the polypeptide is higher in a diseased patient when compared to the level of expression or activity in a healthy patient, the polypeptide, nucleic acid molecule, ligand or compound administered to the patient should be an antagonist. Examples of such antagonists include antisense nucleic acid molecules, ribozymes and ligands, such as antibodies.

25 In a fourteenth aspect, the invention provides transgenic or knockout non-human animals that have been transformed to express higher, lower or absent levels of a polypeptide of the first aspect of the invention. Such transgenic animals are very useful models for the study of disease and may also be used in screening regimes for the identification of compounds that are effective in the treatment or diagnosis of such a
30 disease.

A summary of standard techniques and procedures which may be employed in order to utilize the invention is given below. It will be understood that this invention is not limited

- 9 -

to the particular methodology, protocols, cell lines, vectors and reagents described. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only and it is not intended that this terminology should limit the scope of the present invention. The extent of the invention is limited only by the terms
5 of the appended claims.

Standard abbreviations for nucleotides and amino acids are used in this specification.

The practice of the present invention will employ, unless otherwise indicated, conventional techniques of molecular biology, microbiology, recombinant DNA technology and immunology, which are within the skill of the those working in the art.

10 Such techniques are explained fully in the literature. Examples of particularly suitable texts for consultation include the following: Sambrook Molecular Cloning; A Laboratory Manual, Second Edition (1989); DNA Cloning, Volumes I and II (D.N Glover ed. 1985); Oligonucleotide Synthesis (M.J. Gait ed. 1984); Nucleic Acid Hybridization (B.D. Hames & S.J. Higgins eds. 1984); Transcription and Translation (B.D. Hames & S.J.
15 Higgins eds. 1984); Animal Cell Culture (R.I. Freshney ed. 1986); Immobilized Cells and Enzymes (IRL Press, 1986); B. Perbal, A Practical Guide to Molecular Cloning (1984); the Methods in Enzymology series (Academic Press, Inc.), especially volumes 154 & 155; Gene Transfer Vectors for Mammalian Cells (J.H. Miller and M.P. Calos eds. 1987, Cold Spring Harbor Laboratory); Immunochemical Methods in Cell and
20 Molecular Biology (Mayer and Walker, eds. 1987, Academic Press, London); Scopes, (1987) Protein Purification: Principles and Practice, Second Edition (Springer Verlag, N.Y.); and Handbook of Experimental Immunology, Volumes I-IV (D.M. Weir and C. C. Blackwell eds. 1986).

The first aspect of the invention includes variants of the amino acid sequence recited in
25 SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17 and SEQ ID NO: 18, wherein any amino acid specified in the chosen sequence is non-conservatively substituted, provided that no more than 15% of the amino acid residues in the sequence are so
30 changed. Protein sequences having the indicated number of non-conservative substitutions can be identified using commonly available bioinformatic tools (Mulder NJ and Apweiler R, 2002; Rehm BH, 2001).

- 10 -

In addition to such sequences, a series of polypeptides forms part of the disclosure of the invention. Being fibrillin-like polypeptides known to go through maturation processes including the proteolytic removal of N-terminal sequences (by signal peptidases and other proteolytic enzymes), the present application also claims the
5 mature form of the polypeptide whose sequences are recited in SEQ ID NO: 2 (SEQ ID NO:3), SEQ ID NO: 6 (SEQ ID NO:8), SEQ ID NO: 7 (SEQ ID NO: 9), SEQ ID NO: 13 (SEQ ID NO: 15), SEQ ID NO: 14 (SEQ ID NO: 16). Mature forms are intended to include any polypeptide showing fibrillin-like activity and resulting from *in vivo* (by the expressing cells or animals) or *in vitro* (by modifying the purified polypeptides with
10 specific enzymes) post-translational maturation processes. Other alternative mature forms can also result from the addition of chemical groups such as sugars or phosphates. The present application also claims the histidine tag form of the polypeptide whose sequences are recited in SEQ ID NO: 2 (SEQ ID NO:4), SEQ ID NO: 6 (SEQ ID NO:10), SEQ ID NO: 7 (SEQ ID NO: 11), SEQ ID NO: 13 (SEQ ID NO: 17), SEQ ID NO: 14 (SEQ ID NO: 18).

Other claimed polypeptides are the active variants of the amino acid sequences given by SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17 or SEQ ID NO: 18, wherein
20 any amino acid specified in the chosen sequence is non-conservatively substituted, provided that no more than 15%, preferably no more than 10%, 5%, 3%, or 1%, of the amino acid residues in the sequence are so changed. The indicated percentage has to be measured over the novel amino acid sequences disclosed.

In accordance with the present invention, any substitution should be preferably a
25 "conservative" or "safe" substitution, which is commonly defined as a substitution introducing an amino acid having sufficiently similar chemical properties (eg a basic, positively charged amino acid should be replaced by another basic, positively charged amino acid), in order to preserve the structure and the biological function of the molecule.

30 The literature provides many models on which the selection of conservative amino acid substitutions can be performed on the basis of statistical and physico-chemical studies on the sequence and/or the structure of proteins (Rogov SI and Nekrasov AN, 2001). Protein design experiments have shown that the use of specific subsets of amino acids

can produce foldable and active proteins, helping in the classification of amino acid "synonymous" substitutions which can be more easily accommodated in protein structure, and which can be used to detect functional and structural homologs and paralogs (Murphy LR *et al.*, 2000). The groups of synonymous amino acids and the groups of more preferred synonymous amino acids are shown in Table I.

Active variants having comparable, or even improved, activity with respect of corresponding fibrillin-like polypeptides may result from conventional mutagenesis technique of the encoding DNA, from combinatorial technologies at the level of encoding DNA sequence (such as DNA shuffling, phage display/selection), or from computer-aided design studies, followed by the validation for the desired activities as described in the prior art.

Specific, non-conservative mutations can be also introduced in the polypeptides of the invention with different purposes. Mutations reducing the affinity of the fibrillin-like polypeptide may increase its ability to be reused and recycled, potentially increasing its therapeutic potency (Robinson CR, 2002). Immunogenic epitopes eventually present in the polypeptides of the invention can be exploited for developing vaccines (Stevanovic S, 2002), or eliminated by modifying their sequence following known methods for selecting mutations for increasing protein stability, and correcting them (van den Burg B and Eijnsink V, 2002; WO 02/05146, WO 00/34317, WO 98/52976).

Further alternative polypeptides of the invention are active fragments, precursors, salts, or functionally-equivalent derivatives of the amino acid sequences described above.

Fragments should present deletions of terminal or internal amino acids not altering their function, and should involve generally a few amino acids, e.g., under ten, and preferably under three, without removing or displacing amino acids which are critical to the functional conformation of the proteins. Small fragments may form an antigenic determinant.

The "precursors" are compounds which can be converted into the compounds of present invention by metabolic and enzymatic processing prior or after the administration to the cells or to the body.

The term "salts" herein refers to both salts of carboxyl groups and to acid addition salts of amino groups of the polypeptides of the present invention. Salts of a carboxyl group may be formed by means known in the art and include inorganic salts, for example,

- 12 -

sodium, calcium, ammonium, ferric or zinc salts, and the like, and salts with organic bases as those formed, for example, with amines, such as triethanolamine, arginine or lysine, piperidine, procaine and the like. Acid addition salts include, for example, salts with mineral acids such as, for example, hydrochloric acid or sulfuric acid, and salts
5 with organic acids such as, for example, acetic acid or oxalic acid. Any of such salts should have substantially similar activity to the peptides and polypeptides of the invention or their analogs.

The term "derivatives" as herein used refers to derivatives which can be prepared from the functional groups present on the lateral chains of the amino acid moieties or on the
10 amino- or carboxy-terminal groups according to known methods. Such molecules can result also from other modifications which do not normally alter primary sequence, for example *in vivo* or *in vitro* chemical derivatization of polypeptides (acetylation or carboxylation), those made by modifying the pattern of phosphorylation (introduction of phosphotyrosine, phosphoserine, or phosphothreonine residues) or glycosylation (by
15 exposing the polypeptide to mammalian glycosylating enzymes) of a peptide during its synthesis and processing or in further processing steps. Alternatively, derivatives may include esters or aliphatic amides of the carboxyl-groups and N-acyl derivatives of free amino groups or O-acyl derivatives of free hydroxyl-groups and are formed with acyl-groups as for example alkanoyl- or aryl-groups.

20 The generation of the derivatives may involve a site-directed modification of an appropriate residue, in an internal or terminal position. The residues used for attachment should they have a side-chain amenable for polymer attachment (i.e., the side chain of an amino acid bearing a functional group, e.g., lysine, aspartic acid, glutamic acid, cysteine, histidine, etc.). Alternatively, a residue having a side chain
25 amenable for polymer attachment can replace an amino acid of the polypeptide, or can be added in an internal or terminal position of the polypeptide. Also, the side chains of the genetically encoded amino acids can be chemically modified for polymer attachment, or unnatural amino acids with appropriate side chain functional groups can be employed. The preferred method of attachment employs a combination of peptide
30 synthesis and chemical ligation. Advantageously, the attachment of a water-soluble polymer will be through a biodegradable linker, especially at the amino-terminal region of a protein. Such modification acts to provide the protein in a precursor (or "pro-drug")

form, that, upon degradation of the linker releases the protein without polymer modification.

Polymer attachment may be not only to the side chain of the amino acid naturally occurring in a specific position of the antagonist or to the side chain of a natural or
5 unnatural amino acid that replaces the amino acid naturally occurring in a specific position of the antagonist, but also to a carbohydrate or other moiety that is attached to the side chain of the amino acid at the target position. Rare or unnatural amino acids can be also introduced by expressing the protein in specifically engineered bacterial strains (Bock A, 2001).

10 All the above indicated variants can be natural, being identified in organisms other than humans, or artificial, being prepared by chemical synthesis, by site-directed mutagenesis techniques, or any other known technique suitable thereof, which provide a finite set of substantially corresponding mutated or shortened peptides or polypeptides which can be routinely obtained and tested by one of ordinary skill in the
15 art using the teachings presented in the prior art.

The novel amino acid sequences disclosed in the present patent application can be used to provide different kind of reagents and molecules. Examples of these compounds are binding proteins or antibodies that can be identified using their full sequence or specific fragments, such as antigenic determinants. Peptide libraries can
20 be used in known methods (Tribbick G, 2002) for screening and characterizing antibodies or other proteins binding the claimed amino acid sequences, and for identifying alternative forms of the polypeptides of the invention having similar binding properties.

The present patent application discloses also fusion proteins comprising any of the
25 polypeptides described above. These polypeptides should contain protein sequence heterologous to the one disclosed in the present patent application, without significantly impairing the fibrillin-like activity of the polypeptide and possibly providing additional properties. Examples of such properties are an easier purification procedure, a longer lasting half-life in body fluids, an additional binding moiety, the maturation by means of
30 an endoproteolytic digestion, or extracellular localization. This latter feature is of particular importance for defining a specific group of fusion or chimeric proteins included in the above definition since it allows the claimed molecules to be localized in

- 14 -

the space where not only isolation and purification of these polypeptides is facilitated, but also where generally fibrillin-like polypeptides and their receptor interact.

Design of the moieties, ligands, and linkers, as well methods and strategies for the construction, purification, detection and use of fusion proteins are disclosed in the literature (Nilsson J *et al.*, 1997; Methods Enzymol, Vol. 326-328, Academic Press, 2000). The preferred one or more protein sequences which can be comprised in the fusion proteins belong to these protein sequences: membrane-bound protein, immunoglobulin constant region, multimerization domains, extracellular proteins, signal peptide-containing proteins, export signal-containing proteins. Features of these sequences and their specific uses are disclosed in a detailed manner, for example, for albumin fusion proteins (WO 01/77137), fusion proteins including multimerization domain (WO 01/02440, WO 00/24782), immunoconjugates (Garnett MC, 2001), or fusion protein providing additional sequences which can be used for purifying the recombinant products by affinity chromatography (Constans A, 2002; Burgess RR and Thompson NE, 2002; Lowe CR *et al.*, 2001; J. Bioch. Biophys. Meth., vol. 49 (1-3), 2001; Sheibani N, 1999).

The polypeptides of the invention can be used to generate and characterize ligands binding specifically to them. These molecules can be natural or artificial, very different from the chemical point of view (binding proteins, antibodies, molecularly imprinted polymers), and can be produced by applying the teachings in the art (WO 02/74938; Kuroiwa Y *et al.*, 2002; Haupt K, 2002; van Dijk MA and van de Winkel JG, 2001; Gavilondo JV and Larrick JW, 2000). Such ligands can antagonize or inhibit the fibrillin-like activity of the polypeptide against which they have been generated. In particular, common and efficient ligands are represented by extracellular domain of a membrane-bound protein or antibodies, which can be in the form monoclonal, polyclonal, humanized antibody, or an antigen binding fragment.

The polypeptides and the polypeptide-based derived reagents described above can be in alternative forms, according to the desired method of use and/or production, such as active conjugates or complexes with a molecule chosen amongst radioactive labels, fluorescent labels, biotin, or cytotoxic agents.

Specific molecules, such as peptide mimetics, can be also designed on the sequence and/or the structure of a polypeptide of the invention. Peptide mimetics (also called

- 15 -

peptidomimetics) are peptides chemically modified at the level of amino acid side chains, of amino acid chirality, and/or of the peptide backbone. These alterations are intended to provide agonists or antagonists of the polypeptides of the invention with improved preparation, potency and/or pharmacokinetics features.

- 5 For example, when the peptide is susceptible to cleavage by peptidases following injection into the subject is a problem, replacement of a particularly sensitive peptide bond with a non-cleavable peptide mimetic can provide a peptide more stable and thus more useful as a therapeutic. Similarly, the replacement of an L -amino acid residue is a standard way of rendering the peptide less sensitive to proteolysis, and finally more
- 10 similar to organic compounds other than peptides. Also useful are amino-terminal blocking groups such as t-butyloxycarbonyl, acetyl, theryl, succinyl, methoxysuccinyl, suberyl, adipyl, azelanyl, dansyl, benzyloxycarbonyl, fluorenylmethoxycarbonyl, methoxyazelanyl, methoxyadipyl, methoxysuberyl, and 2,4-dinitrophenyl. Many other modifications providing increased potency, prolonged activity, easiness of purification,
- 15 and/or increased half-life are disclosed in the prior art (WO 02/10195; Villain M *et al.*, 2001).

Preferred alternative, synonymous groups for amino acids derivatives included in peptide mimetics are those defined in Table II. A non-exhaustive list of amino acid derivatives also include aminoisobutyric acid (Aib), hydroxyproline (Hyp), 1,2,3,4 -

20 tetrahydro-isoquinoline-3-COOH, indoline-2carboxylic acid, 4-difluoro-proline, L-thiazolidine-4-carboxylic acid, L-homoproline, 3,4-dehydro-proline, 3,4-dihydroxy-phenylalanine, cyclohexyl-glycine, and phenylglycine.

By "amino acid derivative" is intended an amino acid or amino acid -like chemical entity other than one of the 20 genetically encoded naturally occurring amino acids. In

25 particular, the amino acid derivative may contain substituted or non -substituted, linear, branched, or cyclic alkyl moieties, and may include one or more heteroatoms. The amino acid derivatives can be made de novo or obtained from commercial sources (Calbiochem-Novabiochem AG, Switzerland; Bachem, USA).

Various methodologies for incorporating unnatural amino acids derivatives into

30 proteins, using both *in vitro* and *in vivo* translation systems, to probe and/or improve protein structure and function are disclosed in the literature (Dougherty DA, 2000). Techniques for the synthesis and the development of peptide mimetics, as well as non -

- 16 -

peptide mimetics, are also well known in the art (Golebiowski A *et al.*, 2001; Hruby VJ and Balse PM, 2000; Sawyer TK, in "Structure Based Drug Design", edited by Veerapandian P, Marcel Dekker Inc., pg. 557-663, 1997).

Another object of the present invention are isolated nucleic acids encoding for the polypeptides of the invention having fibrillin-like activity, the polypeptides binding to an antibody or a binding protein generated against them, the corresponding fusion proteins, or mutants having antagonistic activity as disclosed above. Preferably, these nucleic acids should comprise a DNA sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 5, SEQ ID NO: 12, the coding sequence of SEQ ID NO: 1 (the coding region starts at nucleotide 205 and ends at 1425), of SEQ ID NO: 5 (the coding region starts at nucleotide 205 and ends at 1590), of SEQ ID NO: 12 (the coding region starts at nucleotide 205 and ends at 1503), or the complement of said DNA sequences.

Alternatively, the nucleic acids of the invention should hybridize under high stringency conditions, or exhibit at least about 85% identity over a stretch of at least about 30 nucleotides, with a nucleic acid consisting of SEQ ID NO: 1, SEQ ID NO: 5, SEQ ID NO: 12, or be a complement of said DNA sequences.

The wording "high stringency conditions" refers to conditions in a hybridization reaction that facilitate the association of very similar molecules and consist in the overnight incubation at 60-65°C in a solution comprising 50 % formamide, 5X SSC (150 m M NaCl, 15 m M trisodium citrate), 50 mM sodium phosphate (p H 7.6), 5x Denhardt's solution, 10 % dextran sulphate, and 20 microgram/ml denatured, sheared salmon sperm DNA, followed by washing the filters in 0.1X SSC at the same temperature.

These nucleic acids, including nucleotide sequences substantially the same, can be comprised in plasmids, vectors and any other DNA construct which can be used for maintaining, modifying, introducing, or expressing the encoding polypeptide. In particular, vectors wherein said nucleic acid molecule is operatively linked to expression control sequences can allow expression in prokaryotic or eukaryotic host cells of the encoded polypeptide.

The wording "nucleotide sequences substantially the same" includes all other nucleic acid sequences which, by virtue of the degeneracy of the genetic code, also code for

- 17 -

the given amino acid sequences. In this sense, the literature provides indications on preferred or optimized codons for recombinant expression (Kane JF *et al.*, 1995).

The nucleic acids and the vectors can be introduced into cells with different purposes, generating transgenic cells and organisms. A process for producing cells capable of
5 expressing a polypeptide of the invention comprises genetically engineering cells with such vectors and nucleic acids.

In particular, host cells (e.g. bacterial cells) can be modified by transformation for allowing the transient or stable expression of the polypeptides encoded by the nucleic acids and the vectors of the invention. Alternatively, said molecules can be used to
10 generate transgenic animal cells or non-human animals (by non- / homologous recombination or by any other method allowing their stable integration and maintenance), having enhanced or reduced expression levels of the polypeptides of the invention, when the level is compared with the normal expression levels. Such precise modifications can be obtained by making use of the nucleic acids of the
15 inventions and of technologies associated, for example, to gene therapy (Meth. Enzymol., vol. 346, 2002) or to site-specific recombinases (Kolb AF, 2002). Model systems based on the fibrillin-like polypeptides disclosed in the present patent application for the systematic study of their function can be also generated by gene targeting into human cell lines (Bunz F, 2002).

20 Gene silencing approaches may also be undertaken to down-regulate endogenous expression of a gene encoding a polypeptide of the invention. RNA interference (RNAi) (Elbashir, SM *et al.*, Nature 2001, 411, 494-498) is one method of sequence specific post-transcriptional gene silencing that may be employed. Short dsRNA oligonucleotides are synthesised *in vitro* and introduced into a cell. The sequence
25 specific binding of these dsRNA oligonucleotides triggers the degradation of target mRNA, reducing or ablating target protein expression.

Efficacy of the gene silencing approaches assessed above may be assessed through the measurement of polypeptide expression (for example, by Western blotting), and at the RNA level using TaqMan-based methodologies.

30 The polypeptides of the invention can be prepared by any method known in the art, including recombinant DNA-related technologies, and chemical synthesis technologies. In particular, a method for making a polypeptide of the invention may comprise

- 18 -

culturing a host or transgenic cell as described above under conditions in which the nucleic acid or vector is expressed, and recovering the polypeptide encoded by said nucleic acid or vector from the culture. For example, when the vector expresses the polypeptide as a fusion protein with an extracellular or signal-peptide containing proteins, the recombinant product can be secreted in the extracellular space, and can be more easily collected and purified from cultured cells in view of further processing or, alternatively, the cells can be directly used or administered.

The DNA sequence coding for the proteins of the invention can be inserted and ligated into a suitable episomal or non- / homologously integrating vectors, which can be introduced in the appropriate host cells by any suitable means (transformation, transfection, conjugation, protoplast fusion, electroporation, calcium phosphate-precipitation, direct microinjection, etc.). Factors of importance in selecting a particular plasmid or viral vector include: the ease with which recipient cells that contain the vector, may be recognized and selected from those recipient cells which do not contain the vector; the number of copies of the vector which are desired in a particular host; and whether it is desirable to be able to "shuttle" the vector between host cells of different species.

The vectors should allow the expression of the isolated or fusion protein including the polypeptide of the invention in the Prokaryotic or Eukaryotic host cells under the control of transcriptional initiation / termination regulatory sequences, which are chosen to be constitutively active or inducible in said cell. A cell line substantially enriched in such cells can be then isolated to provide a stable cell line.

For Eukaryotic hosts (e.g. yeasts, insect, plant, or mammalian cells), different transcriptional and translational regulatory sequences may be employed, depending on the nature of the host. They may be derived from viral sources, such as adenovirus, bovine papilloma virus, Simian virus or the like, where the regulatory signals are associated with a particular gene which has a high level of expression. Examples are the TK promoter of the Herpes virus, the SV40 early promoter, the yeast gal4 gene promoter, etc. Transcriptional initiation regulatory signals may be selected which allow for repression and activation, so that expression of the genes can be modulated. The cells stably transformed by the introduced DNA can be selected by introducing one or more markers allowing the selection of host cells which contain the expression vector. The marker may also provide for phototrophy to an auxotrophic host, biocide resistance,

- 19 -

e.g. antibiotics, or heavy metals such as copper, or the like. The selectable marker gene can either be directly linked to the DNA gene sequences to be expressed, or introduced into the same cell by co-transfection.

Host cells may be either prokaryotic or eukaryotic. Preferred are eukaryotic hosts, e.g.
5 mammalian cells, such as human, monkey, mouse, and Chinese Hamster Ovary (CHO) cells, because they provide post-translational modifications to proteins, including correct folding and glycosylation. Also yeast cells can carry out post-translational peptide modifications including glycosylation. A number of recombinant DNA strategies exist which utilize strong promoter sequences and high copy number of
10 plasmids which can be utilized for production of the desired proteins in yeast. Yeast recognizes leader sequences in cloned mammalian gene products and secretes peptides bearing leader sequences (i.e., pre-peptides).

The above mentioned embodiments of the invention can be achieved by combining the disclosure provided by the present patent application on the sequence of novel fibrillin -
15 like polypeptides with the knowledge of common molecular biology techniques.

Many books and reviews provides teachings on how to clone and produce recombinant proteins using vectors and Prokaryotic or Eukaryotic host cells, such as some titles in the series "A Practical Approach" published by Oxford University Press ("DNA Cloning 2: Expression Systems", 1995; "DNA Cloning 4: Mammalian Systems", 1996; "Protein
20 Expression", 1999; "Protein Purification Techniques", 2001).

Moreover, updated and more focused literature provides an overview of the technologies for expressing polypeptides in a high-throughput manner (Chambers SP, 2002; Coleman TA, *et al.*, 1997), of the cell systems and the processes used industrially for the large-scale production of recombinant proteins having therapeutic
25 applications (Andersen DC and Krummen L, 2002, Chu L and Robinson DK, 2001), and of alternative eukaryotic expression systems for expressing the polypeptide of interest, which may have considerable potential for the economic production of the desired protein, such the ones based on transgenic plants (Giddings G, 2001) or the yeast *Pichia pastoris* (Lin Cereghino GP *et al.*, 2002). Recombinant protein products
30 can be rapidly monitored with various analytical technologies during purification to verify the amount and the quantity of the expressed polypeptides (Baker KN *et al.*,

- 20 -

2002), as well as to check if there is problem of bioequivalence and immunogenicity (Schellekens H, 2002; Gendel SM, 2002).

Totally synthetic fibrillin-like polypeptides are disclosed in the literature and many examples of chemical synthesis technologies, which can be effectively applied for the
5 fibrillin-like polypeptides of the invention given their short length, are available in the literature, as solid phase or liquid phase synthesis technologies. for example, the amino acid corresponding to the carboxy-terminus of the peptide to be synthesized is bound to a support which is insoluble in organic solvents, and by alternate repetition of reactions, one wherein amino acids with their amino groups and side chain functional
10 groups protected with appropriate protective groups are condensed one by one in order from the carboxy-terminus to the amino-terminus, and one where the amino acids bound to the resin or the protective group of the amino groups of the peptides are released, the peptide chain is thus extended in this manner. Solid phase synthesis methods are largely classified by the tBoc method and the Fmoc method, depending
15 on the type of protective group used. Typically used protective groups include tBoc (t-butoxycarbonyl), Cl-Z (2-chlorobenzyloxycarbonyl), Br-Z (2-bromobenzyloxycarbonyl), Bzl (benzyl), Fmoc (9-fluorenylmethoxycarbonyl), Mbh (4,4'-dimethoxydibenzhydryl), Mtr (4-methoxy-2,3,6-trimethylbenzenesulphonyl), Trt (trityl), Tos (tosyl), Z (benzyloxycarbonyl) and Cl2-Bzl (2,6-dichlorobenzyl) for the amino groups; NO2 (nitro)
20 and Pmc (2,2,5,7,8-pentamethylchromane-6-sulphonyl) for the guanidino groups; and tBu (t-butyl) for the hydroxyl groups). After synthesis of the desired peptide, it is subjected to the de-protection reaction and cut out from the solid support. Such peptide cutting reaction may be carried with hydrogen fluoride or tri-fluoromethane sulfonic acid for the Boc method, and with TFA for the Fmoc method.

25 The purification of the polypeptides of the invention can be carried out by any one of the methods known for this purpose, i.e. any conventional procedure involving extraction, precipitation, chromatography, electrophoresis, or the like. A further purification procedure that may be used in preference for purifying the protein of the invention is affinity chromatography using monoclonal antibodies or affinity groups,
30 which bind the target protein and which are produced and immobilized on a gel matrix contained within a column. Impure preparations containing the proteins are passed through the column. The protein will be bound to the column by heparin or by the specific antibody while the impurities will pass through. After washing, the protein is

eluted from the gel by a change in pH or ionic strength. Alternatively, HPLC (High Performance Liquid Chromatography) can be used. The elution can be carried using a water-acetonitrile-based solvent commonly employed for protein purification.

The disclosure of the novel polypeptides of the invention, and the reagents disclosed in connection to them (antibodies, nucleic acids, cells) allows also to screen and characterize compounds that enhance or reduce their expression level into a cell or in an animal.

"Oligonucleotides" refers to either a single stranded polydeoxynucleotide or two complementary polydeoxynucleotide strands which may be chemically synthesized. Such synthetic oligonucleotides have no 5' phosphate and thus will not ligate to another oligonucleotide without adding a phosphate with an ATP in the presence of a kinase. A synthetic oligonucleotide will ligate to a fragment that has not been dephosphorylated.

The invention includes purified preparations of the compounds of the invention (polypeptides, nucleic acids, cells, etc.). Purified preparations, as used herein, refers to the preparations which contain at least 1%, preferably at least 5%, by dry weight of the compounds of the invention.

The present patent application discloses a series of novel fibrillin-like polypeptides and of related reagents having several possible applications. In particular, whenever an increase in the fibrillin-like activity of a polypeptide of the invention is desirable in the therapy or in the prevention of a disease, reagents such as the disclosed fibrillin-like polypeptides, the corresponding fusion proteins and peptide mimetics, the encoding nucleic acids, the expressing cells, or the compounds enhancing their expression can be used.

Therefore, the present invention discloses pharmaceutical compositions for the treatment or prevention of diseases needing an increase in the fibrillin-like activity of a polypeptide of the invention, which contain one of the disclosed fibrillin-like polypeptides, the corresponding fusion proteins and peptide mimetics, the encoding nucleic acids, the expressing cells, or the compounds enhancing their expression, as active ingredient. The process for the preparation of these pharmaceutical compositions comprises combining the disclosed fibrillin-like polypeptides, the corresponding fusion proteins and peptide mimetics, the encoding nucleic acids, the

- 22 -

expressing cells, or the compounds enhancing their expression, together with a pharmaceutically acceptable carrier. Methods for the treatment or prevention of diseases needing an increase in the fibrillin-like activity of a polypeptide of the invention, comprise the administration of a therapeutically effective amount of the disclosed fibrillin-like polypeptides, the corresponding fusion proteins and peptide mimetics, the encoding nucleic acids, the expressing cells, or the compounds enhancing their expression.

Amongst the reagents disclosed in the present patent application, the ligands, the antagonists or the compounds reducing the expression or the activity of polypeptides of the invention have several applications, and in particular they can be used in the therapy or in the diagnosis of a disease associated to the excessive fibrillin-like activity of a polypeptide of the invention.

Therefore, the present invention discloses pharmaceutical compositions for the treatment or prevention of diseases associated to the excessive fibrillin-like activity of a polypeptide of the invention, which contain one of the ligands, antagonists, or compounds reducing the expression or the activity of such polypeptides, as active ingredient. The process for the preparation of these pharmaceutical compositions comprises combining the ligand, the antagonist, or the compound, together with a pharmaceutically acceptable carrier. Methods for the treatment or prevention of diseases associated to the excessive fibrillin-like activity of the polypeptide of the invention, comprise the administration of a therapeutically effective amount of the antagonist, the ligand or of the compound.

The pharmaceutical compositions of the invention may contain, in addition to fibrillin-like polypeptide or to the related reagent, suitable pharmaceutically acceptable carriers, biologically compatible vehicles and additives which are suitable for administration to an animal (for example, physiological saline) and eventually comprising auxiliaries (like excipients, stabilizers, adjuvants, or diluents) which facilitate the processing of the active compound into preparations which can be used pharmaceutically.

The pharmaceutical compositions may be formulated in any acceptable way to meet the needs of the mode of administration. For example, of biomaterials, sugar-macromolecule conjugates, hydrogels, polyethylene glycol and other natural or synthetic polymers can be used for improving the active ingredients in terms of drug

- 23 -

delivery efficacy. Technologies and models to validate a specific mode of administration are disclosed in literature (Davis BG and Robinson MA, 2002; Gupta P *et al.*, 2002; Luo B and Prestwich GD, 2001; Cleland JL *et al.*, 2001; Pillai O and Panchagnula R, 2001).

- 5 Polymers suitable for these purposes are biocompatible, namely, they are non-toxic to biological systems, and many such polymers are known. Such polymers may be hydrophobic or hydrophilic in nature, biodegradable, non-biodegradable, or a combination thereof. These polymers include natural polymers (such as collagen, gelatin, cellulose, hyaluronic acid), as well as synthetic polymers (such as polyesters, polyorthoesters, polyanhydrides). Examples of hydrophobic non-degradable polymers include polydimethyl siloxanes, polyurethanes, polytetrafluoroethylenes, polyethylenes, polyvinyl chlorides, and polymethyl methacrylates. Examples of hydrophilic non-degradable polymers include poly(2-hydroxyethyl methacrylate), polyvinyl alcohol, poly(N-vinyl pyrrolidone), polyalkylenes, polyacrylamide, and copolymers thereof.
- 15 Preferred polymers comprise as a sequential repeat unit ethylene oxide, such as polyethylene glycol (PEG).

Any accepted mode of administration can be used and determined by those skilled in the art to establish the desired blood levels of the active ingredients. For example, administration may be by various parenteral routes such as subcutaneous, intravenous, intradermal, intramuscular, intraperitoneal, intranasal, transdermal, oral, or buccal routes. The pharmaceutical compositions of the present invention can also be administered in sustained or controlled release dosage forms, including depot injections, osmotic pumps, and the like, for the prolonged administration of the polypeptide at a predetermined rate, preferably in unit dosage forms suitable for single administration of precise dosages.

25

Parenteral administration can be by bolus injection or by gradual perfusion over time. Preparations for parenteral administration include sterile aqueous or non-aqueous solutions, suspensions, and emulsions, which may contain auxiliary agents or excipients known in the art, and can be prepared according to routine methods. In addition, suspension of the active compounds as appropriate oily injection suspensions may be administered. Suitable lipophilic solvents or vehicles include fatty oils, for example, sesame oil, or synthetic fatty acid esters, for example, sesame oil, or synthetic fatty acid esters, for example, ethyl oleate or triglycerides. Aqueous injection

30

- 24 -

suspensions that may contain substances increasing the viscosity of the suspension include, for example, sodium carboxymethyl cellulose, sorbitol, and/or dextran. Optionally, the suspension may also contain stabilizers. Pharmaceutical compositions include suitable solutions for administration by injection, and contain from about 0.01 to 5 99.99 percent, preferably from about 20 to 75 percent of active compound together with the excipient.

The wording "therapeutically effective amount" refers to an amount of the active ingredients that is sufficient to affect the course and the severity of the disease, leading to the reduction or remission of such pathology. The effective amount will depend on 10 the route of administration and the condition of the patient.

The wording "pharmaceutically acceptable" is meant to encompass any carrier, which does not interfere with the effectiveness of the biological activity of the active ingredient and that is not toxic to the host to which is administered. For example, for parenteral administration, the above active ingredients may be formulated in unit dosage form for 15 injection in vehicles such as saline, dextrose solution, serum albumin and Ringer's solution. Carriers can be selected also from starch, cellulose, talc, glucose, lactose, sucrose, gelatin, malt, rice, flour, chalk, silica gel, magnesium stearate, sodium stearate, glycerol monostearate, sodium chloride, dried skim milk, glycerol, propylene glycol, water, ethanol, and the various oils, including those of petroleum, animal, 20 vegetable or synthetic origin (peanut oil, soybean oil, mineral oil, sesame oil).

It is understood that the dosage administered will be dependent upon the age, sex, health, and weight of the recipient, kind of concurrent treatment, if any, frequency of treatment, and the nature of the effect desired. The dosage will be tailored to the individual subject, as is understood and determinable by one of skill in the art. The total 25 dose required for each treatment may be administered by multiple doses or in a single dose. The pharmaceutical composition of the present invention may be administered alone or in conjunction with other therapeutics directed to the condition, or directed to other symptoms of the condition. Usually a daily dosage of active ingredient is comprised between 0.01 to 100 milligrams per kilogram of body weight per day. 30 Ordinarily 1 to 40 milligrams per kilogram per day given in divided doses or in sustained release form is effective to obtain the desired results. Second or subsequent administrations can be performed at a dosage, which is the same, less than, or greater than the initial or previous dose administered to the individual.

- 25 -

Apart from methods having a therapeutic or a production purpose, several other methods can make use of the fibrillin-like polypeptides and of the related reagents disclosed in the present patent application.

In a first example, a method is provided for screening candidate compounds effective
5 to treat a disease related to a fibrillin-like polypeptide of the invention, said method comprising:

- (a) contacting host cells expressing such polypeptide, transgenic non-human animals, or transgenic animal cells having enhanced or reduced expression levels of the polypeptide, with a candidate compound and
- 10 (b) determining the effect of the compound on the animal or on the cell.

In a second example there is provided a method for identifying a candidate compound as an antagonist/inhibitor or agonist/activator of a polypeptide of the invention, the method comprising:

- 15 (a) contacting the polypeptide, the compound, and a mammalian cell or a mammalian cell membrane; and
- (b) measuring whether the molecule blocks or enhances the interaction of the polypeptide, or the response that results from such interaction, with the mammalian cell or the mammalian cell membrane.

In a third example, a method for determining the activity and/or the presence of the
20 polypeptide of the invention in a sample, can detect either the polypeptide or the encoding RNA/DNA. Thus, such a method comprises:

- (a) providing a protein-containing sample;
- (b) contacting said sample with a ligand of the invention; and
- 25 (c) determining the presence of said ligand bound to said polypeptide, thereby determining the activity and/or the presence of polypeptide in said sample.

In an alternative, the method comprises:

- (a) providing a nucleic acids-containing sample;
- (b) contacting said sample with a nucleic acid of the invention; and

- 26 -

(c) determining the hybridization of said nucleic acid with a nucleic acid into the sample, thereby determining the presence of the nucleic acid in the sample.

In this sense, a primer sequence derived from the nucleotide sequences presented in SEQ ID NO:1, SEQ ID NO: 5, or SEQ ID NO: 12, can be used as well for determining
5 the presence or the amount of a transcript or of a nucleic acid encoding a polypeptide of invention in a sample by means of Polymerase Chain Reaction amplification.

A further object of the present invention are kits for measuring the activity and/or the presence of fibrillin-like polypeptide of the invention in a sample comprising one or more of the reagents disclosed in the present patent application: a fibrillin-like
10 polypeptide of the invention, an antagonist, ligand or peptide mimetic, an isolated nucleic acid or the vector, a pharmaceutical composition, an expressing cell, or a compound increasing or decreasing the expression levels.

Such kits can be used for *in vitro* diagnostic or screenings methods, and their actual composition should be adapted to the specific format of the sample (e.g. biological
15 sample tissue from a patient), and the molecular species to be measured. For example, if it is desired to measure the concentration of the fibrillin-like polypeptide, the kit may contain an antibody and the corresponding protein in a purified form to compare the signal obtained in Western blot. Alternatively, if it is desired to measure the concentration of the transcript for the fibrillin-like polypeptide, the kit may contain a
20 specific nucleic acid probe designed on the corresponding ORF sequence, or may be in the form of nucleic acid array containing such probe. The kits can be also in the form of protein-, peptide mimetic-, or cell-based microarrays (Templin MF *et al.*, 2002; Pellois JP *et al.*, 2002; Blagoev B and Pandey A, 2001), allowing high-throughput proteomics studies, by making use of the proteins, peptide mimetics and cells
25 disclosed in the present patent application.

The present patent application discloses novel fibrillin-like polypeptides and a series of related reagents that may be useful, as active ingredients in pharmaceutical compositions appropriately formulated, in the treatment or prevention of diseases and conditions such as those in which skin is damaged from aging, injuries or the sun, or
30 for restoring skin damaged from the same, also multiple sclerosis, cancer, bone, joint or ligament reconstruction after fractures or lesions, osteoarthritis, rheumatoid arthritis,

- 27 -

osteoporosis, cardiovascular diseases and fibrosis (including liver fibrosis, kidney fibrosis, hepatitis and renal disorders).

The therapeutic applications of the polypeptides of the invention and of the related reagents can be evaluated (in terms of safety, pharmacokinetics and efficacy) by the means of the *in vivo* / *in vitro* assays making use of animal cell, tissues and or by the means of *in silico* / computational approaches (Johnson DE and Wolfgang GH, 2000), known for the validation of fibrillin-like polypeptides and other biological products during drug discovery and preclinical development.

10 Therapeutic Uses

SCS0008 nucleic acid molecules, polypeptides, and agonists and antagonists thereof can be used to treat, diagnose, ameliorate, or prevent a number of diseases, disorders, or conditions, including those recited herein.

SCS0008 polypeptides' agonists and antagonists include those molecules which regulate SCS0008 polypeptides' activity and either increase or decrease at least one activity of the mature form of the SCS0008 polypeptides. Agonists or antagonists may be co-factors, such as a protein, peptide, carbohydrate, lipid, or small molecular weight molecule, which interact with the SCS0008 polypeptides and thereby regulate their activity.

Potential polypeptide agonists or antagonists include antibodies that react with either soluble or membrane-bound forms of the SCS0008 polypeptides. Molecules that regulate SCS0008 polypeptides' expression typically include nucleic acids encoding the SCS0008 polypeptides that can act as anti-sense regulators of expression.

Morimura et al. and Brandenberger et al. showed that nephronectin (or POEM) is a novel ligand for $\alpha_8\beta_1$ integrin. As SCS0008 seems to be the human ortholog of nephronectin and as the SCS0008 polypeptides do contain a RGD integrin binding tripeptide (example 4), SCS0008 is likely to act as well as a ligand for $\alpha_8\beta_1$ integrin. Since it is known that $\alpha_8\beta_1$ integrin is implicated in the development of particular diseases, SCS0008, as a ligand, might as well act as a contributing factor in the onset of those disorders. In addition, Morimura et al. And Brandenberger et al. have showed that nephronectin is expressed in many tissues, including the organs and tissues mentioned below. As such, the present SCS0008 polypeptides are not only useful in

- 28 -

the tissues and organs mentioned hereafter, but also in all of those mentioned by Morimura et al. and Brandenberger et al.

Morimura et al. and Brandenberger et al. indicate that nephronectin is involved in kidney morphogenesis and function. Miner Jeffrey suggests that nephronectin may play a role in establishing or maintaining the filtration barrier (Jeffrey H. Miner, The Journal of Cell Biology, Volume 154, Number 2, July 23, 2001, pp.257-259). Miner Jeffrey further points out that nephronectin and laminin-10 appear to be colocalized in the Wolffian duct and ureteric bud basement membranes, it is possible that they may interact with each other and/or cooperate to promote efficient ureteric bud outgrowth and branching, either through parallel or common pathways. In addition, Brandenberger et al. points out that secreted Glial cell-derived neurotrophic factor (GDNF) has been shown to control in-growth of the ureteric bud and that $\alpha_6\beta_1$ is likely involved in this signalling pathway. It is also known that GDNF-ligand GFR α 1 (also GFR α 2, GFR α 3, GFR α 4) activates RET tyrosine kinase. Schuchardt et al. found that RET is required for ureteric bud growth and branching. Eya1 has also been involved in human congenital anomalies of the kidney and urinary tract (CAKUT; Nakanishi K, Yoshikawa N. *Pediatr Int.* 2003 Oct;45(5):610-6.). As such SCS0008 nucleic acid molecules, polypeptides, and agonists and antagonists thereof and particularly SCS0008-SV2 (which is a splice variant from the kidney, example 3) may be useful in diagnosing or treating renal branching defects, cysts, stone formation, vesicoureteral reflux, renal failure, defects in the filtration barrier, hydronephrosis, congenital anomalies of the kidney and urinary tract (CAKUT) such as bilateral renal agenesis. SCS0008 antagonists and particularly SCS0008-SV2 antagonists (e.g. antibodies targeted to SCS0008) may be useful in diagnosing or treating kidney cancers, particularly renal dysplasia or hypoplasia.

Lu et al. (Lu M et al., *J Cell Sci.* 2002 Dec 1 ;115(Pt 23) :4641-8) showed that $\alpha_6\beta_1$ integrin is expressed by alveolar interstitial cells in normal lung and is upregulated during the development of fibrosis. Binding of another ligand of $\alpha_6\beta_1$ integrin, Latency-Associated Peptide (LAP)-Transforming Growth Factor-beta 1 (TGFbeta1), results in cell proliferation and might therefore be important in the development of fibrosis.

Without wishing to be bound to theory, SCS0008, likewise, by $\alpha_6\beta_1$ integrin binding, might be involved in this cell proliferation. In addition, Sparrow and Lamb state that GDNF appears to be produced by airway smooth muscle (ASM), which is an integral component of the primordial lung, and might be involved in smooth muscle myogenesis

(Sparrow MP, Lam JP. *Respir Physiol Neurobiol*.2003 Sep 16;137(2-3):361-72). Levine et al. demonstrated de novo expression of $\alpha_8\beta_1$ integrin in activated hepatic stellate cells, the myofibroblasts equivalent in liver (Levine D et al., *Am J Pathol*. 2000 Jun;156(6):1927-35). Bauhau et al. suggest a role of GDNF in neurofibromatosis type 1 (see Online Mendelian Inheritance in Man (OMIM), <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM>, OMIM*600837). Neurofibromatosis shows a diverse variety of clinical synopsis including Macrocephaly, Sphenoid dysplasia, Lisch nodules (iris hamartomas), Glaucoma, Hypertelorism, Renal artery stenosis, Hypertension, Scoliosis, Spina bifida, Pseudoarthrosis, Thinning of long bone cortex, Local bony overgrowth, Absent patellae, Neurofibromas, Plexiform neurofibroma, Cafe-au-lait spots, Axillary freckling, Inguinal freckling, Mental retardation, 30% learning disabilities, 10% mild mental retardation, Aqueductal stenosis, Hydrocephalus, neoplasia including Optic glioma, Meningioma, Hypothalamic tumor, Neurofibrosarcoma, Rhabdomyosarcoma, Duodenal carcinoid, Somatostatinoma, Parathyroid adenoma, Pheochromocytoma, Tumors at multiple other sites including CNS. As such, the SCS0008 polypeptides of the invention, agonists or antagonists thereof may also be useful in reducing or in the diagnosis or treatment of the individual or combined symptoms or clinical outcome of neurofibromatosis. In addition, SCS0008 nucleic acid molecules, polypeptides, and agonists and antagonists thereof and particularly SCS0008-SV1 (which is a splice variant from the lung, example 3) may be useful in diagnosing or treating fibrosis, preferably pulmonary fibrosis, liver fibrosis, neurofibromatosis type 1, Watson syndrome, or persistent acute respiratory distress syndrome (ARDS). In addition, SCS0008 nucleic acid molecules, polypeptides, and agonists and antagonists thereof and particularly SCS0008-SV1 may be useful in smooth muscle myogenesis.

Brandenberger et al. suggests that nephronectin acting through $\alpha_8\beta_1$ could induce a signalling pathway that results in enhanced expression of GDNF or, alternatively, it could synergize with GDNF by promoting secretion of this factor or its localization in the extracellular matrix. Since it is known that GDNF is implicated in the development of particular diseases, SCS0008, as a potential ligand or at least by enhancing GDNF's expression, might as well act as a contributing factor in the onset of those disorders. Matti Airaksinen and Mart Saarma have reviewed the role of the GDNF family in the brain (Matti S. Airaksinen and Mart Saarma, *Nature Reviews*, Volume 3, May 2002, pp.383-394). They point out that GDNF is implicated in motor

- 30 -

neuron disease or motor neuron injury (GDNF can support long-term motor neuron survival and axon regeneration after peripheral nerve injury), in sensory regeneration (GDNF produces potent analgesic effect in neuropathic models of pain and can cause sensory axons to regenerate back into the spinal chord; local GDNF expression induces Schwann cell migration to the lesion site, leading to remyelination of regenerating axons after injury, Blesch A et al. J Comp Neurol. 2003 Dec 15;467(3):403-17), sensory disorders including diabetic neuropathy, neuropathic pain, in ischaemia or stroke (GDNF administered before or just after anoxia can reduce ischaemic brain injury), preferably GDNF should be administered in the early phase of stroke, in epilepsy (GDNF might modulate seizure susceptibility), in neurodegenerative disorders such as Parkinson's disease or multiple sclerosis (GDNF can prevent the neuro toxin-induced death of dopamine neurons and can promote functional recovery; preferably GDNF is co-administered with heparin; see also Jeffrey H. Kordower, Ann Neurol 2003, 53 (suppl 3):S120-134, and Blesch et al.), and finally in addiction (infusion of GDNF into the VTA blocks certain biochemical adaptations to chronic cocaine and morphine, as well as the rewarding effects of cocaine). Jollivet et al. show that GDNF-releasing microspheres could be a promising strategy in the treatment of Parkinson's disease (Jollivet C et al. Biomaterials. 2004 Feb;25(5):933-42). In a similar way, SCS0008-releasing microspheres could be used. Nagano et al. suggest that GDNF could be used for sciatic nerve ligation (chronic constrictive injury) or spinal nerve ligation (Nagano N et al. Br J Pharmacol. 2003 Dec;140(7):1252-60). Tai et al. suggest that GDNF could be used to treat spinal cord injury (SCI) or spinal cord contusion (Tai MH et al. Exp Neurol. 2003 Oct;183(2):508-15). Ochiai H et al. state that local administration of GDNF in a neonatal preganglionic Erb's palsy model result in significant improvement in deficits (Ochiai H et al. Neurosurgery. 2003 Oct;53(4):973-7). Aszman et al. state that exogenous trophic support of motoneurons (e.g. GDNF or/and BDNF) might have a role in the treatment of all types of severe neonatal plexopathies (e.g. obstetric brachial plexus lesions; Aszmann OC et al. Plast Reconstr Surg. 2002 Sep 15;110(4):1066-72). Tolbert and Clark indicate that GDNF can delay the onset of hereditary Purkinje cell degeneration and gait ataxia (Tolbert DL, Clark BR. Exp Neurol. 2003 Sep;183(1):205-19). Schwann cells are predisposed to develop schwannoma (Bartolami S et al. J Neurobiol. 2003 Dec;57(3):270-90). As such, antagonists (antibodies) of GDNF or SCS0008 could be used to treat schwannoma or, by extension, other brain-related cancers. Perez-Garcia et al. state that changes of

[Ca²⁺], induced by GDNF, promote neuronal survival through a mechanism that involves a direct regulation of PI 3-kinase activation by calmodulin thus suggesting a central role for Ca²⁺ and calmodulin in the signalling cascade for neuronal survival mediated by neurotrophic factors (Perez-Garcia MJ et al. J Biol Chem.2003 Nov 20, 5 Epub; see also Wang J et al. Neurosignals.2003 Mar-Apr;12(2):78-88). As SCS0008 polypeptides contain three Calcium-binding EGF-like domains, the involvement of SCS0008 polypeptides in the above mechanism and thus in neuronal survival is strengthened. McBride et al. show that the viral-mediated gene transfer of GDNF into the striatum provides neuroanatomical and behavioural protection in a rodent model of 10 Huntington's disease (McBride JL et al. Exp Neurol. 2003 Jun;181(2):213-23). Alberch et al. suggest that neurotrophic factors (e.g. GDNF) may be able to provide selective protection for basal ganglia circuits, which are affected in striatonigral degenerative disorders, such as Huntington's disease or multisystem atrophy (Alberch J et al. Brain Res Bull. 2002 Apr;57(6):817-22). Laminin-10 and β 1 integrins have also been involved 15 in neuronal survival (Chen Zu-Lin et al. Molecular Biology of the Cell. 2003 Jul.Vol.14:2665-2676). Chen et al. state that laminin-10 protects against neuronal death. As such SCS0008 nucleic acid molecules, polypeptides, and agonists and antagonists thereof may be useful in diagnosing or treating degenerative disorders, striatonigral degenerative disorders, Huntington's disease, multisystem atrophy.

20 As such SCS0008 nucleic acid molecules, polypeptides, and agonists and antagonists thereof may be useful in the diagnosing or treating motor neuron diseases, motor neuron injuries, in favorising sensory regeneration, in sensory disorders including diabetic neuropathy, neuropathic pain, in ischaemia or stroke (preferably SCS0008 should be administered in the early phase of stroke), in epilepsy, in 25 neurodegenerative disorders such as Parkinson's disease or multiple sclerosis (preferably SCS0008 is co-administered with heparin and/or by the means of microspheres), in striatonigral degenerative disorders, Huntington's disease, multisystem atrophy, in sciatic nerve ligation (chronic constrictive injury), spinal nerve ligation, spinal cord injury (SCI) or spinal cord contusion, in neonatal preganglionic 30 Erb's palsy, in neonatal plexopathies such as obstetric brachial plexus lesions; in hereditary Purkinje cell degeneration and gait ataxia. SCS0008 could be administered alone or in combination with neurotrophic factors (e.g. neurturin, artemin, persephin, nerve growth factor, brain-derived neurotrophic factor, GDNF), heparin or other

therapeutic agent. Antagonists (e.g. antibodies) of SCS0008 could be used to treat schwannoma or other brain-related cancers.

GDNF seems to be involved in Hirschsprung disease (HSCR), which is a congenital disorder characterized by an absence of ganglion cells in the nerve plexus of the lower digestive tract (Garcia-Barcelo M et al. Clin Chem. 2003 Nov 18. Epub; see also Benailly HK et al. Clin Genet. 2003 Sep;64(3):204-9; see also OMIM*600837). As such SCS0008 nucleic acid molecules, polypeptides, and agonists and antagonists thereof may be useful in diagnosing or treating Hirschsprung disease (HSCR), central hypoventilation syndrome (ondine curse).

Signal from the epidermal growth factor (EGF) receptor family are thought to combine with integrin-dependent adhesion to laminins to contribute to disease progression and metastasis in cancer (Zamurs I et al. Biomed Chromatogr. 2003 Mar-Apr;17(2-3):201-11). Zamurs et al. showed that a colon cancer cell line secretes and adheres to laminin-10 through multiple integrin receptors, and that EGF stimulates spreading and migration of these cells on the same substrate. Laminin-10/11 was also shown to be present at the invasive edge of moderately differentiated colon cancers. Gu et al report that laminin 10-11 strongly promotes migration of human lung carcinoma cells by activating a $\beta 1$ integrin pathway (Gu J et al. Exp Dermatol. 2002 Oct;11(5):387-97). Laminin and integrins have also been implicated in pancreas function. Funahashi et al. showed that expression of some of the integrin subunits in pancreatic cancer cells was enhanced by GDNF. This enhancement and associated increase in adhesive and invasive ability by GDNF were inhibited by blocking the GDNF receptor or the integrin beta1 subunit (Funahashi H et al. Pancreas. 2003 Aug;27(2):190-6). Okada et al. report also that GDNF promotes pancreatic cancer cell invasion in vitro (Okada Y et al. Surgery. 2003 Aug;134(2):293-9). Japon et al. (2002) suggest that GDNF could be involved in pituitary tumors, more specifically adenomas, corticotropinomas, somatotropinomas, prolactinomas (see OMIM*600837). The RET and eventually GDNF genes have been associated with multiple endocrine neoplasia, type IIA and type IIB as well as medullary thyroid carcinoma (OMIM*164761 and OMIM*162300). Similar to nephronectin, Eyes absent 1 (EYA1) homozygosity results in an absence of ureteric bud outgrowth and a subsequent failure of metanephric induction (OMIM*601653). In addition, GDNF expression was not detected in Eya1 -/- metanephric mesenchyme (OMIM*601653). As such, nephronectin and Eya1 might interact. Eya1 is involved in branchiootorenal dysplasia and branchiootic syndrome. As

- 33 -

such SCS0008 antagonists (e.g. antibodies targeted to SCS0008) as well as SCS0008-SV1 and SCS0008-SV2 antagonists (which are preferably used for lung or kidney cancers respectively) may be useful in diagnosing or treating pancreatic cancer, pituitary tumors, more specifically adenomas, corticotropinomas, somatotropinomas, prolactinomas, and multiple endocrine neoplasia, type IIA and type IIB, medullary thyroid carcinoma, lung carcinoma, papillary thyroid carcinoma, colonic aganglionosis, colon cancer, MEN2-associated tumors, pheochromocytoma, amyotrophic lateral sclerosis, branchiootorenal dysplasia and branchiootic syndrome.

In view of the possible interaction or cooperation between laminin-10 and nephronectin (see above), disorders where laminin-10 is involved may be treated or diagnosed with the SCS0008 polypeptides of the present invention. Spessotto et al. showed that the preferential locomotion of leukemic cells was directed towards laminin isoform 8 and 10 (Spessotto P et al. *Matrix Biol.* 2003 Jun;22(4):351-61). The motility-promoting interaction with these laminins was mediated by $\beta 1$ integrins. In addition, Yu and Talts point out that both integrins and α -dystroglycan are expressed on cancer cells, and binding to laminin-10/11 may therefore be important for cell movement through endothelial and epithelial basement membranes and thus for metastasis (Yu H and Talts JF. *Biochem J.* 2003. 371:289-299). They also add that Ca^{2+} ion appears to be in some way involved in the interaction between α -dystroglycan and laminin 10/11 through $\alpha 5$ LG4-5. Without wishing to be bound to theory, nephronectin could have a role in this interaction via its calcium-binding EGF domains. Li et al. showed that laminin-10 can restore hair follicle development, thus enabling correction of cutaneous developmental defects (Li J et al. *Embo J.* 2003 May 15;22(10):2400-10). Laminins and $\beta 1$ integrins have also been implicated in hematopoiesis (Gu YC et al. *Blood.* 2003 Feb 1;101(3):877-85) as well as in angiogenesis in wound repair (Li J et al. *Microsc Res Tech.* 2003 Jan 1;60(1):107-14). Kikkawa et al. point out that the Lutheran blood group glycoprotein (Lu) is an Ig superfamily transmembrane receptor for lamin $\alpha 5$ and that Lu is thought to be involved in both normal and disease processes, including sickle cell disease and cancer (Kikkawa Y et al. *J Biol Chem.* 2002 Nov 22;277(47):44864-9). They showed that Lu binds specifically to laminin 10/11 in vivo and in vitro. In addition, Laminin-beta1 has been implicated in neonatal cutis laxa, marfanoid or marfan syndrome and arachnodactyly (OMIM*150240).

- 34 -

As such SCS0008 antagonists (e.g. antibodies targeted to SCS0008) as well as SCS0008-SV1 and SCS0008-SV2 antagonists (which are preferably used for lung or kidney cancers respectively) may be useful in diagnosing or treating leukaemia, in the reduction or inhibition of metastasis in cancer.

5 As such SCS0008 nucleic acid molecules, polypeptides, and agonists and antagonists thereof may be useful in diagnosing or treating cutaneous developmental defects, in hematopoiesis-related diseases, in sickle cell disease, in neonatal cutis laxa, marfanoid or marfan syndrome and arachnodactyly and in favourising angiogenesis in wound repair.

10 Yomogida et al. showed a dramatic amplification of germinal stem cells (GSC) in mouse testis following transfection of human GDNF cDNA into Sertoli cells (Yomogida K et al. Biol Reprod. 2003 Oct;69(4):1303-7). As such SCS0008 nucleic acid molecules, polypeptides, and agonists and antagonists thereof may be useful in the amplification of germinal stem cells.

15 Sariola and Meng state that GDNF is expressed by Sertoli cells and showed that GDNF regulates the cell fate decision of undifferentiated spermatogonia. In addition, they show that in mice overexpressing GDNF in testes, undifferentiated spermatogonia accumulate in the tubules, no sperm is produced, and the mice are infertile. After a year, they observe that GDNF overexpressing mice frequently (89%)
20 develop testicular tumors, and most of them are bilateral (56%), the tumors mimicking classical seminomas in men. As such SCS0008 antagonists (e.g. antibodies targeted to SCS0008) may be useful in the diagnosis or treatment of male infertility, seminomas. SCS0008 nucleic acid molecules, polypeptides, and agonists and antagonists thereof may be useful in the development, proliferation, and in the differentiation of
25 spermatogonia.

Thibault et al. (Thibault G et al., Am J Physiol Cell Physiol. 2001 Nov ;281(5) :C1457-67) observed that rat cardiac fibroblasts harbor $\alpha_8\beta_1$ integrin. In addition, stimulation of cardiac fibroblasts by angiotensin II (ANG II) or TGF β_1 resulted in an increase of protein and heightening by 50% of the receptor density $\alpha_8\beta_1$ integrin.
30 Low nephron number, inherited or acquired, has been linked to increased risk, not only of renal failure, but also of hypertension (Cullen-McEwen LA et al. Hypertension.2003 Feb;41(2):335-40). Cullen et al. show that GDNF heterozygous mice have elevated arterial pressure, glomerular hypertrophy and hyperfiltration. Wintour et al. point out that GDNF is a factor that predispose to the onset of adult hypertension and of

- 35 -

cardiovascular disease (Wintour EM et al. Placenta.2003 Apr;24 Suppl A:S65-71). As such SCS0008 nucleic acid molecules, polypeptides, and agonists and antagonists thereof may be useful in diagnosing or treating hypertension or cardiovascular diseases.

5 Chen et al. show that increased levels of GDNF exclusively in the target rescue 30% of oculomotor neurons that would normally die during developmental cell death, a rate of rescue similar to that with systemic GDNF application (Chen J et al. Mol Cell Neurosci.2003 Sep;24(1):41-56). Wordinger et al. state that a potential for autocrine and/or paracrine GDNF signalling exists within the lamina cribrosa, a tissue involved in
10 glaucoma pathogenesis (Wordinger RJ et al. Mol Vis.2003 Jun 16;9:249-56). In addition, Rothermel et Layer suggest that GDNF, at least in vitro, affects rod receptors (Rothermel A, Layer PG. Invest Ophtalmol Vis Sci.2003 May;44(5):2221-8). Depending on the developmental stage, GDNF regulates their proliferation, differentiation, and survival. Karlsson et al. show that GDNF labelling is mainly found in chicken embryonic
15 day 4-5 retina but weak labelling could also be found over scattered retinal cells at later cells (Karlsson M et al. Mech Dev.2002 Jun;114(1-2):161-5). They also show that c-ret labelling is found over ganglion cells, amacrine and horizontal cells; GRF alpha 1 over amacrine and horizontal cells; and GFR alpha 2 over ganglion cells, amacrine cells and photoreceptors. Ljubimov et al. showed that basement membrane (BM) components,
20 especially fibronectin and laminin-10, in postcataract surgery (PCS) corneas were altered to a great extent (40-60%). In addition, they found that tenascin-C and fibrillin-1 were mostly found in diseased but not in PCS corneas (Ljubimov AV et al. Cornea. 2002 Jan;21(1):74-80). Eya1 has also been implicated in cataract (see OMIM*601653). As such SCS0008 nucleic acid molecules, polypeptides, and agonists and antagonists
25 thereof may be useful in the development, proliferation, differentiation and survival of oculomotor neurons, photoreceptors and particularly rod receptors (rosetted spheroids), ganglion cells, amacrine and horizontal cells or useful in diagnosing or treating corneal edematous diseases, Fuch's dystrophy corneas, cataract or eye injuries. SCS0008 antagonists (e.g. antibodies targeted to SCS0008) may be useful in
30 diagnosing or treating glaucoma.

Gladson et al. established that $\alpha_8\beta_1$ integrin is expressed on neonatal rat astrocytes. They demonstrate also that unstimulated primary neonatal rat astrocytes attach to vitronectin, utilizing $\alpha_8\beta_1$ and $\alpha_5\beta_5$ integrins, and that this attachment is regulated by the type 1 Plasminogen Activator Inhibitor (PAI-1). As such SCS0008

- 36 -

nucleic acid molecules, polypeptides, and agonists and antagonists thereof may be useful in diagnosing or treating germinal matrix hemorrhage and infarction.

Littlewood et al. suggest that $\alpha_6\beta_1$ integrin regulates hair cell differentiation and stereocilia maturation and that mutations in integrin genes may lead to inner-ear diseases (Littlewood Evans A et al., Nat Genet.2000 Apr;24(4):424-8). Furthermore, GDNF is suggested to play a role in the postnatal inner ear (GDNF expression found in the cochlea) in addition to its role during development (Stankovic KM, Corfas G. Hear Res. 2003 Nov;185(1-2):97-108). Kawamoto et al. suggest that adenovirus-mediated overexpression of GDNF and TGF-beta1 can be used in combination to protect cochlear hair and hearing from ototoxic trauma (Kawamoto K et al. Mol Ther. 2003 Apr;7(4):484-92). They also point out that GDNF overexpression in the inner ear can protect hair cells against degeneration induced by ototoxicity. As such SCS0008 nucleic acid molecules, polypeptides, and agonists and antagonists thereof alone or in combination with TGF-beta1 may be useful in diagnosing or treating inner-ear diseases, inner-ear injuries (e.g. ototoxic trauma), scala tympani fibrosis, or for the development, differentiation, and survival of inner hair cells (cochlear hair cells).

Moursi et al. showed that direct osteoblast interactions with the extracellular matrix are mediated by a selected group of integrin receptors that includes $\alpha_5\beta_1$, $\alpha_3\beta_1$, $\alpha_8\beta_1$. Furthermore, Morimura et al. suggest a role of POEM in the early stage of osteoblastic cell differentiation and that POEM may play important roles in osteoblastic function by sending survival signals via $\alpha_6\beta_1$ integrin and mediating cell-cell interaction. POEM was expressed in the endocrine organs (parathyroid gland, thyroid gland, hypophysis, and pineal organ), which are closely related to growth, bone metabolism, and calcium and phosphorus homeostasis. They add that the data suggest a relationship between POEM and calcium metabolism. In addition, the SCS0008 contain three EGF-calcium binding motifs. As such SCS0008 nucleic acid molecules, polypeptides, and agonists and antagonists thereof may be useful in diagnosing or treating osteoporosis-pseudoglioma syndrome, osteopetrosis, endosteal hyperostosis, osteosclerosis, high bone mass disorder, dolichostenomelia, micrognathia, congenital kyphoscoliosis, retrognathiascoliosis, thoracic lordosis, spondylolisthesis, lumbosacral dural ectasia, in growth defects, bone metabolism, and calcium and phosphorus homeostasis as well as osteoblastic cell differentiation.

Several domains have been identified in the SCS0008 protein, which are indicators of SCS0008's associated diseases. Using SMART (Simple Modular

- 37 -

Architecture Research Tool, which allows the identification and annotation of genetically mobile domains and the analysis of domain architectures, <http://smart.embl-heidelberg.de/>) and its OMIM (Online Mendelian Inheritance in Man, which is a database cataloging human genes and genetic disorders, <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM>) curated human diseases associated with missense mutations within domains, it can be retrieved human diseases associated with domains found in SCS0008 (see example 4). Search for OMIM human diseases associated with the EGF-like or the calcium-binding EGF domains were performed. Some of the diseases have already been mentioned above and strengthens the likelihood that the SCS0008 polypeptides of the invention may be useful for the diagnosis or treatment of these diseases.

As such SCS0008 nucleic acid molecules, polypeptides, and agonists and antagonists (antagonists of SCS0008 as antibodies being useful for all types of cancer mentioned below) thereof may be useful in diagnosing or treating Severe classic or mild variable or neonatal form or atypical Marfan syndrome, aortic aneurysm, marfanoid skeletal syndrome, Weill-Marchesani syndrome, MASP2 deficiency, Factor IX deficiency, megaloblastic anaemia, pseudoachondroplasia, epiphyseal dysplasia, cutis laxa, lissencephaly syndrome (preferably Norman-Roberts type), osteoporosis-pseudoglioma syndrome, osteopetrosis, endosteal hyperostosis, osteosclerosis, high bone mass disorder, dolichostenomelia, micrognathia, congenital kyphoscoliosis, recurrent venous or mesenteric or cerebral venous or arterial thrombosis, superficial thrombophlebitis, Warfarin-induced skin necrosis, retrognathiascoliosis, thoracic lordosis, sphondylolisthesis, lumbosacral dural ectasia, Beals syndrome, coronary heart disease, non-small cell lung cancer, ulcerative colitis or proctitis, endometriosis, chronic otitis, benign prostatic hyperplasia, advanced brain tumor, bladder cancer, duodenal ulcers, metastatic breast cancer, prostate cancer, psoriasis, human uterine leiomyoma, gastric ulcer, discoid lupus erythematosus, lichen planus, colorectal cancer, acute myocardial infarction, advanced malignant glioma, endometrial cancer, malignant gliomas, head and neck or lung cancer, breast cancer, laryngeal and hypopharyngeal carcinoma, ovarian tumours, herpetic corneal ulcers and venous ulcers.

The invention will now be described with reference to the specific embodiments by

- 38 -

means of the following Examples, which should not be construed as in any way limiting the present invention. The content of the description comprises all modifications and substitutions which can be practiced by a person skilled in the art in light of the above teachings and, therefore, without extending beyond the meaning and purpose of the

5 claims.

TABLE 1

Amino Acid	Synonymous Groups	More Preferred Synonymous Groups
Ser	Gly, Ala, Ser, Thr, Pro	Thr, Ser
Arg	Asn, Lys, Gln, Arg, His	Arg, Lys, His
Leu	Phe, Ile, Val, Leu, Met	Ile, Val, Leu, Met
Pro	Gly, Ala, Ser, Thr, Pro	Pro
Thr	Gly, Ala, Ser, Thr, Pro	Thr, Ser
Ala	Gly, Thr, Pro, Ala, Ser	Gly, Ala
Val	Met, Phe, Ile, Leu, Val	Met, Ile, Val, Leu
Gly	Ala, Thr, Pro, Ser, Gly	Gly, Ala
Ile	Phe, Ile, Val, Leu, Met	Ile, Val, Leu, Met
Phe	Trp, Phe, Tyr	Tyr, Phe
Tyr	Trp, Phe, Tyr	Phe, Tyr
Cys	Ser, Thr, Cys	Cys
His	Asn, Lys, Gln, Arg, His	Arg, Lys, His
Gln	Glu, Asn, Asp, Gln	Asn, Gln
Asn	Glu, Asn, Asp, Gln	Asn, Gln
Lys	Asn, Lys, Gln, Arg, His	Arg, Lys, His
Asp	Glu, Asn, Asp, Gln	Asp, Glu
Glu	Glu, Asn, Asp, Gln	Asp, Glu
Met	Phe, Ile, Val, Leu, Met	Ile, Val, Leu, Met
Trp	Trp, Phe, Tyr	Trp

TABLE 2

Amino Acid	Synonymous Groups
Ser	D-Ser, Thr, D-Thr, allo-Thr, Met, D-Met, Met(O), D-Met(O), L-Cys, D-Cys
Arg	D-Arg, Lys, D-Lys, homo-Arg, D-homo-Arg, Met, Ile, D-.Met, D-Ile, Orn, D-Orn
Leu	D-Leu, Val, D-Val, AdaA, AdaG, Leu, D-Leu, Met, D-Met
Pro	D-Pro, L-l-thioazolidine-4-carboxylic acid, D-or L-1-oxazolidine-4-carboxylic acid
Thr	D-Thr, Ser, D-Ser, allo-Thr, Met, D-Met, Met(O), D-Met(O), Val, D-Val
Ala	D-Ala, Gly, Aib, B-Ala, Acp, L-Cys, D-Cys
Val	D-Val, Leu, D-Leu, Ile, D-Ile, Met, D-Met, AdaA, AdaG
Gly	Ala, D-Ala, Pro, D-Pro, Aib, .beta.-Ala, Acp
Ile	D-Ile, Val, D-Val, AdaA, AdaG, Leu, D-Leu, Met, D-Met
Phe	D-Phe, Tyr, D-Thr, L-Dopa, His, D-His, Trp, D-Trp, Trans-3,4, or 5-phenylproline, AdaA, AdaG, cis-3,4, or 5-phenylproline, Bpa, D-Bpa
Tyr	D-Tyr, Phe, D-Phe, L-Dopa, His, D-His
Cys	D-Cys, S--Me--Cys, Met, D-Met, Thr, D-Thr
Gln	D-Gln, Asn, D-Asn, Glu, D-Glu, Asp, D-Asp
Asn	D-Asn, Asp, D-Asp, Glu, D-Glu, Gln, D-Gln
Lys	D-Lys, Arg, D-Arg, homo-Arg, D-homo-Arg, Met, D-Met, Ile, D-Ile, Orn, D-Orn
Asp	D-Asp, D-Asn, Asn, Glu, D-Glu, Gln, D-Gln
Glu	D-Glu, D-Asp, Asp, Asn, D-Asn, Gln, D-Gln
Met	D-Met, S--Me--Cys, Ile, D-Ile, Leu, D-Leu, Val, D-Val

EXAMPLES

Example 1 :

Sequences of EGF protein domains from the ASTRAL database (Brenner SE *et al.*
5 "The ASTRAL compendium for protein structure and sequence analysis" Nucleic Acids
Res. 2000 Jan 1; 28 (1): 254-6) were used to search for homologous protein
sequences in genes predicted from human genome sequence (Celera database). The
protein sequences were obtained from the gene predictions and translations thereof as
generated by one of three programs: the Genescan (Burge C, Karlin S., "Prediction of
10 complete gene structures in human genomic DNA, J Mol Biol. 1997 Apr 25;268(1):78 -
94) Grail (Xu Y, Uberbacher EC., « Automated gene identification in large-scale
genomic sequences", J Comput Biol. 1997 Fall;4(3):325-38) and Fgenesh (Proprietary
Celera software).

The sequence profiles of the EGF domains were generated using PIMAll (Profile
15 Induced Multiple Alignment; Boston University software, version II, Das S and Smith TF
2000), an algorithm that aligns homologous sequences and generates a sequence
profile. The homology was detected using PIMAll that generates global-local
alignments between a query profile and a hit sequence. In this case the algorithm was
used with the profile of the EGF functional domain as a query. PIMAll compares the
20 query profile to the database of gene predictions translated into protein sequence and
can therefore identify a match to a DNA sequence that contains that domain. Further
comparison by BLAST (Basic Local Alignment Search Tool; NCBI version 2) of the
sequence with known EFG containing proteins identified the closets homolog (Gish W,
States DJ. "Identification of protein coding regions by database similarity search.", Nat
25 Genet. 1993 Mar;3(3):266-72; Pearson WR, Miller W., "Dynamic programming
algorithms for biological sequence comparison.", Methods Enzymol. 1992;210:575-601;
Altschul SF *et al.*, "Basic local alignment search tool", J Mol Biol. 1990 Oct
5;215(3):403-10). PIMAll parameters used for the detection were the PIMA prior amino
acids probability matrix and a Z-cutoff score of 10. BLAST parameters used were:
30 Comparison matrix = BLOSUM62; word length = 3; .E value cutoff = 10; Gap opening
and extension = default; No filter.

- 42 -

Once the functional domain was identified in the sequence, the genes were re-predicted with the genewise algorithm using the sequence of the closets homolog (Birney E *et al.*, " PairWise and SearchWise: finding the optimal alignment in a simultaneous comparison of a protein profile against all DNA translation frames.",
5 Nucleic Acids Res. 1996 Jul 15;24(14):2730-9)

The profiles for homologous EGF domains were generated automatically using the PSI-BLAST (Altschul *et al* 1997) scripts written in PERL (Practical Extraction and Report Language) and PIMAIL.

A total of 55 predicted genes out of the 464 matching the original query generated on
10 the basis of EGF domain profiles were selected since they were judged as potentially novel.

The novelty of the protein sequences was finally assessed by searching protein databases (SwissProt/Trembl, Human IPI and Derwent GENESEQ) using BLAST and a specific annotation has been attributed on the basis of amino acid sequence
15 homology.

Example 2:

One sequence isolated by the methodology set out in Example 1 is that referred to herein as the SCS0008 polypeptide sequence.

This sequence, even though it presents approximately 50% of homology with human
20 Fibrillin 2 (as well as with other EGF domain containing proteins) in the central region, should be rather considered a human shorter splicing variant of a mouse gene coding for a different extracellular matrix protein recently cloned by two distinct groups and called nephronectin or POEM (see Morimura *et al.*, J Biol Chem 2001 Nov 9;276(45):42172-81 ; Brandenberger *et al.*, J Cell Biol 2001 Jul 23;154(2):447-58
25 (Comment in: J Cell Biol. 2001 Jul 23;154(2):257-9). This protein has been characterized so far as an adhesion molecule acting as a ligand for $\alpha(8)\beta(1)$ integrin and involved in the development and function of various organs, in particular kidney.

Larger Human variants have also been discovered (see SCEP-39 (INCYTE WO0248337); SEQID34 (CURAGEN WO0110902); SEQID82 (CURAGEN
30 WO0110902); NOV5 (CURAGEN WO0257452); PRO334 (GENENTECH WO0104311); PRO26 (GENENTECH WO0104311); and PRO334 (GENENTECH WO9914328)), but none of them, as for the mouse sequences, corresponds to the

- 43 -

submitted sequence which, in particular, lacks a central and the C-terminal region common to all relevant human / mouse sequence identified (see Figure 1). The first sequence corresponds to a region located between the last EGF domain and a proline-rich region.

5 **Example 3 - Identification and cloning of splice variants of SCS0008**

3.1 Introduction:

SCS0008 is a 1700 cDNA prediction encoding an EGF domain-containing protein of 406 amino acids with homology to fibrillin. Using nested PCR primers based on the full coding sequence of the SCS0008 prediction, two splice variants of SCS0008 from
10 human lung (SV1) and human kidney (SV2) respectively. The alignments of SV1 and 2 with SCS0008 are shown in figure 2.

SV1 contains a 55 amino acid extension of exon 8 and a mutation at Q159H. SV2 also contains a 55 amino acid extension of exon 8, a mutation at Q159H but is missing exon 9. The cloned version of SV2 also contains a PCR induced mutation F3L, which was
15 later corrected during subcloning. Both splice variants were subcloned with a C-terminal 6HIS tag using the Gateway cloning methodology into expression vectors pEAK12d and pDEST12.2.

3.2 Cloning of SCS0008-SV1

3.2.1 Preparation of human lung cDNA

20 First strand cDNA was prepared from normal lung total RNA (Clontech) using Superscript II RNase H⁻ Reverse Transcriptase (Invitrogen) according to the manufacturer's protocol. 1 µl Oligo (dT)₁₅ primer (500 µg/ml, Promega), 2 µg human lung total RNA, 1 µl of 10 mM dNTP mix (10 mM each of dATP, dGTP, dCTP and dTTP at neutral pH) and sterile distilled water to a final
25 volume of 12 µl were combined in a 1.5 ml eppendorf tube, heated to 65 °C for 5 min and then chilled on ice. The contents were collected by brief centrifugation and 4 µl of 5X First-Strand Buffer, 2 µl of 0.1 M DTT, and 1 µl of RNaseOUT Recombinant Ribonuclease Inhibitor (40 units/µl, Invitrogen) were added. The contents of the tube were mixed gently and incubated at 42 °C for 2
30 min, then 1 µl (200 units) of SuperScript II enzyme was added and mixed gently by pipetting. The mixture was incubated at 42 °C for 50 min and then inactivated by heating at 70 °C for 15 min. To remove RNA complementary to

- 44 -

the cDNA, 1 μ l (2 units) of *E. coli* RNase H (Invitrogen) was added and the reaction mixture incubated at 37 °C for 20 min. The final 21 μ l reaction mix was diluted by adding 179 μ l sterile water to give a total volume of 200 μ l.

3.2.2 Gene specific cloning primers for PCR

5 A pair of PCR primers having a length of between 18 and 25 bases was designed for amplifying the complete coding sequence of the virtual cDNA using Primer Designer Software (Scientific & Educational Software, PO Box 72045, Durham, NC 27722-2045, USA). PCR primers were optimized to have a T_m close to 55 ± 10 °C and a GC content of 40-60%. Primers were selected
10 which had high selectivity for the target sequence (SCS0008) with little or no non-specific priming.

3.2.3 PCR amplification of SCS0008 from human lung cDNA

Gene-specific cloning primers (SCS0008-CP1 and SCS0008 -CP2, Figure 3, Figure 4 and Table 3) were designed to amplify a cDNA fragment of 1323 bp
15 covering the entire coding sequence of the SCS0008 prediction. The gene-specific cloning primers SCS0008-CP1 and SCS0008 -CP2 were used with human lung cDNA as the template. The PCR was performed in a final volume of 50 μ l containing 1X AmpliTaq™ buffer, 200 μ M dNTPs, SCS0008-CP1, 50 pmoles of SCS0008-CP2, 2.5 units of AmpliTaq™ (Perkin Elmer) and 100 ng of
20 lung cDNA using an MJ Research DNA Engine, programmed as follows: 94 °C, 2 min; 40 cycles of 94 °C, 1 min, 50 °C, 1 min, and 72 °C, 1 min; followed by 1 cycle at 72 °C for 7 min and a holding cycle at 4 °C. PCR products were purified directly using the Wizard PCR Preps DNA Purification System (Promega). PCR products were eluted in 50 μ l of sterile water and 10 μ l of the amplification
25 products were then used as template in a 2nd PCR reaction using the same conditions as described above except that the primers used were nested primers SCS0008-CP1 nest and SCS0008 -CP2 nest. Amplification products were visualized on 0.8 % agarose gels in 1 X TAE buffer (Invitrogen) and a single PCR product migrating near the predicted molecular mass was purified
30 from the gel using the Wizard PCR Preps DNA Purification System (Promega). The PCR product was eluted in 50 μ l of sterile water and subcloned directly.

3.2.4 Subcloning of PCR Products

PCR products were subcloned into the topoisomerase I modified cloning vector (pCR4-TOPO) using the TOPO cloning kit purchased from the Invitrogen Corporation using the conditions specified by the manufacturer. Briefly, 4 µl of gel purified PCR product was incubated for 15 min at room temperature with 1 µl of TOPO vector and 1 µl salt solution. The reaction mixture was then transformed into *E. coli* strain TOP10 (Invitrogen) as follows: a 50 µl aliquot of One Shot TOP10 cells was thawed on ice and 2 µl of TOPO reaction was added. The mixture was incubated for 15 min on ice and then heat shocked by incubation at 42 °C for exactly 30 s. Samples were returned to ice and 250 µl of warm SOC media (room temperature) was added. Samples were incubated with shaking (220 rpm) for 1 h at 37 °C. The transformation mixture was then plated on L-broth (LB) plates containing ampicillin (100 µg/ml) and incubated overnight at 37 °C. Ampicillin resistant colonies containing inserts were identified by colony PCR.

3.2.5 Colony PCR

Colonies were inoculated into 50 µl sterile water using a sterile toothpick. A 10 µl aliquot of the inoculum was then subjected to PCR in a total reaction volume of 20 µl as described above, except the primers used were T7 and T3. The cycling conditions were as follows: 94 °C, 2 min; 30 cycles of 94 °C, 30 sec, 48 °C, 30 sec and 72 °C, 1 min. Samples were then maintained at 4 °C (holding cycle) before further analysis.

PCR reaction products were analyzed on 1 % agarose gel in 1 X TAE buffer. Colonies which gave the expected PCR product size (+ 105 bp due to the multiple cloning site or MCS) were grown up overnight at 37 °C in 5 ml L-Broth (LB) containing ampicillin (100 µg /ml), with shaking (220 rpm).

3.2.6 Plasmid DNA preparation and sequencing

Miniprep plasmid DNA was prepared from the 5 ml culture using a Qiaprep Turbo 9600 robotic system (Qiagen) or Wizard Plus SV Minipreps kit (Promega cat. no. 1460) according to the manufacturer's instructions. Plasmid DNA was eluted in 100 µl of sterile water. The DNA concentration was measured using an Eppendorf BO photometer. Plasmid DNA (200-500 ng) was subjected to DNA sequencing with the T7 primer and SP6 primer using the BigDyeTerminator

- 46 -

system (Applied Biosystems cat. no. 4390246) according to the manufacturer's instructions. The primer sequences are shown in Table 3. Sequencing reactions were purified using Dye-Ex columns (Qiagen) or Montage SEQ 96 cleanup plates (Millipore cat. no. LSKS09624) then analyzed on an Applied Biosystems 3700 sequencer.

Sequence analysis of the cloned cDNA insert revealed a match to the predicted SCS0008 sequence over exons 1-8 except for a single point mutation at nucleotide 681 which lead to an amino acid substitution Q159H. There was also a 165 bp insertion at the end of exon 8 leading to an in frame insertion of 55 amino acids detected in all clones sequenced. The sequence of the cloned cDNA fragment is shown in Figure 4.

3.3 Construction of mammalian cell expression vectors for SCS00008-SV1

A pCR4-TOPO clone containing the coding sequence (ORF) of SCS0008-SV1 identified by DNA sequencing (pCR4-TOPO-SCS0008SV1, plasmid ID.14630) (figure 5) was then used to subclone the insert into the mammalian cell expression vectors pEAK12d (figure 6) and pDEST12.2 (figure 8) using the Gateway™ cloning methodology (Invitrogen).

3.3.1 Generation of Gateway compatible SCS0008-SV1 ORF fused to an in frame 6HIS tag sequence.

The first stage of the Gateway cloning process involves a two step PCR reaction which generates the ORF of SCS0008-SV1 flanked at the 5' end by an attB1 recombination site and Kozak sequence, and flanked at the 3' end by a sequence encoding an in frame 6 histidine (6HIS) tag, a stop codon and the attB2 recombination site (Gateway compatible cDNA). The first PCR reaction (in a final volume of 50 µl) contains: 1 µl (40 ng) of pCR4-TOPO-SCS0008-SV1 (plasmid ID 14630), 1.5 µl dNTPs (10 mM), 10 µl of 10X Pfx polymerase buffer, 1 µl MgSO4 (50 mM), 0.5 µl each of gene specific primer (100 µM) (SCS0008-SV1-EX1 and SCS0008-SV1-EX2), 2.5 µl 10X Enhancer™ solution (Invitrogen) and 0.5 µl Platinum Pfx DNA polymerase (Invitrogen). The PCR reaction was performed using an initial denaturing step of 95 °C for 2 min, followed by 12 cycles of 94 °C for 15 s; 55 °C for 30 s and 68 °C for 2 min; and a holding cycle of 4 °C. The amplification products were visualized on 0.8 % agarose gel in 1 X

- 47 -

TAE buffer (Invitrogen) and a product migrating at the predicted molecular mass was purified from the gel using the Wizard PCR Preps DNA Purification System (Promega) and recovered in 50 µl sterile water according to the manufacturer's instructions.

5 The second PCR reaction (in a final volume of 50 µl) contained 10 µl purified PCR 1 product, 1.5 µl dNTPs (10 mM), 5 µl of 10X Pfx polymerase buffer, 1 µl MgSO₄ (50 mM), 0.5 µl of each Gateway conversion primer (100 µM) (GCP forward and GCP reverse) and 0.5 µl of Platinum Pfx DNA polymerase. The conditions for the 2nd PCR reaction were: 95 °C for 1 min; 4 cycles of 94 °C, 15
10 sec; 50 °C, 30 sec and 68 °C for 2 min; 25 cycles of 94 °C, 15 sec; 55 °C, 30 sec and 68 °C, 2 min; followed by a holding cycle of 4 °C. PCR products were gel purified using the Wizard PCR prep DNA purification system (Promega) according to the manufacturer's instructions.

15 3.3.2 Subcloning of Gateway compatible SCS0008-SV1 ORF into Gateway entry vector pDONR221 and expression vectors pEAK12d and pDEST12.2

The second stage of the Gateway cloning process involves subcloning of the Gateway modified PCR product into the Gateway entry vector pDONR221 (Invitrogen, figure 7) as follows: 5 µl of purified product from PCR2 were incubated with 1.5 µl pDONR221 vector (0.1 µg/µl), 2 µl BP buffer and 1.5 µl of
20 BP clonase enzyme mix (Invitrogen) in a final volume of 10 µl at RT for 1 h. The reaction was stopped by addition of proteinase K 1 µl (2 µg/µl) and incubated at 37 °C for a further 10 min. An aliquot of this reaction (1 µl) was used to transform *E. coli* DH10B cells by electroporation as follows: a 25 µl aliquot of DH10B electrocompetent cells (Invitrogen) was thawed on ice and 1 µl of the
25 BP reaction mix was added. The mixture was transferred to a chilled 0.1 cm electroporation cuvette and the cells electroporated using a BioRad Gene-Pulser™ according to the manufacturer's recommended protocol. SOC media (0.5 ml), which had been pre-warmed to room temperature, was added immediately after electroporation. The mixture was transferred to a 15 ml snap-
30 cap tube and incubated, with shaking (220 rpm) for 1 h at 37 °C. Aliquots of the transformation mixture (10 µl and 50 µl) were then plated on L-broth (LB) plates containing kanamycin (40 µg/ml) and incubated overnight at 37 °C.

- 48 -

Plasmid mini-prep DNA was prepared from 5 ml cultures from 6 of the resultant colonies using a Qiaprep Turbo 9600 robotic system (Qiagen). Plasmid DNA (150-200 ng) was subjected to DNA sequencing with 21M13 and M13Rev primers using the BigDyeTerminator system (Applied Biosystems cat. no. 4390246) according to the manufacturer's instructions. The primer sequences are shown in Table 3. Sequencing reactions were purified using Dye-Ex columns (Qiagen) or Montage SEQ 96 cleanup plates (Millipore cat. no. LSKS09624) then analyzed on an Applied Biosystems 3700 sequencer.

Plasmid eluate (2 µl or approx. 150 ng) from one of the clones which contained the correct sequence (pENTR-SCS0008-SV1-6HIS, plasmid ID 14877, figure 9) was then used in a recombination reaction containing 1.5 µl of either pEAK12d vector or pDEST12.2 vector (figures 4 & 5) (0.1 µg / µl), 2 µl LR buffer and 1.5 µl of LR clonase (Invitrogen) in a final volume of 10 µl. The mixture was incubated at RT for 1 h, stopped by addition of proteinase K (1 µl at 2 µg/µl) and incubated at 37 °C for a further 10 min. An aliquot of this reaction (1 µl) was used to transform *E. coli* DH10B cells by electroporation as follows: a 25 µl aliquot of DH10B electrocompetent cells (Invitrogen) was thawed on ice and 1 µl of the LR reaction mix was added. The mixture was transferred to a chilled 0.1 cm electroporation cuvette and the cells electroporated using a BioRad Gene-Pulser™ according to the manufacturer's recommended protocol. SOC media (0.5 ml), which had been pre-warmed to room temperature, was added immediately after electroporation. The mixture was transferred to a 15 ml snap-cap tube and incubated, with shaking (220 rpm) for 1 h at 37 °C. Aliquots of the transformation mixture (10 µl and 50 µl) were then plated on L-broth (LB) plates containing ampicillin (100 µg/ml) and incubated overnight at 37 °C.

Plasmid mini-prep DNA was prepared from 5 ml cultures from 6 of the resultant colonies subcloned in each vector using a Qiaprep Turbo 9600 robotic system (Qiagen). Plasmid DNA (200-500 ng) in the pEAK12d vector was subjected to DNA sequencing with pEAK12F, pEAK12R, SCS0008SV1-SP1 and SCS0008SV1-SP2 primers as described above. Plasmid DNA (200-500 ng) in the pDEST12.2 vector was subjected to DNA sequencing with 21M13 and M13Rev, SCS0008SV1-SP1 and SCS0008SV1-SP2 primers as described above. Primer sequences are shown in Table 3

- 49 -

5 CsCl gradient purified maxi-prep DNA was prepared from a 500 ml culture of one of each of the sequence verified clones (pEAK12d-SCS0008-SV1-6HIS, plasmid ID number 14883, figure 10, and pDEST12.2-SCS0008-SV1-6HIS, plasmid ID 14887, figure 11) using the method described by Sambrook J. et al., 1989 (in Molecular Cloning, a Laboratory Manual, 2nd edition, Cold Spring Harbor Laboratory Press), Plasmid DNA was resuspended at a concentration of 1 µg/µl in sterile water (or 10 mM Tris-HCl pH 8.5) and stored at -20 °C.

10

Table 3 **SCS0008SV1 cloning and sequencing primers**

Primer	Sequence (5' -3')
SCS0008-CP1	GCT GCC CAA CAT GGA TTT TC
SCS0008-CP2	CCA TTC TCT GCA GGC TGA TT
SCS0008-CP1nest	ATT TTC TCC TGG CGC TGG TG
SCS0008-CP2nest	CACCTTGTCTCCAGTTGCTACA
SCS0008SV1-EX1	AA GCA GGC TTC <u>GCC ACC</u> ATG GAT TTT CTC CTG GCG CTG GTG
SCS0008SV1-EX2	<i>GTG ATG GTG ATG GTG</i> CAA GTC AAG TCC ACC TGC AT
GCP Forward	G GGG ACA AGT TTG TAC AAA AAA GCA GGC TTC <u>GCC ACC</u>
GCP Reverse	GGG GAC CAC TTT GTA CAA GAA AGC TGG GTT TCA <i>ATG GTG ATG GTG ATG GTG</i>
SCS0008SV1-SP1	ATA CGG TGC CAG TGC CCA TC
SCS0008SV1-SP2	CAA CAG GCC TAC TTC TAA GC
pEAK12F	GCC AGC TTG GCA CTT GAT GT
pEAK12R	GAT GGA GGT GGA CGT GTC AG
21M13	TGT AAA ACG ACG GCC AGT
M13REV	CAG GAA ACA GCT ATG ACC
T7 primer	TAA TAC GAC TCA CTA TAG GG
SP6 primer	ATT TAG GTG ACA CTA TAG

5

Underlined sequence = Kozak sequence

Bold = Stop codon*Italic* sequence = His tag10 3.4 Cloning of SCS0008-SV23.4.1 Preparation of human kidney cDNA

First strand cDNA was prepared from normal kidney total RNA (Clontech) using Superscript II RNase H⁻ Reverse Transcriptase (Invitrogen) according to the manufacturer's protocol. 1 µl Oligo (dT)₁₅ primer (500 µg/ml, Promega), 2 µg

15 human lung total RNA, 1 µl of 10 mM dNTP mix (10 mM each of dATP, dGTP, dCTP and dTTP at neutral pH) and sterile distilled water to a final volume of 12 µl were combined in a 1.5 ml eppendorf tube, heated to 65 °C for 5 min and then chilled on ice. The contents were collected by brief centrifugation and 4 µl of 5X First-Strand Buffer, 2 µl of 0.1 M DTT, and 1 µl of

20 RNaseOUT Recombinant Ribonuclease Inhibitor (40 units/µl, Invitrogen) were added. The contents of the tube were mixed gently and incubated at 42 °C for 2

- 51 -

min, then 1 μ l (200 units) of SuperScript II enzyme was added and mixed gently by pipetting. The mixture was incubated at 42 °C for 50 min and then inactivated by heating at 70 °C for 15 min. To remove RNA complementary to the cDNA, 1 μ l (2 units) of *E. coli* RNase H (Invitrogen) was added and the reaction mixture incubated at 37 °C for 20 min. The final 21 μ l reaction mix was diluted by adding 179 μ l sterile water to give a total volume of 200 μ l.

3.4.2 Gene specific cloning primers for PCR

A pair of PCR primers having a length of between 18 and 25 bases was designed for amplifying the complete coding sequence of the virtual cDNA using Primer Designer Software (Scientific & Educational Software, PO Box 72045, Durham, NC 27722-2045, USA). PCR primers were optimized to have a T_m close to 55 ± 10 °C and a GC content of 40-60%. Primers were selected which had high selectivity for the target sequence (SCS0008) with little or no none specific priming.

3.4.3 PCR amplification of SCS0008 from human kidney cDNA

Gene-specific cloning primers (SCS0008-CP1 and SCS0008 -CP2, Figure 12, Figure 13 and Table 4) were designed to amplify a cDNA fragment of 1323 covering the entire coding sequence of the SCS0008 prediction. The gene-specific cloning primers SCS0008-CP1 and SCS0008-CP2 were used with human kidney cDNA as the template. The PCR was performed in a final volume of 50 μ l containing 1X AmpliTaq™ buffer, 200 μ M dNTPs, SCS0008-CP1, 50 pmoles of SCS0008-CP2, 2.5 units of AmpliTaq™ (Perkin Elmer) and 100 ng of lung cDNA using an MJ Research DNA Engine, programmed as follows: 94 °C, 2 min; 40 cycles of 94 °C, 1 min, 50 °C, 1 min, and 72 °C, 1 min; followed by 1 cycle at 72 °C for 7 min and a holding cycle at 4 °C. PCR products were purified directly using the Wizard PCR Preps DNA Purification System (Promega). PCR products were eluted in 50 μ l of sterile water and 10 μ l of the amplification products were then used as template in a 2nd PCR reaction using the same conditions as described above except that the primers used were nested primers SCS0008-CP1 nest and SCS0008 -CP2 nest. Amplification products were visualized on 0.8 % agarose gels in 1 X TAE buffer (Invitrogen) and a single PCR product migrating near the predicted molecular mass was purified

- 52 -

from the gel using the Wizard PCR Preps DNA Purification System (Promega). The PCR product was eluted in 50 µl of sterile water and subcloned directly.

3.4.4 Subcloning of PCR Products

5 PCR products were subcloned into the topoisomerase I modified cloning vector (pCR4-TOPO) using the TOPO cloning kit purchased from the Invitrogen Corporation using the conditions specified by the manufacturer. Briefly, 4 µl of
10 gel purified PCR product was incubated for 15 min at room temperature with 1 µl of TOPO vector and 1 µl salt solution. The reaction mixture was then transformed into *E. coli* strain TOP10 (Invitrogen) as follows: a 50 µl aliquot of One Shot TOP10 cells was thawed on ice and 2 µl of TOPO reaction was added. The mixture was incubated for 15 min on ice and then heat shocked by incubation at 42 °C for exactly 30 s. Samples were returned to ice and 250 µl of warm SOC media (room temperature) was added. Samples were incubated with shaking (220 rpm) for 1 h at 37 °C. The transformation mixture was then
15 plated on L-broth (LB) plates containing ampicillin (100 µg/ml) and incubated overnight at 37 °C. Ampicillin resistant colonies containing inserts were identified by colony PCR.

3.4.5 Colony PCR

20 Colonies were inoculated into 50 µl sterile water using a sterile toothpick. A 10 µl aliquot of the inoculum was then subjected to PCR in a total reaction volume of 20 µl as described above, except the primers used were T7 and T3. The cycling conditions were as follows: 94 °C, 2 min; 30 cycles of 94 °C, 30 sec, 48 °C, 30 sec and 72 °C, 1 min. Samples were then maintained at 4 °C (holding cycle) before further analysis.

25 PCR reaction products were analyzed on 1 % agarose gel in 1 X TAE buffer. Colonies which gave the expected PCR product size (+ 105 bp due to the multiple cloning site or MCS) were grown up overnight at 37 °C in 5 ml L-Broth (LB) containing ampicillin (100 µg /ml), with shaking (220 rpm).

3.4.6 Plasmid DNA preparation and sequencing

30 Miniprep plasmid DNA was prepared from the 5 ml culture using a Qiaprep Turbo 9600 robotic system (Qiagen) or Wizard Plus SV Minipreps kit (Promega cat. no. 1460) according to the manufacturer's instructions. Plasmid DNA was eluted in 100 µl of sterile water. The DNA concentration was measured using an

- 53 -

5 Eppendorf BO photometer. Plasmid DNA (200-500 ng) was subjected to DNA sequencing with the T7 primer and SP6 primer using the BigDyeTerminator system (Applied Biosystems cat. no. 4390246) according to the manufacturer's instructions. The primer sequences are shown in Table 4. Sequencing reactions were purified using Dye-Ex columns (Qiagen) or Montage SEQ 96 cleanup plates (Millipore cat. no. LSKS09624) then analyzed on an Applied Biosystems 3700 sequencer.

10 Sequence analysis of the cloned cDNA insert revealed a match to the predicted SCS0008 sequence over exons 1-8 except for single point mutations at nucleotide 213 (T to G) which lead to an amino acid substitution of F3L, and at nucleotide 681 (G to C) which lead to an amino acid substitution Q159H. There was also a 165 bp insertion at the end of exon 8 leading to an in frame insertion of 55 amino acids detected in all clones sequenced. In 3/9 clones sequenced an 87 bp deletion (corresponding to exon 9) was also detected. These 3 clones 15 likely represent a splice variant of SCS0008 (SCS0008-SV2) which lacks exon 9. The sequence of the cloned SCS0008-SV2 cDNA fragment is shown in Figure 13.

3.5 Construction of mammalian cell expression vectors for SCS00008 -SV2

20 A pCR4-TOPO clone containing the coding sequence (ORF) of SCS0008-SV2 identified by DNA sequencing (pCR4-TOPO-SCS0008SV2, plasmid ID.14631) (figure 14) was then used to subclone the insert into the mammalian cell expression vectors pEAK12d (figure 15) and pDEST12.2 (figure 16) using the Gateway™ cloning methodology (Invitrogen).

25 3.5.1 Generation of Gateway compatible SCS0008-SV2 ORF fused to an in frame 6HIS tag sequence.

The first stage of the Gateway cloning process involves a two step PCR reaction which generates the ORF of SCS0008-SV2 flanked at the 5' end by an attB1 recombination site and Kozak sequence, and flanked at the 3' end by a sequence encoding an in frame 6 histidine (6HIS) tag, a stop codon and the attB2 recombination site (Gateway compatible cDNA). The first PCR reaction 30 (in a final volume of 50 µl) contains: 1 µl (40 ng) of pCR4-TOPO-SCS0008-SV2 (plasmid ID 14631), 1.5 µl dNTPs (10 mM), 10 µl of 10X Pfx polymerase buffer, 1 µl MgSO4 (50 mM), 0.5 µl each of gene specific primer (100 µM) (SCS0008-SV2-EX1 and SCS0008-SV2-EX2), 2.5 µl 10X Enhancer™ solution (Invitrogen)

- 54 -

and 0.5 µl Platinum Pfx DNA polymerase (Invitrogen). The PCR reaction was performed using an initial denaturing step of 95 °C for 2 min, followed by 12 cycles of 94 °C for 15 s; 55 °C for 30 s and 68 °C for 2 min; and a holding cycle of 4 °C. The amplification products were visualized on 0.8 % agarose gel in 1 X TAE buffer (Invitrogen) and a product migrating at the predicted molecular mass was purified from the gel using the Wizard PCR Preps DNA Purification System (Promega) and recovered in 50 µl sterile water according to the manufacturer's instructions.

The second PCR reaction (in a final volume of 50 µl) contained 10 µl purified PCR 1 product, 1.5 µl dNTPs (10 mM), 5 µl of 10X Pfx polymerase buffer, 1 µl MgSO₄ (50 mM), 0.5 µl of each Gateway conversion primer (100 µM) (GCP forward and GCP reverse) and 0.5 µl of Platinum Pfx DNA polymerase. The conditions for the 2nd PCR reaction were: 95 °C for 1 min; 4 cycles of 94 °C, 15 sec; 50 °C, 30 sec and 68 °C for 2 min; 25 cycles of 94 °C, 15 sec; 55 °C, 30 sec and 68 °C, 2 min; followed by a holding cycle of 4 °C. PCR products were gel purified using the Wizard PCR prep DNA purification system (Promega) according to the manufacturer's instructions.

3.5.2 Subcloning of Gateway compatible SCS0008-SV2 ORF into Gateway entry vector pDONR221 and expression vectors pEAK12d and pDEST12.2

The second stage of the Gateway cloning process involves subcloning of the Gateway modified PCR product into the Gateway entry vector pDONR221 (Invitrogen, figure 17) as follows: 5 µl of purified product from PCR2 were incubated with 1.5 µl pDONR221 vector (0.1 µg/µl), 2 µl BP buffer and 1.5 µl of BP clonase enzyme mix (Invitrogen) in a final volume of 10 µl at RT for 1 h. The reaction was stopped by addition of proteinase K 1 µl (2 µg/µl) and incubated at 37 °C for a further 10 min. An aliquot of this reaction (1 µl) was used to transform *E. coli* DH10B cells by electroporation as follows: a 25 µl aliquot of DH10B electrocompetent cells (Invitrogen) was thawed on ice and 1 µl of the BP reaction mix was added. The mixture was transferred to a chilled 0.1 cm electroporation cuvette and the cells electroporated using a BioRad Gene-Pulser™ according to the manufacturer's recommended protocol. SOC media (0.5 ml), which had been pre-warmed to room temperature, was added immediately after electroporation. The mixture was transferred to a 15 ml snap-

- 55 -

cap tube and incubated, with shaking (220 rpm) for 1 h at 37 °C. Aliquots of the transformation mixture (10 µl and 50 µl) were then plated on L-broth (LB) plates containing kanamycin (40 µg/ml) and incubated overnight at 37 °C.

Plasmid mini-prep DNA was prepared from 5 ml cultures from 6 of the resultant colonies using a Qiaprep Turbo 9600 robotic system (Qiagen). Plasmid DNA (150-200 ng) was subjected to DNA sequencing with 21M13 and M13Rev primers using the BigDyeTerminator system (Applied Biosystems cat. no. 4390246) according to the manufacturer's instructions. The primer sequences are shown in Table 4. Sequencing reactions were purified using Dye-Ex columns (Qiagen) or Montage SEQ 96 cleanup plates (Millipore cat. no. LSKS09624) then analyzed on an Applied Biosystems 3700 sequencer.

Plasmid eluate (2 µl or approx. 150 ng) from one of the clones which contained the correct sequence (pENTR-SCS0008-SV2-6HIS, plasmid ID 14878, figure 18) was then used in a recombination reaction containing 1.5 µl of either pEAK12d vector or pDEST12.2 vector (figures 4 & 5) (0.1 µg / µl), 2 µl LR buffer and 1.5 µl of LR clonase (Invitrogen) in a final volume of 10 µl. The mixture was incubated at RT for 1 h, stopped by addition of proteinase K (1 µl at 2 µg/µl) and incubated at 37 °C for a further 10 min. An aliquot of this reaction (1 µl) was used to transform *E. coli* DH10B cells by electroporation as follows: a 25 µl aliquot of DH10B electrocompetent cells (Invitrogen) was thawed on ice and 1 µl of the LR reaction mix was added. The mixture was transferred to a chilled 0.1 cm electroporation cuvette and the cells electroporated using a BioRad Gene-Pulser™ according to the manufacturer's recommended protocol. SOC media (0.5 ml), which had been pre-warmed to room temperature, was added immediately after electroporation. The mixture was transferred to a 15 ml snap-cap tube and incubated, with shaking (220 rpm) for 1 h at 37 °C. Aliquots of the transformation mixture (10 µl and 50 µl) were then plated on L-broth (LB) plates containing ampicillin (100 µg/ml) and incubated overnight at 37 °C.

Plasmid mini-prep DNA was prepared from 5 ml cultures from 6 of the resultant colonies subcloned in each vector using a Qiaprep Turbo 9600 robotic system (Qiagen). Plasmid DNA (200-500 ng) in the pEAK12d vector was subjected to DNA sequencing with pEAK12F, pEAK12R, SCS0008SV1-SP1 and SCS0008SV1-SP2 primers as described above. Plasmid DNA (200-500 ng) in

- 56 -

the pDEST12.2 vector was subjected to DNA sequencing with 21M13 and M13Rev, SCS0008SV1-SP1 and SCS0008SV1-SP2 primers as described above. Primer sequences are shown in Table 4.

5 CsCl gradient purified maxi-prep DNA was prepared from a 500 ml culture of one of each of the sequence verified clones (pEAK12d-SCS0008-SV2-6HIS, plasmid ID number 14884, figure 19, and pDEST12.2-SCS0008-SV2-6HIS, plasmid ID 14888, figure 20) using the method described by Sambrook J. et al., 1989 (in Molecular Cloning, a Laboratory Manual, 2nd edition, Cold Spring Harbor Laboratory Press), Plasmid DNA was resuspended at a concentration of
10 1 µg/µl in sterile water (or 10 mM Tris-HCl pH 8.5) and stored at -20 °C.

Table 4 SCS0008-SV2 cloning and sequencing primers

Primer	Sequence (5' -3')
SCS0008-CP1	GCT GCC CAA CAT GGA TTT TC
SCS0008-CP2	CCA TTC TCT GCA GGC TGA TT
SCS0008-CP1nest	ATT TTC TCC TGG CGC TGG TG
SCS0008-CP2nest	CACCTTGCTCTCCAGTTGCTACA
SCS0008SV2-EX1	AA GCA GGC TTC <u>GCC ACC</u> ATG GAT TTT CTC CTG GCG CTG GTG
SCS0008SV2-EX2	<i>GTG ATG GTG ATG GTG</i> CAA GTC AAG TCC ACC TGC AT
GCP Forward	G GGG ACA AGT TTG TAC AAA AAA GCA GGC TTC <u>GCC ACC</u>
GCP Reverse	GGG GAC CAC TTT GTA CAA GAA AGC TGG GTT TCA <i>ATG GTG ATG GTG ATG GTG</i>
SCS0008SV1-SP1	ATA CGG TGC CAG TGC CCA TC
SCS0008SV1-SP2	CAA CAG GCC TAC TTC TAA GC
pEAK12F	GCC AGC TTG GCA CTT GAT GT
pEAK12R	GAT GGA GGT GGA CGT GTC AG
21M13	TGT AAA ACG ACG GCC AGT
M13REV	CAG GAA ACA GCT ATG ACC
T7 primer	TAA TAC GAC TCA CTA TAG GG
SP6 primer	ATT TAG GTG ACA CTA TAG

- 5 Underlined sequence = Kozak sequence
 Bold = Stop codon
 Italic sequence = His tag

Example 4: Identification and description of SCS0008 domains

10 4.1 Identification

- A bioinformatic tool called SMART (<http://smart.embl-heidelberg.de/>) was used to identify the putative domains of SCS0008 and of the splice variants SCS0008-SV1 and SCS0008-SV2. Results are shown in Figure 21, which also display the mouse ortholog nephronectin, other known human splice variants of SCS0008, and one
 15 protein displaying similar domain organization, i.e. EGFL6. In addition, Prosite was also run on the sequences (<http://us.expasy.org/prosite/>).

- 58 -

Table 5 Domains identified by SMART within the query sequence SCS0008 of 406 residues

Confidently predicted domains, repeats, motifs and features:

name	begin	end	E-value
<u>signal peptide</u>	1	18	-
<u>EGF</u>	59	87	1.01e-01
<u>EGF_CA</u>	89	128	2.13e-09
<u>EGF</u>	132	168	2.11e+01
<u>EGF_CA</u>	169	213	6.50e-05
<u>EGF_CA</u>	214	254	2.85e-10
<u>low complexity</u>	255	301	-

5

Table 6 Domains identified by SMART within the query sequence SCS0008-SV1 of 459 residues

Confidently predicted domains, repeats, motifs and features:

name	begin	end	E-value
<u>signal peptide</u>	1	9	-
<u>EGF</u>	57	85	1.01e-01
<u>EGF_CA</u>	87	126	2.13e-09
<u>EGF</u>	130	166	2.43e+01
<u>EGF_CA</u>	167	211	6.50e-05
<u>EGF_CA</u>	212	252	2.85e-10
<u>low complexity</u>	291	305	-
<u>low complexity</u>	307	354	-

10

- 59 -

Table 7 Domains identified by SMART within the query sequence SCS0008-SV2 of 430 residues

Confidently predicted domains, repeats, motifs and features:

name	begin	end	E-value
<u>signal peptide</u>	1	8	-
<u>EGF</u>	57	85	1.01e-01
<u>EGF_CA</u>	87	126	2.13e-09
<u>EGF</u>	130	166	2.43e+01
<u>EGF_CA</u>	167	211	6.50e-05
<u>EGF_CA</u>	212	252	2.85e-10
<u>low complexity</u>	291	305	-
<u>low complexity</u>	307	354	-

5

Partial ScanProsite results:

- SCS0008:
10 >PDOC00016 PS00016 **RGD** Cell attachment sequence [pattern] [Warning: pattern with a high probability of occurrence].
327 - 329 RGD
- SCS0008-SV1:
15 >PDOC00016 PS00016 **RGD** Cell attachment sequence [pattern] [Warning: pattern with a high probability of occurrence].
380 - 382 RGD
- SCS0008-SV2:
20 >PDOC00016 PS00016 **RGD** Cell attachment sequence [pattern] [Warning: pattern with a high probability of occurrence].
380 - 382 RGD

4.2 Description of the domains

- **EGF. Epidermal growth factor-like domain. Interpro annotation:**

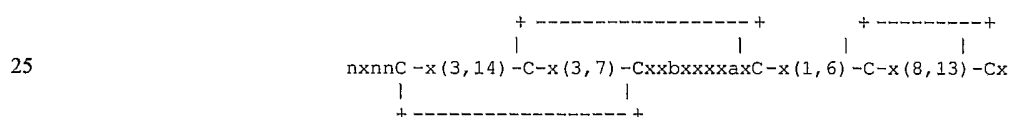
25 A sequence of about thirty to forty amino-acid residues long found in the sequence of epidermal growth factor (EGF) has been shown MEDLINE:,
MEDLINE:88196363, MEDLINE:84117505, MEDLINE:91145344,
MEDLINE:85063790, MEDLINE: to be present, in a more or less conserved form, in a large number of other, mostly animal proteins. The list of proteins currently known to contain one or more copies of an EGF-like pattern is large
30 and varied. The functional significance of EGF domains in what appear to be unrelated proteins is not yet clear. However, a common feature is that these repeats are found in the extracellular domain of membrane-bound proteins or

- 60 -

in proteins known to be secreted (exception: prostaglandin G/H synthase). The EGF domain includes six cysteine residues which have been shown (in EGF) to be involved in disulphide bonds. The main structure is a two-stranded β -sheet followed by a loop to a C-terminal short two-stranded sheet. Subdomains between the conserved cysteines vary in length.

• **GF_CA. Calcium-binding EGF-like domain. Interpro annotation:**

A sequence of about forty amino-acid residues long found in the sequence of epidermal growth factor (EGF) has been shown to be present in a large number of membrane-bound and extracellular, mostly animal proteins (see IPR000561). Many of these proteins require calcium for their biological function and a calcium-binding site has been found to be located at the N-terminus of some EGF-like domains. Calcium-binding may be crucial for numerous protein-protein interactions. For human coagulation factor IX it has been shown that the calcium-ligands form a pentagonal bipyramid. The first, third and fourth conserved negatively charged or polar residues are side chain ligands. Latter is possibly hydroxylated (see IPR000152). A conserved aromatic residue as well as the second conserved negative residue are thought to be involved in stabilizing the calcium-binding site. Like in non-calcium binding EGF-like domains there are six conserved cysteines and the structure of both types is very similar as calcium-binding induces only strictly local structural changes.



'n': negatively charged or polar residue [DEQN]

'b': possibly beta-hydroxylated residue [DN]

'a': aromatic amino acid

'C': cysteine, involved in disulfide bond

'x': any amino acid.

• **MAM. Domain in meprin, A5, receptor protein tyrosine phosphatase mu (and others). SMART and Interpro annotations:**

Likely to have an adhesive function. Mutations in the meprin MAM domain affect noncovalent associations within meprin oligomers. In receptor tyrosine

phosphatase mu-like molecules the MAM domain is important for homophilic cell-cell interactions. A 170 amino acid domain, the so-called MAM domain, has been recognised in the extracellular region of functionally diverse proteins. These proteins have a modular, receptor-like architecture comprising a signal peptide, an N-terminal extracellular domain, a single transmembrane domain and an intracellular domain. Such proteins include meprin (a cell surface glycoprotein); A5 antigen (a developmentally-regulated cell surface protein); and receptor-like tyrosine protein phosphatase. The MAM domain is thought to have an adhesive function. It contains 4 conserved cysteine residues, which probably form disulphide bridges.

- PS00016; RGD

The sequence Arg-Gly-Asp, found in fibronectin, is crucial for its interaction with its cell surface receptor, an integrin. What has been called the 'RGD' tripeptide is also found in the sequences of a number of other proteins, where it has been shown to play a role in cell adhesion. These proteins are: some forms of collagens, fibrinogen, vitronectin, von Willebrand factor (VWF), snake disintegrins, and slime mold discoidins. The 'RGD' tripeptide is also found in other proteins where it may also, but not always, serve the same purpose.

4.2 Conclusion

Results show that SCS0008, SCS0008-SV1 and SCS0008-SV2 display a particular domain organization, being characterized in that they all lack the MAM domain in comparison with the other aligned sequences (see figure 21). As described above, the MAM domain has an adhesive function. In addition, Morimura et al. showed that a mutant POEM-Fc molecule without the MAM domain detached from the cell surface and was detected mainly in the culture medium rather than in cell extracts when expressed in COS-7 cells. They conclude by suggesting that the MAM domain plays a significant role for cell surface localization. Interestingly, they point out that the MAM domain family, including meprin, A5 protein, neuropilin-1, neuropilin-2, and receptor protein-tyrosine phosphatases, mediates cell adhesion activities via homo- or heterophilic MAM domain interactions, which supports the hypothesis that the MAM domain is involved in the cell surface binding via protein-protein

- 62 -

interaction, and that these molecules could serve as candidate receptor molecules for POEM. SCS0008 and the splice variants SCS0008-SV1 and SCS0008-SV2 could therefore be unable to form homo- or heterophilic MAM domain interactions. As such, the proteins of the invention, possibly being
5 unable to bind to the cell surface, display unique properties. Another possibility is that the sequences of the invention contain a different MAM domain, enabling them to bind to another MAM receptor molecule than the one binding to nephronectin.

In addition, Morimura et al. identified an RGD integrin binding motif in POEM
10 and they showed, as well as Brandenberger et al., that POEM is a novel specific ligand molecule for $\alpha_8\beta_1$ integrin. The RGD integrin binding motif is present in SCS0008, SCS0008-SV1 and SCS0008-SV2 (see 4.2) suggesting that $\alpha_8\beta_1$ integrin could be a specific ligand for these proteins as well. However, Brandenberger et al. showed that the binding site of $\alpha_8\beta_1$ integrin
15 appears to be contained in amino acids 382-561 of POEM. As such, SCS0008-SV2, lacking exon 9, might not bind to $\alpha_8\beta_1$ integrin.

Example 5: Identification of diseases related to the identified domains

Using SMART and its OMIM curated human diseases associated with missense
20 mutations within domains, human diseases associated with the domains found in SCS0008 or SCS0008-SVS1 or SCS0008-SV2 can be retrieved (tables 8 and 9). Complementary searches were performed with OMIM on the annotated domains.

[SMART Annotation]

Diseases found associated with domains

25 1) SwissProt sequences and OMIM curated human diseases associated with missense mutations within the EGF-like domain.

Table 8

Protein	Disease
Coagulation factor IX precursor (EC 3.4.21.22) (Christmas factor). (SRS)(SMART)	(OMIM:306900): Hemophilia B ; Warfarin sensitivity
Coagulation factor VII precursor (EC 3.4.21.21) (Serum prothrombin conversion accelerator) (Eptacog alfa). (SRS)(SMART)	(OMIM:227500): Factor VII deficiency ; {Myocardial infarction, decreased susceptibility to}
Fibrillin 2 precursor. (SRS)(SMART)	(OMIM:121050): Contractural arachnodactyly, congenital
Delta-like protein 3 precursor (Drosophila Delta homolog 3). (SRS)(SMART)	(OMIM:602768): Spondylocostal dysostosis, autosomal recessive, 1 (OMIM:277300):
Vitamin-K-dependent protein C precursor (EC 3.4.21.69) (Autoprothrombin IIA) (Anticoagulant protein C) (Blood coagulation factor XIV). (SRS)(SMART)	(OMIM:176860): Thrombophilia due to protein C deficiency ; Purpura fulminans, neonatal
Vitamin K-dependent protein S precursor. (SRS)(SMART)	(OMIM:176880): Protein S deficiency
Low-density lipoprotein receptor precursor (LDL receptor). (SRS)(SMART)	(OMIM:143890): Hypercholesterolemia, familial
EGF-containing fibulin-like extracellular matrix protein 1 precursor (Fibulin -3) (FIBL-3) (Fibrillin-like protein) (Extracellular protein S1-5). (SRS)(SMART)	(OMIM:601548): Doyme honeycomb degeneration of retina (OMIM:126600): (OMIM:126600): Doyme honeycomb retinal dystrophy
Thyroid peroxidase precursor (EC 1.11.1.8) (TPO). (SRS)(SMART)	(OMIM:274500): Thyroid iodine peroxidase deficiency ; Goiter, congenital ; Hypothyroidism, congenital
Fibrillin 1 precursor. (SRS)(SMART)	(OMIM:134797): Marfan syndrome (OMIM:154700): Shprintzen-Goldberg syndrome (OMIM:182212): Ectopia lentis, familial ; MASS syndrome (OMIM:604308):

Search for other OMIM human diseases maybe associated with the EGF-like domain:

- Severe classic or classic or mild variable or neonatal form or atypical Marfan syndrome, aortic aneurysm, marfanoid skeletal syndrome, Weill-Marchesani syndrome,
- 5 MASP2 deficiency, Factor IX deficiency, megaloblastic anaemia,

- 64 -

pseudoachondroplasia, epiphyseal dysplasia, cutis laxa, lissencephaly syndrome (preferably Norman-Roberts type), osteoporosis-pseudoglioma syndrome, osteopetrosis, endosteal hyperostosis, osteosclerosis, high bone mass disorder, dolichostenomelia, micrognathia, congenital kyphoscoliosis, recurrent venous or
5 mesenteric or cerebral venous or arterial thrombosis, superficial thrombophlebitis, Warfarin-induced skin necrosis, retrognathiascoliosis, thoracic lordosis, sphondylolisthesis, lumbosacral dural ectasia, Beals syndrome, coronary heart disease, non-small cell lung cancer, ulcerative colitis or proctitis, endometriosis, chronic
10 otitis, benign prostatic hyperplasia, advanced brain tumor, bladder cancer, duodenal ulcers, metastatic breast cancer, prostate cancer, psoriasis, human uterine leiomyoma, gastric ulcer, discoid lupus erythematosus, lichen planus, colorectal cancer, acute myocardial infarction, advanced malignant glioma, endometrial cancer, malignant gliomas, head and neck or lung cancer, breast cancer, laryngeal and hypopharyngeal carcinoma, ovarian tumours, herpetic corneal ulcers and venous ulcers.

15

2) SwissProt sequences and OMIM curated human diseases associated with missense mutations within the EGF_CA domain.

Table 9

Protein	Disease
Coagulation factor IX precursor (EC 3.4.21.22) (Christmas factor). (<u>SRS</u>)(<u>SMART</u>)	(<u>OMIM:306900</u>): Hemophilia B ; Warfarin sensitivity
E-selectin precursor (Endothelial leukocyte adhesion molecule 1) (ELAM -1) (Leukocyte-endothelial cell adhesion molecule 2) (LECAM2) (CD62E). (<u>SRS</u>)(<u>SMART</u>)	(<u>OMIM:131210</u>): {Atherosclerosis, susceptibility to}
Coagulation factor VII precursor (EC 3.4.21.21) (Serum prothrombin conversion accelerator) (Eptacog alfa). (<u>SRS</u>)(<u>SMART</u>)	(<u>OMIM:227500</u>): Factor VII deficiency ; {Myocardial infarction, decreased susceptibility to}
Fibrillin 2 precursor. (<u>SRS</u>)(<u>SMART</u>)	(<u>OMIM:121050</u>): Contractural arachnodactyly, congenital
Delta-like protein 3 precursor (Drosophila Delta homolog 3). (<u>SRS</u>)(<u>SMART</u>)	(<u>OMIM:602768</u>): Spondylocostal dysostosis, autosomal recessive, 1 (<u>OMIM:277300</u>):
Vitamin-K-dependent protein C precursor (EC 3.4.21.69) (Autoprothrombin IIA) (Anticoagulant protein C) (Blood coagulation factor XIV). (<u>SRS</u>)(<u>SMART</u>)	(<u>OMIM:176860</u>): Thrombophilia due to protein C deficiency ; Purpura fulminans, neonatal
Vitamin K-dependent protein S precursor. (<u>SRS</u>)(<u>SMART</u>)	(<u>OMIM:176880</u>): Protein S deficiency
Low-density lipoprotein receptor precursor (LDL receptor). (<u>SRS</u>)(<u>SMART</u>)	(<u>OMIM:143890</u>): Hypercholesterolemia, familial
EGF-containing fibulin-like extracellular matrix protein 1 precursor (Fibulin -3) (FIBL-3) (Fibrillin-like protein) (Extracellular protein S1-5). (<u>SRS</u>)(<u>SMART</u>)	(<u>OMIM:601548</u>): Doyne honeycomb degeneration of retina (<u>OMIM:126600</u>): (<u>OMIM:126600</u>): Doyne honeycomb retinal dystrophy
Thyroid peroxidase precursor (EC 1.11.1.8) (TPO). (<u>SRS</u>)(<u>SMART</u>)	(<u>OMIM:274500</u>): Thyroid iodine peroxidase deficiency ; Goiter, congenital ; Hypothyroidism, congenital
Fibrillin 1 precursor. (<u>SRS</u>)(<u>SMART</u>)	(<u>OMIM:134797</u>): Marfan syndrome (<u>OMIM:154700</u>): Shprintzen-Goldberg syndrome (<u>OMIM:182212</u>): Ectopia lentis, familial ; MASS syndrome (<u>OMIM:604308</u>):

- 66 -

Search for other OMIM human diseases maybe associated with the EGF_CA domain:

Severe classic or classic or mild variable or neonatal form or atypical Marfan syndrome, aortic aneurysm, marfanoid skeletal syndrome, Weill-Marchesani syndrome, MASP2 deficiency, Factor IX deficiency, megaloblastic anaemia, pseudoachondroplasia, epiphyseal dysplasia, cutis laxa, lissencephaly syndrome (preferably Norman-Roberts type), osteoporosis-pseudoglioma syndrome, osteopetrosis, endosteal hyperostosis, osteosclerosis, high bone mass disorder, dolichostenomelia, micrognathia, congenital kyphoscoliosis, recurrent venous or mesenteric or cerebral venous or arterial thrombosis, superficial thrombophlebitis, Warfarin-induced skin necrosis, retrognathiascoliosis, thoracic lordosis, sphondylolisthesis, lumbosacral dural ectasia and coron ary heart disease.

Example 6: Neurological Assays Suitable for Exploration of the Biological Relevance of proteins Function.

A number of neurological assays have been developed by the Applicant and are of use in the investigation of the biological relevance of protein function. Examples of neurological assays that have been developed by the Applicant include four types of assays. These are discussed below.

A. Oligodendrocytes-based assays

Oligodendrocytes are responsible for myelin formation in the CNS. In multiple sclerosis they are the first cells attacked and their loss leads to major behavioral impairment. In addition to curbing inflammation, enhancing the incomplete remyelination of lesions that occurs in MS has been proposed as a therapeutic strategy for MS. Like neurons, mature oligodendrocytes do not divide but the new oligodendrocytes can arise from progenitors. There are very few of these progenitor cells in adult brain and even in embryos the number of progenitor cells is inadequate for HTS.

Oli-neu is a murine cell line obtained by an immortalization of an oligodendrocyte precursor by the t-neu oncogene. They are well studied and accepted as a representative cell line to study young oligodendrocyte biology.

These cells can be used in two types of assays.

One, to identify factors stimulating oligodendrocytes proliferation, and the other to find factors promoting their differentiation. Both events are key in the perspective of helping renewal and repairing demyelinating diseases.

Another possible cell line is the human cell line, MO3-13. MO3-13 results from the fusion of rabdo-myosarcoma cells with adult human oligodendrocytes. However these cells have a reduced ability to differentiate into oligodendrocytes and their proliferating rate is not sufficient to allow a proliferation assay. Nevertheless, they express certain features of oligodendrocytes and their morphology is well adapted to nuclear translocation studies. Therefore this cell line can be used in assays based on nuclear translocation of three transcription factors, respectively NF- κ B, Stat-1 and Stat-2. The Jak/Stats transcription pathway is a complex pathway activated by many factors such as IFN α , β , γ , cytokines (e.g. IL-2, IL-6; IL-5) or hormones (e.g. GH, TPO, EPO). The specificity of the response depends on the combination of activated Stats. For example, it is noticeable that IFN- β activates Stat1, 2 and 3 nuclear translocations meanwhile IFN- γ only activates Stat1. In the same way, many cytokines and growth factors induced NF- κ B translocation. In these assays the goal is to get a picture of activated pathways for a given protein.

B. Astrocytes-based assays

The biology of astrocytes is very complex, but two general states are recognized. In one state called quiescent, astrocytes regulate the metabolic and excitatory level of neurons by pumping glutamate and providing energetic substratum to neurons and oligodendrocytes. In the activated state, astrocytes produce chemokines and cytokines as well as nitric oxide. The first state could be considered as normal healthy while the second state occurs during inflammation, stroke or neurodegenerative diseases. When this activated state persists it should be regarded as a pathological state.

Many factors and many pathways are known to modulate astrocyte activation. In order to identify new factors modulating astrocyte activation U373 cells, a human cell line of astroglioma origin, can be used. NF- κ B, c-Jun as well as Stats are signaling molecules known to play pivotal roles in astrocyte activation.

A series of screens based on the nuclear translocation of NF- κ B, c-Jun and Stat1, 2 and 3 can be carried out. Prototypical activators of these pathways are IL-1 β , IFN-beta or IFN-gamma. The goal is to identify proteins that could be used as therapeutics in the treatment of CNS diseases.

- 68 -

C. Neurons-based assays

Neurons are very complex and diverse cells but they have all in common two things. First they are post-mitotic cells, secondly they are innervating other cells. Their survival is linked to the presence of trophic factors often produced by the innervated target cells. In many neurodegenerative diseases the lost of target innervation leads to cell body atrophy and apoptotic cell death. Therefore identification of trophic factors supplementing target deficiency is very important in treatment of neurodegenerative diseases.

In this perspective a survival assay using NS1 cells, a sub-clone of rat PC12 cells, can be performed. These cells have been used for years and a lot of neurobiology knowledge has been first acquired on these cells before being confirmed on primary neurons including the pathways involved in neuron survival and differentiation (MEK, PI3K, CREB). In contrast the N2A cells, a mouse neuroblastoma, are not responding to classical neurotrophic factors but Jun-kinase inhibitors prevent apoptosis induced by serum deprivation. Therefore assays on these two cell lines will help to find different types of "surviving promoting" proteins.

It is important to note that in the previous assays we will identify factors that promote both proliferation and differentiation. In order to identify factors specifically promoting neuronal differentiation, a NS1 differentiation assay based on neurite outgrowth can be used. Promoting axonal or dendritic sprouting in neurodegenerative diseases could be advantageous for two reasons. It will first help the degenerating neurons to re-grow and re-establish a contact with the target cells. Secondly, it will potentiate the so-called collateral sprouting from healthy fibers, a compensatory phenomenon that delays terminal phases of neurodegenerative such as Parkinson or AD.

D. Endothelial cells-based assays

The blood brain barrier (BBB) between brain and vessels is responsible of differences between cortical spinal fluid and serum compositions. The BBB results from a tight contact between endothelial cells and astrocytes. It maintains an immunotolerant status by preventing leukocytes penetration in brain, and allows the development of two parallels endocrine systems using the same intracellular signaling pathways. However, in many diseases or traumas, the BBB integrity is altered and leukocytes as well as serum proteins enter the brain inducing neuroinflammation. There is no easy in vitro model of BBB, but cultures of primary endothelial cells such as human embryonic umbilical endothelial cells (HUVEC) could mimic some aspect of BBB biology. For

- 69 -

example, BBB leakiness could be induced by proteins stimulating intracellular calcium release. In the perspective of identifying proteins that modulate BBB integrity, a calcium mobilization assay with or without thrombin can be performed on HUVEC.

5 **Example 7: Fibroblast Assays Suitable for Exploration of the Biological Relevance of the Protein Function**

A number of fibroblasts assays have been developed by the Applicant and are of use in the investigation of the biological relevance of protein function. Examples of fibroblasts assays that have been developed by the Applicant include eight types of assays.

10 These are discussed below.

Activation and pathological proliferation of fibroblasts are the key steps leading to a phenotype known as fibrosis. Fibrosis is characterized by the excessive deposition of extracellular matrix, especially collagen. Stromal cells, including fibroblasts, express specific pro- and anti-fibrotic proteins. Keratinocyte growth factor (KGF) is a well-
15 characterized anti-fibrotic molecule. Additionally, oxidative damage and pro-inflammatory stimuli have been proposed to be among major events leading to myofibroblast phenotype and eventually to fibrosis. NF- κ B is a mediator of oxidative stress and inflammatory reactions. Based on fibroblast biology, we have developed four cell-based assays, namely fibroblast proliferation, collagen production, NF- κ B
20 activation and KGF production assays.

A. Human fibroblast proliferation assay

An activation and pathological proliferation of fibroblasts are the key steps leading to a phenotype known as fibrosis. The assay is based on fluorescence enhancement mediated by CyQUANT GR dye bound to cellular nucleic acids and measures the
25 proliferative responses of human skin-derived fibroblasts to novel proteins and small molecules.

B. Type I collagen production by human fibroblasts

Fibrosis is characterized by the excessive deposition of extracellular matrix, especially collagen. Over production of type I collagen is the main manifestation of systemic
30 sclerosis. TGF β is known to up-regulate production of collagen *in vitro* and *in vivo*. We developed cell-based assay in order to test the ability of novel pro-or anti-fibrotic

- 70 -

molecules to modulate basal or TGF β 1-stimulated levels of type I collagen production by human skin-derived fibroblasts.

C. Keratinocyte growth factor (KGF) production by human fibroblasts

KGF is an important mediator of stroma-to epithelium interactions in many organs
5 (lung, pancreas, kidney, prostate, mammary, gland, uterus) during normal and
pathological growth and development. KGF is specifically produced by stromal cells
and its receptor is specifically expressed by epithelial cells. It is proposed that KGF
might be an important player during pathophysiological reactions in fibrosis and thus
can be used as a marker of these reactions. A KGF ELISA assay has been developed
10 and using human lung-derived fibroblasts it has been shown that the KGF production
can be significantly up-regulated by IL-1 β and TNF α and down-regulated by TGF β .
These cytokines will be used as reference molecules in screening for novel proteins
capable to induce KGF production.

D. NF- κ B transcription activation in Fibroblasts

15 Oxidative damage and pro-inflammatory stimuli have been proposed to be among
major events leading to myofibroblast phenotype and eventually to fibrosis. NF- κ B is a
mediator of oxidative stress and inflammatory reactions. Swiss 3T3 fibroblasts were
generated with a stably integrated NF- κ B-SEAP (secreted alkaline phosphatase)
construct. NF- κ B-SEAP is designed to measure the binding of transcription factors to
20 the κ enhancer allowing a direct measurement of activation of the NF- κ B pathway. The
SEAP enzyme is secreted into the culture medium, so samples can be collected at
various time points to assay for transcription activity without harvesting cells. The Swiss
3T3-NF- κ B-SEAP cell line can be used as a cell-based assay to test novel Functional
Genomics proteins and is very promising for testing small molecules, especially those
25 with predicted pro-/anti-inflammatory activity.

E. Connective tissue growth factor (CTGF) promoter activation/repression in fibroblasts

CTGF, a 38-kD cysteine-rich protein, stimulates the production of extracellular matrix
elements by fibroblasts. CTGF overexpression has reportedly been found in many
30 fibrotic human tissues, including lung, skin, liver, kidney and blood vessels. *In vitro*,
TGF β activates CTGF gene transcription in human lung fibroblasts. A CTGF promoter-

- 71 -

reporter was constructed with secreted alkaline phosphatase (SEAP) as a reporter and Swiss 3T3 fibroblasts were generated with a stably integrated CTGF-SEAP construct. Using these fibroblasts it was shown that CTGF promoter is down-regulated by SARP-1, OPG and FSH and up-regulated by TGF β .

5 F. KL-6 production

KL-6, originally discovered as a pulmonary adenocarcinoma-related protein and later referred to as MUC-1, is a high-molecular-weight glycoprotein, now classified as Cluster 9 antigen. KL-6 is elevated in both sera and BALF of patients with idiopathic pulmonary fibrosis (IPF) and other lung interstitial diseases. In lung tissue from patients
10 suffering from IPF, the majority of cells labelled with KL-6 antibodies are regenerating type II pneumocytes. Two peptides were designed to produce polyclonal antibodies against KL-6. KL-6 ELISA can be used to measure KL-6 production by human lung-derived type II pneumocytes.

15 G. Neutralization of apoptosis of L-929 fibroblasts treated with soluble recombinant TRAIL (TNF-related apoptosis-inducing ligand)

TRAIL has been shown to be one of the cellular ligands for osteoprotegerin (OPG). This assay can be used to measure the biological activity of OPG.

H. RANKL (receptor activator of NF- κ B ligand) production by human fibroblasts

RANKL is another ligand for OPG. This assay can also be used to measure the
20 biological activity of OPG.

Example 8: Reproductive Health Assays Suitable for Exploration of the Biological Relevance of proteins Function:

A number of reproductive health-related assays have been developed by the Applicant
25 and are of use in the investigation of the biological relevance of SCS0008 protein function. In view of the probable implication of SCS0008 in male infertility (see therapeutic uses), such assays seem of particular relevance. Examples of reproductive health-related assays that have been developed by the Applicant include 18 cell-based assays for reproductive health. These are discussed below.

30

- 72 -

A. Primary human uterine smooth muscle proliferation assay:

The proliferation of uterine smooth muscle cells is a precursor for development of tumors in uterine fibroid disease in women. In this assay, the goal is to identify proteins that inhibit proliferation of primary human uterine smooth muscle cells.

5 B. JEG-3 Implantation assay:

JEG-3 cells are a choriotrophoblastic human cancer cell line used as a model for the blastocyst during implantation. Ishikawa cells are a relatively non-differentiated endometrial human cancer cell line that is used as a model for the decidua. JEG-3 cells will "implant" into human decidual tissue. In this assay, a 2-chamber system is used
10 where fluorescently labeled JEG-3 cells invade through a Matrigel-coated porous membrane from an upper chamber into a lower chamber when Ishikawa cells or Ishikawa-conditioned medium are placed into the lower chamber. The cells that migrate are quantified in a plate reader. The goal is to identify proteins that increase invasion of JEG-3 cells for use in aiding implantation *in vivo*.

15 C. Osteopontin bead assay (Ishikawa cells):

Ishikawa human endometrial cancer cells are used as a model for implantation. At the time of implantation in the human, various integrins are expressed by the uterine endometrium that is thought to bind to proteins expressed by the blastocyst. Ishikawa cells have been shown in the literature to express avb3, which is the integrin expressed
20 by the uterine endometrium during the "window of implantation". This integrin is believed to bind the osteopontin expressed by the trophoblast. In this assay, osteopontin-coated fluorescent beads represent the blastocyst, and the Ishikawa cells are primed to accept them for binding by treating them with estradiol. The goal is to identify proteins that increase the ability of the Ishikawa cells to bind the osteopontin -
25 beads as an aid to increase receptivity of the uterine endometrium at the time of implantation.

D. HuF6 assay:

HuF6 cells are primary human uterine fibroblast cells. These cells can be induced to decidualize by treating them with IL-1 β . A marker for decidualization is production of
30 PGE₂, which is measured by ELISA. The goal is to identify proteins that increase production of PGE₂ by the HuF6 cells as a way of enhancing decidualization during early pregnancy.

- 73 -

E. Endometriosis assay:

Peritoneal $\text{TNF}\alpha$ plays a role in endometriosis by inducing the sloughed endometrial cells from the uterus to adhere to and proliferate on peritoneal mesothelial cells. In this assay, BEND cells are treated with $\text{TNF}\alpha$, which increases their ability to bind
5 fibronectin-coated fluorescent beads as an assay for adherence during endometriosis. The goal is to identify proteins that decrease or inhibit the ability of $\text{TNF}\alpha$ to stimulate bead-binding capacity of the cells.

F. Cyclic AMP assay using JC-410 porcine granulosa cells stably transfected with hLHR:

10 In Polycystic Ovary Syndrome, LH from the pituitary is relatively high, and induces androgen output from the ovarian thecal cells. This assay is used to look for an inhibitor of LH signaling which could be used to decrease the action of LH at the ovary during PCOS. The JC-410 porcine granulosa cell line is stably transfected with the human LH receptor. Treatment with LH results in cAMP production.

15 G. Cyclic AMP assay using JC-410 porcine granulosa cells stably transfected with hFSHR:

The JC-410 porcine granulosa cell line was stably transfected with the human FSHR. Treatment with FSH stimulates cAMP production, which is measured in this assay. The goal is to identify proteins that enhance FSH action in the granulosa cells.

20 H. LbetaT2 (mouse) pituitary cells assay:

The LbetaT2 is an immortalized murine pituitary gonadotroph cell line. Stimulation with Activin alone or with GnRH + Activin results in secretion of FSH (stimulation with GnRH alone results in secretion of LH.) The cells can either be treated with GnRH + Bioscreen proteins to find proteins that act in concert with GnRH to stimulate FSH
25 production, or they can be treated with Bioscreen proteins alone to find a protein that can stimulate FSH secretion like activin alone.

I. Cumulus expansion assay:

The cumulus-expansion assay using murine cumulus-oocyte complexes (2/well) has been validated in a 96-well format to assay for proteins that affect oocyte maturation
30 (measured by cumulus expansion). Two 96-well plates can be processed per assay, and 2 assays per week can be performed. If Bioscreen proteins are assayed at only one concentration, all Bioscreen I proteins can be assayed in a month. The read-out

- 74 -

may be a yes/no answer for expansion, or image analysis programs may be used to measure expansion in a quantitative manner.

J. RWPE Proliferation assay:

Benign prostatic hyperplasia is characterized by growth of prostatic epithelium and stroma that is not balanced by apoptosis, resulting in enlargement of the organ. RWPE is a regular human prostatic epithelial cell line that was immortalized with the HPV-18, and may be used in place of primary human prostatic epithelial cells.

K. HT-1080 fibrosarcoma invasion assay:

This assay was developed as a positive cell control for the JEG-3 implantation assay (above). This is a well-established assay as a model for cancer metastasis. Fluorescently-labeled HT-1080 human fibrosarcoma cells are cultured in the upper chamber of a 2-chamber system, and can be stimulated to invade through the porous Matrigel-coated membrane into the bottom chamber where they are quantified. The goal is to identify a protein that inhibits the invasion. The cells are stimulated to invade by adding serum to the bottom chamber and are inhibited with doxycycline.

L. Primary human uterine smooth muscle assay:

One of the hallmarks of uterine fibroid disease is collagen deposition by the uterine smooth muscle cells that have become leiomyomas. Primary human uterine smooth muscle cells are stimulated to produce collagen by treatment with TGF β , which is blocked with Rebif. The goal is to discover proteins that inhibit this fibrotic phenotype.

M. Human leiomyoma cells proliferation assay:

A human leiomyoma cell line may be used as a model for uterine fibroid disease in a proliferation assay. The cells grow very slowly and we are stimulating them to grow at a faster rate by treating them with estradiol and growth factors. The goal is to identify proteins that inhibit estradiol-dependent growth of leiomyoma cells.

N. 937 Migration assay:

Endometriotic lesions secrete cytokines that recruit immune cells to the peritoneal cavity. These immune cells (especially activated macrophages and T lymphocytes) mediate inflammatory symptoms that are common to endometriosis. RANTES has been shown to be produced by endometriotic stromal cells and is present in the peritoneal fluid. In this assay, U937, a monocytic cell line used as a model for activated macrophages, can be induced by treating the lower level of a 2-chamber culture

- 75 -

system to migrate from the upper chamber. If the cells are pre-loaded with fluorescent dye, they can be quantified in the lower chamber. The goal is to identify proteins that inhibit the migration of the U937 cells.

O. JEG3 human trophoblast assay:

- 5 The trophoblast of the blastocyst produces HLA-G, a class I HLA molecule that is believed to be important in preventing immunological rejection of the embryo by the mother. During pre-eclampsia, HLA-G levels are low or non-existent, presumably resulting in hallmark symptoms such as poor invasion of the trophoblast into the endometrium and spiral arteries because of maternal immunological interference. The
- 10 JEG-3 human trophoblast cell line produces HLA-G, which can be increased by treatment with IL-10 or LIF. An ELISA can be used to measure HLA-G production by JEG-3 cells, with the goal being the discovery of other proteins that can increase HLA - G production.

P. Primary rat ovarian dispersate assay:

- 15 Due to the difficulties in measuring appreciable amounts of steroids from the JC-410-FSHR/LHR cell lines, an assay using primary cells from whole ovaries taken from immature rats has been developed. Initially, estradiol production from these cultures is measured after treatment with FSH and/or LH. The goal is then to identify proteins that enhance gonadotropin-stimulated steroidogenesis, or proteins that work alone to
- 20 increase steroidogenesis by these cultures.

Q. Mouse IVF assay:

In this assay, sperm function, measured by ability to fertilize oocytes, is assayed with the goal of finding proteins that stimulate fertilizing potential of sperm.

R. Primary human prostate stromal cells proliferation assay:

- 25 An assay for the epithelial component of BPH has already been described above (see RWPE assay above). This assay uses primary human prostate stromal cells as a model for proliferation of these cells during BPH. The goal is to identify proteins that inhibit proliferation of these cells.

REFERENCES

- Andersen DC and Krummen L, Curr Opin Biotechnol, 13: 117-23, 2002.
- Baker KN *et al.*, Trends Biotechnol, 20: 149-56, 2002.
- Blagoev B and Pandey A, Trends Biochem Sci, 26: 639-41, 2001.
- 5 Bock A, Science, 292: 453-4, 2001.
- Bunz F, Curr Opin Oncol, 14: 73-8, 2002.
- Burgess RR and Thompson NE, Curr Opin Biotechnol, 12: 450-4, 2001.
- Chambers SP, Drug Disc Today, 14: 759-765, 2002.
- Chu L and Robinson DK, Curr Opin Biotechnol, 13: 304-8, 2001.
- 10 Cleland JL *et al.*, Curr Opin Biotechnol, 12: 212-9, 2001.
- Coleman RA *et al.*, Drug Discov Today, 6: 1116-1126, 2001.
- Constans A, The Scientist, 16(4): 37, 2002.
- Davis BG and Robinson MA, Curr Opin Drug Discov Devel, 5: 279-88, 2002.
- Dougherty DA, Curr Opin Chem Biol, 4: 645-52, 2000.
- 15 Garnett MC, Adv Drug Deliv Rev, 53: 171-216, 2001.
- Gavilondo JV and Larrick JW, Biotechniques, 29: 128-136, 2000.
- Gendel SM, Ann N Y Acad Sci, 964: 87-98, 2002.
- Giddings G, Curr Opin Biotechnol, 12: 450-4, 2001.
- Golebiowski A *et al.*, Curr Opin Drug Discov Devel, 4: 428-34, 2001.
- 20 Gupta P *et al.*, Drug Discov Today, 7: 569-579, 2002.
- Haupt K, Nat Biotechnol, 20 : 884-885, 2002.
- Hruby VJ and Balse PM, Curr Med Chem, 7: 945-70, 2000.
- Johnson DE and Wolfgang GH, Drug Discov Today, 5: 445-454, 2000.
- Kane JF, Curr Opin Biotechnol, 6: 494-500, 1995.
- 25 Kolb AF, Cloning Stem Cells, 4: 65-80, 2002.
- Kuroiwa Y *et al.*, Nat Biotechnol, 20: 889-94, 2002.

- 77 -

- Lin Cereghino GP *et al.*, Curr Opin Biotechnol, 13: 329-332, 2001.
- Lowe CR *et al.*, J Biochem Biophys Methods, 49: 561-74, 2001.
- Luo B and Prestwich GD, Exp Opin Ther Patents, 11: 1395-1410, 2001.
- Mulder NJ and Apweiler R, Genome Biol, 3(1):REVIEWS2001, 2002
- 5 Nilsson J *et al.*, Protein Expr Purif, 11: 1-16, 1997.
- Pearson WR and Miller W, Methods Enzymol, 210: 575-601, 1992.
- Pellois JP *et al.*, Nat Biotechnol, 20: 922-6, 2002.
- Pillai O and Panchagnula R, Cur Opin Chem Biol, 5: 447-451, 2001
- Rehm BH, Appl Microbiol Biotechnol, 57: 579-92, 2001.
- 10 Robinson CR, Nat Biotechnol, 20: 879-880, 2002.
- Rogov SI and Nekrasov AN, Protein Eng, 14: 459-463, 2001.
- Schellekens H, Nat Rev Drug Discov, 1: 457-62, 2002
- Sheibani N, Prep Biochem Biotechnol, 29: 77-90, 1999.
- Stevanovic S, Nat Rev Cancer, 2: 514-20, 2002.
- 15 Templin MF *et al.*, Trends Biotechnol, 20: 160-6, 2002.
- Tribbick G, J Immunol Methods, 267: 27-35, 2002.
- van den Burg B and Eijssink V, Curr Opin Biotechnol, 13: 333-337, 2002.
- van Dijk MA and van de Winkel JG, Curr Opin Chem Biol, 5: 368-74, 2001.
- Villain M *et al.*, Chem Biol, 8: 673-9, 2001.

CLAIMS

1. An isolated polypeptide having fibrillin-like activity selected from the group consisting of:
 - 5 a) the amino acid sequences as recited in SEQ ID NO: 2; SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 13, SEQ ID NO: 14;
 - b) the mature form of the polypeptides whose sequences are recited in SEQ ID NO: 2 (SEQ ID NO:3), SEQ ID NO: 6 (SEQ ID NO:8), SEQ ID NO: 7 (SEQ ID NO: 9), SEQ ID NO: 13 (SEQ ID NO: 15), SEQ ID NO: 14 (SEQ ID NO: 16);
 - 10 c) the histidine tag form of the polypeptides whose sequences are recited in SEQ ID NO: 2 (SEQ ID NO:4), SEQ ID NO: 6 (SEQ ID NO:10), SEQ ID NO: 7 (SEQ ID NO: 11), SEQ ID NO: 13 (SEQ ID NO: 17), SEQ ID NO: 14 (SEQ ID NO: 18);
 - 15 d) a variant of the amino acid sequence recited in SEQ ID NO: 2, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 13, SEQ ID NO: 14, wherein any amino acid specified in the chosen sequence is non-conservatively substituted, provided that no more than 15% of the amino acid residues in the sequence are so changed;
 - 20 e) an active fragment, precursor, salt, or derivative of the amino acid sequences given in a) to d).
2. The polypeptide of claim 1 that is a naturally occurring allelic variant of any of the sequences given by SEQ ID NO: 2, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 13, or SEQ ID NO: 14.
- 25 3. The polypeptide of claim 2, wherein the variant is the translation of a single nucleotide polymorphism.
4. The polypeptide of any one of claims 1 to 3, wherein the polypeptide binds specifically an antibody or a binding protein generated against the polypeptide whose sequence is recited in SEQ ID NO: 2, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 13, SEQ ID NO: 14 or a fragment thereof.
- 30

- 79 -

5. A fusion protein comprising a polypeptide according to any of the claims from 1 to 4.
6. The fusion proteins of claim 6 wherein said proteins further comprise one or more amino acid sequence belonging to these protein sequences: membrane-bound protein, immunoglobulin constant region, multimerization domains,
5 extracellular proteins, signal peptide-containing proteins, export signal-containing proteins.
7. An antagonist of a polypeptide of any one of claims 1 to 4, wherein said antagonist comprises an amino acid sequence resulting from the non-conservative substitution and/or the deletion of one or more residues into the
10 corresponding polypeptide.
8. A ligand which binds specifically to a polypeptide according to any one of claims 1 to 4.
9. The ligand of claim 8 that antagonizes or inhibits the fibrillin-like activity of a
15 polypeptide according to any one of claims 1 to 4.
10. A ligand according to claim 9 which is a monoclonal antibody, a polyclonal antibody, a humanized antibody, an antigen binding fragment, or the extracellular domain of a membrane-bound protein.
11. The polypeptide of any one of claims 1 to 6, wherein said polypeptides are in
20 the form of active conjugates or complexes with a molecule chosen amongst radioactive labels, fluorescent labels, biotin, or cytotoxic agents.
12. A peptide mimetic designed on the sequence and/or the structure of a polypeptide according to any one of claims 1 to 4.
13. An isolated nucleic acid encoding for an isolated polypeptide selected from the
25 group consisting of:
 - b) the polypeptides having fibrillin-like activity of any one of claims 1 to 4;
 - c) the fusion proteins of claim 5 or 6; or
 - d) the antagonists of claim 7.
14. The nucleic acid of claim 13, comprising a DNA sequence consisting of SEQ ID
30 NO: 1, nucleotides 205-1425 of SEQ ID NO:1, or a DNA sequence consisting of

- 80 -

SEQ ID NO: 5, nucleotides 205-1590 of SEQ ID NO:5, or a DNA sequence consisting of SEQ ID NO: 12, nucleotides 205-1503 of SEQ ID NO:12, or the complement of either of said DNA sequences.

15. A purified nucleic acid which:

- 5 b) hybridizes under high stringency conditions; or
- c) exhibits at least about 85% identity over a stretch of at least about 30 nucleotides

 with a nucleic acid selected from the group consisting of SEQ ID NO: 1 nucleotides 205-1425 of SEQ ID NO:1, or a DNA sequence consisting of SEQ ID
10 NO: 5, nucleotides 205-1590 of SEQ ID NO:5, or a DNA sequence consisting of SEQ ID NO: 12, nucleotides 205-1503 of SEQ ID NO:12, or a complement of either of said DNA sequences.

16. A vector comprising a nucleic acid as recited in any one of claims 13 to 15.

17. The vector of claim 16, wherein said nucleic acid molecule is operatively linked
15 to expression control sequences allowing expression in prokaryotic or eukaryotic host cells of the encoded polypeptide.

18. A polypeptide encoded by the purified nucleic acid of any one of claims 13-15.

19. A process for producing cells capable of expressing a polypeptide of any one of claims from 1 to 7 or of claim 18, comprising genetically engineering cells with a
20 vector or a nucleic acid according to any of the claims from 13 to 17.

20. A host cell transformed with a vector or a nucleic acid according to any of the claims from 13 to 17.

21. A transgenic animal cell that has been transformed with a vector or a nucleic acid according to any of the claims from 13 to 17, having enhanced or reduced
25 expression levels of a polypeptide according to any one of claims 1 to 4.

22. A transgenic non-human animal that has been transformed to have enhanced or reduced expression levels of a polypeptide according to any one of claims 1 to 4.

23. A method for making a polypeptide of any one of claims from 1 to 7 comprising
30 culturing a cell of claim 20 or 21 under conditions in which the nucleic acid or

- 81 -

vector is expressed, and recovering the polypeptide encoded by said nucleic acid or vector from the culture.

24. A compound that enhances the expression level of a polypeptide according to any one of claims 1 to 4 into a cell or in an animal.
- 5 25. A compound that reduces the expression level of a polypeptide according to any one of claims 1 to 4 into a cell or in an animal.
26. The compound of claim 24 that is an antisense oligonucleotide or a small interfering RNA.
- 10 27. A purified preparation containing a polypeptide of any one of claims 1 to 6 or claim 18, an antagonist of claim 7, a ligand of any one of claims 8 to 10, peptide mimetic of claim 12, a nucleic acid of any one of claims 13 to 17, a cell of claim 20 or 21, or a compound of any one of claims 24 to 26.
- 15 28. Use of a polypeptide of any one of claims 1 to 6 or claim 18, a peptide mimetic of claim 12, a nucleic acid of any one of claims 13 to 17, a cell of claim 20 or 21, or a compound of claim 24, in the therapy or in the prevention of a disease when the increase in the fibrillin-like activity of a polypeptide of any one of claims 1 to 4 is needed.
- 20 29. A pharmaceutical composition for the treatment or prevention of diseases needing an increase in the fibrillin-like activity of a polypeptide of any one of claims 1 to 6 or claim 18, a peptide mimetic of claim 12, a nucleic acid of any one of claims 13 to 17, a cell of claim 20 or 21, or a compound of claim 24, as active ingredient.
- 25 30. Process for the preparation of a pharmaceutical composition, which comprises combining a polypeptide of any one of claims 1 to 6 or claim 18, a peptide mimetic of claim 12, a nucleic acid of any one of claims 13 to 17, a cell of claim 20 or 21, or a compound of claim 24, together with a pharmaceutically acceptable carrier.
- 30 31. Method for the treatment or prevention of a disease needing an increase in the fibrillin-like activity of a polypeptide of any one of claims 1 to 4, comprising the administration of a therapeutically effective amount of a polypeptide of any one

of claims 1 to 6 or claim 18, a peptide mimetic of claim 12, a nucleic acid of any one of claims 13 to 17, a cell of claim 20 or 21, or a compound of claim 24.

32. Use of an antagonist of claim 7, a ligand of any one of claims 8 to 10, or of a compound of claim 25 or claim 26, in the therapy or in the prevention of a disease associated to the excessive fibrillin-like activity of a polypeptide of any one of claims 1 to 4.
33. A pharmaceutical composition for the treatment or prevention of a disease associated to the excessive fibrillin-like activity of a polypeptide of any one of claims 1 to 4, containing an antagonist of claim 7, a ligand of any one of claims 8 to 10, or of a compound of claim 25 or claim 26, as active ingredient.
34. Process for the preparation of pharmaceutical compositions for the treatment or prevention of diseases associated to the excessive fibrillin-like activity of a polypeptide of any one of claims 1 to 4, which comprises combining an antagonist of claim 7, a ligand of any one of claims 8 to 10, or of a compound of claim 25 or claim 26, together with a pharmaceutically acceptable carrier.
35. A method for the treatment or prevention of diseases related to the polypeptide of any one of claims 1 to 4, comprising the administration of a therapeutically effective amount of an antagonist of claim 7, a ligand of any one of claims 8 to 10, or of a compound of claim 25 or claim 26.
36. A method for screening candidate compounds effective to treat a disease related to the fibrillin-like polypeptides of any one of claims 1 to 4, comprising:
- b) contacting a cell of claim 20, a transgenic animal cell of claim 21, or a transgenic non-human animal according to claim 22, having enhanced or reduced expression levels of the polypeptide, with a candidate compound and
 - c) determining the effect of the compound on the animal or on the cell.
37. A method for identifying a candidate compound as an antagonist/inhibitor or agonist/activator of a polypeptide of any one of the claims 1 to 4 comprising:
- a) contacting said polypeptide, said compound, and a mammalian cell or a mammalian cell membrane capable of binding the polypeptide; and

- 83 -

- b) measuring whether the molecule blocks or enhances the interaction of the polypeptide, or the response that results from such interaction, with the mammalian cell or the mammalian cell membrane.
38. A method for determining the activity and/or the presence of the polypeptide of any one of claims from 1 to 4 in a sample, the method comprising:
- a) providing a protein-containing sample;
- b) contacting said sample with a ligand of any one of claims 8 to 10; and
- c) determining the presence or said ligand bound to said polypeptide.
39. A method for determining the presence or the amount of a transcript or of a nucleic acid encoding the polypeptide of any one of claims from 1 to 4 in a sample, the method comprising:
- a) providing a nucleic acids-containing sample;
- b) contacting said sample with a nucleic acid of any one of the claims 13 to 17; and
- c) determining the hybridization of said nucleic acid with a nucleic acid into the sample.
40. Use of a primer derived from a nucleotide sequence as listed in SEQ ID NO:1, SEQ ID NO:5, SEQ ID NO:12 for determining the presence or the amount of a transcript or of a nucleic acid encoding a polypeptide of any one of claims from 1 to 4 in a sample by Polymerase Chain Reaction
41. A kit for measuring the activity and/or the presence of the fibrillin-like polypeptides of any one of claims 1 to 4 in a sample comprising one or more of the following reagents: a polypeptide of any one of claims 1 to 6 or claim 18, an antagonist of claim 7, a ligand of any one of claims 8 to 10, a polypeptide of claim 11, a peptide mimetic of claim 12, a nucleic acid of any one of claims 13 to 17, a cell of claim 20 or 21, a compound of any one of claims 24 to 26, a pharmaceutical composition of claims 29 or 33.

- 1/28 -

Figure 1:

	SCS0008	MDFLALVLVSSLYLQAAAEFDG_RWPRQIVSSIGLCRYGGRIDCCWGWARQSWGQCQP
	POEM (Morimura N)	=AV==A==A=====V==D=====
5	NEPH1A (BRANDENBERGER R)	=AV==A==A=====V==D=====
	NEPH1B (BRANDENBERGER R)	=AV==A==A=====V==D=====
	NEPH1C (BRANDENBERGER R)	=AV==A==A=====V==D=====
	SCEP-39 (INCYTE WO0248337)	=====S=====
	SEQID34 (CURAGEN WO0110902)	=====
10	SEQID82 (CURAGEN WO0110902)	=====
	NOV5 (CURAGEN WO0257452)	=====EY=====
	PRO334 (GENENTECH WO0104311)	=====
	PRO26 (GENENTECH WO0104311)	=====
	PRO334 (GENENTECH WO9914328)	=====
15	SCS0008	VCQPRCKHGECIGPNKCKCHPGYAGKTCN
	POEM (MORIMURA N)	=====Q=====V=====F=====
	NEPH1A (BRANDENBERGER R)	FYVLRQRLARIRCQLKA=====Q=====V=====F=====
	NEPH1B (BRANDENBERGER R)	=====Q=====V=====F=====QDESFHPTPLDQGSSE
20	NEPH1C (BRANDENBERGER R)	FYVLRQRLARIRCQLKA=====Q=====V=====F=====QDESFHPTPLDQGSSE
	SCEP-39 (INCYTE WO0248337)	=====
	SEQID34 (CURAGEN WO0110902)	FYVLRQRIARIRCQLKA=====I=====
	SEQID82 (CURAGEN WO0110902)	=====I=====
	NOV5 (CURAGEN WO0257452)	=====
25	PRO334 (GENENTECH WO0104311)	=====
	PRO26 (GENENTECH WO0104311)	=====
	PRO334 (GENENTECH WO9914328)	=====
30	SCS0008	QDLNECGLKPRPCKHRCMNTYGSYKCYCLNGYMLMPDGSCSSAL
	POEM (MORIMURA N)	=====F=====L=====
	NEPH1A (BRANDENBERGER R)	=====F=====L=====
	NEPH1B (BRANDENBERGER R)	QPLFQPPDHQATNVPSR=====F=====L=====
	NEPH1C (BRANDENBERGER R)	QPLFQPPDHQATNVPSR=====F=====L=====
35	SCEP-39 (INCYTE WO0248337)	=====Q=====
	SEQID34 (CURAGEN WO0110902)	=====V=====
	SEQID82 (CURAGEN WO0110902)	=====V=====
	NOV5 (CURAGEN WO0257452)	=====
	PRO334 (GENENTECH WO0104311)	=====
	PRO26 (GENENTECH WO0104311)	=====
40	PRO334 (GENENTECH WO9914328)	=====
45	SCS0008	TCSMANCQYGCDDVVKQIRQCQCPSPGLQLAPDGRCTCDVDECATGRASCPRFRQCNTFG
	POEM (MORIMURA N)	S=====V=====I=====V=====
	NEPH1A (BRANDENBERGER R)	S=====V=====I=====V=====
	NEPH1B (BRANDENBERGER R)	S=====V=====I=====V=====
	NEPH1C (BRANDENBERGER R)	S=====V=====I=====V=====
	SCEP-39 (INCYTE WO0248337)	=====
	SEQID34 (CURAGEN WO0110902)	=====
	SEQID82 (CURAGEN WO0110902)	=====
50	NOV5 (CURAGEN WO0257452)	=====
	PRO334 (GENENTECH WO0104311)	=====H=====
	PRO26 (GENENTECH WO0104311)	=====H=====
	PRO334 (GENENTECH WO9914328)	=====H=====
55	SCS0008	SYICKCHKGFDMYIGGKYQCHDIDECSLGQYQCSSFARCYNVVRGSYKCKKEGYQGDGL
	POEM (MORIMURA N)	=====T=====H=====Y=====I=====Q=====R=====D=====E=====
	NEPH1A (BRANDENBERGER R)	=====T=====H=====Y=====I=====Q=====R=====D=====E=====
	NEPH1B (BRANDENBERGER R)	=====T=====H=====Y=====I=====Q=====R=====D=====E=====
	NEPH1C (BRANDENBERGER R)	=====T=====H=====Y=====I=====Q=====R=====D=====E=====
60	SCEP-39 (INCYTE WO0248337)	=====
	SEQID34 (CURAGEN WO0110902)	=====
	SEQID82 (CURAGEN WO0110902)	=====
	NOV5 (CURAGEN WO0257452)	=====
	PRO334 (GENENTECH WO0104311)	=====
65	PRO26 (GENENTECH WO0104311)	=====
	PRO334 (GENENTECH WO9914328)	=====

- 2/28 -

[illegible]

- 3/28 -

SEQID34 (CURAGENWO0110902)	SEER
SEQID82 (CURAGENWO0110902)	SEER
NOV5 (CURAGENWO0257452)	SEER

- 4/28 -

Figure 2

SCS0008 MDLLALVLVSSSLYLQAAAEFDGRWPRQIVSSIGLCRYGGRIDCCWGWARQSWGQCQPVC
 5 SV1 --FLLALVLVSSSLYLQAAAEFDGRWPRQIVSSIGLCRYGGRIDCCWGWARQSWGQCQPVC
 SV2 --LLALVLVSSSLYLQAAAEFDGRWPRQIVSSIGLCRYGGRIDCCWGWARQSWGQCQPVC

SCS0008 QPRCKHGECIGPNKCKCHPGYAGKTCNQDLNECGLKPRPCKHRCMNTYGSYKCYCLNGYM
 10 SV1 QPRCKHGECIGPNKCKCHPGYAGKTCNQDLNECGLKPRPCKHRCMNTYGSYKCYCLNGYM
 SV2 QPRCKHGECIGPNKCK CHPGYAGKTCNQDLNECGLKPRPCKHRCMNTYGSYKCYCLNGYM

SCS0008 LMPDGSCSSALTCSMANCQYGCDVVKGQIRCQCPSGLQAPDGRTCDVDDECATGRASC
 15 SV1 LMPDGSCSSALTCSMANCQYGCDVVKGQIRCQCPSGL LLAPDGRTCDVDDECATGRASC
 SV2 LMPDGSCSSALTCSMANCQYGCDVVKGQIRCQCPSGL LLAPDGRTCDVDDECATGRASC

SCS0008 PRFRQCVNTFGSYICKCHKGFDLMYIGGKYQCHDIDECSLGQYQCSSFARCYNVRGSYK
 20 SV1 PRFRQCVNTFGSYICKCHKGFDLMYIGGKYQCHDIDECSLGQYQCSSFARCYNVRGSYK
 SV2 PRFRQCVNTFGSYICKCHKGFDLMYIGGKYQCHDIDECSLGQYQCSSFARCYNVRGSYK

SCS0008 KCKEGYQGDGL TCV-----
 25 SV1 KCKEGYQGDGLTCV YIPKVMIEPSGPIHVPKNGTILKGD TGNNWIPDVGSTWWPPKTP
 SV2 KCKEGYQGDGLTCV YIPKVMIEPSGPIHVPKNGTILKGD TGNNWIPDVGSTWWPPKTP

SCS0008 -----PTSKPTTRPTPKPTPIPTPPPPPLPTELRTPLPPTTPTTGLTTIAPA
 30 SV1 YIPPIITNRPTSKPTTRPTPKPTPIPTPPPPPLPTELRTPLPPTTPTTGLTTIAPA
 SV2 YIPPIITNRPTSKPTTRPTPKPTPIPTPPPPPLPTELRTPLPPTTPTTGLTTIAPA

SCS0008 ASTPPGGITVDNRVQTDPOKPRGDVFIQPSNDLFEIFEIERGVSADDEAKDDPGVLVH
 35 SV1 ASTPPGGITVDNRVQTDPOKPRGDVFIQPSNDLFEIFEIERGVSADDEAKDDPGVLVH
 SV2 ASTPPGG ITVDNRVQTDPOKPRGDVFS VLVH

SCS0008 SCNFDHGLCGWIREKDNDLHWEPIRDPADNC SLHAGGLDL
 40 SV1 SCNFDHGLCGWIREKDNDLHWEPIRDPADNC SLHAGGLDL
 SV2 SCNFDHGLCGWIREKDNDLHWEPIRDPADNC SLHAGGLDL

Single amino acid substitutions are highlighted in yellow
 45 55 amino acid extension of exon 8 is highlighted in yellow
 Region of exon 9 deletion in SV2 is highlighted in lime green

- 5/28 -

Figure 3

Nucleotide sequence of SCS0008 prediction with translation

```

5      1  GCGGGAGGGG GCTCCGGGCG CCGCGCAGCA GACCTGCTCC GGCCGCGCGC CTCGCCGCTG
      61  TCCTCCGGGA GCGGCAGCAG TAGCCCGGGC GCGGAGGGCT GGGGGTTCTT CGAGACTCTC
     121  AGAGGGGCGC CTCCCATCGG CGCCCACCAC CCCAACCTGT TCCTCGCGCG CCACTGCGCT
     181  GCGCCCCAGG ACCCGCTGCC CAACATGGAT TTTCTCTGG CGCTGGTGCT GGTATCCTCG
                                SCS0008-CP1 → SCS0008-CP1nest →
10      M D F L L A L V L V S S

     241  CTCTACCTGC AGGCGGCCGC CGAGTTCGAC GGGAGGTGGC CC AGGCAAAT AGTGTCTATG
          L Y L Q A A A E F D G R W P R Q I V S S
     301  ATTGGCCTAT GTCGTTATGG TGGGAGGATT GACTGCTGCT GGGGCTGGGC TCGCCAGTCT
          I G L C R Y G G R I D C C W G W A R Q S
     361  TGGGGACAGT GTCAGCCTGT GTGCCAACCA CGATGCAAAC ATGGTGAATG TATCGGGCCA
          W G Q C Q P V C Q P R C K H G E C I G P
20      421  AACAAGTGCA AGTGTCTATCC TGGTTATGCT GGAAAAACCT GTAATCAAGA TCTAAATGAG
          N K C K C H P G Y A G K T C N Q D L N E
     481  TGTGGCCTGA AGCCCCGGCC CTGTAAGCAC AGGTGCATGA ACACTTACGG CAGCTACAAG
          C G L K P R P C K H R C M N T Y G S Y K
     541  TGCTACTGTC TCAACGGATA TATGCTCATG CCGGATGGTT CCTGCTCAAG TGCCTGACC
          C Y C L N G Y M L M P D G S C S S A L T
30      601  TGCTCCATGG CAAACTGTCA GTATGGCTGT GATGTTGTTA AAGGACAAAT ACGGTGCCAG
          C S M A N C Q Y G C D V V K G Q I R C Q
     661  TGCCCATCCC CTGGCCTGCA GCTGGCTCCT GATGGGAG GA CCTGTGTAGA TGTTGATGAA
          C P S P G L Q L A P D G R T C V D V D E
35      721  TGTGCTACAG GAAGAGCCTC CTGCCCTAGA TTTAGGCAAT GTGTCAACAC TTTTGGGAGC
          C A T G R A S C P R F R Q C V N T F G S
     781  TACATCTGCA AGTGTCTATA AGGCTTCGAT CTCATGTATA TTGGAGGCAA ATATCAATGT
          Y I C K C H K G F D L M Y I G G K Y Q C
     841  CATGACATAG ACGAATGCTC ACTTGGTCAG TATCAGTGCA GCAGCTTTGC TCGATGTTAT
          H D I D E C S L G Q Y Q C S S F A R C Y
45      901  AACGTACGTG GGTCTTACAA GTGCAAATGT AAAGAAGGAT ACCAGGGTGA TGGACTGACT
          N V R G S Y K C K C K E G Y Q G D G L T
     961  TGTGTGCCTA CTTCTAAGCC AACAACAAGA CCTACACCAA AGCCAACACC AATTCCTACT
          C V P T S K P T T R P T P K P T P I P T
50     1021  CCACCACCAC CACCACCCCT GCCAACAGAG CTCAGAACAC CTCTACCACC TACAACCCCA

```

- 6/28 -

P P P P P P L P T E L R T P L P P T T P

5 1081 GAAAGGCCAA CCACCGGACT GACAACTATA GCA CCAGCTG CCAGTACACC TCCAGGAGGG
E R P T T G L T T I A P A A S T P P G G

1141 ATTACAGTTG ACAACAGGGT ACAGACAGAC CCTCAGAAAC CCAGAGGAGA TGTGTTTCATT
I T V D N R V Q T D P Q K P R G D V F I

10 1201 CCACGGCAAC CTTCAAATGA CTTGTTTGAA ATATTTGAAA TAGAAAGAGG AGTCAGTGCA
P R Q P S N D L F E I F E I E R G V S A

1261 GACGATGAAG CAAAGGATGA TCCAGGTGTT CTGGTACACA GTTGTAAATT TGACCATGGA
D D E A K D D P G V L V H S C N F D H G

15 1321 CTTTGTGGAT GGATCAGGGA GAAAGACAAT GACTTGCACT GGGAACCAAT CAGGGACCCA
L C G W I R E K D N D L H W E P I R D P

1381 GCAGACAATT GTAGCTGTCT GCATGCAGGT GGACTTGACT TGTAGCAACT GGAGACAAGG
← SCS0008-CP2nest

20 A D N C S C L H A G G L D L

1441 TGTAGTTCTC CATGTGTCTG AGATTCCAGA ACTAGACTCA ATTGTGGTTC CAGATCAAAT
←

1501 CAGCCTGCAG AGAATGGCTG TTCCCTGGTG CACATGGAGT TGAGCGGAGC AGCACACTTG
← SCS0008-CP2

1561 GGCACAGGGG TTGGCCGTCA GCTCACTGTC CTGCTGCCAC CCTCCACCAC CTCAGAGGCT

25 1621 GCTGTAACTT TGAGAAAATA ACCTTTGAAC CTCAGTCTCT ACAATGCTTA AATAAAGAGA

1681 ACATTCAATT AGAGAAAATA

Position and sense of PCR primers →

30

- 7/28 -

Figure 4

Nucleotide sequence with translation of SCS0008 product cloned using primers SCS0008-CP1nest and SCS0008-CP2nest.

```

5      1  ATTTTCTCCT GCGCTGGTG CTGGTATCCT CGCTCTACCT GCAGGCGGCC GCCGAGTTCG
        F L L A L V L V S S L Y L Q A A A E F

61  ACGGGAGGTG GCCCAGGCAA ATAGTGTTCAT CGATTGGCCT ATGTCGTTAT GGTGGGAGGA
10      D G R W P R Q I V S S I G L C R Y G G R

121  TTGACTGCTG CTGGGGCTGG GCTCGCCAGT CTTGGGGACA GTGTCAGCCT GTGTGCCAAC
        I D C C W G W A R Q S W G Q C Q P V C Q

15  181  CACGATGCAA ACATGGTGAA TGTATCGGGC CAAACAAGTG CAAGTGTTCAT CCTGGTTATG
        P R C K H G E C I G P N K C K C H P G Y

241  CTGGA AAAAC CTGTAATCAA GATCTAAATG AGTGTGGCCT GAAGCCCCGG CCCTGTAAGC
        A G K T C N Q D L N E C G L K P R P C K

20  301  ACAGGTGCAT GAACACTTAC GGCAGCTACA AGTGCTACTG TCTCAACGGA TATATGCTCA
        H R C M N T Y G S Y K C Y C L N G Y M L

361  TGCCGGATGG TTCCTGCTCA AGTGCCCTGA CCTGCTCCAT GGCAAACTGT CAGTATGGCT
25      M P D G S C S S A L T C S M A N C Q Y G

421  GTGATGTTGT TAAAGGACAA ATACGGTGCC AGTGCCCATC CCCTGGCCTG CACCTGGCTC
        C D V V K G Q I R C Q C P S P G L H L A

30  481  CTGATGGGAG GACCTGTGTA GATGTTGATG AATGTGCTAC AGGAAGAGCC TCCTGCCCTA
        P D G R T C V D V D E C A T G R A S C P

541  GATTTAGGCA ATGTGTCAAC ACTTTTGGGA GCTACATCTG CAAGTGTTCAT AAAGGCTTCG
        R F R Q C V N T F G S Y I C K C H K G F

35  601  ATCTCATGTA TATTGGAGGC AAATATCAAT GTCATGACAT AGACGAATGC TCACCTGGTC
        D L M Y I G G K Y Q C H D I D E C S L G

661  AGTATCAGTG CAGCAGCTTT GCTCGATGTT ATAACGTACG TGGGTCCTAC AAGTGCAAAT
40      Q Y Q C S S F A R C Y N V R G S Y K C K

721  GTAAAGAAGG ATACCAGGGT GATGGACTGA CTTGTGTGTA TATCCCAAAA GTTATGATTG
        C K E G Y Q G D G L T C V Y I P K V M I

45  781  AACCTTCAGG TCCAATTCAT GTACCAAAGG GAAATGGTAC CATTTTAAAG GGTGACACAG
        E P S G P I H V P K G N G T I L K G D T

841  GAAATAATAA TTGGATTCCCT GATGTTGGAA GTA CTGCTGCTG GCCTCCGAAG ACACCATATA
        G N N N W I P D V G S T W W P P K T P Y

50  901  TTCTCTCTAT CATTACCAAC AGGCCTACTT CTAAGCCAAC AACAGACCT ACACCAAAGC
        I P P I I T N R P T S K P T T R P T P K

961  CAACACCAAT TCCTACTCCA CCACCACCAC CACCCCTGCC AACAGAGCTC AGAACACCTC
55      P T P I P T P P P P P P L P T E L R T P

```

- 8/28 -

1021 TACCACCTAC AACCCCAGAA AGGCCAACCA CCGGA CTGAC AACTATAGCA CCAGCTGCCA
L P P T T P E R P T T G L T T I A P A A

5 1081 GTACACCTCC AGGAGGGATT ACAGTTGACA ACAGGGTACA GACAGACCCT CAGAAACCCA
S T P P G G I T V D N R V Q T D P Q K P

1141 GAGGAGATGT GTTCATTCCA CGGCAACCTT CAAATGACTT GTTTGAAATA TTTGAAATAG
R G D V F I P R Q P S N D L F E I F E I

10 1201 AAAGAGGAGT CAGTGCAGAC GATGAAGCAA AGGATGATCC AGGTGTTCTG GTACACAGTT
E R G V S A D D E A K D D P G V L V H S

1261 GTAATTTTGA CCATGGACTT TGTGGATGGA TCAGGGAGAA AGACAATGAC TTGCACTGGG
C N F D H G L C G W I R E K D N D L H W

15 1321 Aaccaatcag GGACCCAGCA GACAATTGTA GCTGTCTGCA TGCAGGTGGA CTTGACTTG
E P I R D P A D N C S C L H A G G L D L

- 9/28 -

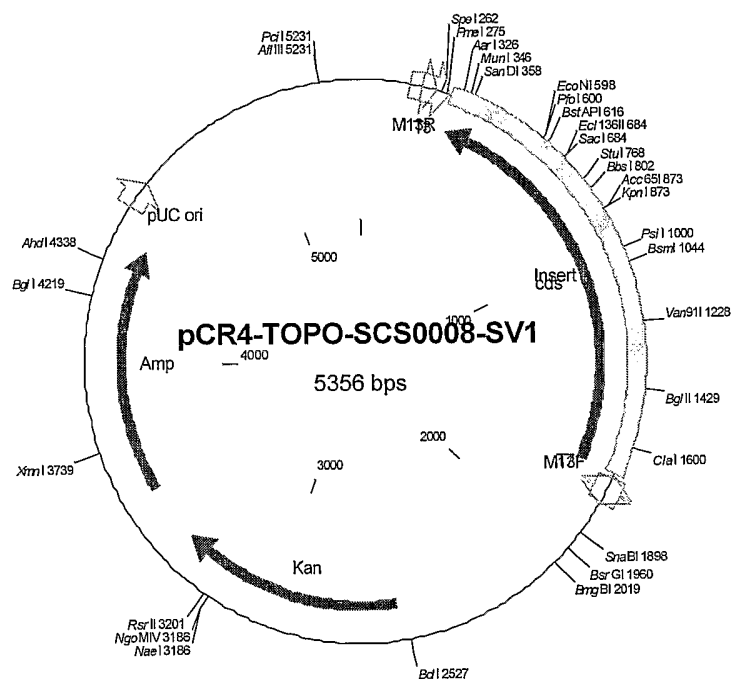
Figure 5**Map of pCR4-TOPO-SCS0008SV1**

Molecule: pCR4-TOPO-SCS0008-SV1, 5356 bps DNA Circular

File Name: 14630[1].cm5

- 5 Description: Ligation of inverted SCS0008 -SV1 PCR product into pCR4-TOPO linear vector*

Type	Start	End	Name	Description
MARKER	205		M13R	M13 rev priming site
MARKER	243		T3	T3 priming site
10 GENE	1693	315	C cds	SCS0008 -SV1 cds missing first 4bp (ATGG)
REGION	1693	295	C Insert	SCS0008 -SV1 F1R1 PCR product
MARKER	1746		C T7	T7 priming site
MARKER	1754		M13F	M13 for priming site
15 GENE	2558	3352	Kan	Kanamycin resistance gene O RF
GENE	3556	4416	Amp	Ampicillin resistance gene ORF
MARKER	4561		pUC ori	pUC origin

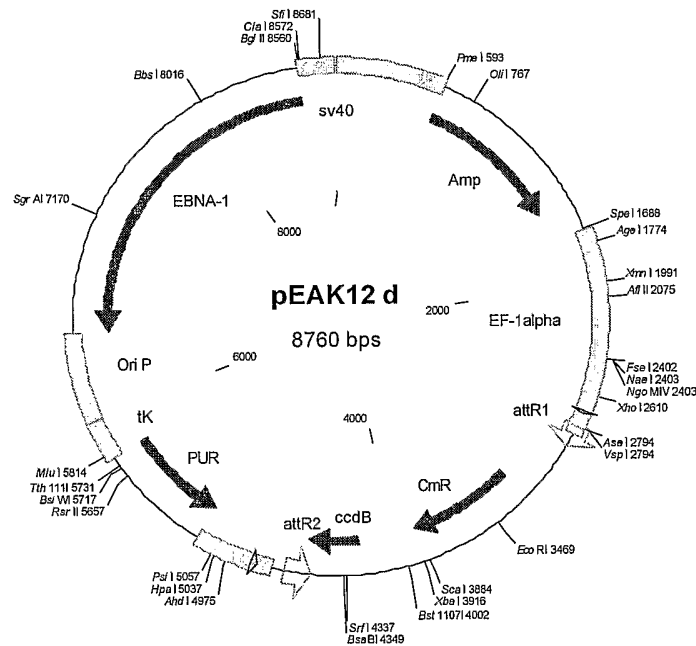


- 10/28 -

Figure 6

Map of expression vector pEAK12d

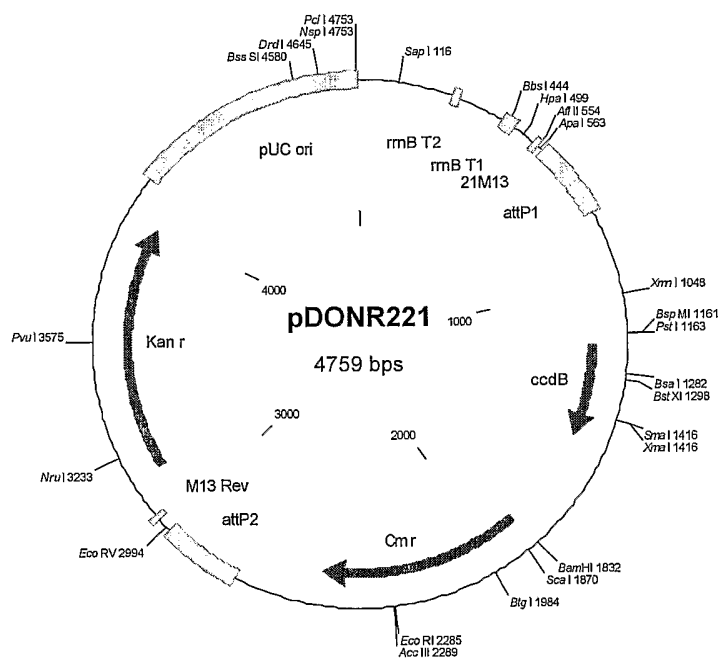
Molecule:	pEAK12 d, 8760 bps DNA Circular			
File Name:	pEAK12DEST.cm5			
Description:	Mammalian cell expression vector (plasmid ID 11345)			
Molecule Features:				
Type	Start	End	Name	Description
REGION	2	595		pmb -ori
GENE	596	1519	Amp	
REGION	1690	2795	EF -1alpha	
REGION	2703	2722		position of pEAK12F primer
REGION	2796	2845		MCS
MARKER	2855		attR1	
GENE	3256	3915	CmR	
GENE	4257	4562	ccdB	
MARKER	4603		C attR2	
REGION	4733	4733		MCS
REGION	4734	5162		poly A/splice
REGION	4819	4848	C	position of pEAK12R primer
GENE	5781	5163	C PUR	PUROMYCIN
REGION	6005	5782	C tK	tK prom oter
REGION	6500	6006	C Ori P	
GENE	8552	6500	C EBNA -1	
REGION	8553	8752	sv40	



- 11/28 -

Figure 7**Map of pDONR 221**

5	Molecule:	pDONR221, 4759 bps DNA Circular			
	File Name:	pDONR221.cm5, dated 03 Jun 2003			
Description:					
	Type	Start	End	Name	Description
10	REGION	295	268	C rrnB T2	transcription termination sequence
	REGION	470	427	C rrnB T1	transcription termination sequence
	REGION	536	553	21M13	M13 Forwa rd primer
	REGION	570	801	attP1	
	GENE	1197	1502	ccdB	
	GENE	1844	2503	Cm r	Chloramphenicol resistance gene
15	REGION	2751	2982	attP2	
	REGION	3040	3023	C M13 Rev	M13 Reverse primer
	GENE	3153	3962	Kan r	
	REGION	4083	4756	pUC ori	



- 12/28 -

Figure 8**Map of Expression vector pDEST12.2**

5	Molecule:	pDEST 12.2, 7278 bps DNA Circular		
	File Name:	pDEST12 -2.cm5		
10	Description:	Eukaryotic expression vector		
	Type	Start	End	Name Description
15	REGION	15	608	CMV CMV promoter
	MARKER	648		M13R M13R primer
20	REGION	687	706	SP6 SP6 promoter
	REGION	730	854	attR1
25	GENE	963	1622	Cm
	GENE	1964	2269	ccdB
30	REGION	2310	2434	attR2
	GENE	2484	2464 C	T7 T7 promoter
35	MARKER	2512		C 21M13 21M13 primer
	REGION	2784	3050	pA SV40 polyadenylation signal
40	REGION	3176	3631	f1 f1 intergenic region
	REGION	3791	4099	P SV40 SV40 ori & early promoter
45	GENE	4158	4952	Neo
	REGION	5016	5064	pA synthetic poly adenylation signal
50	GENE	5475	6335	Amp
	REGION	6480	7153	ori pUC ori

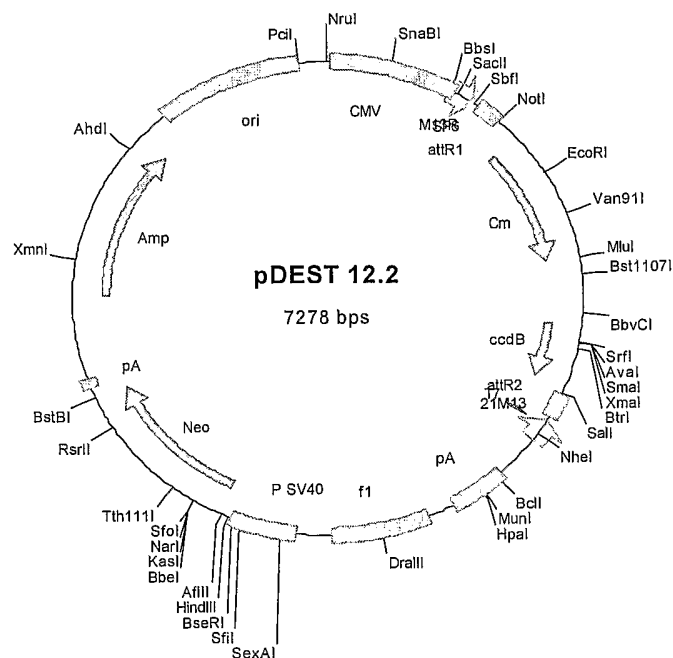


Figure 9
Map of pENTR-SCS0008SV1-6HIS

5	Molecule:	pENTR_SCS0008 -V1-6HIS-V1, 3957 bps DNA Circular			
	File Name:	14877[1].cm5			
10	Description:	Ligation of SCS0008 -V1-6HIS into pdonr221			
	Type	Start	End	Name	Description
	MARKER	536			21M13 Forward primer
	REGION	568	670	attL1	
	GENE	677	2077	SCS0008 -SV1-6HIS	
	REGION	2081	2182	attL2	
	MARKER	2 238		C	M13 Reverse primer
	GENE	2351	3160	Kan	Kanamycin resistance gene
15	REGION	3281	3954	pUC	pUC origin

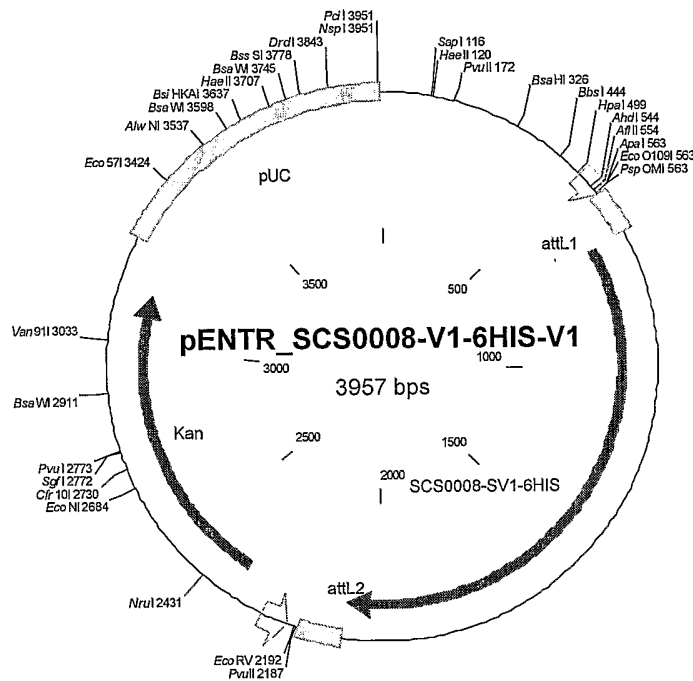
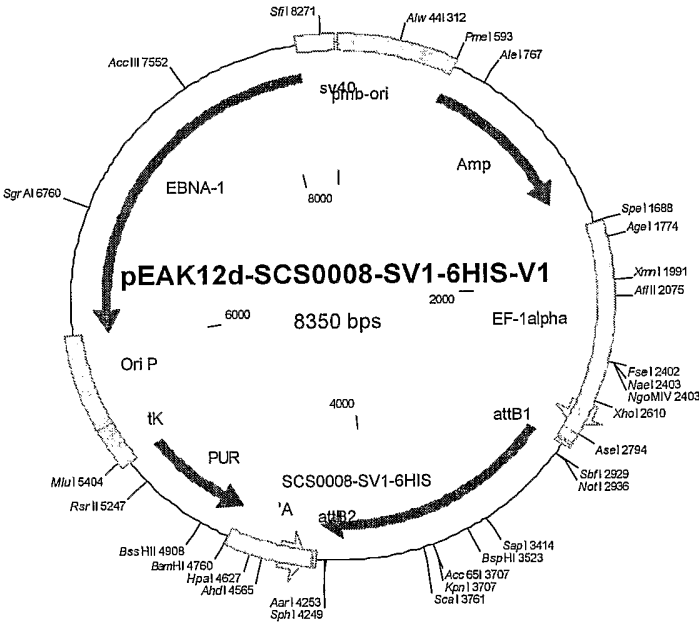


Figure 10

Map of pEAK12d-SCS0008SV1-6HIS

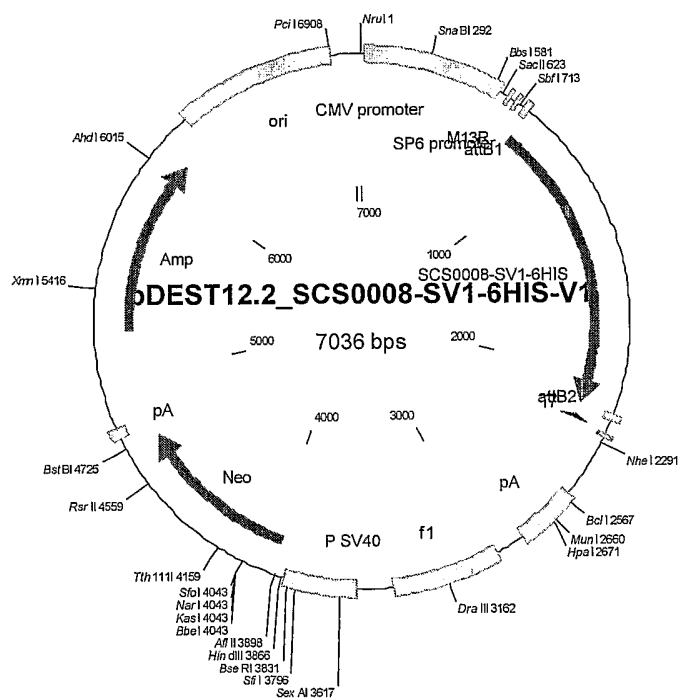
	Molecule:	pEAK12d -SCS0008-SV1-6HIS-V1, 8350 bps DNA Circular			
	File Name:	14883[1].cm5			
5	Description:	pEAK12 DES with two recombination sites attR1 and attR2 between which the cDNA is inserted			
	Type	Start	End	Name	Description
	REGION	2	595	pmb -ori	
	GENE	596	1519	Amp	
10	REGION	1690	2795	EF -1alpha	
	MARKER	2703			pEAK12 F primer
	REGION	2796	2845		MCS''
	REGION	2855	2874	attB1	
	GENE	2888	4291	SCS0008 -SV1-6HIS	
15	REGION	4296	4317	attB2	
	REGION	4323	4323		'MCS
	REGION	4324	4752	'A	poly A/splice
	MARKER	4419		C	pEAK12R primer
	GENE	5371	4753	C PUR	PUROMYCIN
20	REGION	5595	5372	C tK	tK promoter
	REGION	6090	5596	C Ori P	
	GENE	8142	6090	C EBNA -1	
	REGION	8143	8342	sv40	



- 15/28 -

Figure 11**Map of pDEST12.2-SCS0008SV1-6HIS**

5	Molecule:	pDEST12.2_SCS0008-SV1-6HIS-V1, 7036 bps DNA Circular			
	File Name:	14887[1].cm5			
10	Description:	Ligation of SCS0008-SV1-6HIS-V1 into pDEST 12.2			
	Type	Start	End	Name	Description
15	REGION	15	608	CMV promoter	
	REGION	648	665	M13R	Forward primer
20	REGION	687	706	SP6 promoter	
	REGION	730	756	attB1	
	GENE	763	2163	SCS0008-SV1-6HIS	
	REGION	2192	2167 C	attB2	
	GENE	2242	2222 C	T7	T7 promoter
	REGION	2270	2253 C	21M13 reverse primer	
	REGION	2542	2808	pA	SV40 polyadenylation signal
	REGION	2934	3389	f1	f1 intergenic region
	REGION	3549	3857	P SV40	SV40 ori & early promoter
	GENE	3916	4710	Neo	
	REGION	4774	4822	pA	s synthetic poly adenylation signal
	GENE	5233	6093	Amp	
	REGION	6238	6911	ori	pUC ori



- 16/28 -

Figure 12

Nucleotide sequence of SCS0008 prediction with translation

5	2	GCGGGAGGGG GCTCCGGGCG CCGCGCAGCA GACCTGCTCC GGCCGCGCGC CTCGCCGCTG
	62	TCCTCCGGGA GCGGCAGCAG TAGCCCGGGC GGCAGGGGCT GGGGGTTCTT CGAGACTCTC
	122	AGAGGGGCGC CTCCCATCGG CGCCCACCAC CCCAACCTGT TCCTCGCGCG CCACTGCGCT
10	181	GCGCCCCAGG ACCCGCTGCC CAACATGGAT TTTCTCCTGG CGCTGGTGCT GGTATCCTCG
		SCS0008-CP1 → SCS0008-CP1nest
		M D F L L A L V L V S S
15	241	CTCTACCTGC AGGCGGCCGC CGAGTTCGAC GGGAGGTGGC CCAGGCAAAT AGTGTCATCG
		L Y L Q A A A E F D G R W P R Q I V S S
	301	ATTGGCCTAT GTCGTTATGG TGGGAGGATT GACTGCTGCT GGGGCTGGGC TCGCCAG TCT
		I G L C R Y G G R I D C C W G W A R Q S
20	361	TGGGGACAGT GTCAGCCTGT GTGCCAACCA CGATGCAAAC ATGGTGAATG TATCGGGCCA
		W G Q C Q P V C Q P R C K H G E C I G P
	421	AACAAGTGCA AGTGTATCC TG GTTATGCT GGAAAAACCT GTAATCAAGA TCTAAATGAG
		N K C K C H P G Y A G K T C N Q D L N E
25	481	TGTGGCCTGA AGCCCCGGCC CTGTAAGCAC AGGTGCATGA ACACTTACGG CAGCTACAAG
		C G L K P R P C K H R C M N T Y G S Y K
	541	TGCTACTGTC TCAACGGATA TATGCTCATG CCGGATGGTT CCTGCTCAAG TGCCCTGACC
30		C Y C L N G Y M L M P D G S C S S A L T
	601	TGCTCCATGG CAAACTGTCA GTATGGCTGT GATGTTGTTA AAGGACAAAT ACGGTGCCAG
		C S M A N C Q Y G C D V V K G Q I R C Q
35	661	TGCCCATCCC CTGGCCTGCA GCTGGCTCCT GATGGGAGGA CCTGTGTAGA TGTGATGAA
		C P S P G L Q L A P D G R T C V D V D E
	721	TGTGCTACAG GAAGAGCCTC CTGCCCTAGA TTTAGGCAAT GTGTCAACAC TT TTGGGAGC
40		C A T G R A S C P R F R Q C V N T F G S
	781	TACATCTGCA AGTGTATAA AGGCTTCGAT CTCATGTATA TTGGAGGCAA ATATCAATGT
		Y I C K C H K G F D L M Y I G G K Y Q C
	841	CATGACATAG ACGAATGC TC ACTTGGTCAG TATCAGTGCA GCAGCTTTGC TCGATGTTAT
45		H D I D E C S L G Q Y Q C S S F A R C Y
	901	AACGTACGTG GGTCTACAA GTGCAAATGT AAAGAAGGAT ACCAGGGTGA TGGACTGACT
		N V R G S Y K C K C K E G Y Q G D G L T
50	961	TGTGTGCCTA CTTCTAAGCC AACAACAAGA CCTACACCAA AGCCAACACC AATTCCTACT
		C V P T S K P T T R P T P K P T P I P T

- 17/28 -

	1021	CCACCACCAC	CACCACCCCT	GCCAACAGAG	CTCAGAACAC	CTCTACCACC	TACAACCCCA
		P P P	P P P	L P T E	L R T	P L P	P T T P
5	1081	GAAAGGCCAA	CCACCGGACT	GACAACTATA	GCACCAGCTG	CCAGTACACC	TCCAGGAGGG
		E R P	T T G	L T T I	A P A	A S T	P P G G
	1141	ATTACAGTTG	ACAACAGGGT	ACAGACAGAC	CCTCAGAAAC	CCAGAGGA GA	TGTGTTTCATT
		I T V	D N R	V Q T D	P Q K	P R G	D V F I
10	1201	CCACGGCAAC	CTTCAAATGA	CTTGTTTGAA	ATATTTGAAA	TAGAAAGAGG	AGTCAGTGC
		P R Q	P S N	D L F E	I F E	I E R	G V S A
	1261	GACGATGAAG	CAAAGGATGA	TCCAGGTGTT	CTGGTACACA	GTGTGAATTT	TGACCATGGA
		D D E	A K D	D P G V	L V H	S C N	F D H G
15	1321	CTTTGTGGAT	GGATCAGGGA	GAAAGACAAT	GACTTGC	GGGAACCAAT	CAGGGACCCA
		L C G	W I R	E K D N	D L H	W E P	I R D P
20	1381	GCAGACAATT	GTAGCTGTCT	GCATGCAGGT	GGACTTGACT	TCTAGCAACT	GGAGACAAGG
		A D N	C S C	L H A G	G L D	L	SCS0008-CP2 ^{nest}
	1441	TGTAGTTCTC	CATGTGTCTG	AGATTCCAGA	ACTAGACTCA	ATTGTGGTTC	CAGATCAAAT
		◀					◀
	1501	CAGCCTGCAG	AGAATGGCTG	TTCCCTGGTG	CACATGGAGT	TGAGCGGAGC	AGCACACTTG
		◀					
25	1561	GGCACAGGGG	TTGGCCGTCA	GCTCACTGTC	CTGCTGCCAC	CCTCCACCAC	CTCAGAGGCT
	1621	GCTGTAACCT	TGAGAAAATA	ACCTTTGAAC	CTCAGTCTCT	ACAATGCTTA	AATAAAGAGA
	1681	ACATTCAATT	AGAGAAAATA				

Position and sense of PCR primers →

- 18/28 -

Figure 13

Nucleotide sequence with translation of SCS0008 product cloned using primers SCS0008-CP1nest and SCS0008-CP2nest.

```

5
      1  ATTTGCTCCT GGCCTGGTG CTGGTATCCT CGCTCTACCT GCAGGCGGCC GCCGAGTTCG
          L L L A L V L V S S L Y L Q A A A E F
10
     61  ACGGGAGGTG GCCCAGGCAA ATAGTGTTCAT CGATTGGCCT ATGTCGTTAT GGTGGGAGGA
          D G R W P R Q I V S S I G L C R Y G G R
121  TTGACTGCTG CTGGGGCTGG GCTCGCCAGT CTTGGGGACA GTGTCAGCCT GTGTGCCAAC
          I D C C W G W A R Q S W G Q C Q P V C Q
15
     181  CACGATGCAA ACATGGTGAA TGTATCGGGC CAAACAAGTG CAAGTGTTCAT CCTGGTTATG
          P R C K H G E C I G P N K C K C H P G Y
20
     241  CTGGAAAAAC CTGTAATCAA GATCTAAATG AGTGTGGCCT GAAGCCCCGG CCCTGTAAGC
          A G K T C N Q D L N E C G L K P R P C K
301  ACAGGTGCAT GAACACTTAC GGCAGCTACA AGTGTCTACTG TCTCAACGGA TATATGCTCA
          H R C M N T Y G S Y K C Y C L N G Y M L
25
     361  TGCCGGATGG TTCCTGCTCA AGTGGCCTGA CCTGCTCCAT GGCAAACGTG CAGTATGGCT
          M P D G S C S S A L T C S M A N C Q Y G
421  GTGATGTTGT TAAAGGACAA ATACGGTGCC AGTGGCCATC CCCTGGCCTG CACCTGGCTC
          C D V V K G Q I R C Q C P S P G L H L A
30
     481  CTGATGGGAG GACCTGTGTA GATGTTGATG AATGTGCTAC AGGAAGAGCC TCCTGCCCTA
          P D G R T C V D V D E C A T G R A S C P
541  GATTTAGGCA ATGTGTCAAC ACTTTTGGGA GCTACATCTG CAAGTGTTCAT AAAGGCTTCG
          R F R Q C V N T F G S Y I C K C H K G F
35
     601  ATCTCATGTA TATTGGAGGC AAATATCAAT GTCATGACAT AGACGAATGC TCACTTGGTC
          D L M Y I G G K Y Q C H D I D E C S L G
40
     661  AGTATCAGTG CAGCAGCTTT GCTCGATGTT ATAACGTACG TG GGTCCCTAC AAGTGCAAAT
          Q Y Q C S S F A R C Y N V R G S Y K C K
721  GTAAAGAAGG ATACCAGGGT GATGGACTGA CTTGTGTGTA TATCCCAAAA GTTATGATTG
          C K E G Y Q G D G L T C V Y I P K V M I
45
     781  AACCTTCAGG TCCAATTCAT GTACCAAAGG GAAATGGTAC CATTTTAAAG GGTGACACAG
          E P S G P I H V P K G N G T I L K G D T
841  GAAATAATAA TTGGATTCCCT GATGTTGGAA GTACTTGGTG GCCTCCGAAG ACACCATATA
          G N N N W I P D V G S T W W P P K T P Y
50
     901  TTCCTCCTAT CATTACCAAC AGGCCTACTT CTAAGCCAAC AACAAGACCT ACACCAAAGC
          I P P I I T N R P T S K P T T R P T P K
55
     961  CAACACCAAT TCCTACTCCA CCACCACCAC CACCCCTGCC AACAGAGCTC AGAACACCTC
          P T P I P T P P P P P P L P T E L R T P

```

- 19/28 -

1021 TACCACCTAC AACCCAGAA AGGCCAACCA CCGGACTGAC AACTATAGCA CCAGCTGCCA
L P P T T P E R P T T G L T T I A P A A

5 1081 GTACACCTCC AGGAGGGATT ACAGTTGACA ACAGGGTACA GACA GACCCT CAGAAACCCA
S T P P G G I T V D N R V Q T D P Q K P

1141 GAGGAGATGT GTTCAGTGTT CTGGTACACA GTTGTAATTT TGACCATGGA CTTGTGGAT
R G D V F S V L V H S C N F D H G L C G

10 1201 GGATCAGGGA GAAAGACAAT GACTTGCACT GGGAACCAAT CAGGGACCCA GCAGACAATT
W I R E K D N D L H W E P I R D P A D N

1261 GTAGCTGTCT GCATGCAGGT GGACTTGACT TG
15 C S C L H A G G L D L

- 20/28 -

Figure 14
Map of pCR4-TOPO-SCS0008-SV2

5	Molecule:	pCR4 -TOPO-SCS0008-SV2, 5269 bps DNA Circular			
	File Name:	14631[1].cm5			
	Description:	Ligation of inverted SCS0008 -SV2 PCR product into pCR4 -TOPO linear vector*			
	Type	Start	End	Name	Description
	MARKER	205		M13R	M13 rev priming site
10	MARKER	243		T3	T3 priming site
	GENE	1606	315	C cds	SCS0008 -SV2 cds missing first 4bp (ATGG)
	REGION	1606	295	C Insert	SCS0008 -SV2 F1R1 PCR product
	MARKER	1659		C T7	T7 priming site
	MARKER	1667		M13F	M13 for priming site
15	GENE	2471	3265	Kan	Kanamycin resistance gene ORF
	GENE	3469	4329	Amp	Ampicillin resistance gene ORF
	MARKER	4474		pUC ori	pUC origin

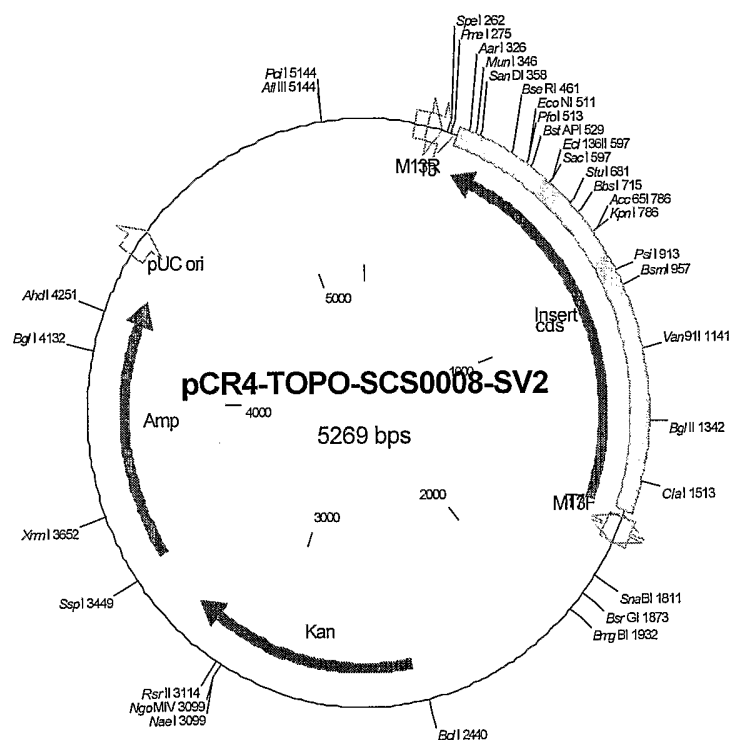
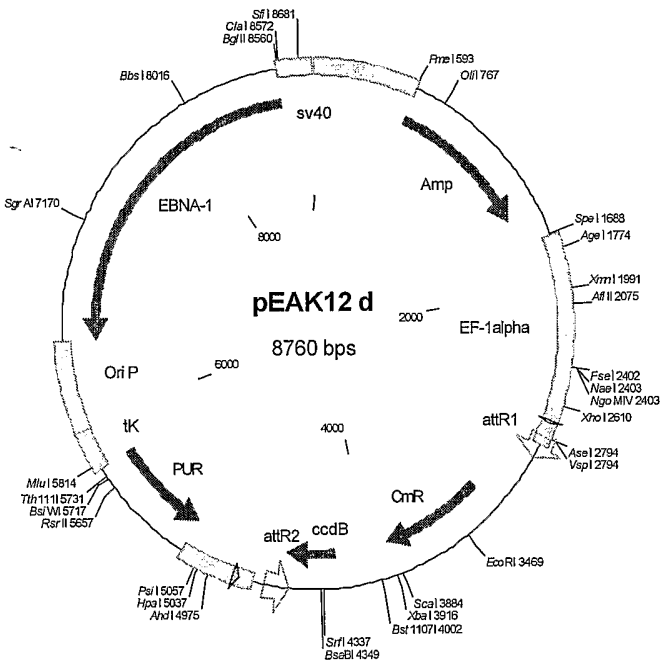


Figure 15

Map of expression vector pEAK12d

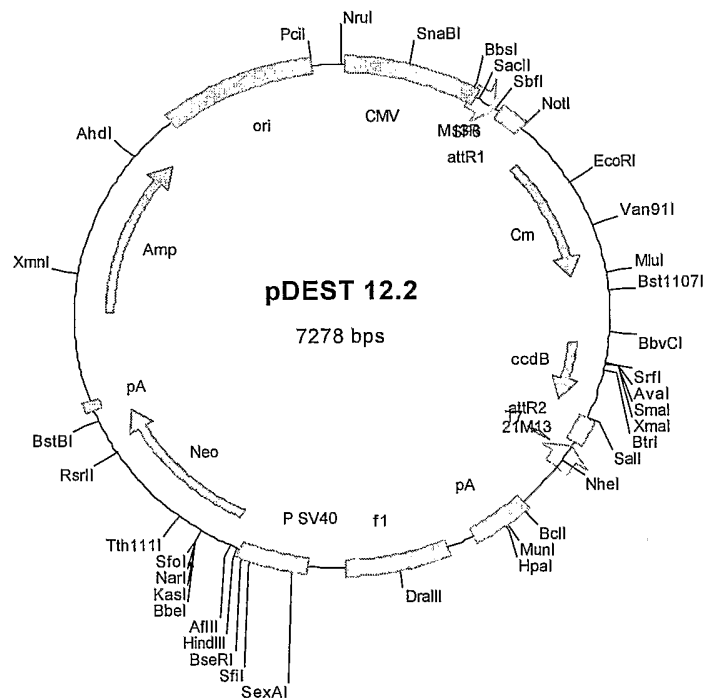
Molecule:	pEAK12 d, 8760 bps DNA Circular			
File Name:	pEAK12DEST.cm5			
Description:	Mammalian cell expression vector (plasmid ID 11345)			
Molecule Features:				
Type	Start	End	Name	Description
REGION	2	595		pmb -ori
GENE	596	1519	Amp	
REGION	1690	2795	EF -1alpha	
REGION	2703	2722		position of pEAK12F primer
REGION	2796	2845		M CS
MARKER	2855		attR1	
GENE	3256	3915	CmR	
GENE	4257	4562	ccdB	
MARKER	4603		C attR2	
REGION	4733	4733		MCS
REGION	4734	5162		poly A/splice
REGION	4819	4848	C	position of pEAK12R primer
GENE	5781	5163	C PUR	PUROMYCIN
REGION	6005	5782	C tK	tK promoter
REGION	6500	6006	C Ori P	
GENE	8552	6500	C EBNA -1	
REGION	8553	8752	sv40	



- 22/28 -

Figure 16**Map of Expression vector pDEST12.2**

Molecule:		pDEST 12.2, 7278 bps DNA Circular			
File Name:		pDEST12-2.cm5			
5	Description:		Eukaryotic expression vector		
	Type	Start	End	Name	Description
	REGION	15	608	CMV	CMV promoter
	MARKER	648		M13R	M13R primer
	REGION	687	706	SP6	SP6 promoter
10	REGION	730	854	attR1	
	GENE	963	1622	Cm	
	GENE	1964	2269	ccdB	
	REGION	2310	2434	attR2	
	GENE	2484	2464	C T7	T7 promoter
15	MARKER	2512		C 21M13	21M13 primer
	REGION	2784	3050	pA	SV40 polyadenylation signal
	REGION	3176	3631	f1	f1 inte rgenic region
	REGION	3791	4099	P SV40	SV40 ori & early promoter
	GENE	4158	4952	Neo	
20	REGION	5016	5064	pA	synthetic poly adenylation signal.
	GENE	5475	6335	Amp	
	REGION	6480	7153	ori	pUC ori



- 23/28 -

Figure 17**Map of pDONR 221**

Molecule: pDONR221, 4759 bps DNA Circular
 File Name: pDONR221.cm5, dated 03 Jun 2003

5 Description:

Type	Start	End	Name	Description
REGION	295	268	C rrnB T2	transcription termination sequence
REGION	470	427	C rrnB T1	transcription termination sequence
REGION	536	553	21M13	M13 Forward primer
10 REGION	570	801	attP1	
GENE	1197	1502	ccdB	
GENE	1844	2503	Cm r	Chloramphenicol resistance gene
REGION	2751	2982	attP2	
REGION	3040	3023	C M13 Rev	M13 Reverse primer
15 GENE	3153	3962	Kan r	
REGION	4083	4756	pUC ori	

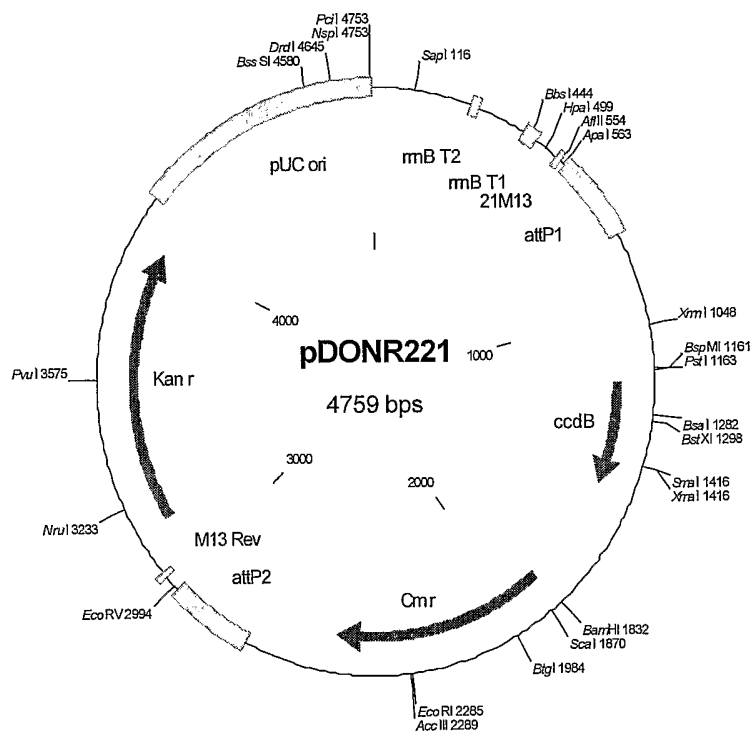


Figure 18

Map of pENTR-SCS0008SV2-6HIS

		Molecule: pENTR_SCS0008 -SV2-6HIS-V1, 3870 bps DNA Circular			
		File Name: 14878[1].cm5			
5	Description:	Ligation of SCS0008 -SV2-6HIS into pdonr221			
	Type	Start	End	Name	Description
	MARKER	536			21M13 Forward primer
	REGION	568	670	attL1	
	GENE	677	1990	SCS0008 -SV2-6HIS	
10	REGION	1994	2095	attL2	
	MARKER	2151		C	M13 Reverse primer
	GENE	2264	3073	Kan	Kanamycin resistance gene
	REGION	3194	3867	pUC	pUC origin

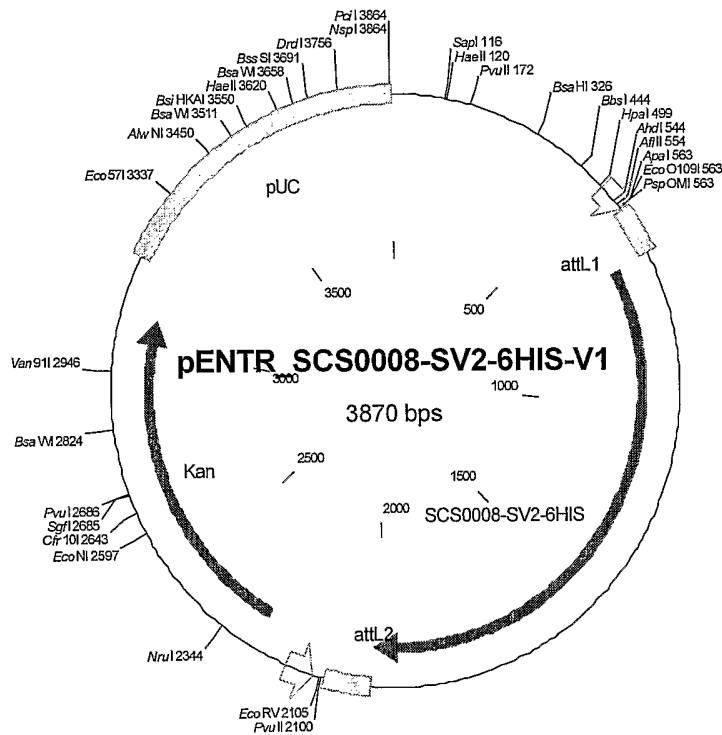
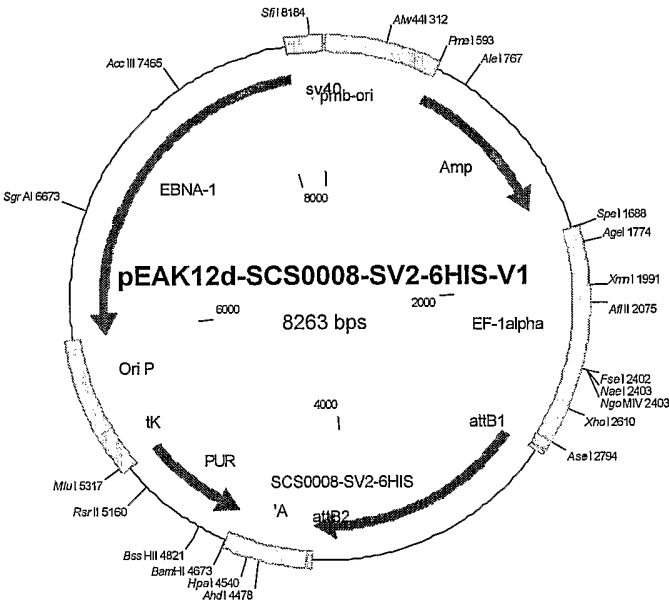


Figure 19

Map of pEAK12d-SCS0008SV2-6HIS

	Molecule:	pEAK12d-SCS0008-SV2-6HIS-V1, 8263 bps DNA Circular			
	File Name:	14884[1].cm5			
5	Description:	pEAK12 DES with two recombination sites a ttR1 and attR2 between which the cDNA is inserted			
	Type	Start	End	Name	Description
	REGION	2	595	pmb-ori	
	GENE	596	1519	Amp	
10	REGION	1690	2795	EF-1alpha	
	REGION	2796	2845		M CS''
	REGION	2855	2874	attB1	
	GENE	2888	4204	SCS0008 -SV2-6HIS	
	REGION	4209	4230	attB2	
15	REGION	4236	4236		'MCS
	REGION	4237	4665	'A	poly A/splice
	GENE	5284	4666	C PUR	PUROMYCIN
	REGION	5508	5285	C tK	tK promoter
	REGION	6003	5509	C Ori P	
20	GENE	8055	6003	C EBNA -1	
	REGION	8056	8255	sv40	



- 26/28 -

Figure 20**Map of pDEST12.2-SCS0008SV2-6HIS**

Molecule: pDEST12.2_SCS0008-SV2-6HIS-V1, 6949 bps DNA Circular
 File Name: 14888[1].cm5

5	Description: Ligation of SCS0008 -SV2-6HIS-V1 into pDEST 12.2				
	Type	Start	End	Name	Description
10	REGION	15	608	CMV promoter	
	REGION	648	665	M13R	Forward primer
	REGION	687	706	SP6 promoter	
15	REGION	730	756	attB1	
	GENE	763	2076	SCS0008 -SV2-6HIS	
	REGION	2105	2080	attB2	
20	GENE	2155	2135	T7	T7 promoter
	REGION	2183	2166	C	21M13 reverse primer
	REGION	2455	2721	pA	SV40 polyadenylation signal
	REGION	2847	3302	f1	f1 intergenic region
	REGION	3462	3770	P SV40	SV40 ori & early promoter
	GENE	3829	4623	Neo	
	REGION	4687	4735	pA	synthetic poly adenylation signal
	GENE	5146	6006	Amp	
	REGION	6151	6824	ori	pUC ori

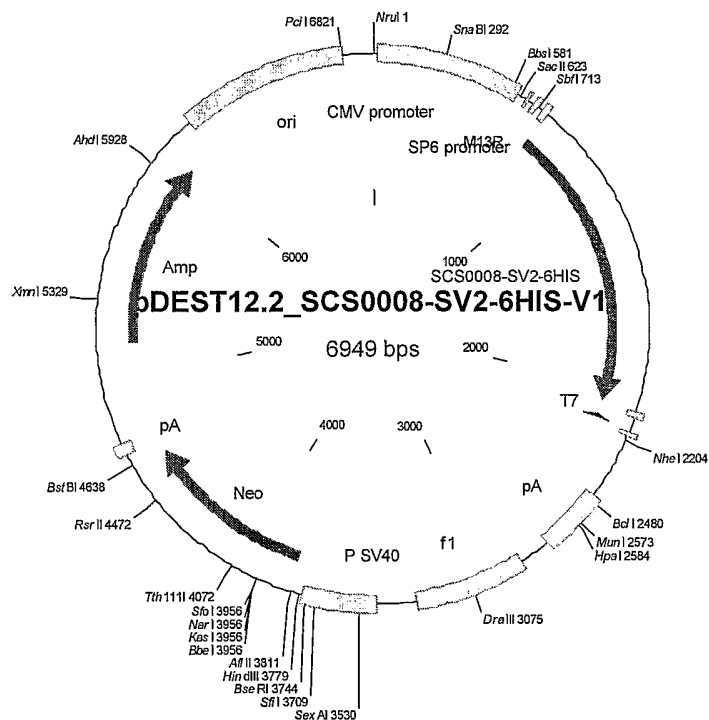
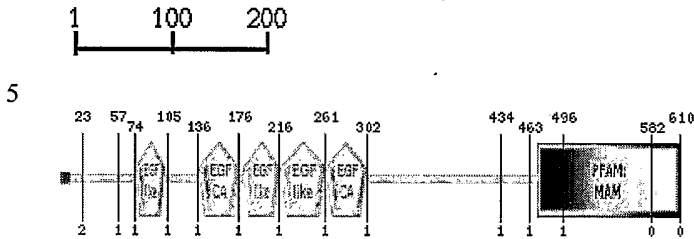
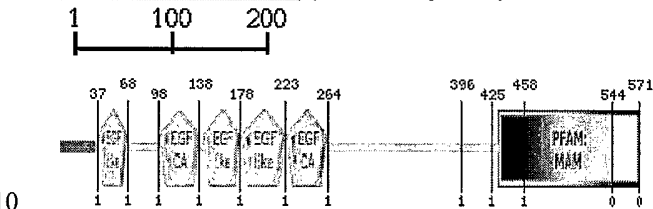


Figure 21

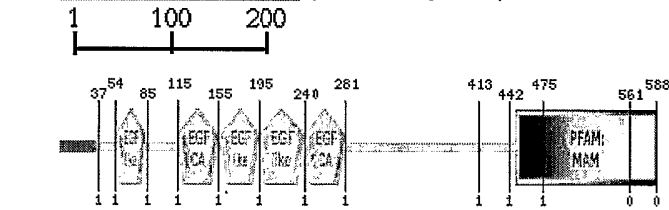
Nephronectin (mus musculus)



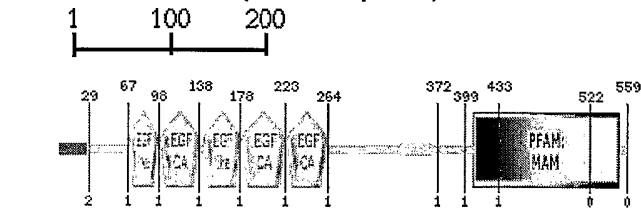
ENSP00000302557 (homo sapiens)



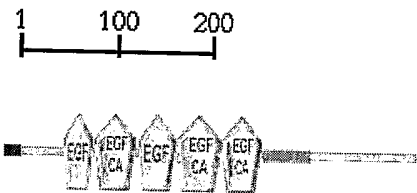
ENSP00000307392 (homo sapiens)



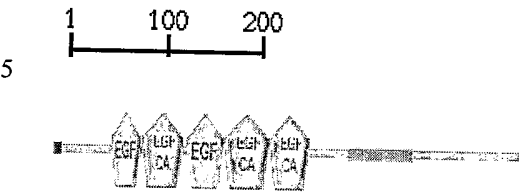
ENSP00000315342, EPIDERMAL GROWTH FACTOR-LIKE PROTEIN 6 PRECURSOR (homo sapiens)



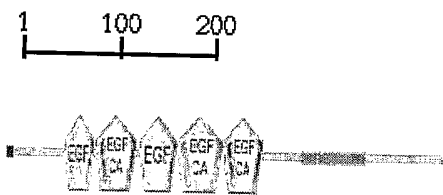
Domains within the query sequence SCS0008 of 406 residues



Domains within the query sequence SCS0008-sv1 of 459 residues



Domains within the query sequence SCS0008-sv2 of 430 residues



10 Transmembrane segments as predicted by the *TMHMM2* program () , coiled coil regions determined by the *Coils2* program () and Segments of low compositional complexity, determined by the *SEG* program () , signal peptides determined by the *Sigcleave* program () , GPI anchors are indicated by () . Regions containing repeats detected by *Prospero*, but not covered by domains are indicated by

15 

- 1 -

SEQUENCE LISTING

<110> Applied Research Systems ARS Holding N.V.

5 <120> NOVEL FIBRILLIN-LIKE POLYPEPTIDES

<130> PCT/811

<150> 60/436,835

10 <151> 2002-12-27

<160> 18

<170> PatentIn version 3.2

15 <210> 1

<211> 1700

<212> DNA

<213> homo sapiens

20

<220>

<221> CDS

<222> (205)..(1425)

25

<400> 1

gcgggagggg gctccggggc ccgcgcagca gacctgctcc ggccgcgcgc ctccgctg 60

tcctccggga gcggcagcag tagcccgggc ggcgagggtt gggggttcct cgagactctc 120

30 agagggggcg ctcccatcgg cgcccaccac cccaacctgt tcctcgcgcg ccaactgcgt 180

gcgccccagg acccgctgcc caac atg gat ttt ctc ctg gcg ctg gtg ctg 231

35 Met Asp Phe Leu Leu Ala Leu Val Leu

1 5

gta tcc tcg ctc tac ctg cag gcg gcc gcc gag ttc gac ggg agg tgg 279

Val Ser Ser Leu Tyr Leu Gln Ala Ala Ala Glu Phe Asp Gly Arg Trp

10 15 20 25

40 ccc agg caa ata gtg tca tcg att ggc cta tgt cgt tat ggt ggg agg 327

Pro Arg Gln Ile Val Ser Ser Ile Gly Leu Cys Arg Tyr Gly Gly Arg

30 35 40

45 att gac tgc tgc tgg ggc tgg gct cgc cag tct tgg gga cag tgt cag 375

Ile Asp Cys Cys Trp Gly Trp Ala Arg Gln Ser Trp Gly Gln Cys Gln

45 50 55

50 cct gtg tgc caa cca cga tgc aaa cat ggt gaa tgt atc ggg cca aac 423

Pro Val Cys Gln Pro Arg Cys Lys His Gly Glu Cys Ile Gly Pro Asn

60 65 70

aag tgc aag tgt cat cct ggt tat gct gga aaa acc tgt aat caa gat 471

55 Lys Cys Lys Cys His Pro Gly Tyr Ala Gly Lys Thr Cys Asn Gln Asp

75 80 85

- 2 -

	cta aat gag tgt ggc ctg aag ccc cgg ccc tgt aag cac agg tgc atg	519
	Leu Asn Glu Cys Gly Leu Lys Pro Arg Pro Cys Lys His Arg Cys Met	
	90 95 100 105	
5	aac act tac ggc agc tac aag tgc tac tgt ctc aac gga tat atg ctc	567
	Asn Thr Tyr Gly Ser Tyr Lys Cys Tyr Cys Leu Asn Gly Tyr Met Leu	
	110 115 120	
10	atg ccg gat ggt tcc tgc tca agt gcc ctg acc tgc tcc atg gca aac	615
	Met Pro Asp Gly Ser Cys Ser Ser Ala Leu Thr Cys Ser Met Ala Asn	
	125 130 135	
15	tgt cag tat ggc tgt gat gtt gtt aaa gga caa ata cgg tgc cag tgc	663
	Cys Gln Tyr Gly Cys Asp Val Val Lys Gly Gln Ile Arg Cys Gln Cys	
	140 145 150	
20	cca tcc cct ggc ctg cag ctg gct cct gat ggg agg acc tgt gta gat	711
	Pro Ser Pro Gly Leu Gln Leu Ala Pro Asp Gly Arg Thr Cys Val Asp	
	155 160 165	
25	gtt gat gaa tgt gct aca gga aga gcc tcc tgc cct aga ttt agg caa	759
	Val Asp Glu Cys Ala Thr Gly Arg Ala Ser Cys Pro Arg Phe Arg Gln	
	170 175 180 185	
30	tgt gtc aac act ttt ggg agc tac atc tgc aag tgt cat aaa ggc ttc	807
	Cys Val Asn Thr Phe Gly Ser Tyr Ile Cys Lys Cys His Lys Gly Phe	
	190 195 200	
35	gat ctc atg tat att gga ggc aaa tat caa tgt cat gac ata gac gaa	855
	Asp Leu Met Tyr Ile Gly Gly Lys Tyr Gln Cys His Asp Ile Asp Glu	
	205 210 215	
40	tgc tca ctt ggt cag tat cag tgc agc agc ttt gct cga tgt tat aac	903
	Cys Ser Leu Gly Gln Tyr Gln Cys Ser Ser Phe Ala Arg Cys Tyr Asn	
	220 225 230	
45	gta cgt ggg tcc tac aag tgc aaa tgt aaa gaa gga tac cag ggt gat	951
	Val Arg Gly Ser Tyr Lys Cys Lys Cys Lys Glu Gly Tyr Gln Gly Asp	
	235 240 245	
50	gga ctg act tgt gtg cct act tct aag cca aca aca aga cct aca cca	999
	Gly Leu Thr Cys Val Pro Thr Ser Lys Pro Thr Thr Arg Pro Thr Pro	
	250 255 260 265	
55	aag cca aca cca att cct act cca cca cca cca cca ccc ctg cca aca	1047
	Lys Pro Thr Pro Ile Pro Thr Pro Pro Pro Pro Pro Pro Leu Pro Thr	
	270 275 280	
60	gag ctc aga aca cct cta cca cct aca acc cca gaa agg cca acc acc	1095
	Glu Leu Arg Thr Pro Leu Pro Pro Thr Thr Pro Glu Arg Pro Thr Thr	
	285 290 295	
65	gga ctg aca act ata gca cca gct gcc agt aca cct cca gga ggg att	1143
	Gly Leu Thr Thr Ile Ala Pro Ala Ala Ser Thr Pro Pro Gly Gly Ile	
	300 305 310	
70	aca gtt gac aac agg gta cag aca gac cct cag aaa ccc aga gga gat	1191
	Thr Val Asp Asn Arg Val Gln Thr Asp Pro Gln Lys Pro Arg Gly Asp	

- 3 -

	315	320	325	
	gtg ttc att cca cgg caa cct tca aat gac ttg ttt gaa ata ttt gaa			1239
	Val Phe Ile Pro Arg Gln Pro Ser Asn Asp Leu Phe Glu Ile Phe Glu			
5	330	335	340	345
	ata gaa aga gga gtc agt gca gac gat gaa gca aag gat gat cca ggt			1287
	Ile Glu Arg Gly Val Ser Ala Asp Asp Glu Ala Lys Asp Asp Pro Gly			
		350	355	360
10	gtt ctg gta cac agt tgt aat ttt gac cat gga ctt tgt gga tgg atc			1335
	Val Leu Val His Ser Cys Asn Phe Asp His Gly Leu Cys Gly Trp Ile			
		365	370	375
15	agg gag aaa gac aat gac ttg cac tgg gaa cca atc agg gac cca gca			1383
	Arg Glu Lys Asp Asn Asp Leu His Trp Glu Pro Ile Arg Asp Pro Ala			
		380	385	390
	gac aat tgt agc tgt ctg cat gca ggt gga ctt gac ttg tag			1425
20	Asp Asn Cys Ser Cys Leu His Ala Gly Gly Leu Asp Leu			
		395	400	405
	caactggaga caaggtgtag ttotccatgt gtctgagatt ccagaactag actcaattgt			1485
25	ggttccagat caaatcagcc tgcagagaat ggctgttccc tgggtgcacat ggagttgagc			1545
	ggagcagcac acttgggcac aggggttggc cgtcagctca ctgtcctgct gccaccctcc			1605
	accacctcag aggctgctgt aactttgaga aaataacctt tgaacctcag totctacaat			1665
30	gcttaaataa agagaacatt caattagaga aaata			1700
	<210> 2			
35	<211> 406			
	<212> PRT			
	<213> homo sapiens			
	<400> 2			
40	Met Asp Phe Leu Leu Ala Leu Val Leu Val Ser Ser Leu Tyr Leu Gln			
	1 5 10 15			
45	Ala Ala Ala Glu Phe Asp Gly Arg Trp Pro Arg Gln Ile Val Ser Ser			
	20 25 30			
	Ile Gly Leu Cys Arg Tyr Gly Gly Arg Ile Asp Cys Cys Trp Gly Trp			
50	35 40 45			
	Ala Arg Gln Ser Trp Gly Gln Cys Gln Pro Val Cys Gln Pro Arg Cys			
	50 55 60			
55	Lys His Gly Glu Cys Ile Gly Pro Asn Lys Cys Lys Cys His Pro Gly			
	65 70 75 80			

- 4 -

Tyr Ala Gly Lys Thr Cys Asn Gln Asp Leu Asn Glu Cys Gly Leu Lys
 85 90 95
 5

Pro Arg Pro Cys Lys His Arg Cys Met Asn Thr Tyr Gly Ser Tyr Lys
 100 105 110
 10

Cys Tyr Cys Leu Asn Gly Tyr Met Leu Met Pro Asp Gly Ser Cys Ser
 115 120 125
 15

Ser Ala Leu Thr Cys Ser Met Ala Asn Cys Gln Tyr Gly Cys Asp Val
 130 135 140
 20

Val Lys Gly Gln Ile Arg Cys Gln Cys Pro Ser Pro Gly Leu Gln Leu
 145 150 155 160
 25

Ala Pro Asp Gly Arg Thr Cys Val Asp Val Asp Glu Cys Ala Thr Gly
 165 170 175
 30

Arg Ala Ser Cys Pro Arg Phe Arg Gln Cys Val Asn Thr Phe Gly Ser
 180 185 190
 35

Tyr Ile Cys Lys Cys His Lys Gly Phe Asp Leu Met Tyr Ile Gly Gly
 195 200 205
 40

Lys Tyr Gln Cys His Asp Ile Asp Glu Cys Ser Leu Gly Gln Tyr Gln
 210 215 220
 45

Cys Ser Ser Phe Ala Arg Cys Tyr Asn Val Arg Gly Ser Tyr Lys Cys
 225 230 235 240
 50

Lys Cys Lys Glu Gly Tyr Gln Gly Asp Gly Leu Thr Cys Val Pro Thr
 245 250 255
 55

Ser Lys Pro Thr Thr Arg Pro Thr Pro Lys Pro Thr Pro Ile Pro Thr
 260 265 270
 60

Pro Pro Pro Pro Pro Pro Leu Pro Thr Glu Leu Arg Thr Pro Leu Pro
 275 280 285
 65

Pro Thr Thr Pro Glu Arg Pro Thr Thr Gly Leu Thr Thr Ile Ala Pro
 290 295 300

- 5 -

Ala Ala Ser Thr Pro Pro Gly Gly Ile Thr Val Asp Asn Arg Val Gln
305 310 315 320

5 Thr Asp Pro Gln Lys Pro Arg Gly Asp Val Phe Ile Pro Arg Gln Pro
325 330 335

10 Ser Asn Asp Leu Phe Glu Ile Phe Glu Ile Glu Arg Gly Val Ser Ala
340 345 350

15 Asp Asp Glu Ala Lys Asp Asp Pro Gly Val Leu Val His Ser Cys Asn
355 360 365

Phe Asp His Gly Leu Cys Gly Trp Ile Arg Glu Lys Asp Asn Asp Leu
370 375 380

20 His Trp Glu Pro Ile Arg Asp Pro Ala Asp Asn Cys Ser Cys Leu His
385 390 395 400

25 Ala Gly Gly Leu Asp Leu
405

30 <210> 3
<211> 404
<212> PRT
<213> homo sapiens

35 <400> 3

Phe Leu Leu Ala Leu Val Leu Val Ser Ser Leu Tyr Leu Gln Ala Ala
1 5 10 15

40 Ala Glu Phe Asp Gly Arg Trp Pro Arg Gln Ile Val Ser Ser Ile Gly
20 25 30

45 Leu Cys Arg Tyr Gly Gly Arg Ile Asp Cys Cys Trp Gly Trp Ala Arg
35 40 45

Gln Ser Trp Gly Gln Cys Gln Pro Val Cys Gln Pro Arg Cys Lys His
50 55 60

Gly Glu Cys Ile Gly Pro Asn Lys Cys Lys Cys His Pro Gly Tyr Ala
65 70 75 80

55 Gly Lys Thr Cys Asn Gln Asp Leu Asn Glu Cys Gly Leu Lys Pro Arg
85 90 95

- 6 -

Pro Cys Lys His Arg Cys Met Asn Thr Tyr Gly Ser Tyr Lys Cys Tyr
 100 105 110

5

Cys Leu Asn Gly Tyr Met Leu Met Pro Asp Gly Ser Cys Ser Ser Ala
 115 120 125

10

Leu Thr Cys Ser Met Ala Asn Cys Gln Tyr Gly Cys Asp Val Val Lys
 130 135 140

15

Gly Gln Ile Arg Cys Gln Cys Pro Ser Pro Gly Leu Gln Leu Ala Pro
 145 150 155 160

20

Asp Gly Arg Thr Cys Val Asp Val Asp Glu Cys Ala Thr Gly Arg Ala
 165 170 175

25

Ser Cys Pro Arg Phe Arg Gln Cys Val Asn Thr Phe Gly Ser Tyr Ile
 180 185 190

30

Cys Lys Cys His Lys Gly Phe Asp Leu Met Tyr Ile Gly Gly Lys Tyr
 195 200 205

35

Gln Cys His Asp Ile Asp Glu Cys Ser Leu Gly Gln Tyr Gln Cys Ser
 210 215 220

40

Ser Phe Ala Arg Cys Tyr Asn Val Arg Gly Ser Tyr Lys Cys Lys Cys
 225 230 235 240

45

Lys Glu Gly Tyr Gln Gly Asp Gly Leu Thr Cys Val Pro Thr Ser Lys
 245 250 255

50

Pro Thr Thr Arg Pro Thr Pro Lys Pro Thr Pro Ile Pro Thr Pro Pro
 260 265 270

55

Pro Pro Pro Pro Leu Pro Thr Glu Leu Arg Thr Pro Leu Pro Pro Thr
 275 280 285

Thr Pro Glu Arg Pro Thr Thr Gly Leu Thr Thr Ile Ala Pro Ala Ala
 290 295 300

Ser Thr Pro Pro Gly Gly Ile Thr Val Asp Asn Arg Val Gln Thr Asp
 305 310 315 320

Pro Gln Lys Pro Arg Gly Asp Val Phe Ile Pro Arg Gln Pro Ser Asn

- 7 -

[illegible]

- 8 -

	Cys	Tyr	Cys	Leu	Asn	Gly	Tyr	Met	Leu	Met	Pro	Asp	Gly	Ser	Cys	Ser	
			115					120					125				
5	Ser	Ala	Leu	Thr	Cys	Ser	Met	Ala	Asn	Cys	Gln	Tyr	Gly	Cys	Asp	Val	
		130					135					140					
10	Val	Lys	Gly	Gln	Ile	Arg	Cys	Gln	Cys	Pro	Ser	Pro	Gly	Leu	Gln	Leu	
	145					150					155					160	
15	Ala	Pro	Asp	Gly	Arg	Thr	Cys	Val	Asp	Val	Asp	Glu	Cys	Ala	Thr	Gly	
					165					170					175		
20	Arg	Ala	Ser	Cys	Pro	Arg	Phe	Arg	Gln	Cys	Val	Asn	Thr	Phe	Gly	Ser	
				180					185					190			
25	Tyr	Ile	Cys	Lys	Cys	His	Lys	Gly	Phe	Asp	Leu	Met	Tyr	Ile	Gly	Gly	
			195					200					205				
30	Lys	Tyr	Gln	Cys	His	Asp	Ile	Asp	Glu	Cys	Ser	Leu	Gly	Gln	Tyr	Gln	
		210					215					220					
35	Cys	Ser	Ser	Phe	Ala	Arg	Cys	Tyr	Asn	Val	Arg	Gly	Ser	Tyr	Lys	Cys	
	225					230					235					240	
40	Lys	Cys	Lys	Glu	Gly	Tyr	Gln	Gly	Asp	Gly	Leu	Thr	Cys	Val	Pro	Thr	
				245						250					255		
45	Ser	Lys	Pro	Thr	Thr	Arg	Pro	Thr	Pro	Lys	Pro	Thr	Pro	Ile	Pro	Thr	
				260					265					270			
50	Pro	Pro	Pro	Pro	Pro	Pro	Leu	Pro	Thr	Glu	Leu	Arg	Thr	Pro	Leu	Pro	
			275					280					285				
55	Pro	Thr	Thr	Pro	Glu	Arg	Pro	Thr	Thr	Gly	Leu	Thr	Thr	Ile	Ala	Pro	
		290				295						300					
60	Ala	Ala	Ser	Thr	Pro	Pro	Gly	Gly	Ile	Thr	Val	Asp	Asn	Arg	Val	Gln	
	305					310					315					320	
65	Thr	Asp	Pro	Gln	Lys	Pro	Arg	Gly	Asp	Val	Phe	Ile	Pro	Arg	Gln	Pro	
				325						330					335		
70	Ser	Asn	Asp	Leu	Phe	Glu	Ile	Phe	Glu	Ile	Glu	Arg	Gly	Val	Ser	Ala	
			340					345						350			

- 9 -

Asp Asp Glu Ala Lys Asp Asp Pro Gly Val Leu Val His Ser Cys Asn
 355 360 365
 5
 Phe Asp His Gly Leu Cys Gly Trp Ile Arg Glu Lys Asp Asn Asp Leu
 370 375 380
 10
 His Trp Glu Pro Ile Arg Asp Pro Ala Asp Asn Cys Ser Cys Leu His
 385 390 395 400
 15 Ala Gly Gly Leu Asp Leu His His His His His His
 405 410
 20
 <210> 5
 <211> 1590
 <212> DNA
 <213> homo sapiens
 25
 <220>
 <221> CDS
 <222> (205)..(1590)
 <400> 5
 30 gcgggagggg gctccgggcg ccgcgcagca gacctgctcc ggccgcgcgc ctgcgcgctg 60
 tcctccggga gcggcagcag tagcccgggc ggcgagggct gggggttcct cgagactctc 120
 agagggggcg ctcccatcgg cgcaccac cccaacctgt tcctgcgcgc ccaactgcgt 180
 35 gcgccccagg acccgctgcc caac atg gat ttt ctc ctg gcg ctg gtg ctg 231
 Met Asp Phe Leu Leu Ala Leu Val Leu
 1 5
 40 gta tcc tcg ctc tac ctg cag gcg gcc gcc gag ttc gac ggg agg tgg 279
 Val Ser Ser Leu Tyr Leu Gln Ala Ala Ala Glu Phe Asp Gly Arg Trp
 10 15 20 25
 45 ccc agg caa ata gtg tca tcg att ggc cta tgt cgt tat ggt ggg agg 327
 Pro Arg Gln Ile Val Ser Ser Ile Gly Leu Cys Arg Tyr Gly Gly Arg
 30 35 40
 att gac tgc tgc tgg ggc tgg gct cgc cag tct tgg gga cag tgt cag 375
 50 Ile Asp Cys Cys Trp Gly Trp Ala Arg Gln Ser Trp Gly Gln Cys Gln
 45 50 55
 cct gtg tgc caa cca cga tgc aaa cat ggt gaa tgt atc ggg cca aac 423
 Pro Val Cys Gln Pro Arg Cys Lys His Gly Glu Cys Ile Gly Pro Asn
 60 65 70
 55 aag tgc aag tgt cat cct ggt tat gct gga aaa acc tgt aat caa gat 471
 Lys Cys Lys Cys His Pro Gly Tyr Ala Gly Lys Thr Cys Asn Gln Asp
 75 80 85

- 10 -

	cta aat gag tgt ggc ctg aag ccc cgg ccc tgt aag cac agg tgc atg	519
	Leu Asn Glu Cys Gly Leu Lys Pro Arg Pro Cys Lys His Arg Cys Met	
	90 95 100 105	
5	aac act tac ggc agc tac aag tgc tac tgt ctc aac gga tat atg ctc	567
	Asn Thr Tyr Gly Ser Tyr Lys Cys Tyr Cys Leu Asn Gly Tyr Met Leu	
	110 115 120	
10	atg ccg gat ggt tcc tgc tca agt gcc ctg acc tgc tcc atg gca aac	615
	Met Pro Asp Gly Ser Cys Ser Ser Ala Leu Thr Cys Ser Met Ala Asn	
	125 130 135	
15	tgt cag tat ggc tgt gat gtt gtt aaa gga caa ata cgg tgc cag tgc	663
	Cys Gln Tyr Gly Cys Asp Val Val Lys Gly Gln Ile Arg Cys Gln Cys	
	140 145 150	
20	cca tcc cct ggc ctg cac ctg gct cct gat ggg agg acc tgt gta gat	711
	Pro Ser Pro Gly Leu His Leu Ala Pro Asp Gly Arg Thr Cys Val Asp	
	155 160 165	
25	gtt gat gaa tgt gct aca gga aga gcc tcc tgc cct aga ttt agg caa	759
	Val Asp Glu Cys Ala Thr Gly Arg Ala Ser Cys Pro Arg Phe Arg Gln	
	170 175 180 185	
30	tgt gtc aac act ttt ggg agc tac atc tgc aag tgt cat aaa ggc ttc	807
	Cys Val Asn Thr Phe Gly Ser Tyr Ile Cys Lys Cys His Lys Gly Phe	
	190 195 200	
35	gat ctc atg tat att gga ggc aaa tat caa tgt cat gac ata gac gaa	855
	Asp Leu Met Tyr Ile Gly Gly Lys Tyr Gln Cys His Asp Ile Asp Glu	
	205 210 215	
40	tgc tca ctt ggt cag tat cag tgc agc agc ttt gct cga tgt tat aac	903
	Cys Ser Leu Gly Gln Tyr Gln Cys Ser Ser Phe Ala Arg Cys Tyr Asn	
	220 225 230	
45	gta cgt ggg tcc tac aag tgc aaa tgt aaa gaa gga tac cag ggt gat	951
	Val Arg Gly Ser Tyr Lys Cys Lys Cys Lys Glu Gly Tyr Gln Gly Asp	
	235 240 245	
50	gga ctg act tgt gtg tat atc cca aaa gtt atg att gaa cct tca ggt	999
	Gly Leu Thr Cys Val Tyr Ile Pro Lys Val Met Ile Glu Pro Ser Gly	
	250 255 260 265	
55	cca att cat gta cca aag gga aat ggt acc att tta aag ggt gac aca	1047
	Pro Ile His Val Pro Lys Gly Asn Gly Thr Ile Leu Lys Gly Asp Thr	
	270 275 280	
60	gga aat aat aat tgg att cct gat gtt gga agt act tgg tgg cct ccg	1095
	Gly Asn Asn Asn Trp Ile Pro Asp Val Gly Ser Thr Trp Trp Pro Pro	
	285 290 295	
65	aag aca cca tat att cct cct atc att acc aac agg cct act tct aag	1143
	Lys Thr Pro Tyr Ile Pro Pro Ile Ile Thr Asn Arg Pro Thr Ser Lys	
	300 305 310	
70	cca aca aca aga cct aca cca aag cca aca cca att cct act cca cca	1191

- 11 -

	Pro Thr Thr Arg Pro Thr Pro Lys Pro Thr Pro Ile Pro Thr Pro Pro	
	315 320 325	
5	cca cca cca ccc ctg cca aca gag ctc aga aca cct cta cca cct aca Pro Pro Pro Pro Leu Pro Thr Glu Leu Arg Thr Pro Leu Pro Pro Thr 330 335 340 345	1239
10	acc cca gaa agg cca acc acc gga ctg aca act ata gca cca gct gcc Thr Pro Glu Arg Pro Thr Thr Gly Leu Thr Thr Ile Ala Pro Ala Ala 350 355 360	1287
15	agt aca cct cca gga ggg att aca gtt gac aac agg gta cag aca gac Ser Thr Pro Pro Gly Gly Ile Thr Val Asp Asn Arg Val Gln Thr Asp 365 370 375	1335
20	cct cag aaa ccc aga gga gat gtg ttc att cca cgg caa cct tca aat Pro Gln Lys Pro Arg Gly Asp Val Phe Ile Pro Arg Gln Pro Ser Asn 380 385 390	1383
25	gac ttg ttt gaa ata ttt gaa ata gaa aga gga gtc agt gca gac gat Asp Leu Phe Glu Ile Phe Glu Ile Glu Arg Gly Val Ser Ala Asp Asp 395 400 405	1431
30	gaa gca aag gat gat cca ggt gtt ctg gta cac agt tgt aat ttt gac Glu Ala Lys Asp Asp Pro Gly Val Leu Val His Ser Cys Asn Phe Asp 410 415 420 425	1479
35	cat gga ctt tgt gga tgg atc agg gag aaa gac aat gac ttg cac tgg His Gly Leu Cys Gly Trp Ile Arg Glu Lys Asp Asn Asp Leu His Trp 430 435 440	1527
40	gaa cca atc agg gac cca gca gac aat tgt agc tgt ctg cat gca ggt Glu Pro Ile Arg Asp Pro Ala Asp Asn Cys Ser Cys Leu His Ala Gly 445 450 455	1575
45	gga ctt gac ttg tag Gly Leu Asp Leu 460	1590
50	<210> 6 <211> 461 <212> PRT <213> homo sapiens <400> 6 Met Asp Phe Leu Leu Ala Leu Val Leu Val Ser Ser Leu Tyr Leu Gln 1 5 10 15	
55	Ala Ala Ala Glu Phe Asp Gly Arg Trp Pro Arg Gln Ile Val Ser Ser 20 25 30	
	Ile Gly Leu Cys Arg Tyr Gly Gly Arg Ile Asp Cys Cys Trp Gly Trp 35 40 45	

- 12 -

Ala Arg Gln Ser Trp Gly Gln Cys Gln Pro Val Cys Gln Pro Arg Cys
 50 55 60

5 Lys His Gly Glu Cys Ile Gly Pro Asn Lys Cys Lys Cys His Pro Gly
 65 70 75 80

10 Tyr Ala Gly Lys Thr Cys Asn Gln Asp Leu Asn Glu Cys Gly Leu Lys
 85 90 95

15 Pro Arg Pro Cys Lys His Arg Cys Met Asn Thr Tyr Gly Ser Tyr Lys
 100 105 110

20 Cys Tyr Cys Leu Asn Gly Tyr Met Leu Met Pro Asp Gly Ser Cys Ser
 115 120 125

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

25 Val Lys Gly Gln Ile Arg Cys Gln Cys Pro Ser Pro Gly Leu His Leu
 145 150 155 160

30 Ala Pro Asp Gly Arg Thr Cys Val Asp Val Asp Glu Cys Ala Thr Gly
 165 170 175

35 Arg Ala Ser Cys Pro Arg Phe Arg Gln Cys Val Asn Thr Phe Gly Ser
 180 185 190

40 Tyr Ile Cys Lys Cys His Lys Gly Phe Asp Leu Met Tyr Ile Gly Gly
 195 200 205

Lys Tyr Gln Cys His Asp Ile Asp Glu Cys Ser Leu Gly Gln Tyr Gln
 210 215 220

45 Cys Ser Ser Phe Ala Arg Cys Tyr Asn Val Arg Gly Ser Tyr Lys Cys
 225 230 235 240

50 Lys Cys Lys Glu Gly Tyr Gln Gly Asp Gly Leu Thr Cys Val Tyr Ile
 245 250 255

55 Pro Lys Val Met Ile Glu Pro Ser Gly Pro Ile His Val Pro Lys Gly
 260 265 270

Asn Gly Thr Ile Leu Lys Gly Asp Thr Gly Asn Asn Asn Trp Ile Pro

- 13 -

	275	280	285
5	Asp Val Gly Ser Thr Trp Trp Pro Pro Lys Thr Pro Tyr Ile Pro Pro 290 295 300		
10	Ile Ile Thr Asn Arg Pro Thr Ser Lys Pro Thr Thr Arg Pro Thr Pro 305 310 315 320		
15	Lys Pro Thr Pro Ile Pro Thr Pro Pro Pro Pro Pro Pro Leu Pro Thr 325 330 335		
20	Glu Leu Arg Thr Pro Leu Pro Pro Thr Thr Pro Glu Arg Pro Thr Thr 340 345 350		
25	Gly Leu Thr Thr Ile Ala Pro Ala Ala Ser Thr Pro Pro Gly Gly Ile 355 360 365		
30	Thr Val Asp Asn Arg Val Gln Thr Asp Pro Gln Lys Pro Arg Gly Asp 370 375 380		
35	Val Phe Ile Pro Arg Gln Pro Ser Asn Asp Leu Phe Glu Ile Phe Glu 385 390 395 400		
40	Ile Glu Arg Gly Val Ser Ala Asp Asp Glu Ala Lys Asp Asp Pro Gly 405 410 415		
45	Val Leu Val His Ser Cys Asn Phe Asp His Gly Leu Cys Gly Trp Ile 420 425 430		
50	Arg Glu Lys Asp Asn Asp Leu His Trp Glu Pro Ile Arg Asp Pro Ala 435 440 445		
55	Asp Asn Cys Ser Cys Leu His Ala Gly Gly Leu Asp Leu 450 455 460		
	<210> 7		
	<211> 461		
	<212> PRT		
	<213> homo sapiens		
	<400> 7		
55	Met Asp Phe Leu Leu Ala Leu Val Leu Val Ser Ser Leu Tyr Leu Gln 1 5 10 15		

- 14 -

	Ala	Ala	Ala	Glu	Phe	Asp	Gly	Arg	Trp	Pro	Arg	Gln	Ile	Val	Ser	Ser	
				20					25					30			
5	Ile	Gly	Leu	Cys	Arg	Tyr	Gly	Gly	Arg	Ile	Asp	Cys	Cys	Trp	Gly	Trp	
			35					40					45				
10	Ala	Arg	Gln	Ser	Trp	Gly	Gln	Cys	Gln	Pro	Val	Cys	Gln	Pro	Arg	Cys	
		50					55					60					
15	Lys	His	Gly	Glu	Cys	Ile	Gly	Pro	Asn	Lys	Cys	Lys	Cys	His	Pro	Gly	
	65					70					75					80	
	Tyr	Ala	Gly	Lys	Thr	Cys	Asn	Gln	Asp	Leu	Asn	Glu	Cys	Gly	Leu	Lys	
					85					90					95		
20	Pro	Arg	Pro	Cys	Lys	His	Arg	Cys	Met	Asn	Thr	Tyr	Gly	Ser	Tyr	Lys	
				100					105					110			
25	Cys	Tyr	Cys	Leu	Asn	Gly	Tyr	Met	Leu	Met	Pro	Asp	Gly	Ser	Cys	Ser	
			115					120					125				
30	Ser	Ala	Leu	Thr	Cys	Ser	Met	Ala	Asn	Cys	Gln	Tyr	Gly	Cys	Asp	Val	
		130					135						140				
35	Val	Lys	Gly	Gln	Ile	Arg	Cys	Gln	Cys	Pro	Ser	Pro	Gly	Leu	Gln	Leu	
	145					150					155					160	
	Ala	Pro	Asp	Gly	Arg	Thr	Cys	Val	Asp	Val	Asp	Glu	Cys	Ala	Thr	Gly	
					165					170					175		
40	Arg	Ala	Ser	Cys	Pro	Arg	Phe	Arg	Gln	Cys	Val	Asn	Thr	Phe	Gly	Ser	
				180					185						190		
45	Tyr	Ile	Cys	Lys	Cys	His	Lys	Gly	Phe	Asp	Leu	Met	Tyr	Ile	Gly	Gly	
			195					200					205				
50	Lys	Tyr	Gln	Cys	His	Asp	Ile	Asp	Glu	Cys	Ser	Leu	Gly	Gln	Tyr	Gln	
		210					215					220					
55	Cys	Ser	Ser	Phe	Ala	Arg	Cys	Tyr	Asn	Val	Arg	Gly	Ser	Tyr	Lys	Cys	
	225					230					235					240	
	Lys	Cys	Lys	Glu	Gly	Tyr	Gln	Gly	Asp	Gly	Leu	Thr	Cys	Val	Tyr	Ile	
					245					250					255		

- 15 -

[illegible]

- 16 -

<400> 8

5 Phe Leu Leu Ala Leu Val Leu Val Ser Ser Leu Tyr Leu Gln Ala Ala
 1 5 10 15
 Ala Glu Phe Asp Gly Arg Trp Pro Arg Gln Ile Val Ser Ser Ile Gly
 10 20 25 30
 Leu Cys Arg Tyr Gly Gly Arg Ile Asp Cys Cys Trp Gly Trp Ala Arg
 15 35 40 45
 Gln Ser Trp Gly Gln Cys Gln Pro Val Cys Gln Pro Arg Cys Lys His
 50 55 60
 20 Gly Glu Cys Ile Gly Pro Asn Lys Cys Lys Cys His Pro Gly Tyr Ala
 65 70 75 80
 Gly Lys Thr Cys Asn Gln Asp Leu Asn Glu Cys Gly Leu Lys Pro Arg
 25 85 90 95
 Pro Cys Lys His Arg Cys Met Asn Thr Tyr Gly Ser Tyr Lys Cys Tyr
 30 100 105 110
 Cys Leu Asn Gly Tyr Met Leu Met Pro Asp Gly Ser Cys Ser Ser Ala
 115 120 125
 35 Leu Thr Cys Ser Met Ala Asn Cys Gln Tyr Gly Cys Asp Val Val Lys
 130 135 140
 40 Gly Gln Ile Arg Cys Gln Cys Pro Ser Pro Gly Leu His Leu Ala Pro
 145 150 155 160
 Asp Gly Arg Thr Cys Val Asp Val Asp Glu Cys Ala Thr Gly Arg Ala
 45 165 170 175
 Ser Cys Pro Arg Phe Arg Gln Cys Val Asn Thr Phe Gly Ser Tyr Ile
 50 180 185 190
 Cys Lys Cys His Lys Gly Phe Asp Leu Met Tyr Ile Gly Gly Lys Tyr
 195 200 205
 55 Gln Cys His Asp Ile Asp Glu Cys Ser Leu Gly Gln Tyr Gln Cys Ser
 210 215 220

- 17 -

Ser Phe Ala Arg Cys Tyr Asn Val Arg Gly Ser Tyr Lys Cys Lys Cys
 225 230 235 240
 5 Lys Glu Gly Tyr Gln Gly Asp Gly Leu Thr Cys Val Tyr Ile Pro Lys
 245 250 255
 10 Val Met Ile Glu Pro Ser Gly Pro Ile His Val Pro Lys Gly Asn Gly
 260 265 270
 15 Thr Ile Leu Lys Gly Asp Thr Gly Asn Asn Asn Trp Ile Pro Asp Val
 275 280 285
 20 Gly Ser Thr Trp Trp Pro Pro Lys Thr Pro Tyr Ile Pro Pro Ile Ile
 290 295 300
 Thr Asn Arg Pro Thr Ser Lys Pro Thr Thr Arg Pro Thr Pro Lys Pro
 305 310 315 320
 25 Thr Pro Ile Pro Thr Pro Pro Pro Pro Pro Pro Leu Pro Thr Glu Leu
 325 330 335
 30 Arg Thr Pro Leu Pro Pro Thr Thr Pro Glu Arg Pro Thr Thr Gly Leu
 340 345 350
 35 Thr Thr Ile Ala Pro Ala Ala Ser Thr Pro Pro Gly Gly Ile Thr Val
 355 360 365
 40 Asp Asn Arg Val Gln Thr Asp Pro Gln Lys Pro Arg Gly Asp Val Phe
 370 375 380
 Ile Pro Arg Gln Pro Ser Asn Asp Leu Phe Glu Ile Phe Glu Ile Glu
 385 390 395 400
 45 Arg Gly Val Ser Ala Asp Asp Glu Ala Lys Asp Asp Pro Gly Val Leu
 405 410 415
 50 Val His Ser Cys Asn Phe Asp His Gly Leu Cys Gly Trp Ile Arg Glu
 420 425 430
 55 Lys Asp Asn Asp Leu His Trp Glu Pro Ile Arg Asp Pro Ala Asp Asn
 435 440 445
 Cys Ser Cys Leu His Ala Gly Gly Leu Asp Leu

- 18 -

450 455

<210> 9
5 <211> 459
<212> PRT
<213> homo sapiens

<400> 9

10 Phe Leu Leu Ala Leu Val Leu Val Ser Ser Leu Tyr Leu Gln Ala Ala
1 5 10 15

15 Ala Glu Phe Asp Gly Arg Trp Pro Arg Gln Ile Val Ser Ser Ile Gly
 20 25 30

20 Leu Cys Arg Tyr Gly Gly Arg Ile Asp Cys Cys Trp Gly Trp Ala Arg
 35 40 45

Gln Ser Trp Gly Gln Cys Gln Pro Val Cys Gln Pro Arg Cys Lys His
25 50 55 60

Gly Glu Cys Ile Gly Pro Asn Lys Cys Lys Cys His Pro Gly Tyr Ala
65 70 75 80

30 Gly Lys Thr Cys Asn Gln Asp Leu Asn Glu Cys Gly Leu Lys Pro Arg
 85 90 95

35 Pro Cys Lys His Arg Cys Met Asn Thr Tyr Gly Ser Tyr Lys Cys Tyr
 100 105 110

40 Cys Leu Asn Gly Tyr Met Leu Met Pro Asp Gly Ser Cys Ser Ser Ala
 115 120 125

Leu Thr Cys Ser Met Ala Asn Cys Gln Tyr Gly Cys Asp Val Val Lys
45 130 135 140

Gly Gln Ile Arg Cys Gln Cys Pro Ser Pro Gly Leu Gln Leu Ala Pro
145 150 155 160

50 Asp Gly Arg Thr Cys Val Asp Val Asp Glu Cys Ala Thr Gly Arg Ala
 165 170 175

55 Ser Cys Pro Arg Phe Arg Gln Cys Val Asn Thr Phe Gly Ser Tyr Ile
 180 185 190

- 19 -

Cys Lys Cys His Lys Gly Phe Asp Leu Met Tyr Ile Gly Gly Lys Tyr
 195 200 205

5 Gln Cys His Asp Ile Asp Glu Cys Ser Leu Gly Gln Tyr Gln Cys Ser
 210 215 220

10 Ser Phe Ala Arg Cys Tyr Asn Val Arg Gly Ser Tyr Lys Cys Lys Cys
 225 230 235 240

15 Lys Glu Gly Tyr Gln Gly Asp Gly Leu Thr Cys Val Tyr Ile Pro Lys
 245 250 255

Val Met Ile Glu Pro Ser Gly Pro Ile His Val Pro Lys Gly Asn Gly
 260 265 270

20 Thr Ile Leu Lys Gly Asp Thr Gly Asn Asn Asn Trp Ile Pro Asp Val
 275 280 285

25 Gly Ser Thr Trp Trp Pro Pro Lys Thr Pro Tyr Ile Pro Pro Ile Ile
 290 295 300

30 Thr Asn Arg Pro Thr Ser Lys Pro Thr Thr Arg Pro Thr Pro Lys Pro
 305 310 315 320

35 Thr Pro Ile Pro Thr Pro Pro Pro Pro Pro Pro Leu Pro Thr Glu Leu
 325 330 335

Arg Thr Pro Leu Pro Pro Thr Thr Pro Glu Arg Pro Thr Thr Gly Leu
 340 345 350

40 Thr Thr Ile Ala Pro Ala Ala Ser Thr Pro Pro Gly Gly Ile Thr Val
 355 360 365

45 Asp Asn Arg Val Gln Thr Asp Pro Gln Lys Pro Arg Gly Asp Val Phe
 370 375 380

50 Ile Pro Arg Gln Pro Ser Asn Asp Leu Phe Glu Ile Phe Glu Ile Glu
 385 390 395 400

55 Arg Gly Val Ser Ala Asp Asp Glu Ala Lys Asp Asp Pro Gly Val Leu
 405 410 415

Val His Ser Cys Asn Phe Asp His Gly Leu Cys Gly Trp Ile Arg Glu
 420 425 430

- 20 -

Lys Asp Asn Asp Leu His Trp Glu Pro Ile Arg Asp Pro Ala Asp Asn
 435 440 445
 5
 Cys Ser Cys Leu His Ala Gly Gly Leu Asp Leu
 450 455
 10
 <210> 10
 <211> 467
 <212> PRT
 <213> homo sapiens
 15
 <400> 10
 Met Asp Phe Leu Leu Ala Leu Val Leu Val Ser Ser Leu Tyr Leu Gln
 1 5 10 15
 20
 Ala Ala Ala Glu Phe Asp Gly Arg Trp Pro Arg Gln Ile Val Ser Ser
 20 25 30
 25
 Ile Gly Leu Cys Arg Tyr Gly Gly Arg Ile Asp Cys Cys Trp Gly Trp
 35 40 45
 30
 Ala Arg Gln Ser Trp Gly Gln Cys Gln Pro Val Cys Gln Pro Arg Cys
 50 55 60
 35
 Lys His Gly Glu Cys Ile Gly Pro Asn Lys Cys Lys Cys His Pro Gly
 65 70 75 80
 Tyr Ala Gly Lys Thr Cys Asn Gln Asp Leu Asn Glu Cys Gly Leu Lys
 85 90 95
 40
 Pro Arg Pro Cys Lys His Arg Cys Met Asn Thr Tyr Gly Ser Tyr Lys
 100 105 110
 45
 Cys Tyr Cys Leu Asn Gly Tyr Met Leu Met Pro Asp Gly Ser Cys Ser
 115 120 125
 50
 Ser Ala Leu Thr Cys Ser Met Ala Asn Cys Gln Tyr Gly Cys Asp Val
 130 135 140
 55
 Val Lys Gly Gln Ile Arg Cys Gln Cys Pro Ser Pro Gly Leu His Leu
 145 150 155 160
 Ala Pro Asp Gly Arg Thr Cys Val Asp Val Asp Glu Cys Ala Thr Gly

- 21 -

	165	170	175
5	Arg Ala Ser Cys Pro Arg Phe Arg Gln Cys Val Asn Thr Phe Gly Ser 180 185 190		
10	Tyr Ile Cys Lys Cys His Lys Gly Phe Asp Leu Met Tyr Ile Gly Gly 195 200 205		
15	Lys Tyr Gln Cys His Asp Ile Asp Glu Cys Ser Leu Gly Gln Tyr Gln 210 215 220		
20	Cys Ser Ser Phe Ala Arg Cys Tyr Asn Val Arg Gly Ser Tyr Lys Cys 225 230 235 240		
25	Lys Cys Lys Glu Gly Tyr Gln Gly Asp Gly Leu Thr Cys Val Tyr Ile 245 250 255		
30	Pro Lys Val Met Ile Glu Pro Ser Gly Pro Ile His Val Pro Lys Gly 260 265 270		
35	Asn Gly Thr Ile Leu Lys Gly Asp Thr Gly Asn Asn Asn Trp Ile Pro 275 280 285		
40	Asp Val Gly Ser Thr Trp Trp Pro Pro Lys Thr Pro Tyr Ile Pro Pro 290 295 300		
45	Ile Ile Thr Asn Arg Pro Thr Ser Lys Pro Thr Thr Arg Pro Thr Pro 305 310 315 320		
50	Lys Pro Thr Pro Ile Pro Thr Pro Pro Pro Pro Pro Leu Pro Thr 325 330 335		
55	Glu Leu Arg Thr Pro Leu Pro Pro Thr Thr Pro Glu Arg Pro Thr Thr 340 345 350		
	Gly Leu Thr Thr Ile Ala Pro Ala Ala Ser Thr Pro Pro Gly Gly Ile 355 360 365		
	Thr Val Asp Asn Arg Val Gln Thr Asp Pro Gln Lys Pro Arg Gly Asp 370 375 380		
	Val Phe Ile Pro Arg Gln Pro Ser Asn Asp Leu Phe Glu Ile Phe Glu 385 390 395 400		

- 22 -

```

      Ile Glu Arg Gly Val Ser Ala Asp Asp Glu Ala Lys Asp Asp Pro Gly
                        405                      410                      415
- 5      Val Leu Val His Ser Cys Asn Phe Asp His Gly Leu Cys Gly Trp Ile
                        420                      425                      430
10      Arg Glu Lys Asp Asn Asp Leu His Trp Glu Pro Ile Arg Asp Pro Ala
                        435                      440                      445
15      Asp Asn Cys Ser Cys Leu His Ala Gly Gly Leu Asp Leu His His His
                        450                      455                      460
      His His His
      465
20
      <210> 11
      <211> 467
      <212> PRT
25      <213> homo sapiens
      <400> 11
30      Met Asp Phe Leu Leu Ala Leu Val Leu Val Ser Ser Leu Tyr Leu Gln
      1                      5                      10                      15
      Ala Ala Ala Glu Phe Asp Gly Arg Trp Pro Arg Gln Ile Val Ser Ser
                        20                      25                      30
35
      Ile Gly Leu Cys Arg Tyr Gly Gly Arg Ile Asp Cys Cys Trp Gly Trp
                        35                      40                      45
40
      Ala Arg Gln Ser Trp Gly Gln Cys Gln Pro Val Cys Gln Pro Arg Cys
                        50                      55                      60
45      Lys His Gly Glu Cys Ile Gly Pro Asn Lys Cys Lys Cys His Pro Gly
      65                      70                      75                      80
50      Tyr Ala Gly Lys Thr Cys Asn Gln Asp Leu Asn Glu Cys Gly Leu Lys
                        85                      90                      95
      Pro Arg Pro Cys Lys His Arg Cys Met Asn Thr Tyr Gly Ser Tyr Lys
                        100                      105                      110
55
      Cys Tyr Cys Leu Asn Gly Tyr Met Leu Met Pro Asp Gly Ser Cys Ser
                        115                      120                      125

```

- 23 -

5 Ser Ala Leu Thr Cys Ser Met Ala Asn Cys Gln Tyr Gly Cys Asp Val
 130 135 140

10 Val Lys Gly Gln Ile Arg Cys Gln Cys Pro Ser Pro Gly Leu Gln Leu
 145 150 155 160

15 Ala Pro Asp Gly Arg Thr Cys Val Asp Val Asp Glu Cys Ala Thr Gly
 165 170 175

20 Arg Ala Ser Cys Pro Arg Phe Arg Gln Cys Val Asn Thr Phe Gly Ser
 180 185 190

25 Tyr Ile Cys Lys Cys His Lys Gly Phe Asp Leu Met Tyr Ile Gly Gly
 195 200 205

30 Lys Tyr Gln Cys His Asp Ile Asp Glu Cys Ser Leu Gly Gln Tyr Gln
 210 215 220

35 Cys Ser Ser Phe Ala Arg Cys Tyr Asn Val Arg Gly Ser Tyr Lys Cys
 225 230 235 240

40 Lys Cys Lys Glu Gly Tyr Gln Gly Asp Gly Leu Thr Cys Val Tyr Ile
 245 250 255

45 Pro Lys Val Met Ile Glu Pro Ser Gly Pro Ile His Val Pro Lys Gly
 260 265 270

50 Asn Gly Thr Ile Leu Lys Gly Asp Thr Gly Asn Asn Asn Trp Ile Pro
 275 280 285

55 Asp Val Gly Ser Thr Trp Trp Pro Pro Lys Thr Pro Tyr Ile Pro Pro
 290 295 300

60 Ile Ile Thr Asn Arg Pro Thr Ser Lys Pro Thr Thr Arg Pro Thr Pro
 305 310 315 320

65 Lys Pro Thr Pro Ile Pro Thr Pro Pro Pro Pro Pro Leu Pro Thr
 325 330 335

70 Glu Leu Arg Thr Pro Leu Pro Pro Thr Thr Pro Glu Arg Pro Thr Thr
 340 345 350

- 24 -

Gly Leu Thr Thr Ile Ala Pro Ala Ala Ser Thr Pro Pro Gly Gly Ile
 355 360 365

5 Thr Val Asp Asn Arg Val Gln Thr Asp Pro Gln Lys Pro Arg Gly Asp
 370 375 380

10 Val Phe Ile Pro Arg Gln Pro Ser Asn Asp Leu Phe Glu Ile Phe Glu
 385 390 395 400

15 Ile Glu Arg Gly Val Ser Ala Asp Asp Glu Ala Lys Asp Asp Pro Gly
 405 410 415

Val Leu Val His Ser Cys Asn Phe Asp His Gly Leu Cys Gly Trp Ile
 420 425 430

20 Arg Glu Lys Asp Asn Asp Leu His Trp Glu Pro Ile Arg Asp Pro Ala
 435 440 445

25 Asp Asn Cys Ser Cys Leu His Ala Gly Gly Leu Asp Leu His His His
 450 455 460

30 His His His
 465

35 <210> 12
 <211> 1503
 <212> DNA
 <213> homo sapiens

40 <220>
 <221> CDS
 <222> (205)..(1503)

45 <400> 12
 gcgggagggg gctccgggcg ccgcgcagca gacctgctcc ggccgcgcgc ctgcgcgctg 60
 tcctccggga gcggcagcag tagccggg gcgcgaggct gggggttcct cgagactctc 120
 agaggggagc ctcccatcgg cgcaccac cccaacctgt tcctcgcgcg cactgcgct 180
 50 gcgccccagg acccgctgcc caac atg gat ttg ctc ctg gcg ctg gtg ctg 231
 Met Asp Leu Leu Leu Ala Leu Val Leu
 1 5

55 gta tcc tcg ctc tac ctg cag gcg gcc gcc gag ttc gac ggg agg tgg 279
 Val Ser Ser Leu Tyr Leu Gln Ala Ala Ala Glu Phe Asp Gly Arg Trp
 10 15 20 25

ccc agg caa ata gtg tca tcg att ggc cta tgt cgt tat ggt ggg agg 327

- 25 -

	Pro	Arg	Gln	Ile	Val	Ser	Ser	Ile	Gly	Leu	Cys	Arg	Tyr	Gly	Gly	Arg	
				30						35					40		
5	att	gac	tgc	tgc	tgg	ggc	tgg	gct	cgc	cag	tct	tgg	gga	cag	tgt	cag	375
	Ile	Asp	Cys	Cys	Trp	Gly	Trp	Ala	Arg	Gln	Ser	Trp	Gly	Gln	Cys	Gln	
			45						50					55			
10	cct	gtg	tgc	caa	cca	cga	tgc	aaa	cat	ggc	gaa	tgt	atc	ggg	cca	aac	423
	Pro	Val	Cys	Gln	Pro	Arg	Cys	Lys	His	Gly	Glu	Cys	Ile	Gly	Pro	Asn	
			60					65					70				
15	aag	tgc	aag	tgt	cat	cct	ggc	tat	gct	gga	aaa	acc	tgt	aat	caa	gat	471
	Lys	Cys	Lys	Cys	His	Pro	Gly	Tyr	Ala	Gly	Lys	Thr	Cys	Asn	Gln	Asp	
			75				80					85					
20	cta	aat	gag	tgt	ggc	ctg	aag	ccc	cgg	ccc	tgt	aag	cac	agg	tgc	atg	519
	Leu	Asn	Glu	Cys	Gly	Leu	Lys	Pro	Arg	Pro	Cys	Lys	His	Arg	Cys	Met	
			90			95					100					105	
25	aac	act	tac	ggc	agc	tac	aag	tgc	tac	tgt	ctc	aac	gga	tat	atg	ctc	567
	Asn	Thr	Tyr	Gly	Ser	Tyr	Lys	Cys	Tyr	Cys	Leu	Asn	Gly	Tyr	Met	Leu	
					110					115					120		
30	atg	cgc	gat	ggc	tcc	tgc	tca	agt	gcc	ctg	acc	tgc	tcc	atg	gca	aac	615
	Met	Pro	Asp	Gly	Ser	Cys	Ser	Ser	Ala	Leu	Thr	Cys	Ser	Met	Ala	Asn	
				125					130					135			
35	tgt	cag	tat	ggc	tgt	gat	gtt	gtt	aaa	gga	caa	ata	cgg	tgc	cag	tgc	663
	Cys	Gln	Tyr	Gly	Cys	Asp	Val	Val	Lys	Gly	Gln	Ile	Arg	Cys	Gln	Cys	
			140					145					150				
40	cca	tcc	cct	ggc	ctg	cac	ctg	gct	cct	gat	ggg	agg	acc	tgt	gta	gat	711
	Pro	Ser	Pro	Gly	Leu	His	Leu	Ala	Pro	Asp	Gly	Arg	Thr	Cys	Val	Asp	
			155				160					165					
45	gtt	gat	gaa	tgt	gct	aca	gga	aga	gcc	tcc	tgc	cct	aga	ttt	agg	caa	759
	Val	Asp	Glu	Cys	Ala	Thr	Gly	Arg	Ala	Ser	Cys	Pro	Arg	Phe	Arg	Gln	
			170			175					180					185	
50	tgt	gtc	aac	act	ttt	ggg	agc	tac	atc	tgc	aag	tgt	cat	aaa	ggc	ttc	807
	Cys	Val	Asn	Thr	Phe	Gly	Ser	Tyr	Ile	Cys	Lys	Cys	His	Lys	Gly	Phe	
					190					195					200		
55	gat	ctc	atg	tat	att	gga	ggc	aaa	tat	caa	tgt	cat	gac	ata	gac	gaa	855
	Asp	Leu	Met	Tyr	Ile	Gly	Gly	Lys	Tyr	Gln	Cys	His	Asp	Ile	Asp	Glu	
				205				210						215			
60	tgc	tca	ctt	ggc	cag	tat	cag	tgc	agc	agc	ttt	gct	cga	tgt	tat	aac	903
	Cys	Ser	Leu	Gly	Gln	Tyr	Gln	Cys	Ser	Ser	Phe	Ala	Arg	Cys	Tyr	Asn	
			220					225					230				
65	gta	cgt	ggg	tcc	tac	aag	tgc	aaa	tgt	aaa	gaa	gga	tac	cag	ggc	gat	951
	Val	Arg	Gly	Ser	Tyr	Lys	Cys	Lys	Cys	Lys	Glu	Gly	Tyr	Gln	Gly	Asp	
			235				240					245					
70	gga	ctg	act	tgt	gtg	tat	atc	cca	aaa	gtt	atg	att	gaa	cct	tca	ggc	999
	Gly	Leu	Thr	Cys	Val	Tyr	Ile	Pro	Lys	Val	Met	Ile	Glu	Pro	Ser	Gly	
						255					260					265	

- 26 -

cca att cat gta cca aag gga aat ggt acc att tta aag ggt gac aca 1047
 Pro Ile His Val Pro Lys Gly Asn Gly Thr Ile Leu Lys Gly Asp Thr
 270 275 280

5

gga aat aat aat tgg att cct gat gtt gga agt act tgg tgg cct ccg 1095
 Gly Asn Asn Asn Trp Ile Pro Asp Val Gly Ser Thr Trp Trp Pro Pro
 285 290 295

10

aag aca cca tat att cct cct atc att acc aac agg cct act tct aag 1143
 Lys Thr Pro Tyr Ile Pro Pro Ile Ile Thr Asn Arg Pro Thr Ser Lys
 300 305 310

15

cca aca aca aga cct aca cca aag cca aca cca att cct act cca cca 1191
 Pro Thr Thr Arg Pro Thr Pro Lys Pro Thr Pro Ile Pro Thr Pro Pro
 315 320 325

20

cca cca cca ccc ctg cca aca gag ctc aga aca cct cta cca cct aca 1239
 Pro Pro Pro Pro Leu Thr Thr Glu Leu Arg Thr Pro Leu Pro Pro Thr
 330 335 340 345

25

acc cca gaa agg cca acc acc gga ctg aca act ata gca cca gct gcc 1287
 Thr Pro Glu Arg Thr Thr Thr Gly Leu Thr Thr Ile Ala Pro Ala Ala
 350 355 360

30

agt aca cct cca gga ggg att aca gtt gac aac agg gta cag aca gac 1335
 Ser Thr Pro Pro Gly Gly Ile Thr Val Asp Asn Arg Val Gln Thr Asp
 365 370 375

35

cct cag aaa ccc aga gga gat gtg ttc agt gtt ctg gta cac agt tgt 1383
 Pro Gln Lys Pro Arg Gly Asp Val Phe Ser Val Leu Val His Ser Cys
 380 385 390

40

aat ttt gac cat gga ctt tgt gga tgg atc agg gag aaa gac aat gac 1431
 Asn Phe Asp His Gly Leu Cys Gly Trp Ile Arg Glu Lys Asp Asn Asp
 395 400 405

45

ttg cac tgg gaa cca atc agg gac cca gca gac aat tgt agc tgt ctg 1479
 Leu His Trp Glu Pro Ile Arg Asp Pro Ala Asp Asn Cys Ser Cys Leu
 410 415 420 425

cat gca ggt gga ctt gac ttg tag 1503
 His Ala Gly Gly Leu Asp Leu
 430

50

<210> 13
 <211> 432
 <212> PRT
 <213> homo sapiens
 <400> 13

55

Met Asp Leu Leu Leu Ala Leu Val Leu Val Ser Ser Leu Tyr Leu Gln
 1 5 10 15

Ala Ala Ala Glu Phe Asp Gly Arg Trp Pro Arg Gln Ile Val Ser Ser

- 27 -

	20	25	30
5	Ile Gly Leu Cys Arg Tyr Gly Gly Arg Ile Asp Cys Cys Trp Gly Trp 35 40 45		
10	Ala Arg Gln Ser Trp Gly Gln Cys Gln Pro Val Cys Gln Pro Arg Cys 50 55 60		
15	Lys His Gly Glu Cys Ile Gly Pro Asn Lys Cys Lys Cys His Pro Gly 65 70 75 80		
20	Tyr Ala Gly Lys Thr Cys Asn Gln Asp Leu Asn Glu Cys Gly Leu Lys 85 90 95		
25	Pro Arg Pro Cys Lys His Arg Cys Met Asn Thr Tyr Gly Ser Tyr Lys 100 105 110		
30	Cys Tyr Cys Leu Asn Gly Tyr Met Leu Met Pro Asp Gly Ser Cys Ser 115 120 125		
35	Ser Ala Leu Thr Cys Ser Met Ala Asn Cys Gln Tyr Gly Cys Asp Val 130 135 140		
40	Val Lys Gly Gln Ile Arg Cys Gln Cys Pro Ser Pro Gly Leu His Leu 145 150 155 160		
45	Ala Pro Asp Gly Arg Thr Cys Val Asp Val Asp Glu Cys Ala Thr Gly 165 170 175		
50	Arg Ala Ser Cys Pro Arg Phe Arg Gln Cys Val Asn Thr Phe Gly Ser 180 185 190		
55	Tyr Ile Cys Lys Cys His Lys Gly Phe Asp Leu Met Tyr Ile Gly Gly 195 200 205		
	Lys Tyr Gln Cys His Asp Ile Asp Glu Cys Ser Leu Gly Gln Tyr Gln 210 215 220		
	Cys Ser Ser Phe Ala Arg Cys Tyr Asn Val Arg Gly Ser Tyr Lys Cys 225 230 235 240		
	Lys Cys Lys Glu Gly Tyr Gln Gly Asp Gly Leu Thr Cys Val Tyr Ile 245 250 255		

- 28 -

Pro Lys Val Met Ile Glu Pro Ser Gly Pro Ile His Val Pro Lys Gly
 260 265 270

5 Asn Gly Thr Ile Leu Lys Gly Asp Thr Gly Asn Asn Asn Trp Ile Pro
 275 280 285

10 Asp Val Gly Ser Thr Trp Trp Pro Pro Lys Thr Pro Tyr Ile Pro Pro
 290 295 300

15 Ile Ile Thr Asn Arg Pro Thr Ser Lys Pro Thr Thr Arg Pro Thr Pro
 305 310 315 320

20 Lys Pro Thr Pro Ile Pro Thr Pro Pro Pro Pro Pro Pro Leu Pro Thr
 325 330 335

Glu Leu Arg Thr Pro Leu Pro Pro Thr Thr Pro Glu Arg Pro Thr Thr
 340 345 350

25 Gly Leu Thr Thr Ile Ala Pro Ala Ala Ser Thr Pro Pro Gly Gly Ile
 355 360 365

30 Thr Val Asp Asn Arg Val Gln Thr Asp Pro Gln Lys Pro Arg Gly Asp
 370 375 380

35 Val Phe Ser Val Leu Val His Ser Cys Asn Phe Asp His Gly Leu Cys
 385 390 395 400

Gly Trp Ile Arg Glu Lys Asp Asn Asp Leu His Trp Glu Pro Ile Arg
 405 410 415

40 Asp Pro Ala Asp Asn Cys Ser Cys Leu His Ala Gly Gly Leu Asp Leu
 420 425 430

45 <210> 14
 <211> 432
 <212> PRT
 <213> homo sapiens

50 <400> 14

Met Asp Phe Leu Leu Ala Leu Val Leu Val Ser Ser Leu Tyr Leu Gln
 1 5 10 15

55 Ala Ala Ala Glu Phe Asp Gly Arg Trp Pro Arg Gln Ile Val Ser Ser
 20 25 30

- 29 -

Ile Gly Leu Cys Arg Tyr Gly Gly Arg Ile Asp Cys Cys Trp Gly Trp
 35 40 45
 5
 Ala Arg Gln Ser Trp Gly Gln Cys Gln Pro Val Cys Gln Pro Arg Cys
 50 55 60
 10
 Lys His Gly Glu Cys Ile Gly Pro Asn Lys Cys Lys Cys His Pro Gly
 65 70 75 80
 15 Tyr Ala Gly Lys Thr Cys Asn Gln Asp Leu Asn Glu Cys Gly Leu Lys
 85 90 95
 20 Pro Arg Pro Cys Lys His Arg Cys Met Asn Thr Tyr Gly Ser Tyr Lys
 100 105 110
 25 Cys Tyr Cys Leu Asn Gly Tyr Met Leu Met Pro Asp Gly Ser Cys Ser
 115 120 125
 Ser Ala Leu Thr Cys Ser Met Ala Asn Cys Gln Tyr Gly Cys Asp Val
 130 135 140
 30 Val Lys Gly Gln Ile Arg Cys Gln Cys Pro Ser Pro Gly Leu Gln Leu
 145 150 155 160
 35 Ala Pro Asp Gly Arg Thr Cys Val Asp Val Asp Glu Cys Ala Thr Gly
 165 170 175
 40 Arg Ala Ser Cys Pro Arg Phe Arg Gln Cys Val Asn Thr Phe Gly Ser
 180 185 190
 45 Tyr Ile Cys Lys Cys His Lys Gly Phe Asp Leu Met Tyr Ile Gly Gly
 195 200 205
 Lys Tyr Gln Cys His Asp Ile Asp Glu Cys Ser Leu Gly Gln Tyr Gln
 210 215 220
 50 Cys Ser Ser Phe Ala Arg Cys Tyr Asn Val Arg Gly Ser Tyr Lys Cys
 225 230 235 240
 55 Lys Cys Lys Glu Gly Tyr Gln Gly Asp Gly Leu Thr Cys Val Tyr Ile
 245 250 255

- 30 -

	Pro	Lys	Val	Met	Ile	Glu	Pro	Ser	Gly	Pro	Ile	His	Val	Pro	Lys	Gly
				260					265					270		
5	Asn	Gly	Thr	Ile	Leu	Lys	Gly	Asp	Thr	Gly	Asn	Asn	Asn	Trp	Ile	Pro
			275					280					285			
10	Asp	Val	Gly	Ser	Thr	Trp	Trp	Pro	Pro	Lys	Thr	Pro	Tyr	Ile	Pro	Pro
		290					295					300				
15	Ile	Ile	Thr	Asn	Arg	Pro	Thr	Ser	Lys	Pro	Thr	Thr	Arg	Pro	Thr	Pro
	305					310					315					320
20	Lys	Pro	Thr	Pro	Ile	Pro	Thr	Pro	Pro	Pro	Pro	Pro	Pro	Leu	Pro	Thr
					325					330					335	
25	Glu	Leu	Arg	Thr	Pro	Leu	Pro	Pro	Thr	Thr	Pro	Glu	Arg	Pro	Thr	Thr
				340					345					350		
30	Gly	Leu	Thr	Thr	Ile	Ala	Pro	Ala	Ala	Ser	Thr	Pro	Pro	Gly	Gly	Ile
			355					360					365			
35	Thr	Val	Asp	Asn	Arg	Val	Gln	Thr	Asp	Pro	Gln	Lys	Pro	Arg	Gly	Asp
		370					375					380				
40	Val	Phe	Ser	Val	Leu	Val	His	Ser	Cys	Asn	Phe	Asp	His	Gly	Leu	Cys
	385					390					395					400
45	Gly	Trp	Ile	Arg	Glu	Lys	Asp	Asn	Asp	Leu	His	Trp	Glu	Pro	Ile	Arg
					405					410					415	
50	Asp	Pro	Ala	Asp	Asn	Cys	Ser	Cys	Leu	His	Ala	Gly	Gly	Leu	Asp	Leu
				420					425					430		
55	<210>	15														
	<211>	430														
	<212>	PRT														
	<213>	homo sapiens														
60	<400>	15														
	Leu	Leu	Leu	Ala	Leu	Val	Leu	Val	Ser	Ser	Leu	Tyr	Leu	Gln	Ala	Ala
	1				5					10					15	
65	Ala	Glu	Phe	Asp	Gly	Arg	Trp	Pro	Arg	Gln	Ile	Val	Ser	Ser	Ile	Gly
				20					25					30		

- 31 -

Leu Cys Arg Tyr Gly Gly Arg Ile Asp Cys Cys Trp Gly Trp Ala Arg
 35 40 45

5 Gln Ser Trp Gly Gln Cys Gln Pro Val Cys Gln Pro Arg Cys Lys His
 50 55 60

10 Gly Glu Cys Ile Gly Pro Asn Lys Cys Lys Cys His Pro Gly Tyr Ala
 65 70 75 80

15 Gly Lys Thr Cys Asn Gln Asp Leu Asn Glu Cys Gly Leu Lys Pro Arg
 85 90 95

20 Pro Cys Lys His Arg Cys Met Asn Thr Tyr Gly Ser Tyr Lys Cys Tyr
 100 105 110

Cys Leu Asn Gly Tyr Met Leu Met Pro Asp Gly Ser Cys Ser Ser Ala
 115 120 125

25 Leu Thr Cys Ser Met Ala Asn Cys Gln Tyr Gly Cys Asp Val Val Lys
 130 135 140

30 Gly Gln Ile Arg Cys Gln Cys Pro Ser Pro Gly Leu His Leu Ala Pro
 145 150 155 160

35 Asp Gly Arg Thr Cys Val Asp Val Asp Glu Cys Ala Thr Gly Arg Ala
 165 170 175

40 Ser Cys Pro Arg Phe Arg Gln Cys Val Asn Thr Phe Gly Ser Tyr Ile
 180 185 190

Cys Lys Cys His Lys Gly Phe Asp Leu Met Tyr Ile Gly Gly Lys Tyr
 195 200 205

45 Gln Cys His Asp Ile Asp Glu Cys Ser Leu Gly Gln Tyr Gln Cys Ser
 210 215 220

50 Ser Phe Ala Arg Cys Tyr Asn Val Arg Gly Ser Tyr Lys Cys Lys Cys
 225 230 235 240

55 Lys Glu Gly Tyr Gln Gly Asp Gly Leu Thr Cys Val Tyr Ile Pro Lys
 245 250 255

Val Met Ile Glu Pro Ser Gly Pro Ile His Val Pro Lys Gly Asn Gly

- 32 -

	260	265	270
5	Thr Ile Leu Lys Gly Asp Thr Gly Asn Asn Asn Trp Ile Pro Asp Val 275 280 285		
10	Gly Ser Thr Trp Trp Pro Pro Lys Thr Pro Tyr Ile Pro Pro Ile Ile 290 295 300		
15	Thr Asn Arg Pro Thr Ser Lys Pro Thr Thr Arg Pro Thr Pro Lys Pro 305 310 315 320		
20	Thr Pro Ile Pro Thr Pro Pro Pro Pro Pro Pro Leu Pro Thr Glu Leu 325 330 335		
25	Arg Thr Pro Leu Pro Pro Thr Thr Pro Glu Arg Pro Thr Thr Gly Leu 340 345 350		
30	Thr Thr Ile Ala Pro Ala Ala Ser Thr Pro Pro Gly Gly Ile Thr Val 355 360 365		
35	Asp Asn Arg Val Gln Thr Asp Pro Gln Lys Pro Arg Gly Asp Val Phe 370 375 380		
40	Ser Val Leu Val His Ser Cys Asn Phe Asp His Gly Leu Cys Gly Trp 385 390 395 400		
45	Ile Arg Glu Lys Asp Asn Asp Leu His Trp Glu Pro Ile Arg Asp Pro 405 410 415		
50	Ala Asp Asn Cys Ser Cys Leu His Ala Gly Gly Leu Asp Leu 420 425 430		
55	<210> 16 <211> 430 <212> PRT <213> homo sapiens <400> 16 Phe Leu Leu Ala Leu Val Leu Val Ser Ser Leu Tyr Leu Gln Ala Ala 1 5 10 15 Ala Glu Phe Asp Gly Arg Trp Pro Arg Gln Ile Val Ser Ser Ile Gly 20 25 30		

- 33 -

Leu Cys Arg Tyr Gly Gly Arg Ile Asp Cys Cys Trp Gly Trp Ala Arg
 35 40 45

5 Gln Ser Trp Gly Gln Cys Gln Pro Val Cys Gln Pro Arg Cys Lys His
 50 55 60

10 Gly Glu Cys Ile Gly Pro Asn Lys Cys Lys Cys His Pro Gly Tyr Ala
 65 70 75 80

15 Gly Lys Thr Cys Asn Gln Asp Leu Asn Glu Cys Gly Leu Lys Pro Arg
 85 90 95

20 Pro Cys Lys His Arg Cys Met Asn Thr Tyr Gly Ser Tyr Lys Cys Tyr
 100 105 110

25 Cys Leu Asn Gly Tyr Met Leu Met Pro Asp Gly Ser Cys Ser Ser Ala
 115 120 125

30 Leu Thr Cys Ser Met Ala Asn Cys Gln Tyr Gly Cys Asp Val Val Lys
 130 135 140

35 Gly Gln Ile Arg Cys Gln Cys Pro Ser Pro Gly Leu Gln Leu Ala Pro
 145 150 155 160

40 Asp Gly Arg Thr Cys Val Asp Val Asp Glu Cys Ala Thr Gly Arg Ala
 165 170 175

45 Ser Cys Pro Arg Phe Arg Gln Cys Val Asn Thr Phe Gly Ser Tyr Ile
 180 185 190

50 Cys Lys Cys His Lys Gly Phe Asp Leu Met Tyr Ile Gly Gly Lys Tyr
 195 200 205

55 Gln Cys His Asp Ile Asp Glu Cys Ser Leu Gly Gln Tyr Gln Cys Ser
 210 215 220

60 Ser Phe Ala Arg Cys Tyr Asn Val Arg Gly Ser Tyr Lys Cys Lys Cys
 225 230 235 240

65 Lys Glu Gly Tyr Gln Gly Asp Gly Leu Thr Cys Val Tyr Ile Pro Lys
 245 250 255

70 Val Met Ile Glu Pro Ser Gly Pro Ile His Val Pro Lys Gly Asn Gly
 260 265 270

- 34 -

Thr Ile Leu Lys Gly Asp Thr Gly Asn Asn Asn Trp Ile Pro Asp Val
 275 280 285
 5
 Gly Ser Thr Trp Trp Pro Pro Lys Thr Pro Tyr Ile Pro Pro Ile Ile
 290 295 300
 10
 Thr Asn Arg Pro Thr Ser Lys Pro Thr Thr Arg Pro Thr Pro Lys Pro
 305 310 315 320
 15 Thr Pro Ile Pro Thr Pro Pro Pro Pro Pro Pro Leu Pro Thr Glu Leu
 325 330 335
 20 Arg Thr Pro Leu Pro Pro Thr Thr Pro Glu Arg Pro Thr Thr Gly Leu
 340 345 350
 25 Thr Thr Ile Ala Pro Ala Ala Ser Thr Pro Pro Gly Gly Ile Thr Val
 355 360 365
 Asp Asn Arg Val Gln Thr Asp Pro Gln Lys Pro Arg Gly Asp Val Phe
 370 375 380
 30 Ser Val Leu Val His Ser Cys Asn Phe Asp His Gly Leu Cys Gly Trp
 385 390 395 400
 35 Ile Arg Glu Lys Asp Asn Asp Leu His Trp Glu Pro Ile Arg Asp Pro
 405 410 415
 40 Ala Asp Asn Cys Ser Cys Leu His Ala Gly Gly Leu Asp Leu
 420 425 430
 45 <210> 17
 <211> 438
 <212> PRT
 <213> homo sapiens
 <400> 17
 50 Met Asp Leu Leu Leu Ala Leu Val Leu Val Ser Ser Leu Tyr Leu Gln
 1 5 10 15
 55 Ala Ala Ala Glu Phe Asp Gly Arg Trp Pro Arg Gln Ile Val Ser Ser
 20 25 30
 Ile Gly Leu Cys Arg Tyr Gly Gly Arg Ile Asp Cys Cys Trp Gly Trp

- 35 -

	35	40	45
5	Ala Arg Gln Ser Trp Gly	Gln Cys Gln Pro Val	Cys Gln Pro Arg Cys
	50	55	60
10	Lys His Gly Glu Cys Ile Gly	Pro Asn Lys Cys Lys Cys His Pro Gly	
	65	70	75 80
15	Tyr Ala Gly Lys Thr Cys Asn Gln Asp Leu Asn Glu Cys Gly Leu Lys		
		85 90	95
20	Pro Arg Pro Cys Lys His Arg Cys Met Asn Thr Tyr Gly Ser Tyr Lys		
		100 105	110
25	Cys Tyr Cys Leu Asn Gly Tyr Met Leu Met Pro Asp Gly Ser Cys Ser		
		115 120	125
30	Ser Ala Leu Thr Cys Ser Met Ala Asn Cys Gln Tyr Gly Cys Asp Val		
		130 135	140
35	Val Lys Gly Gln Ile Arg Cys Gln Cys Pro Ser Pro Gly Leu His Leu		
		145 150	155 160
40	Ala Pro Asp Gly Arg Thr Cys Val Asp Val Asp Glu Cys Ala Thr Gly		
		165 170	175
45	Arg Ala Ser Cys Pro Arg Phe Arg Gln Cys Val Asn Thr Phe Gly Ser		
		180 185	190
50	Tyr Ile Cys Lys Cys His Lys Gly Phe Asp Leu Met Tyr Ile Gly Gly		
		195 200	205
55	Lys Tyr Gln Cys His Asp Ile Asp Glu Cys Ser Leu Gly Gln Tyr Gln		
		210 215	220
60	Cys Ser Ser Phe Ala Arg Cys Tyr Asn Val Arg Gly Ser Tyr Lys Cys		
		225 230	235 240
65	Lys Cys Lys Glu Gly Tyr Gln Gly Asp Gly Leu Thr Cys Val Tyr Ile		
		245 250	255
70	Pro Lys Val Met Ile Glu Pro Ser Gly Pro Ile His Val Pro Lys Gly		
		260 265	270

- 36 -

Asn Gly Thr Ile Leu Lys Gly Asp Thr Gly Asn Asn Asn Trp Ile Pro
 275 280 285

5 Asp Val Gly Ser Thr Trp Trp Pro Pro Lys Thr Pro Tyr Ile Pro Pro
 290 295 300

10 Ile Ile Thr Asn Arg Pro Thr Ser Lys Pro Thr Thr Arg Pro Thr Pro
 305 310 315 320

15 Lys Pro Thr Pro Ile Pro Thr Pro Pro Pro Pro Pro Leu Pro Thr
 325 330 335

20 Glu Leu Arg Thr Pro Leu Pro Pro Thr Thr Pro Glu Arg Pro Thr Thr
 340 345 350

25 Gly Leu Thr Thr Ile Ala Pro Ala Ala Ser Thr Pro Pro Gly Gly Ile
 355 360 365

Thr Val Asp Asn Arg Val Gln Thr Asp Pro Gln Lys Pro Arg Gly Asp
 370 375 380

30 Val Phe Ser Val Leu Val His Ser Cys Asn Phe Asp His Gly Leu Cys
 385 390 395 400

35 Gly Trp Ile Arg Glu Lys Asp Asn Asp Leu His Trp Glu Pro Ile Arg
 405 410 415

40 Asp Pro Ala Asp Asn Cys Ser Cys Leu His Ala Gly Gly Leu Asp Leu
 420 425 430

His His His His His His
 435

45 <210> 18
 <211> 438
 <212> PRT
 <213> homo sapiens

50 <400> 18

Met Asp Phe Leu Leu Ala Leu Val Leu Val Ser Ser Leu Tyr Leu Gln
 1 5 10 15

55 Ala Ala Ala Glu Phe Asp Gly Arg Trp Pro Arg Gln Ile Val Ser Ser
 20 25 30

- 37 -

Ile Gly Leu Cys Arg Tyr Gly Gly Arg Ile Asp Cys Cys Trp Gly Trp
 35 40 45
 5

Ala Arg Gln Ser Trp Gly Gln Cys Gln Pro Val Cys Gln Pro Arg Cys
 50 55 60
 10

Lys His Gly Glu Cys Ile Gly Pro Asn Lys Cys Lys Cys His Pro Gly
 65 70 75 80
 15

Tyr Ala Gly Lys Thr Cys Asn Gln Asp Leu Asn Glu Cys Gly Leu Lys
 85 90 95
 20

Pro Arg Pro Cys Lys His Arg Cys Met Asn Thr Tyr Gly Ser Tyr Lys
 100 105 110
 25

Cys Tyr Cys Leu Asn Gly Tyr Met Leu Met Pro Asp Gly Ser Cys Ser
 115 120 125
 30

Ser Ala Leu Thr Cys Ser Met Ala Asn Cys Gln Tyr Gly Cys Asp Val
 130 135 140
 35

Val Lys Gly Gln Ile Arg Cys Gln Cys Pro Ser Pro Gly Leu Gln Leu
 145 150 155 160
 40

Ala Pro Asp Gly Arg Thr Cys Val Asp Val Asp Glu Cys Ala Thr Gly
 165 170 175
 45

Arg Ala Ser Cys Pro Arg Phe Arg Gln Cys Val Asn Thr Phe Gly Ser
 180 185 190
 50

Tyr Ile Cys Lys Cys His Lys Gly Phe Asp Leu Met Tyr Ile Gly Gly
 195 200 205
 55

Lys Tyr Gln Cys His Asp Ile Asp Glu Cys Ser Leu Gly Gln Tyr Gln
 210 215 220
 60

Cys Ser Ser Phe Ala Arg Cys Tyr Asn Val Arg Gly Ser Tyr Lys Cys
 225 230 235 240
 65

Lys Cys Lys Glu Gly Tyr Gln Gly Asp Gly Leu Thr Cys Val Tyr Ile
 245 250 255

	Pro	Lys	Val	Met	Ile	Glu	Pro	Ser	Gly	Pro	Ile	His	Val	Pro	Lys	Gly
				260				265						270		
5	Asn	Gly	Thr	Ile	Leu	Lys	Gly	Asp	Thr	Gly	Asn	Asn	Asn	Trp	Ile	Pro
	275						280			285						
10	Asp	Val	Gly	Ser	Thr	Trp	Trp	Pro	Pro	Lys	Thr	Pro	Tyr	Ile	Pro	Pro
	290						295			300						
15	Ile	Ile	Thr	Asn	Arg	Pro	Thr	Ser	Lys	Pro	Thr	Thr	Arg	Pro	Thr	Pro
	305			310						315			320			
20	Lys	Pro	Thr	Pro	Ile	Pro	Thr	Pro	Pro	Pro	Pro	Pro	Pro	Leu	Pro	Thr
				325						330			335			
25	Glu	Leu	Arg	Thr	Pro	Leu	Pro	Pro	Thr	Thr	Pro	Glu	Arg	Pro	Thr	Thr
				340						345			350			
30	Gly	Leu	Thr	Thr	Ile	Ala	Pro	Ala	Ala	Ser	Thr	Pro	Pro	Gly	Gly	Ile
	355						360			365						
35	Thr	Val	Asp	Asn	Arg	Val	Gln	Thr	Asp	Pro	Gln	Lys	Pro	Arg	Gly	Asp
	370						375			380						
40	Val	Phe	Ser	Val	Leu	Val	His	Ser	Cys	Asn	Phe	Asp	His	Gly	Leu	Cys
	385			390						395			400			
45	Gly	Trp	Ile	Arg	Glu	Lys	Asp	Asn	Asp	Leu	His	Trp	Glu	Pro	Ile	Arg
				405						410			415			
50	Asp	Pro	Ala	Asp	Asn	Cys	Ser	Cys	Leu	His	Ala	Gly	Gly	Leu	Asp	Leu
				420						425			430			
55	His	His	His	His	His	His										
	435															