

(19) World Intellectual Property Organization  
International Bureau(43) International Publication Date  
7 August 2008 (07.08.2008)

PCT

(10) International Publication Number  
**WO 2008/093982 A1**

## (51) International Patent Classification:

C07K 7/04 (2006.01) C07K 7/08 (2006.01)  
C07K 7/06 (2006.01)

## (21) International Application Number:

PCT/KR2008/000525

(22) International Filing Date: 29 January 2008 (29.01.2008)

(25) Filing Language: English

(26) Publication Language: English

## (30) Priority Data:

60/887,060 29 January 2007 (29.01.2007) US

(71) Applicant (for all designated States except US): **INDUSTRY FOUNDATION OF CHONNAM NATIONAL UNIVERSITY** [KR/KR]; 300 Yongbong-dong, Buk-gu, Gwangju 500-757 (KR).

## (72) Inventors; and

(75) Inventors/Applicants (for US only): **JO, Dae Woong** [KR/KR]; Kumho Town Apt. 105-303, 1130 Pungam-dong, Seo-gu, Gwangju 502-776 (KR). **KO, Jae Sun** [KR/KR]; Chungam Apt. 1013-1210, Yangsan-dong, Buk-gu, Gwangju 500-200 (KR). **KIM, Jin Sook**[KR/KR]; Geumgwang Apt. 103-904, Wolgye-dong, Gwangsan-gu, Gwangju 506-708 (KR). **PARK, Kyung Mi** [KR/KR]; Dundok Jugong Apt. 14-501, 464 Dundok-dong, Yeosu-si, Jeollanam-do 550-250 (KR). **SONG, Jin Kyung** [KR/KR]; Dongcheonmaeul 6-danji Jugong Apt. 609-508, 1662 Unam-dong, Buk-gu, Gwangju 500-921 (KR). **LIM, Jung Hee** [KR/KR]; 176-9 Tongsa-ri, Gaejeong-myeon, Gunsan-si, Jeollabuk-do 573-923 (KR). **DO, Thi Thuy Nga** [VN/KR]; Booyoung Apt. 607-310, Singi-ri, Hwasun-eup, Hwasun-gun, Jeollanam-do 519-808 (KR). **DO, Thi Lan Phuong** [VN/KR]; Booyoung Apt. 607-310, Singi-ri, Hwasun-eup, Hwasun-gun, Jeollanam-do 519-808 (KR). **DUONG, Minh Tam** [VN/KR]; Booyoung Apt. 607-310, Singi-ri, Hwasun-eup, Hwasun-gun, Jeollanam-do 519-808 (KR).(74) Agents: **CHANG, Soo Kil** et al.; Kim & Chang, Seyang B/D, 223 Naeja-dong, Jongno-gu, Seoul 110-720 (KR).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN,

[Continued on next page]

(54) Title: NOVEL MACROMOLECULE TRANSDUCTION DOMAINS AND METHODS FOR IDENTIFICATION AND USES THEREOF

Fig. 1a

| Red: hydrophobic Ala<br>Black: hydrophilic Ala<br>Green: nonpolar Ala<br>Blue: basic Ala<br>No: No helix structure |                                                                                              | HRSS: Hydrophobic Region of Signal Sequence<br>Helix: Bird's eye view of peptide structure (determined by SOSUI program)<br>SOSUI analysis requires at least 20 amino acids-length peptide. To make 20 Ala-length, additional Ala derived from N-terminus of E-GFP was added to the end of the peptide.<br>The numbers of HRSS represent the amino acid number of the origin protein.<br>The number of preliminarily modified and finally modified sequences represents the amino acid-length of the peptide. |                                    |       |                                     |       |                               |       |
|--------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------|-------|-------------------------------------|-------|-------------------------------|-------|
| #                                                                                                                  | Origin                                                                                       | Hydropathy                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | HRSS<br>(Candidate Domain)         | Helix | Preliminarily<br>modified Sequences | Helix | finally modified<br>Sequences | Helix |
| JO-001                                                                                                             | CACG4038 putative<br>NLPIP3-family<br>saccharide protein<br>[Strophomena<br>aculeolata/3(2)] |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | 7<br>VLTTAVTVCAITVL<br>AAPG 25     |       | 14<br>VLAVVCAIVLAAP                 | No    | AVVCAIVLAAP 12                | No    |
| JO-002                                                                                                             | NP_057021<br>phosphatidylinositol<br>glycan, class T<br>precursor [Homo<br>sapiens]          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | 1<br>MAAAMPLALLVLLL<br>GPGGWCLA 23 |       | 17<br>AAAPLALLVLLGLP<br>LA 17       |       | PLALLVLLGLP 12                | No    |
| JO-003                                                                                                             | NP_072471<br>oligodactylin<br>somatomedinotropin<br>homone-2 isoform 3<br>[Homo sapiens]     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | 1<br>MAAGSRSTLLAFAL<br>LCLPWLGE 23 |       | 14<br>AAGLLAFALLCLP                 | No    | LLAFALLCLP 11                 |       |

(57) Abstract: The present invention discloses novel macromolecule transduction domain (MTD) peptides which facilitate the traverse of a biologically active molecule across the cell membrane. Also disclosed are polynucleotides encoding the MTD peptides, methods of identifying the MTD peptides; methods of genetically engineering a biologically active molecule to have cell permeability by using the MTD peptides, methods of importing a biologically active molecule into a cell by using the MTD peptides, and uses thereof.



MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, NO, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

**(84) Designated States** (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI,

**Published:**

- *with international search report*
- *with sequence listing part of description published separately in electronic form and available upon request from the International Bureau*

## NOVEL MACROMOLECULE TRANSDUCTION DOMAINS AND METHODS FOR IDENTIFICATION AND USES THEREOF

### 【Technical Field】

5           The present invention relates to novel macromolecule transduction domain (MTD) peptides which facilitate the traverse of a biologically active molecule across the cell membrane, polynucleotides encoding the same, methods of identifying the same, methods of genetically engineering a biologically active molecule having cell permeability by using the same, 10 methods of importing a biologically active molecule into a cell by using the same, and uses thereof.

### 【Background Art】

Cellular internalization of macromolecules, such as DNA, RNA, 15 proteins, oligonucleotides, and peptides, is still a challenging task because of the presence of the plasma membrane, which constitutes an impermeable barrier for such molecules. Numerous difficulties have been encountered in delivering such molecules to a desired target, including poor penetration into a tissue or cell, toxicity when delivered systemically due to the insufficient 20 specificity of targeting to a particular tissue or cell, side effects when delivered in a high concentration in order to achieve an adequate local concentration at a particular target cell or tissue, and degradation such that inadequate amounts are delivered to the target and/or such that byproducts of degradation result in undesirable side effects.

25           In order to circumvent these problems, several carrier-mediated delivery systems have been developed. Among them, much attention has recently been given to the use of peptide-based delivery systems. The use of peptides with cell permeability has several advantages, which are mainly due to the various modifications that can be made to the peptide sequence. This allows the 30 engineering of carriers that can address different cellular subdomains and/or are

able to transport various types of cargo molecules.

Many cell permeable peptides are designed from sequences of membrane-interacting proteins, such as recombinant proteins, signal peptides, transmembrane domains, and antimicrobial peptides. Within these sequences, short sequences called protein transduction domains (PTDs) have been proved to efficiently cross biological membranes without the need of a carrier or a receptor and to deliver peptides or proteins into intracellular compartments. A number of studies have suggested that the use of PTD-based peptides could be of major importance for therapies against viral diseases or cancers. Among the PTD-based peptides, the third helix of the homeodomain of antennapedia called penetratin (Joliot, A. and A. Prochiantz, *Nat. Cell Biol.* 6(3):189-96 (2004)), the Tat peptide derived from the transactivating protein Tat of HIV-1 (Wadia, J.S. and S.F. Dowdy, *Curr. Opin. Biotechnol.* 13(1):52-6 (2002)), transportan (Pooga et al., *Faseb J.* 12(1):67-77 (1998)), and VP22 (Elliott, G. and P. O'Hare, *Cell* 88(2):223-33 (1997)) have been shown to improve the cellular uptake of peptides, proteins, and oligonucleotides.

A second category of cell-penetrating peptides, called amphipathic peptides, has also been described. An amphipathic molecule can be defined, in short, as consisting of two domains: a hydrophilic (polar) domain and a hydrophobic (non-polar) domain. For peptides, the amphipathic character can arise from either the primary structure or the secondary structure. Primary amphipathic peptides can be defined as the sequential assembly of a domain of hydrophobic residues with a domain of hydrophilic residues. Secondary amphipathic peptides are generated by the conformational state which allows the positioning of the hydrophobic and hydrophilic residues on opposite sides of the molecule.

Other peptides, such as polyarginine-based peptides, calcitonin-derived peptides, and oligomers, have also been proposed as tools for intracellular delivery of therapeutics.

However, the currently known delivery systems appear to be limited

due to their lack of efficiency and/or their toxicity, and little is known about the pathway of their cellular uptake, constituting a handicap for improving their efficiency. In addition, a number of delivery systems are limited in their ability to cross cellular and nuclear membranes. Even where such delivery  
5 peptides do cross cell membranes, they are often limited in their efficacies due to their entrapment in endosomes.

The present invention is directed to overcoming these deficiencies in the art.

10 **【Disclosure】**

**【Technical Solution】**

The present invention relates to isolated macromolecule transduction domain (MTD) peptides capable of mediating the transport of a biologically active molecule into a cell.

15 Another aspect of the present invention relates to isolated polynucleotides encoding such MTD peptides.

The present invention also relates to a method of identifying a MTD peptide having cell permeability.

Another aspect of the present invention relates to a method of  
20 genetically engineering a biologically active molecule having cell permeability by attaching a MTD peptide to a biologically active molecule.

Still another aspect of the present invention relates to an isolated recombinant protein having cell permeability comprising a MTD peptide having cell permeability and a biologically active molecule.

25 The present invention also relates to a pharmaceutical composition for delivery of a biologically active molecule to a cell comprising a cell permeable recombinant protein where a MTD is attached to a biologically active molecule.

Another aspect of the present invention relates to a method of transporting a biologically active molecule into a cell in a subject comprising  
30 administering to a subject a cell permeable recombinant protein comprising a

MTD peptide attached to a biologically active molecule.

Other aspects of the present invention relate to the use of MTD peptides for drug delivery, vaccine administration, protein therapy, and gene therapy.

5     **【Advantageous Effects】**

The present invention provides MTD peptides having cell permeability which can facilitate the transport of a biologically active molecule into a cell and, thus, can be effectively used for drug delivery, vaccine administration, peptide therapy, and gene therapy.

10

**【Description of Drawings】**

Figs. 1a to 1o show tables illustrating the amino acid sequences and structural characteristics of macromolecule transduction domain (MTD) peptides identified according to the method of the present invention.

15     Fig. 2 is a schematic diagram illustrating the His-MTD-EGFP recombinant protein having cell permeability and constructed according to the present invention.

20     Figs. 3a to 3c are photographs of an agarose gel electrophoresis analysis showing DNA fragments encoding MTDs-EGFP amplified by PCR according to the present invention.

25     Fig. 4a is a schematic diagram illustrating the subcloning of a DNA fragment encoding MTD-EGFP into the pGEM-Teasy vector according to the present invention, whereas Figs. 4b to 4d are photographs of an agarose gel electrophoresis analysis showing DNA fragments encoding MTDs-EGFP subcloned into pGEM-Teasy vector according to the present invention.

30     Fig. 5a is a schematic diagram illustrating the cloning of a DNA fragment encoding MTD-EGFP into the pET 28(+) vector according to the present invention, whereas Figs. 5b to 5c are photographs of an agarose gel electrophoresis analysis showing DNA fragments encoding MTDs-EGFP cloned into the pET 28(+) vector according to the present invention.

Figs. 6a and 6b are photographs of a SDS-PAGE analysis illustrating the inducible expression of His-MTD-EGFP recombinant proteins according to the present invention.

5 Figs. 7a and 7b are photographs of a SDS-PAGE analysis showing the purity of His-MTD-EGFP recombinant proteins under denaturing conditions according to the present invention.

Fig. 8 is a photograph of a SDS-PAGE analysis showing the purification of His-MTD-EGFP recombinant proteins under renaturing conditions according to the present invention.

10 Figs. 9a to 9g are graphs illustrating the cell permeabilities of His-MTD-EGFP recombinant proteins analyzed by flow cytometry according to the present invention.

Figs. 10a to 10i are photographs visualizing the cell permeabilities of His-MTD-EGFP recombinant proteins by confocal laser scanning microscopy according to the present invention.

Figs. 11a and 11b are graphs comparing the cell permeabilities of His-MTD-EGFP recombinant proteins according to the present invention with a positive control.

20 Figs. 12a to 12i are photographs visualizing the *in vivo* distribution of His-MTD-EGFP recombinant proteins by fluorescence microscopy according to the present invention.

Figs. 13a to 13k show tables summarizing the characteristics of the MTD peptides identified according to the method of the present invention.

## 25 **【Detailed Description】**

The present invention relates to novel macromolecule transduction domain (MTD) peptides having cell permeability that facilitate the traverse of a biologically active molecule across the cell membrane. The MTD peptides of the present invention are cell permeable polypeptides capable of mediating the import of a biologically active molecule, including polypeptides, protein

30

domains, or full-length proteins, through the cell membrane. The MTD peptides of the present invention are characterized as having a single hydrophobic region at their N-terminus, forming a helix structure, exhibiting flexibility, and having relatively short amino acids (7 to 17 amino acids) in length (see Figs. 1a to 1o).

One embodiment of the present invention relates to isolated MTD peptides having cell permeability, where the MTD peptide has an amino acid sequence selected from the group consisting of SEQ ID NOS: 1-193, analogs, derivatives, amidated variations, and conservative variations thereof.

A person having ordinary skill in the art can make similar substitutions to obtain peptides having greater cell permeability and a broader host range. For example, the present invention provides peptides corresponding to amino acid sequences SEQ ID NOS: 1 to 193, as well as analogues, derivatives, and amidated derivatives thereof, as long as the cell permeability of the peptide remains. Minor modifications to the primary amino acid sequence of the peptides of the present invention may result in peptides which have substantially equivalent or enhanced cell permeability, as compared to the specific peptides described herein. Such modifications may be deliberate, as by site-directed mutagenesis, or may be spontaneous. All of the peptides produced by these modifications are included herein, as long as the cell permeability of the original peptide still exists or, in the case of amidated versions of the peptide, the cell permeability of the original peptide is enhanced or altered such that the amidated peptide is therapeutically useful. It is envisioned that such modifications are useful for altering or enhancing the cell permeability of a particular peptide.

Further, the deletion of one or more amino acids can also result in a modification to the structure of the resultant molecule without any significant change in its cell permeability. This can lead to the development of a smaller active molecule which may also have utility. For example, amino- or carboxy-terminal amino acids which may not be required for the cell permeability of a

particular peptide can be removed. The peptides of the present invention include any analog, homolog, mutant, isomer, or derivative of the peptides disclosed in the present application, so long as the cell permeability, as described herein, remains. All peptides were synthesized using L-amino acids; however, D forms of all of the peptides can be synthetically produced. In addition, C-terminal derivatives, such as C-terminal methyl esters and C-terminal amidates, can be produced, in order to increase the cell permeability of the peptide of the present invention.

A "peptide" of the present invention includes amino acid sequences that are conservative variations of those peptides specifically exemplified herein. The term "conservative variation" as used herein denotes the replacement of an amino acid residue by another, biologically similar residue. Examples of conservative variations include the substitution of one hydrophobic residue, such as isoleucine, valine, leucine, alanine, cysteine, glycine, phenylalanine, proline, tryptophan, tyrosine, norleucine, or methionine for another, or the substitution of one polar residue for another, such as the substitution of arginine for lysine, glutamic for aspartic acids, or glutamine for asparagine, and the like. Neutral hydrophilic amino acids which can be substituted for one another include asparagine, glutamine, serine, and threonine. The term "conservative variation" also includes the use of a substituted amino acid in place of an unsubstituted parent amino acid, provided that antibodies raised to the substituted polypeptide also immunoreact with the unsubstituted polypeptide. Such conservative substitutions are within the definition of the classes of the peptides of the present invention.

Another aspect of the present invention relates to isolated polynucleotides encoding the above MTD peptides of the present invention. Exemplary polynucleotides encode the MTD peptides having an amino acid sequence selected from the group consisting of SEQ ID NOS: 1 to 193 and analogs, derivatives, amidated variations, and conservative variations thereof. The polynucleotides encoding the MTD peptides of the present invention have

nucleotide sequences represented by SEQ ID NOs: 194 to 386.

The polynucleotides of the present invention may be in the form of a RNA or DNA, where DNA includes cDNA and synthetic DNA. The DNA may be single stranded or double stranded. If it is single stranded, it may be the coding strand or non-coding (antisense) strand. The coding sequence may be identical to the nucleotide sequence selected from SEQ ID NOs:194 to 386 or may be a different coding sequence, where the coding sequence, as a result of degeneracy or redundancy of the genetic code, encodes for the same polypeptide.

The polynucleotides of the present invention also include variants of the above-described polynucleotides, which encode fragments, analogs, and derivatives of the polynucleotides characterized by the deduced amino acid sequences of SEQ ID NOs: 1 to 193. The variant of the polynucleotide may be a naturally occurring allelic variant of the polynucleotide or a non-naturally occurring variant of the polynucleotide.

The polynucleotide of the present invention may have a coding sequence which is a naturally occurring allelic variant of the coding sequence characterized by the nucleotide sequence selected from SEQ ID NOs: 194 to 386. An allelic variant is an alternate form of a polynucleotide sequence which may have a substitution, deletion, or addition of one or more nucleotides, which does not substantially alter the function of the encoded polypeptide.

Sequences similar or homologous (e.g., at least about 70% sequence identity) to the sequences disclosed herein are also part of the present invention. In other embodiments of the present invention, the sequence identity at the amino acid level can be about 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or higher. At the polynucleotide level, the sequence identity can be about 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or higher. Alternatively, substantial identity exists when the polynucleotide segments hybridize under selective hybridization conditions (e.g., very high stringency hybridization conditions), to the

complement of the strand. The polynucleotides may be present in whole cells, in a cell lysate, or in a partially purified or substantially pure form.

It is well known in the art that a single amino acid may be encoded by more than one nucleotide codon and that the polynucleotide may be easily  
5 modified to produce an alternate polynucleotide that encodes the same peptide. Therefore, other embodiments of the present invention include alternate nucleotide sequences encoding peptides containing the amino acid sequences as previously described. Nucleic acid molecules encoding peptides containing the claimed amino acid sequence include nucleotide sequences which encode  
10 any combination of the claimed sequence and any other amino acids located N-terminal or C-terminal to the claimed amino acid sequence.

It is to be understood that amino acid and nucleic acid sequences of the present invention may include additional residues, particularly N- or C-terminal amino acids or 5' or 3' polynucleotides, and still be essentially as set forth in the  
15 sequences disclosed herein, as long as the sequence confers membrane permeability upon the polypeptide or protein moiety of the recombinant protein.

The present invention also relates to a method of identifying the MTD peptides having cell permeability.

The method of the present invention involves the following steps:

- 20 1) identifying secreted proteins having a signal sequence-like domain from multiple amino acid sequence databases;
- 2) selecting from said identified secreted proteins hydrophobic peptides having a single hydrophobic region at their N-terminus and forming a helix structure; and
- 25 3) optimizing and minimizing said selected hydrophobic peptides under conditions effective to produce peptides suitable for use as an MTD peptide.

In order to identify certain candidate peptides that can potentially penetrate the plasma membrane, secreted proteins that have a signal sequence-  
30 like domain are identified from multiple amino acid sequence databases, as

described in step 1). In a specific embodiment of the present invention, these secreted proteins are selected from the PubMed Entrez Protein Database by using several key words, such as "hydrophobic region of signal sequence," "signal sequence hydrophobic region," "signal sequence of secreted protein," "hydrophobic signal sequence," and "hydrophobic domain of secreted protein." As a result, more than 1,500 secreted proteins that have a signal sequence-like domain are identified.

As used herein, the term "signal sequence-like domain" refers to a peptide capable of mediating macromolecule translocation across the plasma membrane.

The term "signal peptide" as used herein is a short (3-60 amino acids long) peptide chain that directs the post-translational transport of a protein. Signal peptides may also be called targeting signals, signal sequences, transit peptides, or localization signals. The amino acid sequences of signal peptides direct proteins (which are synthesized in the cytoplasm) to certain organelles such as the nucleus, mitochondrial matrix, endoplasmic reticulum, chloroplast, apoplast and peroxisome. Some signal peptides are cleaved from the protein by signal peptidase after the proteins are transported. The N-terminal sequence of a secreted protein, which is required for transport through the cell membrane.

Examples of signal peptides may include, but are not limited to, transport to the nucleus (NLS: -Pro-Pro-Lys-Lys-Lys-Arg-Lys-Val-), transport to the endoplasmic reticulum (H<sub>2</sub>N-Met-Met-Ser-Phe-Val-Ser-Leu-Leu-Leu-Val-Gly-Ile-Leu-Phe-Trp-Ala-Thr-Glu-Ala-Glu-Gln-Leu-Thr-Lys-Cys-Glu-Val-Phe-Gln-), retention to the endoplasmic reticulum (-Lys-Asp-Glu-Leu-COOH), transport to the mitochondrial matrix (H<sub>2</sub>N-Met-Leu-Ser-Leu-Arg-Gln-Ser-Ile-Arg-Phe-Phe-Lys-Pro-Ala-Thr-Arg-Thr-Leu-Cys-Ser-Ser-Arg-Tyr-Leu-Leu-), transport to the peroxisome (PTS1: -Ser-Lys-Leu-COOH), transport to the peroxisome (PTS2: H<sub>2</sub>N-----Arg-Leu-X5-His-Leu-) (wherein, H<sub>2</sub>N is the N-terminus of a protein, and COOH is the C-Terminus of a protein).

Examples of signal sequence-like domains may include, but are not limited to, penetratin (Arg-Gln-Ile-Lys-Ile-Trp-Phe-Gln-Asn-Arg-Arg-Met-Lys-Trp-Lys-Lys), the Tat peptide derived from the transactivating protein Tat of HIV-1 (Arg-Lys-Lys-Arg-Arg-Gln-Arg-Arg-Pro), transportan (Gly-Trp-Thr-Leu-Asn-Ser-Ala-Gly-Tyr-Leu-Leu-Gly-Lys-Ile-Asn-Lys-Ala-Leu-Ala-Ala-Leu-Ala-Lys-Lys-Ile-Leu), Buforin II (Thr-Arg-Ser-Ser-Arg-Ala-Gly-Leu-Gln-Phe-Arg-Val-Gly-Arg-Val-His-Arg-Leu-Leu-Arg-Lys), MAP (model amphiphatic peptide: Lys-Leu-Ala-Leu-Lys-Ala-Ser-Leu-Lys-Ala-Leu-Lys-Ala-Ala-Leu-Lys-Leu-Ala), k-FGF (Ala-Ala-Val-Ala-Leu-Leu-Pro-Ala-Val-Leu-Leu-Ala-Leu-Leu-Ala-Pro), Ku 70 (Val-Pro-Met-Leu-Lys-Pro-Met-Leu-Lys-Glu), prion (Met-Ala-Asn-Leu-Gly-Tyr-Trp-Leu-Leu-Ala-Leu-Phe-Val-Thr-Met-Trp-Thr-Asp-Val-Gly-Leu-Cys-Lys-Lys-Arg-Pro-Lys-Pro), pVEC (Leu-Leu-Ile-Ile-Leu-Arg-Arg-Arg-Ile-Arg-Lys-Gln-Ala-His-Ala-His-Ser-Lys), pep-1 (Lys-Glu-Thr-Trp-Trp-Glu-Thr-Trp-Trp-Thr-Glu-Trp-Ser-Gln-Pro-Lys-Lys-Lys-Arg-Lys-Val), SynB1 (Arg-Gly-Gly-Arg-Leu-Ser-Tyr-Ser-Arg-Arg-Arg-Phe-Ser-Thr-Ser-Thr-Gly-Arg), pep-7 (Ser-Asp-Leu-Trp-Glu-Met-Met-Met-Val-Ser-Leu-Ala-Cys-Gln-Tyr), and HN-1 (Thr-Ser-Pro-Leu-Leu-Ile-His-Asn-Gly-Gln-Lys-Leu).

Next, as described in step 2), hydrophobic peptides that have a single hydrophobic region at their N-terminus and form a helix structure are selected from the secreted proteins identified in step 1). All of the secreted proteins having a signal sequence-like domain that were identified in step 1) are subjected to hydropathy analysis to determine whether they contain a single hydrophobic region (H-region) at their N-terminus, followed by computer-aided genomic and proteomic information analysis to determine whether they form a helix structure. The hydrophobic region usually forms a helix structure, which imparts the protein with membrane-translocating activity.

In one embodiment of the present invention, the SOSUI system, a classification and secondary prediction system for membrane proteins, can be used for the computer-aided genomic and proteomic information analysis.

The SOSUI system is a useful tool for secondary structure prediction of membrane proteins from a protein sequence and is freely available on-line (see <http://bp.nuap.nagoya-u.ac.jp/sosui>). Since the SOSUI system currently requires peptides to be analyzed to have at least about 20 amino acids in length, additional amino acids derived from the N-terminal domain of enhanced green fluorescent protein are added to the end of the signal sequence-like domain. When the SOSUI system is used, 220 hydrophobic peptides that have a single hydrophobic region at their N-terminus and form a transmembrane helix structure are selected.

Finally, the hydrophobic peptides selected in step 2) are optimized and minimized so that they are suitable for use as an MTD peptide, as described in step 3). To optimize the selected hydrophobic peptides, empirical modifications are made to them as follows: i) hydrophilic, nonpolar, and positively or negatively charged amino acids are removed from the selected hydrophobic peptides; and ii) various combinations of five hydrophobic amino acids, i.e., alanine (Ala or A), valine (Val or V), proline (Pro or P), leucine (Leu or L), and isoleucine (Ile or I) are created empirically. Said hydrophobic amino acids are known to provide flexibility to the H-region which may allow the hydrophobic region of the signal sequences to form a hairpin-like loop, resulting in destabilized conformational change of the phospholipid bilayer. Such local and transient formation of the non-bilayer lipid structure induced by contact with the flexible H-region leads to topological transformation.

To minimize the length of the hydrophobic peptides optimized as described above, the following principles are employed for determining which amino acids should be eliminated therefrom: i) non-hydrophobic amino acids at the left or the right side of the hydrophobic peptides are eliminated to minimize the size, while maintaining the hydrophobic region thereof; ii) non-hydrophobic amino acids at the middle or the right side of the hydrophobic peptides are replaced with a hydrophobic amino acid, proline (Pro or P) to provide flexibility or a bending potential; and iii) the number of repeated hydrophobic

amino acids is minimized to reduce the size of the hydrophobic peptides. Thus, the MTD candidates are characterized as modified peptides exhibiting hydrophobicity and flexibility, having relatively short amino acids (7 to 17) in length, and forming a helix structure.

5 According to the method of the present invention described above, novel MTD peptides having an amino acid sequence selected from the group consisting of SEQ ID NOS: 1 to 193 and capable of facilitating the traverse of biologically active molecules across the cell membrane can be identified.

10 In another embodiment of the present invention, one of the novel MTDs, JO-98 (SEQ ID NO: 98) can be identified by the following process:

i) immunoglobulin domain containing 4 from *Homo sapiens* is selected from the PubMed Entrez Protein Database as a secreted protein having a signal sequence-like domain;

15 ii) through a hydropathy analysis, it is confirmed that said protein contains a single N-terminal hydrophobic region;

iii) the hydrophobic region (H-region) is selected from the full-length signal sequences of said protein, which has the amino acid sequence represented by MGILLGLLLLGHILTVDY  
GRPIL (23 amino acids in length);

20 iv) through a structural analysis utilizing the SOSUI system, it is confirmed that the selected hydrophobic region forms a helix structure;

v) non-hydrophobic amino acids (methionine: M; glycine: G) at the left side of the hydrophobic region are deleted to make the first amino acid of the hydrophobic region a hydrophobic amino acid, hydrophobic peptide that has an  
25 amino acid sequence represented by ILLGLLLLGHILTVDYGRPIL (21 amino acids in length);

vi) basic (histidine: H; arginine: R) and hydrophilic (threonine: T; aspartic acid: D; tyrosine: Y) amino acids are removed from the hydrophobic region to reduce the peptide length to no more than 15 amino acids, resulting in  
30 a peptide that has an amino acid sequence represented by ILLGLLLLGLVGPIL

(15 amino acids in length);

vii) the structural analysis shows that the modified peptide, as described above, forms a helix structure;

viii) in order to impart flexibility to the peptide, nonpolar amino acids (glycine: G) at the middle and the right sides of the peptide are replaced with a hydrophobic amino acid, proline (P), respectively, resulting in a peptide has an amino acid sequence represented by ILLPLLLLGLVPPIL (15 amino acids in length); and

ix) some of the repeated hydrophobic amino acids at the far right side of the peptide sequence are deleted to shorten the peptide length to 10 or less amino acids.

Therefore, the finally optimized peptide as a macromolecule transduction domain has an amino acid sequence represented by ILLPLLLLP (9 amino acids in length; SEQ ID NO: 98) and forms a helix structure, which has been designated as "JO-98." The last amino acid, proline, of the JO-98 peptide is a flexible site for interaction with a biologically active molecule to form a recombinant protein.

A further aspect of the present invention relates to a method of genetically engineering a biologically active molecule to have cell permeability by using the MTD peptides.

The therapeutic use of biologically active molecules is often hindered by their low cell permeability. Although biologically active molecules have been shown to be taken up by cells via an endocytic process, the molecules that enter the cell in this manner are usually trapped in endocytic vesicles and degraded in lysosomes.

The MTD of the present invention provides efficient transport of biologically active molecules having a high molecular weight across the cell membrane, whereas other membrane transport peptides previously tested have not been shown to transport molecules of a size greater than approximately 25 amino acids. The MTD of the present invention can be attached to a peptide

or polypeptide using conventional methods in the art to enhance the peptide's or polypeptide's membrane permeability. The MTD can be provided in the form of a kit, which includes the necessary components known to those skilled in the art to facilitate linkage of a peptide to a target polypeptide. A target protein  
5 attached to the MTD in this manner can then be delivered to the cell either *in vitro* or *in vivo* for intracellular import.

According to the method of the present invention, the polynucleotide described above can be inserted into a protein expression vector to produce a protein which can be imported from the exterior to the interior of a cell by the  
10 action of the MTD peptides described herein.

An expression vector is genetically engineered to incorporate a nucleic acid sequence encoding a MTD in an orientation either N-terminal and/or C-terminal to a nucleic acid sequence encoding a peptide, polypeptide, protein domain, or full-length protein of interest as a biologically active molecule, and  
15 in the correct reading frame so that a recombinant protein consisting of the macromolecule transduction domain and the target biologically active molecule may be expressed. Expression vectors may be chosen from among those readily available for use in prokaryotic or eukaryotic expression systems.

As used herein, an MTD is a macromolecule transduction domain of the  
20 present invention, which directs cellular transport of a target protein from the exterior to the interior of a cell. In another embodiment of the present invention, the MTD may comprise an alternate sequence which mediates the import of a peptide or polypeptide through the cell membrane to the interior of a cell.

25 A target protein is a protein which normally exhibits less than optimal permeability through the cell membrane, but which, when attached either N-terminal and/or C-terminal to an MTD of the present invention, is transported from the exterior to the interior of the cell.

In another embodiment of the present invention, a cleavage site is  
30 located between the MTD and the target polypeptide, protein domain, or full-

length protein. This site may alternatively be a factor X site or another site that is known to those skilled in the art to relate to the cleavage of the recombinant protein to physically remove the MTD from the subject peptide or polypeptide.

5           The method of the present invention provides a means for producing proteins having cell permeability for introduction into the interior of the cell, where their actions help to further elucidate cellular control and biosynthesis mechanisms. This method also provides a means for introducing intracellular proteins into cells to produce targeted cellular changes, such as the inhibition of  
10       apoptosis by the introduction of Bcl-2. Cell cycle control, for example, can be altered by the introduction of a functional p53 protein product into those cells that have become tumorigenic due to an abnormal p53 protein.

          As used herein, the term "biologically active molecule" includes any molecule which, if imported into a cell, is capable of exhibiting a biological  
15       effect. Since very large proteins having molecular weights ranging from about 100,000 to about 1 million are exported by cells (e.g., antibodies, fibrinogen, and macroglobulin), very large proteins can be imported into cells by the method of the present invention. Therefore, proteins having sizes ranging from a few amino acids to about a thousand amino acids can be used. A  
20       preferable size range for proteins is from a few amino acids to about 250 amino acids. For any molecule, the size can be up to a molecular weight of about 1 million, specifically up to a molecular weight of about 25,000, and more specifically up to a molecular weight of about 3,000. In addition, only those molecules which can be attached to an MTD peptide as a signal peptide, either  
25       directly or indirectly, are within the scope of the present invention.

          Examples of biologically active molecules include, but are not limited to, proteins, polypeptides, and peptides, which include functional domains of biologically active molecules, such as growth factors, enzymes, transcription factors, toxins, antigenic peptides (for vaccines), antibodies, and antibody  
30       fragments. Additional examples of biologically active molecules include

nucleic acids, such as plasmids, coding nucleic acid sequences, mRNAs and antisense RNA molecules, carbohydrates, lipids, and glycolipids. Further examples of biologically active molecules include therapeutic agents, in particular those with low cell membrane permeability. Some examples of these therapeutic agents include cancer drugs, such as Daunorubicin, and toxic chemicals which, because of the lower dosage that can be used when administered by this method, can now be more safely administered. Yet another example of a biologically active molecule is an antigenic peptide. Antigenic peptides can be administered to provide immunological protection when imported by cells involved in the immune response. Other examples include immunosuppressive peptides (e.g., peptides that block autoreactive T cells, which are known in the art). Numerous other examples will be apparent to the skilled artisan.

Representative examples of biologically active molecules suitable for the present invention may include enzymes, hormones, transport proteins, immunoglobulin or antibodies, structural proteins, motor function proteins, receptors, signaling proteins and storage proteins in terms of their function; and membrane or transmembrane proteins, internal proteins, external or secret proteins, virus proteins, native proteins, glycoproteins, cleaved proteins, proteins with disulfide bonds, protein complexes, chemically modified proteins and prions in terms of their location and roles.

Standard recombinant nucleic acid methods can be used to express a genetically engineered recombinant protein. In one embodiment of the present invention, a nucleic acid sequence encoding the macromolecule transduction domain is cloned into a nucleic acid expression vector, e.g., with appropriate signal and processing sequences and regulatory sequences for transcription and translation. In another embodiment, the protein can be synthesized using automated organic synthetic methods. Synthetic methods for producing proteins are described, for example, in *Methods in Enzymology*, Volume 289: Solid-Phase Peptide Synthesis by Gregg B. Fields (Editor), Sidney

P. Colowick, Melvin I. Simon (Editor), Academic Press (1997).

In order to obtain high level expression of a cloned gene or nucleic acid, such as a cDNA encoding a MTD peptide, a MTD sequence is typically subcloned into an expression vector that contains a strong promoter for directing transcription, a transcription/translation terminator, and, in the case of a nucleic acid encoding a protein, a ribosome binding site for translational initiation. Suitable bacterial promoters are well known in the art and are described, e.g., in Sambrook & Russell, *Molecular Cloning: A Laboratory Manual*, 3d Edition, Cold Spring Harbor Laboratory, N.Y. (2001) and Ausubel et al., *Current Protocols in Molecular Biology*, Greene Publishing Associates and Wiley Interscience, N. Y. (1989). Bacterial expression systems for expressing the MTD peptide of the present invention are available in, e.g., *E. coli*, *Bacillus* sp., and *Salmonella* (Palva et al., *Gene* 22: 229-235 (1983); Mosbach et al., *Nature* 302: 543-545 (1983)). Kits for such expression systems are commercially available. Eukaryotic expression systems for mammalian cells, yeast, and insect cells are well known in the art and are also commercially available. In another embodiment of the present invention, the eukaryotic expression vector is an adenoviral vector, an adeno-associated vector, or a retroviral vector.

The expression vector for expressing the cell permeable recombinant protein can include regulatory sequences, including for example, a promoter, operably attached to a sequence encoding the macromolecule transduction domain. Non-limiting examples of inducible promoters that can be used include steroid-hormone responsive promoters (e.g., ecdysone-responsive, estrogen-responsive, and glucocorticoid-responsive promoters), the tetracycline "Tet-On" and "Tet-Off" systems, and metal-responsive promoters. The construct can be introduced into an appropriate host cell, e.g., a bacterial cell, yeast cell, insect cell, or tissue culture cell. The construct can also be introduced into embryonic stem cells to generate a transgenic organism as a model subject. Large numbers of suitable vectors and promoters are known to

those skilled in the art and are commercially available for generating the recombinant constructs of the present invention.

Known methods can be used to construct vectors containing the polynucleotide of the present invention and appropriate transcriptional/translational control signals. These methods include *in vitro* recombinant DNA techniques, synthetic techniques, and *in vivo* recombination/genetic recombination. See, for example, the techniques described in Sambrook & Russell, *Molecular Cloning: A Laboratory Manual*, 3d Edition, Cold Spring Harbor Laboratory, N. Y. (2001) and Ausubel et al., *Current Protocols in Molecular Biology* Greene Publishing Associates and Wiley Interscience, N.Y. (1989).

Host cells suitable for producing a cell permeable recombinant protein include bacterial cells and eukaryotic cells (e.g., fungal, insect, plant, and mammalian cells). Host cells can be disrupted by any convenient method, including freeze-thaw cycling, sonication, mechanical disruption, or the use of cell lysing agents. Scopes, *Protein Purification: Principles and Practice*, New York: Springer-Verlag (1994) describes a number of general methods for purifying recombinant (and non-recombinant) proteins. The methods can include, e.g., ion-exchange chromatography, size-exclusion chromatography, affinity chromatography, selective precipitation, dialysis, and hydrophobic interaction chromatography. These methods can be adapted to devise a purification strategy for the cell permeable recombinant protein. If the cell permeable recombinant protein includes a purification handle, such as an epitope tag or a metal chelating sequence, affinity chromatography can be used to purify the protein greatly.

The amount of protein produced can be evaluated by detecting the macromolecule transduction domain directly (e.g., using Western analysis) or indirectly (e.g., by assaying materials from the cells for specific DNA binding activity, such as by electrophoretic mobility shift assay). Proteins can be detected prior to purification, during any stage of purification, or after

purification. In some implementations, purification or complete purification may not be necessary.

In a specific embodiment of the present invention, the expression vector is pEGFP-C1 (Clontech, Mountain View, CA), which comprises a polynucleotide encoding a recombinant protein including an enhanced green fluorescent protein (EGFP). The insertion of a polynucleotide encoding an MTD according to the present invention into vector pEGFP-C1, 5' and/or 3' to the EGFP gene of the vector enables the expression of a recombinant protein incorporating both the MTD and the EGFP. The MTD, connected N-terminal and/or C-terminal to the EGFP, carries the EGFP protein to the interior of the cell.

The genetically engineered recombinant proteins prepared by the method of the present invention are cell permeable proteins and can be used as protein-based vaccines, particularly where killed or attenuated whole organism vaccines are impractical. The cell permeable proteins prepared by the method of the present invention can also be used for the treatment of diseases, particularly cancer. Cell permeable proteins can be delivered to the interior of the cell, eliminating the need to transfect or transform the cell with a recombinant vector. The cell permeable proteins of the present invention can be used *in vitro* to investigate protein function or can be used to maintain cells in a desired state.

The MTD of the present invention can be used to deliver peptides, polypeptides, protein domains, or proteins to the interior of a target cell either *in vitro* or *in vivo*. The MTD can be attached to the target protein through a peptide linkage formed by the expression of the recombinant protein from a recombinant DNA or RNA molecule or can be attached to the target protein by means of a linker covalently attached to the MTD. A covalent linkage can be used to attach the MTD of the present invention to a non-protein molecule, such as a polynucleotide, for import into the cell.

Another aspect of the present invention relates to a pharmaceutical

composition comprising the cell permeable recombinant proteins of the present invention, where the MTD peptide is operably attached to a biologically active molecule.

The cell permeable recombinant protein produced by the method of the present invention may be administered *in vitro* by any of the standard methods known to those skilled in the art, such as by addition of the recombinant protein to culture medium. Furthermore, it will be appreciated by those skilled in the art that recombinant proteins produced by this method may be delivered *in vivo* by standard methods utilized for protein/drug delivery, including parenteral administration, intravenous administration, topical administration, aerosol administration or inhalation, oral administration (particularly when provided in an encapsulated form), or by rectal or vaginal administration (particularly when provided in a suppository form).

Examples of administration include, but are not limited to, parenteral administration, e.g., by intravenous injection including regional perfusion, inhalation of an aerosol, subcutaneous or intramuscular injection, topical administration, such as to skin wounds and lesions, direct transfection into, e.g., bone marrow cells prepared for transplantation and subsequent transplantation into the subject, and direct transfection into an organ that is subsequently transplanted into the subject. Parenteral administration, e.g., regional perfusion, if used, is generally characterized by injection. Injectables can be prepared in conventional forms, such as liquid solutions, suspensions, or emulsions. A slow release or sustained release system can also be used, allowing the maintenance of a constant level of dosage.

Further administration methods include oral administration, particularly when the complex is encapsulated, or rectal administration, particularly when the complex is in a suppository form. A pharmaceutically acceptable carrier includes any material that is not biologically or otherwise undesirable, i.e., the material may be administered to an individual along with the selected complex without causing any undesirable biological effects or interacting in a deleterious

manner with any of the other components of the pharmaceutical composition in which it is administered.

Depending on the intended mode of administration (e.g., including, but not limited to, intravenous, parenteral, transcutaneous, subcutaneous, intramuscular, intracranial, intraorbital, ophthalmic, intraventricular, intracapsular, intraspinal, intracisternal, intraperitoneal, intranasal, intrarectal, intravaginal, aerosol, or oral), the pharmaceutical compositions may be in the form of solid, semi-solid or liquid dosage forms, such as, for example, tablets, suppositories, pills, capsules, powders, liquids, suspensions, lotions, creams, gels, or the like, preferably in unit dosage form suitable for single administration of a precise dosage. The pharmaceutical compositions of the present invention will include, as noted above, an effective amount of the cell permeable recombinant protein in combination with a pharmaceutically acceptable carrier and, in addition, may include other medicinal agents, pharmaceutical agents, carriers, adjuvants, diluents, etc.

For solid compositions, conventional nontoxic solid carriers include, for example, pharmaceutical grades of mannitol, lactose, starch, magnesium stearate, sodium saccharin, talc, cellulose, glucose, sucrose, magnesium carbonate, and the like. Liquid pharmaceutically administrable compositions can, for example, be prepared by dissolving or dispersing an active compound as described herein, and optional pharmaceutical adjuvants in an excipient, such as, for example, water, saline, aqueous dextrose, glycerol, ethanol, and the like, to thereby form a solution or suspension. If desired, the pharmaceutical composition to be administered may also contain minor amounts of nontoxic auxiliary substances, such as wetting or emulsifying agents, pH buffering agents, and the like. Actual methods of preparing such dosage forms are known, or will be apparent, to those skilled in the art.

Administration of a cell permeable recombinant protein produced by the method of the present invention may be performed for a period of time ranging from 10 minutes to 72 hours, particularly when the administration is carried

out by the addition of recombinant protein to culture media for *in vitro* use. For *in vivo* or *in vitro* use, the effective administration time for a cell permeable recombinant protein produced by the method of the present invention may be readily determined by those skilled in the relevant art.

5           For either *in vitro* or *in vivo* use, the cell permeable recombinant protein can be administered at any effective concentration. An effective concentration is that amount that results in the importation of the biologically active molecule into the cell. Such a concentration - culture medium concentration (*in vitro*) or blood serum concentration (*in vivo*) - will typically be from about 0.1 nM to  
10   about 500  $\mu$ M. Optimal concentrations for the specific recombinant protein and/or the specific target cell can be readily determined according to the teachings herein. Thus, the *in vivo* dosage for the recombinant protein is the concentration which will cause the blood serum concentration of the recombinant protein to be about 0.1 nM to about 500  $\mu$ M, more specifically,  
15   about 0.5 nM to about 100  $\mu$ M. The amount of the recombinant protein administered will, of course, depend upon the subject being treated, the subject's age and weight, the manner of administration, and the judgment of the skilled administrator. The exact amount of the complex will further depend upon the general condition of the subject, the severity of the disease/condition  
20   being treated by the administration, and the particular complex chosen. However, the appropriate amount can be determined by one of ordinary skill in the art, using routine optimization, given the teachings provided herein.

          Further, the cell permeable recombinant proteins produced by the method of the present invention can be used for vaccine administration  
25   (Linqnan et al., *Expert. Rev. Vaccines* 6: 741-746, 2007; O' Haqan et al., *Methods* 40: 10-19, 2009).

          Vaccines using cell permeable recombinant proteins provide an advantage over the typical peptide vaccines. The immune system recognition of an antigen depends upon appropriate antigen processing. Previously, entire  
30   proteins or protein domains could not be delivered to the interior of the cell for

processing to occur. As a consequence, peptides representing antigenic epitopes had to be identified prior to the delivery to the cell of small peptides representing those epitopes. The method of the present invention allows whole proteins or protein domains to be imported into the cell, where antigenic  
5 processing can occur. This provides multiple antigenic epitopes in one administration, and eliminates the need for experimental identification of specific epitopes for vaccine development.

Typically, such vaccines are prepared for injection into a human or mammalian subject. Injectable vaccines can be prepared as liquid solutions or  
10 suspensions. Solid forms can be prepared which are suitable for solution in, or suspension in, a liquid, prior to injection. The preparation may also be emulsified. The active immunogenic ingredient is often mixed with a pharmaceutically acceptable carrier which is compatible with the active ingredient. Suitable carriers include, but are not limited to, water, dextrose,  
15 glycerol, saline, ethanol, and combinations thereof. The vaccine may contain additional agents, such as wetting or emulsifying agents, pH buffering agents, or adjuvants which enhance the effectiveness of the vaccine.

The vaccine may be conventionally administered parenterally. Either subcutaneous or intramuscular injection is appropriate. Other modes of  
20 administration may include oral administration, nasal administration, rectal administration, and vaginal administration, which may involve combining the peptide immunogen with pharmaceutically acceptable carriers, such as mannitol, lactose, starch, magnesium stearate, sodium saccharine, cellulose, magnesium carbonate, or other carriers. Compositions for oral administration  
25 may form solutions, suspensions, tablets, pills, capsules, sustained release formulations, or powders. A protein-based vaccine of the present invention can be administered by an enteric-coated capsule for release of the polypeptide into the lumen of the intestine.

The cell permeable recombinant proteins of the present invention may  
30 be formulated into the vaccine as neutral or salt forms. Pharmaceutically

acceptable salts include the acid addition salts (formed with the free amino groups of the polypeptide) and which are formed with inorganic acids, such as, for example, hydrochloric or phosphoric acids, or such organic acids as acetic, mandelic, oxalic, and tartaric acids. Salts formed with the free carboxyl  
5 groups of the polypeptide may also be derived from inorganic bases such as, for example, sodium, potassium ammonium, calcium, or ferric hydroxides, and such organic bases as isopropylamine, trimethylamine, 2-ethylamino ethanol, and histidine.

The vaccine is administered in a manner compatible with the dosage  
10 formulation, and in such an amount as will be therapeutically effective and immunogenic. The quantity to be administered depends on the subject to be treated, taking into account, for example, the capacity of the individual's immune system to synthesize antibodies and the degree of protection desired. The precise amounts of the active ingredient (peptide immunogen) to be  
15 administered depend on the judgment of the practitioner. Suitable dosage ranges generally require several hundred micrograms of active ingredient per vaccination. Also variable are the regimes for initial administration and booster vaccinations, which should be determined according to the judgment of the practitioner. The dosage of the vaccine will depend on the route of  
20 administration and will vary according to the size of the host.

Furthermore, the cell permeable recombinant proteins of the present invention can be effectively used in drug delivery systems (Spencer and Verma, *Proc. Natl. Acad. Sci. USA* 104: 7594-7599, 2007; Choi et al., *Nat. Med.* 12: 574-579, 2006; Vives et al., *J. Biol. Chem.* 272: 16010-16017, 1997; Kumar et  
25 al., *Nature* 448: 39-43, 2007; Jo et al., *Nat. Med.* 11: 892-898, 2005).

The method of genetically engineering proteins with cell membrane permeability according to the present invention provides a means for delivering therapeutic protein products into a cell. Combination of the present invention with previously described methods of extracellular protein delivery provides a  
30 method of delivering proteins for import into a cell in a stabilized, functional

form in a controlled-release fashion.

Polypeptides are produced using an appropriate expression vector and an expression system. Cell permeability is conferred upon the protein or polypeptide by the expression of a recombinant protein with the MTD located  
5 N-terminal and/or C-terminal to the expressed polypeptide. The less stable proteins are stabilized by methods known to those skilled in the art and described previously. Delivery to the extracellular environment is accomplished by providing the stabilized recombinant protein in an appropriate carrier, such as microsphere carriers. The protein of choice will dictate the  
10 appropriate vector and expression system, as well as the appropriate stabilization and delivery techniques. A person of ordinary skill in the art of drug delivery systems can choose the appropriate techniques from among those described.

The method of the present invention provides a means for producing  
15 cell permeable proteins for the treatment of cancer. Regulators of apoptosis and cell cycle control have been found to play a key role in oncogenesis, while gene therapy techniques (Balaggan et al., *Gene Therapy* 13(15): 1153-1165, 2006; Kuo et al., *Proc. Natl. Acad. Sci. USA* 10; 98(8): 4605-4610, 2001), using intratumoral injection of an adenoviral expression vector encoding the p53 gene,  
20 have shown promise for the control of some tumors. However, the delivery of specific protein products through the use of viral vectors has proven to be problematic. The MTDs and methods of the present invention provide a means for producing cell permeable proteins from among the cell cycle regulators and regulators of apoptosis, as well as other proteins identified to  
25 play a role in the development of the cancer state.

For example, in one embodiment of the present invention, the polynucleotide encoding the p53 gene can be inserted into a suitable vector, either 5' or 3' to the sequence of the MTD of the present invention. Under expression conditions appropriate for the vector of choice and known to those  
30 skilled in the art, a recombinant protein comprising an MTD and the p53

protein can be expressed. The attachment of the MTD to the p53 protein renders the p53 protein cell permeable, and protein can be administered to tumor cells to inhibit tumor development. Administration of cell permeable proteins can be accomplished in various ways, including, but not limited to, intratumoral injection, infusion, and intravenous administration. Bax and Bcl-x<sub>L</sub> are other examples of proteins from among a wide variety of proteins that have been determined to affect cell cycle control and apoptosis and, therefore, be effective for cancer therapy. The method of the present invention provides a more efficient, less labor-intensive, and less costly method for the delivery of anti-oncogenic proteins to tumor cells.

The term "cell membrane" as used herein refers to a lipid-containing barrier which separates cells or groups of cells from the extracellular space. Cell membranes include, but are not limited to, plasma membranes, cell walls, intracellular organelle membranes, such as the mitochondrial membrane, nuclear membranes, and the like.

The term "biologically active molecule" as used herein refers to compounds or molecules that are capable of eliciting or modifying a biological response in a system. Non-limiting examples of biologically active molecules include antibodies (e.g., monoclonal, chimeric, humanized etc.), cholesterol, hormones, antivirals, peptides, proteins, chemotherapeutics, small molecules, vitamins, co-factors, nucleosides, nucleotides, oligonucleotides, enzymatic nucleic acids, antisense nucleic acids, triplex forming oligonucleotides, 2,5-A chimeras, siNA, siRNA, miRNA, RNAi inhibitors, dsRNA, allozymes, aptamers, decoys and analogs thereof. Biologically active molecules of the invention also include molecules capable of modulating the pharmacokinetics and/or pharmacodynamics of other biologically active molecules, for example, lipids and polymers such as polyamines, polyamides, polyethylene glycol and other polyethers. In certain embodiments, the term biologically active molecule is used interchangeably with the term "macromolecule".

The term "macromolecule" as used herein refers to large molecules

(molecular weight greater than 1000 daltons) exemplified by, but not limited to, peptides, proteins, and oligonucleotides and polynucleotides of biological or synthetic origin.

The term "peptide" as used herein refers to a compound made up of a single chain of D- or L-amino acids or a mixture of D- and L-amino acids joined by peptide bonds. Preferably, peptides contain at least two amino acid residues and are less than about 50 amino acids in length.

The term "protein" as used herein refers to a compound that is composed of linearly arranged amino acids attached by peptide bonds, but in contrast to peptides, has a well-defined conformation. Proteins, as opposed to peptides, preferably contain chains of 50 or more amino acids.

The term "polypeptide" as used herein refers to a polymer of at least two amino acid residues and which contains one or more peptide bonds. Polypeptides encompass peptides and proteins, regardless of whether the polypeptide has a well-defined conformation.

The term "nucleic acid" as used herein refers to oligonucleotides or polynucleotides, such as deoxyribonucleic acid (DNA) and ribonucleic acid (RNA), as well as analogs of either RNA or DNA, for example made from nucleotide analogs any of which are in single or double stranded form.

Amino acid residues are herein referred to by their standard single-letter or three-letter notations or by their full names: A, Ala, alanine; C, Cys, cysteine; D, Asp, aspartic acid; E, Glu, glutamic acid; F, Phe, phenylalanine; G, Gly, glycine; H, His, histidine; I, Ile, isoleucine; K, Lys, lysine; L, Leu, leucine; M, Met, methionine; N, Asn, asparagine; P, Pro, proline; Q, Gln, glutamine; R, Arg, arginine; S, Ser, serine; T, Thr, threonine; V, Val, valine; W, Trp, tryptophan; X, Hyp, hydroxyproline; and Y, Tyr, tyrosine.

The term "macromolecule transduction domain (MTD)" as used herein refers to a peptide that facilitates the transport of macromolecules into a cell.

The term "isolated" polynucleotide as used herein refers to a polynucleotide that is substantially free of the sequences with which it is

associated in nature. By substantially free is meant at least 50%, preferably at least 70%, more preferably at least 80%, and even more preferably at least 90% free of the materials with which it is associated in nature. As used herein, an "isolated" polynucleotide also refers to recombinant polynucleotides, which, by virtue of origin or manipulation: (1) are not associated with all or a portion of a polynucleotide with which it is associated in nature, (2) are linked to a polynucleotide other than that to which it is linked in nature, or (3) does not occur in nature.

The term "operatively attached" as used herein means that DNA fragments of a promoter, a MTD peptide or other genes are sufficiently connected to direct and regulate the expression of genes.

The term "homology" or "similarity" as used herein refers to the likeness or the percentage of identity between two sequences (e.g., polynucleotide or polypeptide sequences). Typically, a higher similarity in sequences means more similarity in the physical or chemical properties or biological activities.

Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any method and material similar or equivalent to those described herein can also be used in the practice or testing of the present invention, specific methods and materials are now described. All publications mentioned herein are incorporated herein by reference in their entirety to disclose and describe the methods and/or materials in connection with which the publications are cited.

It must be noted that as used herein and in the appended claims, the singular forms "a," "an," and "the" include plural referents unless the context clearly dictates otherwise.

The publications discussed herein are provided solely for their disclosure prior to the filing date of the present application. Nothing herein is to be construed as an admission that the present invention is not entitled to

antedate such publication by virtue of prior invention. Further, the dates of publication provided may be different from the actual publication dates which may need to be independently confirmed.

5 **[Examples]**

The following examples are presented to aid practitioners of the invention, provide experimental support for the invention, and to provide model protocols. In no way are these examples to be understood to limit the invention.

10

**Example 1 - Identification of Novel Macromolecule Transduction Domain Peptides**

In order to identify novel macromolecule transduction domain (MTD) candidates that can potentially penetrate the plasma membrane of live cells, secreted proteins having a signal sequence-like domain were selected from the PubMed Entrez Protein Database by using several key words, including "hydrophobic region of signal sequence," "signal sequence hydrophobic region," "signal sequence of secreted protein," "hydrophobic signal sequence," and "hydrophobic domain of secreted protein." As a result, more than 1,500 secreted proteins having a signal sequence-like domain were selected.

All of the selected secreted proteins having a signal sequence-like domain were subjected to hydropathy analysis to determine whether they contain a single hydrophobic region at their N-terminus. Subsequently, they were subjected to a computer-aided genomic and proteomic information analysis, i.e., the SOSUI system (as described on <http://bp.nuap.nagoya-u.ac.jp/sosui>), to determine whether they form a helix structure. Since the SOSUI system currently requires peptides to be analyzed to have at least about 20 amino acids in length, additional amino acids derived from the N-terminal domain of enhanced green fluorescent protein were added to the end of the signal sequence-liked domain. When the SOSUI system is used, 220 peptide

sequences having a single hydrophobic region at their N-terminus and forming a transmembrane helix structure were selected.

To optimize the selected peptide sequences as a signal sequence, empirical modifications were made to them as follows: first, hydrophilic, nonpolar, and positively or negatively charged amino acids were removed from the selected peptide sequences and, then, various combinations of five hydrophobic amino acids (i.e., alanine: A, valine: V, proline: P, leucine: L, isoleucine: I) were created empirically.

After that, non-hydrophobic amino acids at the left or right side of the peptide sequences were eliminated to minimize the size while maintaining their hydrophobic region, followed by adding or replacing non-hydrophobic amino acids at the middle or right side of the peptide sequences with proline, to thereby provide flexibility to the peptide. Finally, the number of repeated hydrophobic amino acids was minimized to reduce the length of the peptide sequence. The resulting peptides as MTD candidates exhibited hydrophobicity and flexibility, have relatively short amino acids (7 to 17 in length), and form a helix structure.

According to the method described above, novel 193 MTD peptides having amino acid sequences represented by SEQ ID NOs: 1 to 193, respectively, and are capable of facilitating the traverse of biologically active molecules across the cell membrane were identified. The 193 MTDs were designated as JO-001 to JO-193, respectively, and their structural characteristics and sequences are shown in Figs. 1a to 1o.

## **Example 2 - Expression of Recombinant Proteins Fused to MTDs**

Having demonstrated the feasibility of the cellular import of MTD peptides identified in Example 1 above, the enhanced green fluorescent protein (EGFP) was used as a full-length protein cargo molecule. Green fluorescent protein (GFP) is a protein cloned from the jellyfish, *Aequorea Victoria*. GFP is one of the most widely used reporter proteins and produces a green light when

illuminated with blue or UV light (Inouye et al., *FEBS Letters* 351(2): 211-14 (1994)). In one embodiment of the present invention, a commercially available GFP expression vector, pEGFP-C1 (Clontech), is used. The EGFP protein encoded by pEGFP-C1 is a mutant of wild-type GFP, which has been  
 5 modified to produce a stronger green light. Further, when a foreign gene is inserted into the multiple cloning sites of the pEGFP-C1 vector, the inserted foreign gene is expressed in the form of a recombinant protein with EGFP.

In order to construct expression vectors for EGFP proteins fused to each novel MTD (His-MTD-EGFP: HME) (Fig. 2), a polymerase chain reaction  
 10 (PCR) was conducted by using the oligonucleotides described in Tables 1 and 2 below as a primer set (forward primers SEQ ID NOs: 393 to 585 for MTD JO-1 to JO-193, respectively; reverse primer SEQ ID NO: 586 for MTD JO-1 to JO-193) and the pEGFP-C1 plasmid as a template. The conditions for the PCR were 30 cycles of 45 seconds at 95°C, 45 seconds at 68°C, and 1 minute at 72°C  
 15 after initial denaturation of 5 minutes at 95°C, and final extension of 5 minutes at 72°C. The amplified PCR product was digested with restriction enzyme *NdeI* and subjected to agarose gel electrophoresis. An MTD-derived from kFGF-4 16 mer (FGF-V, AAVALLPAVLLALLAP, Lin et al., *J. Biol. Chem.* 275: 16774-16778, 2000; Veatch et al., *J. Biol. Chem.* 279: 11425-11431, 2004),  
 20 an MTD-derived from kFGF-4 12 mer (FGF-J, AAVLLPVLLAAP, Jo et al., *Nat. Biotech.* 19: 929-933, 2001; Jo et al., *Nat. Med.* 11: 892-898, 2005 (SEQ ID NO: 389), a protein transduction domain derived from HIV-Tat (HIV-Tat, YGRKKRRQRRR, Schwarze et al., *Science* 285: 1569-1572, 1999), and HIV-Tat homologous (Hph-1, YARVRRRGPRR) were used as positive controls,  
 25 and a scramble peptide (SEQ ID NO: 387) was used as a negative control. As shown in Figs. 3a to 3c, it was confirmed that the DNA fragments encoding MTDs-EGFP were successfully amplified.

<Table 1>

PCR forward primer sequences for each MTD-EGFP protein

| MTD   | Sequence                                                                                   | SEQ ID NO |
|-------|--------------------------------------------------------------------------------------------|-----------|
| JO-01 | 5'-CCGCATATGGCGGTGGTGGTGTGCGCGATTGTGCTGGCGGCGCCGGT<br>GAGCAAGGGCGAGGAGCTG-3'               | 393       |
| JO-02 | 5'-CCGCATATGCCGCTGGCGCTGCTGGTGTGCTGCTGCTGGGCCCGGT<br>GAGCAAGGGCGAGGAGCTG-3'                | 394       |
| JO-03 | 5'-CCGCATATGCTGCTGCTGGCGTTTTCGCTGCTGTGCCTGCCGGTGAG<br>CAAGGGCGAGGAGCTG-3'                  | 395       |
| JO-04 | 5'-CCGCATATGCTGCTGGGCGCACTGGCGGCGGTGCTGCTGGCGCTGG<br>CGGTGAGCAAGGGCGAGGAGCTG-3'            | 396       |
| JO-05 | 5'-CCGCATATGCCGGTGTGCTGGCGCTGGGCGTGGGCCTGGTGTGCT<br>GGGCCTGGCGGTGGTGAGCAAGGGCGAGGAGCTG-3'  | 397       |
| JO-06 | 5'-CCGCATATGGCGGCGGCGGCGGTGCTGCTGGCGGCGGTGAGCAAGG<br>GCGAGGAGCTG-3'                        | 398       |
| JO-07 | 5'-CCGCATATGATTGTGGTGGCGGTGGTGGTATTGTGAGCAAGGGCG<br>AGGAGCTG-3'                            | 399       |
| JO-08 | 5'-CCGCATATGGCGGTGCTGGCGCCGGTGGTGGCGGTGGTGAGCAAGG<br>GCGAGGAGCTG-3'                        | 400       |
| JO-09 | 5'-CCGCATATGCTGGCGGTGTGCGGCCTGCCGGTGGTGGCGCTGCTGG<br>CGGTGAGCAAGGGCGAGGAGCTG-3'            | 401       |
| JO-10 | 5'-CCGCATATGCTGGGCGGCGCGG348TGGTGGCGGCGCCGGTGGCGCG<br>GCGGTGGCGCCGGTGAGCAAGGGCGAGGAGCTG-3' | 402       |
| JO-11 | 5'-CCGCATATGCTGCTGCTGGTGTGCTGGCGGTGCTGCTGGCGGTGCTGCC<br>GGTGAGCAAGGGCGAGGAGCTG-3'          | 403       |
| JO-12 | 5'-CCGCATATGCTGCTGATTCTGCTGCTGCTGCCGCTGCTGATTGTGGTG<br>AGCAAGGGCGAGGAGCTG-3'               | 404       |
| JO-13 | 5'-CCGCATATGCTGGCGGCGGCGGCGCTGGCGGTGCTGCCGCTGGTGA<br>GCAAGGGCGAGGAGCTG-3'                  | 405       |
| JO-14 | 5'-CCGCATATGTTTCTGATGCTGCTGCTGCCGCTGCTGCTGCTGCTGGTG<br>GCGGTGAGCAAGGGCGAGGAGCTG-3'         | 406       |
| JO-15 | 5'-CCGCATATGGCGGCGGCGGCGGCGGCGCTGGGTCTGGCGGCGGCGG<br>TGCCGGCGGTGAGCAAGGGCGAGGAGCTG-3'      | 407       |
| JO-16 | 5'-CCGCATATGCTGCTGCTGGCGGCGCTGCTGCTGATTGCGTTTTCGCGC<br>GGTGGTGAGCAAGGGCGAGGAGCTG-3'        | 408       |
| JO-17 | 5'-CCGCATATGGCGCTGGCGGCGGTGGTGTGATTCCGCTGGGCATTGC<br>GGCGGTGAGCAAGGGCGAGGAGCTG-3'          | 409       |
| JO-18 | 5'-CCGCATATGGCGGCGCTGATTGGCGCGGTGCTGGCGCCGGTGGTGG<br>CGGTGGTGAGCAAGGGCGAGGAGCTG-3'         | 410       |
| JO-19 | 5'-CCGCATATGGCGGCGGCGGTGGCGGTGGCGGGCCTGGCGCCGCTGG<br>CGCTGGTGAGCAAGGGCGAGGAGCTG-3'         | 411       |
| JO-20 | 5'-CCGCATATGATTGCGGTGGCGATTGCGGCGATTGTGCCGCTGGCGGT<br>GAGCAAGGGCGAGGAGCTG-3'               | 412       |
| JO-21 | 5'-CCGCATATGGCGGCGGCGGCGGTGCTGGCGGCGCCGGCGCTGGCGG<br>TGAGCAAGGGCGAGGAGCTG-3'               | 413       |
| JO-22 | 5'-CCGCATATGGCGGCGCTGGCGCTGGGCGTGGCGGCGGCGCCGGCGG<br>CGGCGCCGGCGGTGAGCAAGGGCGAGGAGCTG-3'   | 414       |
| JO-23 | 5'-CCGCATATGCTGGCGGTGCTGGTGTGCTGGTGTGCTGCCGGTGAG<br>CAAGGGCGAGGAGCTG-3'                    | 415       |
| JO-24 | 5'-CCGCATATGGTGGTGGCGGTGCTGGCGCCGGTGTGGTGAGCAAGG<br>GCGAGGAGCTG-3'                         | 416       |

|       |                                                                                             |     |
|-------|---------------------------------------------------------------------------------------------|-----|
| JO-25 | 5'-CCGCATATGGCGGCGCTGCTGCTGCCGCTGCTGCTGCTGCCGGT<br>GAGCAAGGGCGAGGAGCT G-3'                  | 417 |
| JO-26 | 5'-CCGCATATGCCGGCGGCGGTGGCGGCGCTGCTGGTGATTGTGAGCA<br>AGGGCGAGGAGCTG-3'                      | 418 |
| JO-27 | 5'-CCGCATATGCTGCTGATTGCGGCGCTGCTGCCGGTGAGCAAGGGCG<br>AGGAGCTG-3'                            | 419 |
| JO-28 | 5'-CCGCATATGGCGGCGGTGGTGCTGCTGCCGCTGGCGGCGGCGCCGG<br>TGAGCAAGGGCGAGGAGCTG-3'                | 420 |
| JO-29 | 5'-CCGCATATGGCGGCGGCGGCGGCGGCGCTGCTGGTGCCGGTGAGC<br>AAGGGCGAGGAGCTG-3'                      | 421 |
| JO-30 | 5'-CCGCATATGCTGCCGGTGGTGCGCTGCTGGCGGTGAGCAAGGGCG<br>AGGAGCTG-3'                             | 422 |
| JO-31 | 5'-CCGCATATGGCGGCGGCGCTGGCGGCGCCGCTGGCGCTGCCGGTGA<br>GCAAGGGCGAGGAGCTG -3'                  | 423 |
| JO-32 | 5'-CCGCATATGCTGCTGCTGGCGCTGCTGCTGGCGGCGGTGAGCAAGG<br>GCGAGGAGCTG-3'                         | 424 |
| JO-33 | 5'-CCGCATATGGCGGTGGCGGTGGTGCGCTGCTGGTGAGCAAGGGC<br>GAGGAGCTG-3'                             | 425 |
| JO-34 | 5'-CCGCATATGCTGCTGCTGATTATTGTGCTGCTGATTGTGCCGGTGAG<br>CAAGGGCGAGGAGCTG -3'                  | 426 |
| JO-35 | 5'-CCGCATATGCTGGCGCTGGCGGCGGCGGTGGTGCCGGTGAGCAAG<br>GGCGAGGAGCTG-3'                         | 427 |
| JO-36 | 5'-CCGCATATGCCGGCGGCGCTGGCGCTGCTGCTGGTGCGGTGAGCA<br>AGGGCGAGGAGCTG-3'                       | 428 |
| JO-37 | 5'-CCGCATATGATTGTGGCGCTGCTGCTGGTGCCGCTGGTGCTGGCGAT<br>TGCGGCGGTGCTGGTGAGCAAGGGCGAGGAGCTG-3' | 429 |
| JO-38 | 5'-CCGCATATGATTGTGGCGCTGCTGCTGGTGCCGGTGAGCAAGGGCG<br>AGGAGCTG-3'                            | 430 |
| JO-39 | 5'-CCGCATATGCCGCTGGTGCTGGCGATTGCGGCGGTGCTGGTGAGCA<br>AGGGCGAGGAGCTG-3'                      | 431 |
| JO-40 | 5'-CCGCATATGCCGCTGGTGCTGGCGGCGCTGGTGCGGTGAGCAAGG<br>GCGAGGAGCTG-3'                          | 432 |
| JO-41 | 5'-CCGCATATGGCGGCGGCGCTGCTGGCGGTGGCGGTGAGCAAGGGCG<br>AGGAGCTG-3'                            | 433 |
| JO-42 | 5'-CCGCATATGCCGCTGCTGCTGCTGGCGCTGGCGGTGAGCAAGGGCG<br>AGGAGCTG-3'                            | 434 |
| JO-43 | 5'-CCGCATATGGCGCTGGCGCTGGTGGTGGCGGTGAGCAAGGGCGAGG<br>AGCTG-3'                               | 435 |
| JO-44 | 5'-CCGCATATGGTGCGGCGGTGGTGGTGGCGGCGGTGAGCAAGGGCG<br>AGGAGCTG-3'                             | 436 |
| JO-45 | 5'-CCGCATATGCCGCTGCTGCCGCTGCTGCTGCTGGTGGTGAGCAAGG<br>GCGAGGAGCTG-3'                         | 437 |
| JO-46 | 5'-CCGCATATGGTGGTGCTGGTGGTGCTGCCGCTGGCGGTGCTGGC<br>GGTGAGCAAGGGCGAGGAGCTG-3'                | 438 |
| JO-47 | 5'-CCGCATATGGCGGCGGCGGTGCCGGTGCTGGTGCGGCGGTGAGCA<br>AGGGCGAGGAGCTG-3'                       | 439 |
| JO-48 | 5'-CCGCATATGCCGCGCTGCTGCTGCTGCTGCTGGCGGCGGTGGTGGT<br>GAGCAAGGGCGAGGAGCTG-3'                 | 440 |
| JO-49 | 5'-CCGCATATGCCGCTGGCGATTCTGCTGCTGCTGCTGATTGCGCCGGT<br>GAGCAAGGGCGAGGAGCTG-3'                | 441 |

|       |                                                                                 |     |
|-------|---------------------------------------------------------------------------------|-----|
| JO-50 | 5'-CCGCATATGCCGCTGCTGGCGCTGGTGCTGCTGCTGGCGCTGATTGC<br>GGTGAGCAAGGGCGAGGAGCTG-3' | 442 |
| JO-51 | 5'-CCGCATATGGTGGTGGCGGTGCTGGCGCTGGTGCTGGCGGCGCTGG<br>TGAGCAAGGGCGAGGAGCTG-3'    | 443 |
| JO-52 | 5'-CCGCATATGCCGCTGCTGCTGCTGCTGCCGGCGCTGGTGAGCAAGG<br>GCGAGGAGCTG-3'             | 444 |
| JO-53 | 5'-CCGCATATGCTGGCGGCGGTGGCGGCGCTGGCGGTGGTGGTGCCGG<br>TGAGCAAGGGCGAGGAGCTG-3'    | 445 |
| JO-54 | 5'CCGCATATGCTGCTGCTGCTGGTGCTGATTCTGCCGCTGGCGGCGGTG<br>AGCAAG GCGAGGAGCTG-3'     | 446 |
| JO-55 | 5'-CCGCATATGCTGGCGGTGGTGGTGGTGGCGGCGGTGGTGAGCAAGG<br>GCGAGGAGCTG-3'             | 447 |
| JO-56 | 5'CCGCATATGGTGCTGCTGGCGGCGGCGCTGATTGCGGTGAGCAAGG<br>GCGAGGA GCTG-3'             | 448 |
| JO-57 | 5'-CCGCATATGCTGATTGCGCTGCTGGCGGCGCCGCTGGCGGTGAGCA<br>AGGGCGAGGAGCTG-3'          | 449 |
| JO-58 | 5'-CCGCATATGCTGGCGCTGCTGCTGCTGGCGGCGGTGAGCAAGGGCG<br>AGGAGCTG-3'                | 450 |
| JO-59 | 5'-CCGCATATGCTGCTGGCGGCGGCGCTGCTGCTGCTGCTGGCGGT<br>GAGCAAGGGCGAGGAGCTG-3'       | 451 |
| JO-60 | 5'-CCGCATATGGTGATTATTGCGCTGATTGTGATTGTGGCGGTGAGCAA<br>GGGCGAGGAGCTG-3'          | 452 |
| JO-61 | 5'-CCGCATATGGTGGTGCTGGTGGTGGCGGCGGTGCTGGCGCTGGTGA<br>GCAAGGGCGAGGAGCTG -3'      | 453 |
| JO-62 | 5'-CCGCATATGGTGGCGGTGGCGATTGCGGTGGTGCTGGTGAGCAAGG<br>GCGAGGAGCTG-3'             | 454 |
| JO-63 | 5'-CCGCATATGCCGCTGATTGTGGTGGTGGCGGCGGCGGTGGTGGCGG<br>TGGTGAGCAAGGGCGAGGAGCTG-3' | 455 |
| JO-64 | 5'-CCGCATATGCCGCTGGCGGTGGCGGTGGCGGCGGTGGCGGCGGTGA<br>GCAAGGGCGAGGAGCTG -3'      | 456 |
| JO-65 | 5'-CCGCATATGGCGGCGATTGCGCTGGTGGCGGTGGTGCTGGTGAGCA<br>AGGGCGAGGAGCTG-3'          | 457 |
| JO-66 | 5'-CCGCATATGGCGGCGGCGCTGGCGGCGATTGCGGTGATTGTGAGCA<br>AGGGCGAGGAGCTG-3'          | 458 |
| JO-67 | 5'-CCGCATATGGCGGCGGCGCCGGCGGTGGCGGCGGTGAGCAAGGGC<br>GAGGAGCTG-3'                | 459 |
| JO-68 | 5'-CCGCATATGCTGCTGCTGGCGGCGCTGCCGGTGAGCAAGGGCGAGG<br>AGCTG-3'                   | 460 |
| JO-69 | 5'-CCGCATATGGCGCTGCTGGCGGTGGTGGCGGCGGTGAGCAAGGGCG<br>AGGAGCTG-3'                | 461 |
| JO-70 | 5'-CCGCATATGGCGGTGGTGGTGGTGCTGCCGATTCTGCTGGTGAGCAA<br>GGGCGAGGAGCTG-3'          | 462 |
| JO-71 | 5'-CCGCATATGGCGCTGGCGCTGCTGCTGCTGGTGCCGGTGAGCAAGG<br>GCGAGGAGCTG-3'             | 463 |
| JO-72 | 5'-CCGCATATGCTGGTGGTGCTGCTGGCGGCGCTGCTGGTGCTGGTGAG<br>CAAGGGCGAGGAGCTG -3'      | 464 |
| JO-73 | 5'-CCGCATATGCCGGTGCTGCTGCTGCTGGCGCCGGTGAGCAAGGGCG<br>AGGAGCTG-3'                | 465 |
| JO-74 | 5'-CCGCATATGGCGCTGGCGGTGGTGGCGGCGCCGGTGAGCAAGGGCG<br>AGGAGCTG-3'                | 466 |

|       |                                                                                 |     |
|-------|---------------------------------------------------------------------------------|-----|
| JO-75 | 5'-CCGCATATGGTGATTGTGGCGCTGCTGGCGGTGGTGAGCAAGGGCG<br>AGGAGCTG-3'                | 467 |
| JO-76 | 5'-CCGCATATGGCGCTGGTGCTGCCGCTGGCGCCGGTGAGCAAGGGCG<br>AGGAGCTG-3'                | 468 |
| JO-77 | 5'-CCGCATATGGCGGTGGCGCTGCTGATTCTGGCGGTGGTGAGCAAGG<br>GCGAGGAGCTG-3'             | 469 |
| JO-78 | 5'-CCGCATATGGTGCTGCTGGCGGTGATTCCGGTGAGCAAGGGCGAGG<br>AGCTG-3'                   | 470 |
| JO-79 | 5'-CCGCATATGCTGATTGTGGCGGCGGTGGTGGTGGTGGCGGTGCTGAT<br>TGTGAGCAAGGGCGAGGAGCTG-3' | 471 |
| JO-80 | 5'-CCGCATATGGCGGTGGTGGTGGCGGCGCCGGTGAGCAAGGGCGAGG<br>AGCTG-3'                   | 472 |
| JO-81 | 5'-CCGCATATGCTGGCGGCGGTGCTGCTGCTGATTCCGGTGAGCAAGG<br>GCGAGGAGCTG-3'             | 473 |
| JO-82 | 5'-CCGCATATGCTGCTGCTGCTGCTGCTGGCGGTGGTGCCGGTGAGCAA<br>GGGCGAGGAGCTG-3'          | 474 |
| JO-83 | 5'-CCGCATATGGCGGTGGCGCTGGTGGCGGTGGTGGCGGTGGCGGTGA<br>GCAAGGGCGAGGAGCTG-3'       | 475 |
| JO-84 | 5'-CCGCATATGCTGGTGGCGGCGCTGCTGGCGGTGCTGGTGAGCAAGG<br>GCGAGGAGCTG-3'             | 476 |
| JO-85 | 5'-CCGCATATGCTGCTGGCGGCGGCGGCGGCGCTGCTGCTGGCGGTGA<br>GCAAGGGCGAGGAGCTG-3'       | 477 |
| JO-86 | 5'-CCGCATATGCTGGCGGTGCTGGCGGCGGCGCCGGTGAGCAAGGGCG<br>AGGAGCTG-3'                | 478 |
| JO-87 | 5'-CCGCATATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT<br>GGTGAGCAAGGGCGAGGAGCTG-3' | 479 |
| JO-88 | 5'-CCGCATATGGTGGTGATTGCGGTGGTGCCGGTGAGCAAGGGCGAGG<br>AGCTG-3'                   | 480 |
| JO-89 | 5'-CCGCATATGGTGCTGCTGGTGCTGCTGGCGCTGGTGGTGAGCAAGG<br>GCGAGGAGCTG-3'             | 481 |
| JO-90 | 5'-CCGCATATGGTGCTGCTGGTGCTGCTGGCGCTGGTGGTGAGCAAGG<br>GCGAGGAGCTG-3'             | 482 |
| JO-91 | 5'-CCGCATATGCCGGTGCTGGTGCCGCGCGGTGCCGGTGAGCAAGGGCG<br>AGGAGCTG-3'               | 483 |
| JO-92 | 5'-CCGCATATGCCGGCGCTGGCGCTGGCGCTGGCGGTGAGCAAGGGCG<br>AGGAGCTG-3'                | 484 |
| JO-93 | 5'-CCGCATATGGCGGCGGCGGCGCCGGCGCTGGCGGTGAGCAAGGGCG<br>AGGAGCTG-3'                | 485 |
| JO-94 | 5'-CCGCATATGATTGTGCTGCCGGTGCTGGCGGCGCCGGTGAGCAAGG<br>GCGAGGAGCTG-3'             | 486 |
| JO-95 | 5'-CCGCATATGCTGGTGCTGCTGCTGCTGCCGCTGCTGATTGTGAGCAA<br>GGGCGAGGAGCTG-3'          | 487 |
| JO-96 | 5'-CCGCATATGCTGGCGGCGGTGGCGCCGGCGCTGGCGGTGGTGGTGA<br>GCAAGGGCGAGGAGCTG-3'       | 488 |
| JO-97 | 5'-CCGCATATGATTCTGGTGCTGGTGCTGCCGATTGTGAGCAAGGGCGA<br>GGAGCTG-3'                | 489 |
| JO-98 | 5'-CCGCATATGATTCTGCTGCCGCTGCTGCTGCTGCCGGTGAGCAAGGG<br>CGAGGAGCTG-3'             | 490 |
| JO-99 | 5'-CCGCATATGATTGCGCCGGCGGTGGTGGCGGCGCTGCCGGTGAGCA<br>AGGGCGAGGAGCTG-3'          | 491 |

|        |                                                                                    |     |
|--------|------------------------------------------------------------------------------------|-----|
| JO-100 | 5'-CCGCATATGCTGCTGCTGGTGGCGGTGGTGCCGCTGCTGGTGCCGGT<br>GAGCAAGGGCGAGGAGCTG-3'       | 492 |
| JO-101 | 5'-CCGCATATGCTGATTCTGCTGCTGCTGCCGATTATTGTGAGCAAGGGC<br>GAGGAGCTG-3'                | 493 |
| JO-102 | 5'-CCGCATATGGCGGTGCTGGCGGCGCCGGCGGTGCTGGTGGTGAGCA<br>AGGGCGAGGAGCTG-3'             | 494 |
| JO-103 | 5'-CCGCATATGCTGGCGCTGCCGGTGCTGCTGCTGGCGGTGAGCAAGG<br>GCGAGGAGCTG-3'                | 495 |
| JO-104 | 5'-CCGCATATGCTGGCGCTGGCGCTGCTGCTGGTGAGCAAGGGCGAGG<br>AGCTG-3'                      | 496 |
| JO-105 | 5'-CCGCATATGGTGGCGGTGCCGCTGCTGGTGGTGGCGGTGAGCAAGG<br>GCGAGG AGCTG-3'               | 497 |
| JO-106 | 5'-CCGCATATGGCGGTGGCGGTGGCGCCGGTGGCGGCGGCGGCGGTGA<br>GCAAGGGCGAGGAGCTG-3'          | 498 |
| JO-107 | 5'-CCCATATGGCGGCGGCGGTGGTGGCGGCGGTGCCGGCGGCGGTGAG<br>CAAGGGCGAGGAGCTG-3'           | 499 |
| JO-108 | 5'-CCCATATGGCGCTGCTGGCGGCGCTGCTGGCGCCGGTGAGCAAGGG<br>CGAGGAGCTG-3'                 | 500 |
| JO-109 | 5'-CCCATATGGCGCTGGCGCTGCTGGTGCCGGTGAGCAAGGGCGAGGAGC<br>TG-3'                       | 501 |
| JO-110 | 5'-CCGCATATGGCGCTGCTGGCGGCGCTGCTGGCGCTGCTGGCGCTGCT<br>GGTGGTGAGCAAGGGCGAGGAGCTG-3' | 502 |
| JO-111 | 5'-CCGCATATGGCGGCGGCGCTGCCGCTGCTGGTGCTGCTGCCGGTGA<br>GCAAGGGCGAGGAGCTG-3'          | 503 |
| JO-112 | 5'-CCCATATGGCGGCGGCGGTGCCGGCGGCGCTGGCGCCGGTGAGCAA<br>GGGCGAGGAGCTG-3'              | 504 |
| JO-113 | 5'-CCCATATGGCGGCGCTGGCGGTGGCGGCGCTGGCGGCGGTGAGCAAG<br>GGCGAGGAGCTG-3'              | 505 |
| JO-114 | 5'-CCCATATGGCGGTGCTGGCGGCGGCGGTGCCGGTGAGCAAGGGCGAG<br>GAGCTG-3'                    | 506 |
| JO-115 | 5'-CCGCATATGGTGGCGGCGCTGCCGGCGCCGGCGGTGAGCAAGGGCG<br>AGGAGCTG-3'                   | 507 |
| JO-116 | 5'-CCGCATATGGCGCTGGCGCTGGCGGTGCCGGCGGTGCTGCCGGTGA<br>GCAAGGGCGAGGAGCTG-3'          | 508 |
| JO-117 | 5'-CCGCATATGGCGGCGCTGCTGCCGGCGGCGGTGGCGGTGCCGGTGA<br>GCAAGGGCGAGGAGCTG-3'          | 509 |
| JO-118 | 5'-CCGCATATGGCGGTGGTGGTGGCGCTGGCGCCGGTGAGCAAGGGCG<br>AGGAGCTG-3'                   | 510 |
| JO-119 | 5'-CCGCATATGGCGGCGGCGGTGGCGCTGCCGGCGGCGGCGGCGCTGC<br>TGCGGTGAGCAAGGGCGAGGAGCTG-3'  | 511 |
| JO-120 | 5'-CCGCATATGGCGGTGGTGGTGGCGCTGGCGCTGGTGGCGGTGGCGC<br>CGGTGAGCAAGGGCGAGGAGCTG-3'    | 512 |
| JO-121 | 5'-CCGCATATGCTGGTGGCGCTGCCGCTGCTGCCGGTGAGCAAGGGCG<br>AGGAGCTG-3'                   | 513 |
| JO-122 | 5'-CCGCATATGGTGGTGGTGGCGCTGCTGCTGATTGTGCCGGTGAGCAA<br>GGGCGAGGAGCTG-3'             | 514 |
| JO-123 | 5'-CCGCATATGCTGGCGGTGGTGGTGGCGGTGCCGGTGAGCAAGGGCG<br>AGGAGCTG-3'                   | 515 |
| JO-124 | 5'-CCGCATATGCTGCTGGCGGTGCCGATTCTGCTGGTGCCGGTGAGCAA<br>GGGCGAGGAGCTG-3'             | 516 |

|        |                                                                                     |     |
|--------|-------------------------------------------------------------------------------------|-----|
| JO-125 | 5'-CCGCATATGCTGGTGGCGCTGGTGTCTGCCGGTGAGCAAGGGCG<br>AGGAGCTG-3'                      | 517 |
| JO-126 | 5'-CCGCATATGCTGGTGTCTGTCTGGCGGTGCTGTCTGGCGGTGCT<br>GCCGGTGAGCAAGGGCGAGGAGCTG-3'     | 518 |
| JO-127 | 5'-CCGCATATGCTGTCTGGCGCCGGTGGTGGCGCTGGTGATTCTGCCGGT<br>GAGCAAGGGCGAGGAGCTG-3'       | 519 |
| JO-128 | 5'-CCGCATATGGTGTCTGGCGGTGCTGGCGGTGCCGGTGCTGTCTGCC<br>GGTGAGCAAGGGCGAGGAGCTG-3'      | 520 |
| JO-129 | 5'-CCGCATATGGTGGTGATTGCGGTGGTGGCGGTGGTGGTGGTGAGCA<br>AGGGCGAGGAGCTG-3'              | 521 |
| JO-130 | 5'-CCGCATATGCTGTCTGGTGTCTGTCTGGCGCTGGTGGTGGTGGCGGTGAG<br>CAAGGGCGAGGAGCTG-3'        | 522 |
| JO-131 | 5'-CCGCATATGGTGTCTGTCTGGCGCTGCCGGTGGTGGCGGCGCCGGTGA<br>GCAAGGGCGAGGAGCTG-3'         | 523 |
| JO-132 | 5'-CCGCATATGGCGGTGGTGGTGGCGGCGATTGTGTCTGGCGGCGCCGG<br>TGAGCAAGGGCGAGGAGCTG-3'       | 524 |
| JO-133 | 5'-CCGCATATGGCGGTGCTGGTGGCGGCGGCGGCGCTGGTGGCGGTGA<br>GCAAGGGCGAGGAGCTG-3'           | 525 |
| JO-134 | 5'-CCGCATATGGTGGTGGCGGCGCTGCCGCTGGTGTCTGCCGGTGAGCA<br>AGGGCGAGGAGCTG-3'             | 526 |
| JO-135 | 5'-CCGCATATGGCGGCGGTGGCGCTGCCGGCGGCGGCGCCGGTGAGCA<br>AGGGCGAGGAGCTG-3'              | 527 |
| JO-136 | 5'-CCGCATATGCTGATTGCGCTGCCGCTGTCTGCCGGTGAGCAAGGGCG<br>AGGAGCTG-3'                   | 528 |
| JO-137 | 5'-CCGCATATGCTGTCTGGCGCTGCCGCTGGTGTCTGGTGTCTGGCGCTGCC<br>GGTGAGCAAGGGCGAGGAGCTG-3'  | 529 |
| JO-138 | 5'-CCGCATATGATTGTGCCGCTGTCTGTCTGGCGGCGCCGGTGAGCAAGG<br>GCGAGGAGCTG-3'               | 530 |
| JO-139 | 5'-CCGCATATGCTGTCTGGCGCCGCTGTCTGTCTGGCGCCGGTGAGCA<br>AGGGCGAGGAGCTG-3'              | 531 |
| JO-140 | 5'-CCGCATATGCTGGCGGCGCTGCCGGTGGCGGCGGTGCCGGTGAGCA<br>AGGGCGAGGAGCTG-3'              | 532 |
| JO-141 | 5'-CCGCATATGGCGCTGGCGGTGATTGTGTCTGGTGTCTGGTGTAGCAA<br>GGGCGAGGAGCTG-3'              | 533 |
| JO-142 | 5'-CCGCATATGCTGGCGCTGTCTGTCTGCCGGCGGCGCTGATTCCGGTGAG<br>CAAGGGCGAGGAGCTG-3'         | 534 |
| JO-143 | 5'-CCGCATATGGCGCTGTCTGCCGCTGTCTGGCGGTGGTGTCTGCCGCCGGT<br>GAGCAAGGGCGAGGAGCTG-3'     | 535 |
| JO-144 | 5'-CCGCATATGGCGATTGCGGTGCCGGTGTCTGGCGGCGCCGGTGAGCA<br>AGGGCGAGGAGCTG-3'             | 536 |
| JO-145 | 5'-CCGCATATGGCGGCGGCGCCGGTGTCTGTCTGTCTGTCTGCCGGTGA<br>GCAAGGGCGAGGAGCTG-3'          | 537 |
| JO-146 | 5'-CCGCATATGGCGGCGGCGCCGGTGTCTGTCTGTCTGTCTGCCGGTGA<br>GCAAGGGCGAGGAGCTG-3'          | 538 |
| JO-147 | 5'-CCGCATATGGCGGCGCTGGCGGCGCTGGTGGTGGCGGCGCCGCCGG<br>TGAGCAAGGGCGAGGAGCTG-3'        | 539 |
| JO-148 | 5'-CCGCATATGGCGGCGCTGGCGGCGGTGCCGCTGGCGCTGGCGCCGG<br>TGAGCAAGGGCGAGGAGCTG-3'        | 540 |
| JO-149 | 5'-CCGCATATGGCGCTGGCGGTGGCGGCGCCGGCGCTGGCGCTGTCTGC<br>CGCCGGTGAGCAAGGGCGAGGAGCTG-3' | 541 |

|         |                                                                                     |     |
|---------|-------------------------------------------------------------------------------------|-----|
| JO-150  | 5'-CCGCATATGCTGGTGGCGCTGGTGTCTGCTGCCGGTGAGCAAGGGCG<br>AGGAGC TG-3'                  | 542 |
| JO-151  | 5'-CCGCATATGCTGGTGTCTGCTGCTGGCGGTGCTGCTGCTGGCGGTGC<br>TGCCGGTGAGCAAGGGCGAGGAGCTG-3' | 543 |
| JJO-152 | 5'-CCGCATATGCTGCTGGCGCCGGTGGTGGCGCTGGTGATTCTGCCGG<br>TGAGCAAGGGCGAGGAGCTG-3'        | 544 |
| JO-153  | 5'-CCGCATATGGTGTCTGGCGGTGCTGGCGGTGCCGGTGCTGCTGCTGC<br>CGGTGAGCAAGGGCGAGGAGCTG-3'    | 545 |
| JO-154  | 5'-CCGCATATGGTGGTATTGCGGTGGTGCCGGTGGTGGTGGTGAGCA<br>AGGGCGAGGAGCTG-3'               | 546 |
| JO-155  | 5'-CCGCATATGCTGCTGGTGTCTGCTGGCGCTGGTGGTGGTGCCGGTGA<br>GCAAGGGCGAGGAGCTG-3'          | 547 |
| JO-156  | 5'-CCGCATATGGTGTCTGCTGGCGCTGCCGGTGGTGGCGGCGCCGGTGA<br>GCAAGGGCGAGGAGCTG-3'          | 548 |
| JO-157  | 5'-CCGCATATGGCGGTGGTGGTGCCGGCGATTGTGCTGGCGGCGCCGG<br>TGAGCAAGGGCGAGGAGCTG-3'        | 549 |
| JO-158  | 5'-CCGCATATGGCGGTGCTGGTGGCGGCGGCGGCGCTGGTGCCGGTGA<br>GCAAGGGCGAGGAGCTG-3'           | 550 |
| JO-159  | 5'-CCGCATATGGTGGTGGCGGCGCTGCCGCTGGTGTCTGCCGGTGAGCA<br>AGGGCGAGGAGCTG-3'             | 551 |
| JO-160  | 5'-CCGCATATGGCGGCGGTGGCGCTGCCGGCGGCGGCGCCGGTGAGC<br>AAGGGCGAGGAGCTG-3'              | 552 |
| JO-161  | 5'-CCGCATATGCTGATTGCGCTGCCGCTGCTGCCGGTGAGCAAGGGCG<br>AGGAGCTG-3'                    | 553 |
| JO-162  | 5'-CCGCATATGCTGTCTGGCGCTGCCGCTGGTGTCTGGTGTCTGGCGCTGC<br>CGGTGAGCAAGGGCGAGGAGCTG-3'  | 554 |
| JO-163  | 5'-CCGCATATGATTGTCCGCTGCTGCTGGCGGCGCCGGTGAGCAAGG<br>GCGAGGAGCTG-3'                  | 555 |
| JO-164  | 5'-CCGCATATGCTGTCTGGCGCCGCTGCTGCTGGCGCCGGTGAGCA<br>AGGGCGAGGAGCTG-3'                | 556 |
| JO-165  | 5'-CCGCATATGCTGGCGGCGCTGCCGGTGGCGGCGGTGCCGGTGAGC<br>AAGGGCGAGGAGCTG-3'              | 557 |
| JO-166  | 5'-CCGCATATGGCGCTGGCGGTGATTGTGCTGGTGTCTGGTGTAGCA<br>AGGGCGAGGAGCTG-3'               | 558 |
| JO-167  | 5'-CCGCATATGCTGGCGCTGCTGTCTGCCGGCGGCGCTGATTCCGGTGA<br>GCAAGGGCGAGGAGCTG-3'          | 559 |
| JO-168  | 5'-CCGCATATGCTGGCGGCGGTGGTGCCGGTGGCGGCGGCGGTGCCG<br>GTGAGCAAGGGCGAGGAGCTG-3'        | 560 |
| JO-169  | 5'-CCGCATATGGTGGCGGCGCCGGCGGCGGCGGCGCCGGTGAGCAAG<br>GGCGAGGAGCTG-3'                 | 561 |
| JO-170  | 5'-CCGCATATGGCGGTGCCGGTGCCGGTGCCGCTGGTGAGCAAGGGC<br>GAGGAGCTG-3'                    | 562 |
| JO-171  | 5'-CCGCATATGCTGTCTGATTCTGCCGATTGTGCTGCTGCCGGTGAGCAA<br>GGGCGAGGAGCTG-3'             | 563 |
| JO-172  | 5'-CCGCATATGGCGCTGGCGCTGCCGGCGCTGGCGATTGCGCCGGTGA<br>GCAAGGGCGAGGAGCTG-3'           | 564 |
| JO-173  | 5'-CCGCATATGGCGGTGATTCCGATTCTGGCGGTGCCGGTGAGCAAGG<br>GCGAGGAGCTG-3'                 | 565 |
| JO-174  | 5'-CCGCATATGCTGATTCTGTCTGCTGCCGGCGGTGGCGCTGCCGGTGA<br>GCAAGGGCGAGGAGCTG-3'          | 566 |

|        |                                                                       |     |
|--------|-----------------------------------------------------------------------|-----|
| JO-175 | 5'-CCGCATATGATTGTGCTGGCGCCGGTGCCGGCGGCGGCGGTGAGCAAGGGCGAGGAGCTG-3'    | 567 |
| JO-176 | 5'-CCGCATATGGTGGTGGTGGTGCCGGTGCTGGCGGCGGCGGCGGTGAGCAAGGGCGAGGAGCTG-3' | 568 |
| JO-177 | 5'-CCGCATATGCTGGTGGCGGTGGCGGCGCCGGTGAGCAAGGGCGAGGAGCTG-3'             | 569 |
| JO-178 | 5'-CCGCATATGCTGGTGGTGGCGGCGCCGGCGGCGCTGCCGGTGAGCAAGGGCGAGGAGCTG-3'    | 570 |
| JO-179 | 5'-CCGCATATGCTGATTGCGCCGGCGGCGGCGGTGCCGGTGAGCAAGGGCGAGGAGCTG-3'       | 571 |
| JO-180 | 5'-CCGCATATGGCGCTGGCGGCGCTGCCGATTGCGCTGCCGGTGAGCAAGGGCGAGGAGCTG-3'    | 572 |
| JO-181 | 5'-CCGCATATGGCGGTGCTGCTGCTGCCGGCGGCGGCGGTGAGCAAGGGCGAGGAGCTG-3'       | 573 |
| JO-182 | 5'-CCGCATATGATTGCGCTGGCGCTGCTGCCGCTGCTGGTGAGCAAGGGCGAGGAGCTG-3'       | 574 |
| JO-183 | 5'-CCGCATATGGTGGTGGTGGCGGCGGCGCTGATTGCGCCGGTGAGCAAGGGCGAGGAGCTG-3'    | 575 |
| JO-184 | 5'-CCGCATATGGCGCCGGCGGTGCTGCCGCCGGTGGTGGTGATTGTGAGCAAGGGCGAGGAGCTG-3' | 576 |
| JO-185 | 5'-CCGCATATGGTGGTGGGCGCTGCTGGTGGCGGCGCTGGTGAGCAAGGGCGAGGAGCTG-3'      | 577 |
| JO-186 | 5'-CCGCATATGGCGGCGATTGCGGCGGCGGCGCCGCTGGCGGCGGTGAGCAAGGGCGAGGAGCTG-3' | 578 |
| JO-187 | 5'-CCGCATATGCTGCTGCTGGCGGTGGCGCCGGTGAGCAAGGGCGAGGAGCTG-3'             | 579 |
| JO-188 | 5'-CCGCATATGCTGATTCTGCTGCTGCCGCTGGCGGCGCTGGTGAGCAAGGGCGAGGAGCTG-3'    | 580 |
| JO-189 | 5'-CCGCATATGGCGCTGCTGCTGCTGGTGGTGGCGGTGAGCAAGGGCGAGGAGCTG-3'          | 581 |
| JO-190 | 5'-CCGCATATGCTGCTGCTGCTGCTGCCGCTGGCGGTGAGCAAGGGCGAGGAGCTG-3'          | 582 |
| JO-191 | 5'-CCGCATATGCTGGCGCTGCCGCTGCTGCTGCCGGTGAGCAAGGGCGAGGAGCTG-3'          | 583 |
| JO-192 | 5'-CCGCATATGCTGCTGGTGGTGGCGCTGCTGATTGTGAGCAAGGGCGAGGAGCTG-3'          | 584 |
| JO-193 | 5'-CCGCATATGCTGCCGCTGCTGCCGGCGGCGCTGGTGGTGAGCAAGGGCGAGGAGCTG-3'       | 585 |

<Table 2>

PCR reverse primer sequence for each MTD-EGFP protein

| MTD             | Sequence                                  | SEQ ID NO |
|-----------------|-------------------------------------------|-----------|
| JO-1~<br>JO-193 | 3'-TTA TCT AGA TCC GGT GGA TCC CGG GCC-5' | 586       |

Referring to Fig. 4a, each of the PCR amplified products was subcloned into the pGEM-Teasy vector (Promega, Madison WI, USA) with a T4 DNA ligase, according to the TA cloning method and then *E. coli* DH5 $\alpha$  was transformed, where transformants were selected on LB plate media containing 50  $\mu$ g/mL ampicillin, IPTG and X-gal. After the recombinant fragment inserted into pGEM-Teasy vector was isolated by treating with the restriction enzyme *NdeI*, it was subjected to 0.8 % agarose gel electrophoresis. As a result, DNA fragments of about 1 kb and vector fragments of about 3 kb were detected, which confirms that the insert DNA of MTD-EGFP was appropriately subcloned into pGEM-Teasy vector (Figs. 4b to 4d).

As described in Fig. 5a, each of the isolated insert DNA fragments encoding MTDs-EGFP was cloned into the *E. coli* expression plasmid pET-28a(+) (commercially available from Novagen, Madison, WI). The pET-28a(+) plasmids are designed to facilitate His-tag fusions at either the N- or C-terminus and to provide strong expression of the genes in *E. coli* from the T7 phage promoter. At the 3' end of each MTD-EGFP encoding gene, the coding sequence was fused in frame at the *NdeI* site to the His-tag sequence followed by a translation stop codon, which results in the production of MTD-EGFP recombinant proteins with six histidine residues added to the C-terminus for the sake of easy purification on nickel columns.

After the clones were treated with the restriction enzyme *NdeI* and subjected to agarose gel electrophoresis, it was verified that DNA fragments of about 1 kb and vector fragments of about 5 kb were detected, which confirms the cloning of the insert DNA of MTD-EGFP into pET-28a(+) vector, as shown in Figs. 5b to 5c. Further, out of the 193 novel MTDs, 148 MTDs were found to be successfully cloned for the expression of His-MTD-EGFP recombinant proteins.

### **Example 3 - Inducible Expression and Purification of Recombinant**

### Proteins Fused to MTDs

To express the cell permeable recombinant proteins fused to MTDs prepared as described in Example 2 above, the expression vectors comprising the His-MTD-EGFP recombinant proteins were transfected in the BL21 (DE3), BL21-Gold (DE3), BL21-CodonPlus (DE3) and BL21-GoldpLysS (DE3) strains, respectively. An EGFP expression vector including kFGF4-derived MTD (SEQ ID NO: 387) was used as a positive control, while an EGFP expression vector fused to a scramble peptide (SEQ ID NO: 389) having no function was used as a negative control.

After the transfection, cells were grown at 37°C in an LB medium containing kanamycin (30 µg/ml) with vigorous shaking until the optical density 600 (OD<sub>600</sub>) reached between 0.4 and 0.6. IPTG (isopropyl-β-D-thiogalactoside) was then added thereto at a final concentration of 0.6 mM to induce the expression of the His-MTD-EGFP recombinant proteins. Protein induction was prolonged for 3 hours at 37°C. His-MTD-EGFP recombinant proteins expressed in *E. coli* with IPTG were loaded on a SDS-PAGE gel, stained with Coomassie Brilliant Blue, and then destained. As illustrated in Figs. 6a and 6b, except for some His-MTD-EGFP recombinant proteins expressed in BL21-GoldpLysS (DE3) strains, most His-MTD-EGFP recombinant proteins were expressed at a high level in BL21-CodonPlus (DE3). Several His-MTD-EGFP recombinant proteins were not expressed.

The inducible expression of His-MTD-EGFP recombinant proteins in an *E. coli* system leads to the formation of insoluble aggregates, which are known as inclusion bodies. To completely solubilize these inclusion bodies, all of the above expressed proteins were denatured by dissolving them in 8 M urea. Denatured His-MTD-EGFP recombinant proteins were purified by histidine tag affinity chromatography, using a nickel nitrilotriacetate resin (Qiagen, Hilden, Germany). Since strong denaturants, such as 8 M urea, completely solubilize the inclusion bodies, the purification method was carried out under pH-dependent denaturing conditions.

The *E. coli* culture solutions were harvested by centrifugation at 4,000 ×g for 20 minutes, re-suspended in a lysis buffer (100 mM NaH<sub>2</sub>PO<sub>4</sub>, 10 mM Tris·Cl, 8 M urea, pH 8.0), and subjected to ultrasonication on ice using a sonicator equipped with a probe. The cell lysates were centrifuged at 7,000 ×g for 20 minutes, so as to separate the supernatant and the cellular debris pellet. The supernatant was taken out and then incubated with a Ni-NTA resin equilibrated with the lysis buffer by gently shaking (using a rotary shaker) for 2 hours to overnight. After washing with a washing buffer (100 mM NaH<sub>2</sub>PO<sub>4</sub>, 10 mM Tris·Cl, 8 M urea, pH 6.3) five times, the proteins bound to the resin were eluted with an elution buffer (100 mM NaH<sub>2</sub>PO<sub>4</sub>, 10 mM Tris·Cl, 8 M urea, pH 4.5). The His-MTD-EGFP recombinant proteins purified under the denaturing conditions described above were analyzed on a SDS-PAGE gel and stained with Coomassie Brilliant Blue, where the results thereof are shown in Figs. 7a and 7b.

In order to renature the His-MTD-EGFP recombinant proteins purified above, the denatured proteins were refolded by removing the denaturant. Urea was removed from the proteins by dialyzing them against a refolding buffer (0.55 M Guanidine HCl, 0.44 M L-Arginine, 50 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, 100 mM NDSB, 2 mM Glutathione Oxidized, and 0.2 mM Glutathione Reduced). All of the refolded recombinant proteins were dialyzed for 9 hours against a physiological buffer, such as a cell culture medium (e.g., Dulbecco's modified Eagle's Medium: DMEM). After the replacement of the refolding buffer with DMEM, the cell permeabilities of all of the purified recombinant proteins were ready to be determined *in vitro* and *in vivo*. According to the SDS-PAGE analysis results shown in Fig. 8, out of the 148 transformants established as described in Example 2 above, the inducible expression of His-MTD-EGFP recombinant proteins was visually detected in 112 transformants, whereas 4 His-MTD-EGFP recombinant proteins were not expressed enough to be visually detected but were sufficiently purified for cell permeability analysis. The 116 His-MTD-EGFP recombinant proteins were

prepared in a soluble form.

**Example 4 - Determination of Quantitative Cell Permeability of Recombinant Proteins Fused to MTDs**

5 In order to quantitatively determine the cell permeability of the His-MTD-EGFP recombinant proteins, the 116 recombinant proteins purified in a soluble form as described in Example 3 above were mixed with 0.7  $\mu\text{g}/\mu\text{l}$  of fluorescein isothiocyanate (FITC) and reacted at room temperature for 1 hours by stirring. The reaction solution was subjected to a dialysis against  
10 Dulbecco's modified Eagle's medium (DMEM; WelGENE Inc., Korea) for 2 days until the FITC was completely removed to thereby obtain FITC-conjugated recombinant proteins. RAW 264.7 cells derived from mouse macrophage were treated with 10  $\mu\text{M}$  of the FITC-labeled protein. The RAW 264.7 cells were maintained in DMEM supplemented with 10% fetal bovine  
15 serum and 1% penicillin (500 mg/ml, WelGENE Inc.), and incubated at 37°C in a humidified atmosphere of 5% CO<sub>2</sub> in air. After the incubation, the cells were incubated with 10  $\mu\text{M}$  of each of the FITC-conjugated recombinant protein prepared above for 1 hour at 37°C, followed by treating them with Trypsin/EDTA (T/E, Invitrogen, Carlsbad CA, USA) to remove cell surface  
20 bound proteins and washing with cold PBS three times.

The FITC-conjugated His-MTD-EGFP recombinant proteins were subjected to fluorescence-activated cell sorting (FACS) analysis (FACS Calibur, Beckton-Dickinson, San Diego CA, USA). For each sample, the cells ( $1 \times 10^4$ ) were analyzed by using the CellQuest Pro cytometric analysis software. Each  
25 experiment was conducted at least twice. The cell permeable potency of each His-MTD-EGFP recombinant protein fused to novel MTDs (JO-1 to JO-193) was visually compared to that of a positive control protein fused to kFGF4-derived MTD.

30 Figs. 9a to 9g show the results of a flow cytometry analysis where the gray filled curve represents cell only, the black curve represents FITC only, the

blue curve represents the cell permeability of a negative control (scramble peptide), the red curve represents the cell permeability of a positive control (kFGF4-derived MTD), and the green curve represents the cell permeability of each recombinant protein.

5 In order to evaluate the cell permeability of each MTD in mammalian cells, the cellular uptake efficiencies of the His-MTD-EGFP recombinant proteins were compared with those of the positive control protein fused to kFGF4-derived MTD and the negative control protein fused to a scramble peptide, by evaluating the change in median fluorescence after incubation with  
10 each FITC-conjugated His-MTD-EGFP recombinant protein.

Referring to the results shown in Figs. 9a to 9g, none of the cells treated with 10  $\mu$ M of the FITC-conjugated His-MTD-EGFP proteins for 1 hour at 37°C underwent a change in median fluorescence such that the median  
15 fluorescence after treatment was significantly lower than that of the cells treated with the negative control. In contrast, 80 His-MTD-EGFP recombinant proteins, out of the 111 proteins that were tested, exhibited 0.5-fold (50%) or higher fluorescence signal than that of the positive control. These results suggest that some of the newly developed novel MTDs fused with the cargo molecule EGFP exhibit significantly high levels of plasma  
20 membrane-penetrating ability which is enough to deliver macromolecules, such as protein, into live cells.

#### **Example 5 - Determination of Cell Permeability and Intracellular Localization of Recombinant Proteins Fused to MTDs**

25 Out of the 111 FITC-conjugated His-MTD-EGFP recombinant proteins tested by flow cytometry, 60 His-MTD-EGFP recombinant proteins were selected for visualization of cell permeability and intracellular localization thereof. Among them, 50 His-MTD-EGFP recombinant proteins exhibited higher cell permeability than 0.5-fold as compared with the positive control  
30 protein fused to kFGF4-derived MTD.

NIH 3T3 cells were treated without (cell only) or with FITC (FITC only), or FITC-conjugated recombinant proteins such as a negative control (His-scramble peptide-EGFP), a positive control (His-kFGF4-derived MTD-EGFP) or the recombinant proteins fused to novel MTDs (His-MTDs-EGFP), and visualized by confocal laser scanning microscopy. NIH 3T3 cells were cultured for 24 hours in an 8-well chamber slide (LabTek, Nalgen Nunc, Rochester NY, USA). Cells were maintained in DMEM supplemented with 10% fetal bovine serum, 1% penicillin and streptomycin in 5% CO<sub>2</sub> at 37°C. The cells were washed with PBS three times, and then treated for 1 hour with DMEM, DMEM plus free FITC, or DMEM containing 10 µM FITC-conjugated recombinant proteins in 5% CO<sub>2</sub> at 37°C. One hour after the treatment, the cells were fixed in 4% paraformaldehyde (PFA) for 20 minutes at room temperature for observation.

For the direct detection of FITC-conjugated recombinant proteins that were internalized, the cells were washed with PBS three times and counterstained with a nuclear fluorescent stain solution, propidium iodide (PI, Sigma-Aldrich, St Louis MO, USA), at a concentration of 1 µg/ml. After PI staining for 5 minutes, the cells were washed with PBS three times and fixed by polyvinyl alcohol mountain medium with DABCO (Fluca, St Louis MO, USA). The intracellular distribution of the fluorescence was determined at the middle of a single cell analyzed by confocal laser scanning microscopy, where the results are shown in Figs. 10a to 10i. Parameters specific for each fluorochrome were followed as FITC: excited at 488 nm light, detected with a 530 nm bandpass filter.

Surprisingly, as shown in Figs. 10a to 10i, 11a, and 11b, the FITC-conjugated His-MTD-EGFP recombinant proteins were well distributed largely in either the cytoplasm or the nucleus (JO-13, -18, -49, -58, -68, -101, -108, -116, -118, -122, -123, -127, -132, -133, -136, -138, -140, -148, -162, -169, -170, and -172) or both, as compared with the negative control protein fused to a scramble peptide. Ten MTDs (JO-17, -18, -21, -31, -41, -49, -88, -107, -116,

and -169), which exhibited lower cell permeability than 0.5-fold as compared with the positive control (1.0-fold), showed weak intracellular localization ability after an 1 hour treatment. These results completely coincided with those from the flow cytometry analysis.

5

**Example 6 - Determination of *in vivo* Tissue Distribution of Recombinant Proteins Fused to MTDs**

Forty-three recombinant proteins were selected from 111 FITC-conjugated recombinant proteins tested by flow cytometry for the analysis of *in vivo* tissue distribution ability. FITC only (DMEM plus free FITC), a negative control (FITC-conjugated recombinant protein fused to a scramble peptide), a positive control (FITC-conjugated recombinant protein fused to a kFGF4-derived MTD), or the His-MTD-EGFP recombinant proteins (FITC-conjugated recombinant proteins fused to selected novel MTDs) were intraperitoneally injected into Balb/c mice (6 weeks, female, Central Lab. Animal Inc., Korea) at a dose of 300  $\mu$ g/500  $\mu$ l/mouse, respectively. After 2 hours, six organs (i.e., liver, kidney, spleen, lung, heart, and brain) of each mouse were extracted, washed with PBS, and quickly frozen with an O.C.T. compound (Sakura, Japan) on dry ice. Cryosections (20  $\mu$ m thickness) of each organ were prepared by a microtome cryostat (Sakura), placed on glass slides and, then, mounted in a Vectashield mounting medium (Vecta lab, Burlingame CA, USA). *In vivo* tissue distribution of selected novel MTDs was analyzed by using a fluorescence microscope (Nikon, Japan), where the results are shown in Figs. 12a to 12i.

In the negative control, there was no significant level of fluorescence. In contrast, the positive control protein fused to kFGF4-derived MTD showed well distributed fluorescence activity in all of the organs. With respect to the mice injected with the recombinant proteins fused to selected novel MTDs, most MTDs having 0.5-fold or higher cell permeability than the positive control as determined by flow cytometry showed relatively strong *in vivo* tissue

30

distribution ability in all of the organs, although there were some minor differences among the organs. In particular, JO-13 (1.3-fold) and JO-133 (1.5-fold) MTDs were distributed well in all of the organs except the brain. These results indicate that the novel MTDs demonstrated their strong cell permeability and/or intracellular localization ability by flow cytometry and confocal laser scanning microscopy and can effectively mediate macromolecule intracellular transduction *in vitro* and *in vivo*.

In summary, the macromolecule transduction domain (MTD) peptides identified according to the method of the present invention have the characteristics of inducible protein expression, purification and cell permeability as represented in Figs. 13a to 13k.

The present invention has been described in detail with reference to specific embodiments thereof. However, it will be appreciated by those skilled in the art that changes may be made in these embodiments without departing from the principles and spirit of the invention, the scope of which is defined in the appended claims and their equivalents.

**【Claims】****【Claim 1】**

An isolated macromolecule transduction domain (MTD) peptide capable of mediating the transport of a biologically active molecule into a cell,  
5 wherein the peptide has an amino acid sequence selected from the group consisting of SEQ ID NOS: 1-193.

**【Claim 2】**

An isolated polynucleotide encoding the macromolecule transduction domain (MTD) peptide according to Claim 1.

10 **【Claim 3】**

The isolated polynucleotide according to Claim 2, wherein the polynucleotide has a nucleotide sequence selected from the group consisting of SEQ ID NOS: 194-386.

**【Claim 4】**

15 A method of identifying a macromolecule transduction domain (MTD) peptide capable of mediating the transport of a biologically active molecule into a cell, comprising:

identifying secreted proteins having a signal sequence-like domain from multiple amino acid sequence databases;

20 selecting from said identified secreted proteins hydrophobic peptides having a single hydrophobic region at their N-terminus and forming a helix structure; and

optimizing and minimizing said selected hydrophobic peptides under conditions effective to produce peptides suitable for use as an MTD peptide.

25 **【Claim 5】**

The method according to Claim 4, wherein the signal sequence-like domain is selected from the group consisting of penetratin (Arg-Gln-Ile-Lys-Ile-Trp-Phe-Gln-Asn-Arg-Arg-Met-Lys-Trp-Lys-Lys), the Tat peptide derived from the transactivating protein Tat of HIV-1 (Arg-Lys-Lys-Arg-Arg-Gln-Arg-Arg-Pro), transportan (Gly-Trp-Thr-Leu-Asn-Ser-Ala-Gly-Tyr-Leu-Leu-Gly-

30

Lys-Ile-Asn-Lys-Ala-Leu-Ala-Ala-Leu-Ala-Lys-Lys-Ile-Leu), Buforin II (Thr-Arg-Ser-Ser-Arg-Ala-Gly-Leu-Gln-Phe-Arg-Val-Gly-Arg-Val-His-Arg-Leu-Leu-Arg-Lys), MAP (model amphiphatic peptide: Lys-Leu-Ala-Leu-Lys-Ala-Ser-Leu-Lys-Ala-Leu-Lys-Ala-Ala-Leu-Lys-Leu-Ala), k-FGF (Ala-Ala-Val-Ala-Leu-Leu-Pro-Ala-Val-Leu-Leu-Ala-Leu-Leu-Ala-Pro), Ku 70 (Val-Pro-Met-Leu-Lys-Pro-Met-Leu-Lys-Glu), prion (Met-Ala-Asn-Leu-Gly-Tyr-Trp-Leu-Leu-Ala-Leu-Phe-Val-Thr-Met-Trp-Thr-Asp-Val-Gly-Leu-Cys-Lys-Lys-Arg-Pro-Lys-Pro), pVEC (Leu-Leu-Ile-Ile-Leu-Arg-Arg-Arg-Ile-Arg-Lys-Gln-Ala-His-Ala-His-Ser-Lys), pep-1 (Lys-Glu-Thr-Trp-Trp-Glu-Thr-Trp-Trp-Thr-Glu-Trp-Ser-Gln-Pro-Lys-Lys-Lys-Arg-Lys-Val), SynB1 (Arg-Gly-Gly-Arg-Leu-Ser-Tyr-Ser-Arg-Arg-Arg-Phe-Ser-Thr-Ser-Thr-Gly-Arg), pep-7 (Ser-Asp-Leu-Trp-Glu-Met-Met-Met-Val-Ser-Leu-Ala-Cys-Gln-Tyr), and HN-1 (Thr-Ser-Pro-Leu-Leu-Ile-His-Asn-Gly-Gln-Lys-Leu).

**【Claim 6】**

The method according to Claim 4, wherein the amino acid sequence database is selected from the group consisting of the PubMed Entrez Protein Database.

**【Claim 7】**

The method according to Claim 4, wherein said optimizing and minimizing step is carried out by the following steps:

eliminating non-hydrophobic amino acids at the left or right side of the selected hydrophobic peptide to minimize the peptide length while maintaining the hydrophobic region thereof;

replacing non-hydrophobic amino acids at the middle or right side of the selected hydrophobic peptide with a hydrophobic amino acid, proline, to provide flexibility; and

minimizing the number of repeated hydrophobic amino acids to reduce the length of the peptide.

**【Claim 8】**

The method according to Claim 4, wherein the MTD peptide produced

as a result of said optimizing and minimizing exhibits cell permeability, hydrophobicity, and flexibility, has about 7 to about 17 amino acids in length, and forms a helix structure.

**【Claim 9】**

5       The method according to Claim 8, wherein the MTD peptide has an amino acid sequence selected from the group consisting of SEQ ID NOS: 1-193.

**【Claim 10】**

      An isolated recombinant protein having cell permeability comprising:  
      a macromolecule transduction domain (MTD) peptide having an amino  
10   acid sequence selected from the group consisting of SEQ ID NOS: 1-193; and  
      a biologically active molecule.

**【Claim 11】**

      The isolated recombinant protein according to Claim 10, wherein the biologically active molecule is selected from the group consisting of proteins,  
15   polypeptides, and peptides.

**【Claim 12】**

      The isolated recombinant protein according to Claim 10, wherein the biologically active molecule is selected from the group consisting of growth factors, enzymes, transcription factors, toxins, antigenic peptides, antibodies,  
20   and antibody fragments.

**【Claim 13】**

      The isolated recombinant protein according to Claim 12, wherein the biologically active molecule is selected from the group consisting of enzymes, hormones, transport proteins, immunoglobulin, antibodies, structural proteins,  
25   motor function proteins, receptors, signaling proteins, storage proteins, membrane proteins, transmembrane proteins, internal proteins, external proteins, secret proteins, virus proteins, native proteins, glycoproteins, cleaved proteins, proteins with disulfide bonds, protein complexes, chemically modified proteins and prions.

30   **【Claim 14】**

The isolated recombinant protein according to Claim 10, wherein the biologically active molecule is selected from the group consisting of nucleic acids, coding nucleic acid sequences, mRNAs, antisense RNA molecules, carbohydrates, lipids, and glycolipids.

5     **【Claim 15】**

The isolated recombinant protein according to Claim 10, wherein the biologically active molecule is a therapeutic agent.

**【Claim 16】**

10     The isolated recombinant protein according to Claim 15, wherein the biologically active molecule is selected from the group consisting of therapeutic drugs and toxic chemicals.

**【Claim 17】**

15     A method of genetically engineering a biologically active molecule having cell permeability comprising attaching a macromolecule transduction domain (MTD) peptide to a biologically active molecule under conditions effective to produce a biologically active molecule having cell permeability, wherein the MTD peptide has the amino acid sequence selected from the group consisting of SEQ ID NOS: 1-193.

**【Claim 18】**

20     The method according to Claim 17, wherein said attaching comprises attaching a MTD peptide to the N-terminal, C-terminal, both or middle of the biologically active molecule.

**【Claim 19】**

25     The method according to Claim 17, wherein said attaching comprises attaching a MTD peptide to the biologically active molecule by a peptide bond.

**【Claim 20】**

The method according to Claim 17, wherein said attaching comprises attaching a MTD peptide to the biologically active molecule by a covalent bond.

**【Claim 21】**

30     A method of transporting a biologically active molecule into a cell in a

subject comprising administering to a subject a cell permeable recombinant protein comprising a macromolecule transduction domain (MTD) peptide attached to a biologically active molecule, wherein the MTD peptide has an amino acid sequence selected from the group consisting of SEQ ID NOS: 1-193.

5     **【Claim 22】**

Use of a macromolecule transduction domain (MTD) peptide for drug delivery, wherein the MTD peptide has an amino acid sequence selected from the group consisting of SEQ ID NOS: 1-193.

**【Claim 23】**

10     Use of a macromolecule transduction domain (MTD) peptide for vaccine administration, wherein the MTD peptide has an amino acid sequence selected from the group consisting of SEQ ID NOS: 1-193.

**【Claim 24】**

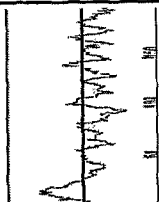
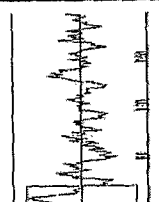
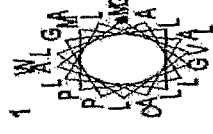
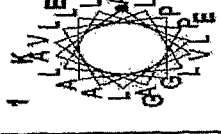
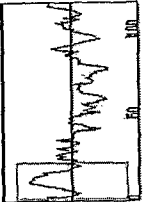
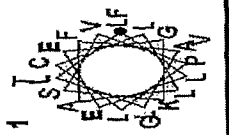
15     Use of a macromolecule transduction domain (MTD) peptide for protein therapy, wherein the MTD peptide has an amino acid sequence selected from the group consisting of SEQ ID NOS: 1-193.

**【Claim 25】**

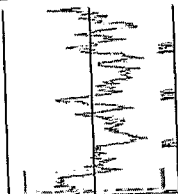
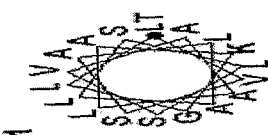
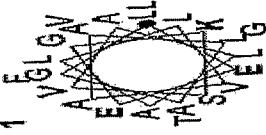
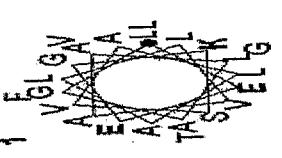
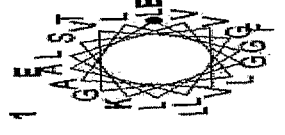
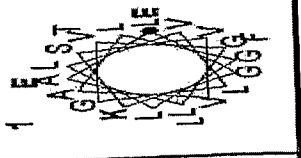
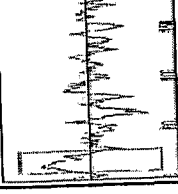
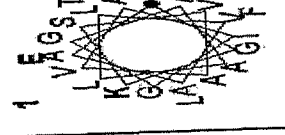
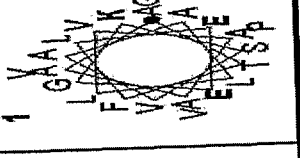
20     Use of a macromolecule transduction domain (MTD) peptide for gene therapy, wherein the MTD peptide has an amino acid sequence selected from the group consisting of SEQ ID NOS: 1-193.

1/134

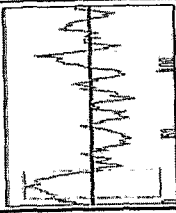
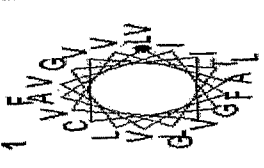
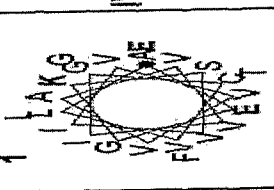
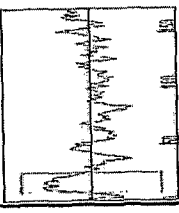
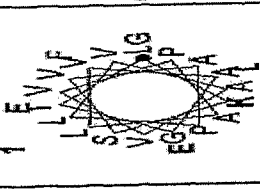
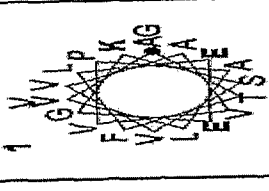
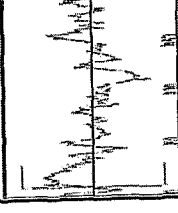
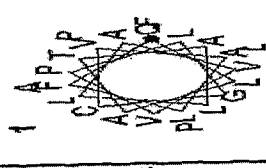
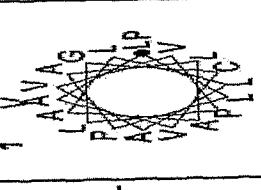
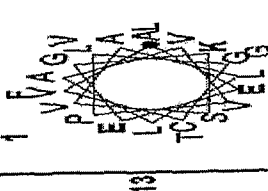
[Fig. 1a]

| Red: hydrophobic Ala<br>Black: hydrophilic Ala<br>Green: nonpolar Ala<br>Blue: basic Ala<br>No: No helix structure |                                                                                               | HRSS: Hydrophobic Region of Signal Sequence<br>Helix: Bird's eye view of peptide structure (determined by SOSUJ program)<br>SOSUJ analysis requires at least 20 amino acids-length peptide. To make 20 Ala-length, additional Ala derived from N-terminus of EGFP was added to the end of the peptide.<br>The numbers of HRSS represent the amino acid number of the origin protein.<br>The number of preliminarily modified and finally modified sequences represents the amino acid-length of the peptide. |                                     |                                                                                       |                                     |                                                                                    |                               |                                                                                     |
|--------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------|---------------------------------------------------------------------------------------|-------------------------------------|------------------------------------------------------------------------------------|-------------------------------|-------------------------------------------------------------------------------------|
| #                                                                                                                  | Origin                                                                                        | Hydropathy                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | HRSS<br>(Candidate Domain)          | Helix                                                                                 | Preliminarily<br>modified Sequences | Helix                                                                              | finally modified<br>Sequences | Helix                                                                               |
| JO-001                                                                                                             | CAC04038 putative<br>NLP/P60-family<br>secreted protein<br>[Streptomyces<br>coelicolorA3(2)]. |                                                                                                                                                                                                                                                                                                                                                                                                                           | 7<br>VLTTAVTWCAIVL<br>AAPG 25       |    | VLAVVCAIVLAAP<br>14                 | No                                                                                 | AVVCAIVLAAP 12                | No                                                                                  |
| JO-002                                                                                                             | NP_057021<br>phosphatidylinositol<br>glycan, class T<br>precursor [Homo<br>sapiens].          |                                                                                                                                                                                                                                                                                                                                                                                                                          | 1<br>MAAAMPLALLVLLLL<br>GPGGWCLA 23 |   | AAAPLALLVLLLLGP<br>LA 17            |  | PLALLVLLLLGP 12               | No                                                                                  |
| JO-003                                                                                                             | NP_072171<br>cholesterol<br>somatomammotropin<br>hormone 2 isoform 3<br>[Homo sapiens].       |                                                                                                                                                                                                                                                                                                                                                                                                                         | 1<br>MAAGSRTSLLLAFAF<br>LCIPWLQE 23 |  | AAGLLAFAFALLCLP<br>14               | No                                                                                 | LLLAFAFALLCLP 11              |  |

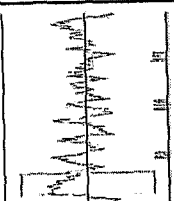
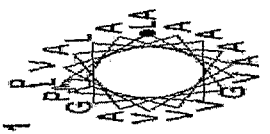
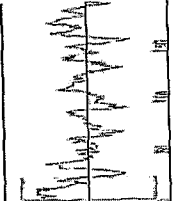
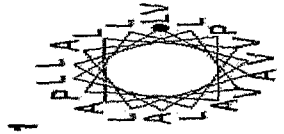
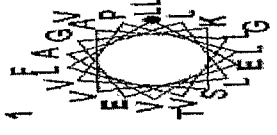
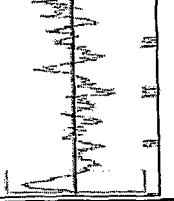
[Fig. 1a] continued

|                                                                                            |                                                                                      |                                    |                                                                                      |                          |                                                                                    |                          |                                                                                    |
|--------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------|------------------------------------|--------------------------------------------------------------------------------------|--------------------------|------------------------------------------------------------------------------------|--------------------------|------------------------------------------------------------------------------------|
| NP_932156 nudix -<br>type motif 9 isoform a<br>[Homo sapiens]                              |   | 5<br>LLGKALAAVSLAL<br>ASVTIRSS 27  |   | LLGALAALLALA 13          |   | LLGALAALLALA 13          |   |
| NP_057327<br>NAD(P)H:quinone<br>oxidoreductase type<br>3, polypeptide A2<br>[Homo sapiens] |   | 5<br>TSPVLLASLGLVLT<br>LLGLAVGS 27 |   | PVLLALGVGLVLLGL<br>AV 17 |   | PVLLALGVGLVLLGL<br>AV 17 |   |
| CAD55300 putative<br>secreted protein.<br>[Streptomyces<br>coelicolor A3(2)]               |  | 8<br>WIAGAAAALLAAG<br>GITYAVAGD 30 |  | AAAALLAAGGIAVA<br>15     |  | AAAALLAAGGIAVA<br>15     |  |

[Fig. 1a] continued

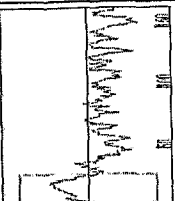
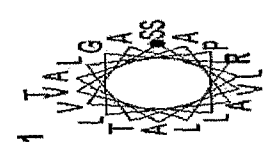
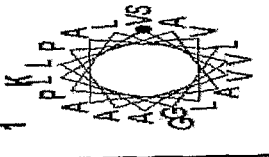
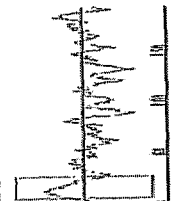
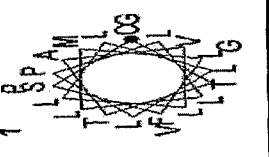
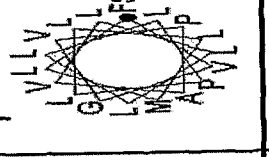
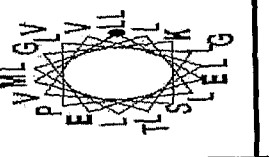
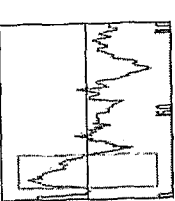
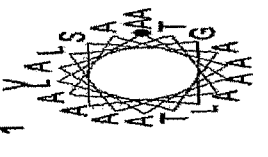
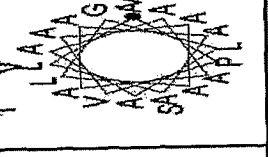
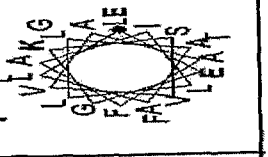
|        |                                                                                       |                                                                                      |                                    |                                                                                      |                           |                                                                                    |                 |                                                                                    |
|--------|---------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------|------------------------------------|--------------------------------------------------------------------------------------|---------------------------|------------------------------------------------------------------------------------|-----------------|------------------------------------------------------------------------------------|
| JO-007 | NP_629514 secreted protein<br>[Streptomyces coelicolor A3(2)].                        |   | 5<br>LALTVCGIIVAVVIG<br>LFVFGL 26  |   | LALVCGIIVAVVVI<br>14      |   | INVAVVVI 8      | No                                                                                 |
| JO-008 | CAB57190 putative secreted chitin binding protein<br>[Streptomyces coelicolor A3(2)]. |   | 7<br>TAALIGAVLAPVAV<br>SLPA 25     | No                                                                                   | AALIGAVLAPVAVL<br>PA 17   |   | AVLAPVAV 9      |   |
| JO-009 | CAB51015 putative secreted protein<br>[Streptomyces coelicolor A3(2)].                |  | 9<br>QELAVCGLPWALLA<br>TALFAPLP 31 |  | LAVCGLPWALLAAL<br>APLP 19 |  | LAVCGLPWALLA 13 |  |

[Fig. 1a] continued

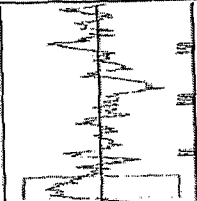
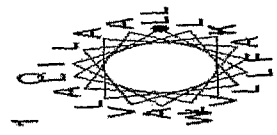
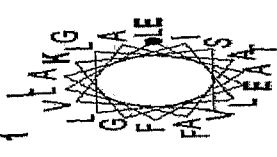
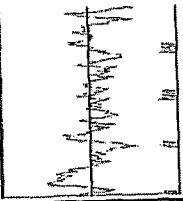
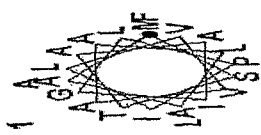
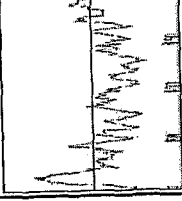
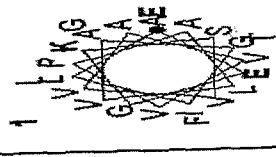
|        |                                                                                   |                                                                                      |                                      |                                                                                      |                            |                                                                                    |                          |                                                                                    |
|--------|-----------------------------------------------------------------------------------|--------------------------------------------------------------------------------------|--------------------------------------|--------------------------------------------------------------------------------------|----------------------------|------------------------------------------------------------------------------------|--------------------------|------------------------------------------------------------------------------------|
| JO-010 | NP_62502 (glycosyl hydrolase (secreted protein) [Streptomyces coelicolor A3(2)]). |   | 14<br>LLVALGGAVVAAPVA<br>AAVAPHAL 36 |   | LLVALGGAVVAAPV<br>AAVAP 20 |   | LGGA VVAAPVAAAV<br>AP 17 |   |
| JO-011 | NP_630686 secreted protein [Streptomyces coelicolor A3(2)].                       |   | 15<br>RLLLVLA VLLAVLT<br>TAPEALA 36  | No                                                                                   | LLLVLAVLLAVLPAP<br>ALA 18  |   | LLLVLAVLLAVLP 13         |   |
| JO-012 | NP_057329 dehydrogenase (reductase) (SDR family) member 8 (Homo sapiens).         |  | 4<br>LLDILLPLLVCSLE<br>SFVKLF 26     |  | LLILLPLLVNL 13             |  | LLILLPLLV 12             |  |

5/134

[Fig. 1b]

| #      | Origin                                                                        | Hydropathy                                                                           | HRSS<br>(Candidate Domain)         | Helix                                                                                | Preliminary<br>modified<br>Sequences | Helix                                                                              | finally modified<br>Sequences | Helix                                                                              |
|--------|-------------------------------------------------------------------------------|--------------------------------------------------------------------------------------|------------------------------------|--------------------------------------------------------------------------------------|--------------------------------------|------------------------------------------------------------------------------------|-------------------------------|------------------------------------------------------------------------------------|
| JO-013 | NP_639877 putative<br>secreted protein<br>[Streptomyces<br>coelicolor A3(2)]  |   | 4<br>SVLRGLAAALAVLP<br>LTVSTPAH 26 |   | VLGLAAAAALAVLPV<br>PA 17             |   | LAAAAALVLP 11<br>No           |                                                                                    |
| JO-014 | NP_699201<br>protease inhibitor 16<br>precursor [Homo<br>sapiens].            |   | 5<br>CSFLMILLPILLLV<br>ATTGPVG 26  |   | FLMILLPILLLVAP<br>VG 17              |   | FLMILLPILLLV 14               |   |
| JO-015 | NP_639877 putative<br>secreted protein<br>[Streptomyces<br>coelicolor A3(2)]. |  | 6<br>ALTASAATAAAALG<br>LAAVPAQ 28  |  | ALAAAAAALGLA<br>AAVPA 19             |  | AAAAAALGLAAAVP<br>A 15        |  |

[Fig. 1b] continued

|        |                                                                                                |                                                                                      |                                    |                                                                                     |                            |                                                                                    |                       |                                                                                    |  |
|--------|------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------|------------------------------------|-------------------------------------------------------------------------------------|----------------------------|------------------------------------------------------------------------------------|-----------------------|------------------------------------------------------------------------------------|--|
| JO-016 | CAB86250 putative<br>secreted protein<br>[Neisseria<br>meningitidis Z2491].                    |   | 7<br>LLLAALLLIIFAFAVKL<br>VLLQW 27 |  | LLLAALLLIIFAFAVLV<br>LL 18 |   | LLLAALLLIIFAFAV<br>14 |   |  |
| JO-017 | NP_626397 small<br>secreted hydrophilic<br>protein<br>[Streptomyces<br>coelicolor A3(2)].      |   | 7<br>MAALAAVVLIPLGIAA<br>TSFAL 27  |  | MAALAAVVLIPLGIAA<br>16     |   | ALAAVLIPLGIAA<br>14   | No                                                                                 |  |
| JO-018 | CAB57190 putative<br>secreted chitin<br>binding protein<br>[Streptomyces<br>coelicolor A3(2)]. |  | 7<br>TAALIGAVLPWAV<br>SLPA 25      | No                                                                                  | AALIGAVLPWAVL<br>PA 17     |  | AALIGAVLPWAV<br>14    |  |  |

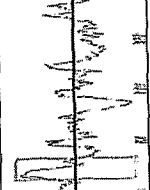
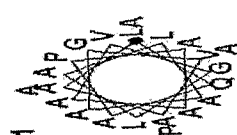
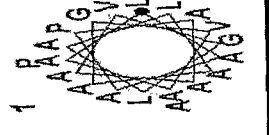
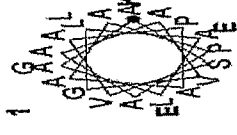
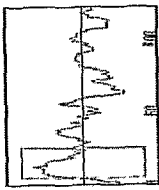
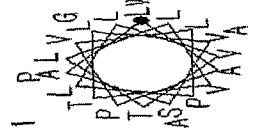
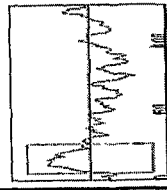
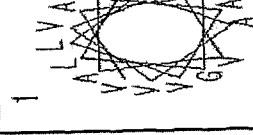
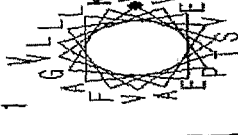
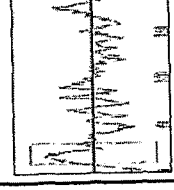
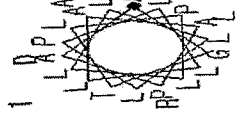
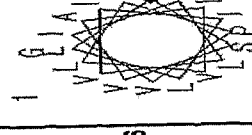
7/134

[Fig. 1b] continued

|                                                                                            |  |                                   |    |                          |    |                        |    |
|--------------------------------------------------------------------------------------------|--|-----------------------------------|----|--------------------------|----|------------------------|----|
| NP_626007<br>secreted cellulose-<br>binding protein<br>(Streptomyces<br>coelicolor A3(2)). |  | 10<br>AAAVAVAGLPLAL<br>23         | No | AAAVAVAGLPLAL<br>14      | No | AAGIAVAIAIVPLA<br>15   |    |
| NP_626632secreted<br>protein<br>(Streptomyces<br>coelicolor A3(2))                         |  | 6<br>GAAGIAVAIAIVPLA<br>DPAPA 26  | No | AAGIAVAIAIVPLA<br>15     | No | IIVAIAIVPLA 12         | No |
| CAC31790 putative<br>secreted protein<br>(Mycobacterium<br>leprae).                        |  | 5<br>GLAANVAMAAAATV<br>LAAPALA 25 |    | LAAVAMAAAALAA<br>PALA 18 |    | VAMAAAIVLAAPAL<br>A 15 |    |

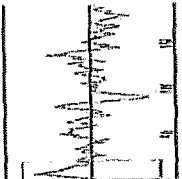
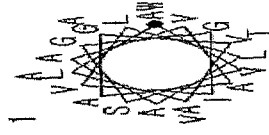
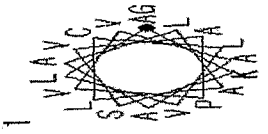
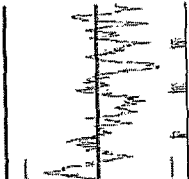
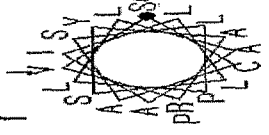
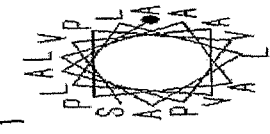
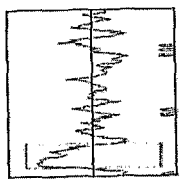
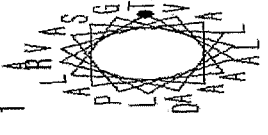
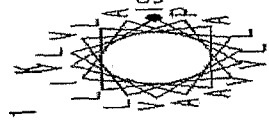
8/134

[Fig. 1b] continued

|        |                                                                             |                                                                                       |                                      |                                                                                       |                                |                                                                                     |                         |                                                                                   |
|--------|-----------------------------------------------------------------------------|---------------------------------------------------------------------------------------|--------------------------------------|---------------------------------------------------------------------------------------|--------------------------------|-------------------------------------------------------------------------------------|-------------------------|-----------------------------------------------------------------------------------|
| JO-022 | CAB38593putative<br>secreted protein<br>[Streptomyces<br>coelicolor A3(2)]. |    | 11<br>LAAVGAALALGVAA<br>APAAQAAPA 32 |    | LAAVGAALALGVAA<br>APAAQAAPA 21 |    | AALALGVAAAPAAA<br>PA 16 |  |
| JO-023 | NP_630266secreted<br>Protein<br>[Streptomyces<br>coelicolor A3(2)].         |    | 5<br>LASVLTVLLTVLLPL<br>VPAWPAAG 27  |    | LLLALLAAGLVLP<br>15            |    | LAVLVLLVLLP 11          | No                                                                                |
| JO-024 | NP_630165secreted<br>Protein<br>[Streptomyces<br>coelicolor A3(2)].         |    | 5<br>TAARTVAVLAPVLF<br>ALQFFAPS 27   |   | ALVAVAVAWALLGV<br>A 16         |    | WAVLAPVL 9              |  |
| JO-025 | NC_003888secreted<br>Protein<br>[Streptomyces<br>coelicolor A3(2)].         |  | 11<br>AAPAALLPLLLLLP<br>LTGCDRLA 33  |  | LLLIIVLLVLPALVIA 16            |  | AALLPLLLLLP 12          | No                                                                                |

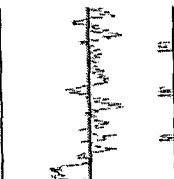
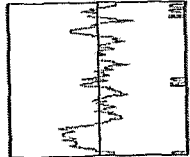
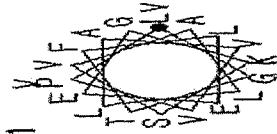
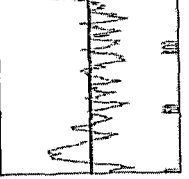
9/134

[Fig. 1c]

| #      | Origin                                                              | Hydropathy                                                                           | HRSS<br>(Candidate Domain)          | Helix                                                                                | Preliminary<br>modified<br>Sequences | Helix                                                                              | finally modified<br>Sequences | Helix |
|--------|---------------------------------------------------------------------|--------------------------------------------------------------------------------------|-------------------------------------|--------------------------------------------------------------------------------------|--------------------------------------|------------------------------------------------------------------------------------|-------------------------------|-------|
| JO-026 | NP_627363secreted<br>Protein<br>[Streptomyces<br>coelicolor A3(2)]. |   | 18<br>ALAVGAIVAALLVIG<br>GSWAVTA 40 |   | ALVLCALAAWPA<br>15                   |   | PAAVAALLVI 10<br>No           |       |
| JO-027 | NP_631288secreted<br>Protein<br>[Streptomyces<br>coelicolor A3(2)]. |   | 4<br>SYRAYSLIAALLPLS<br>ACGIPET 26  |   | AAPVPPAALALLVA<br>15                 |   | LLIAALLP 8<br>No              |       |
| JO-028 | NP_630325secreted<br>Protein<br>[Streptomyces<br>coelicolor A3(2)]. |  | 4<br>TRALSAAVLLGLAA<br>APAAADSS 26  |  | IVALLVPLVLAIAAV<br>L 17              |  | AAVVLLPLAASP 12<br>No         |       |

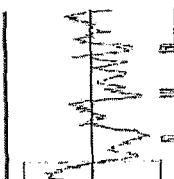
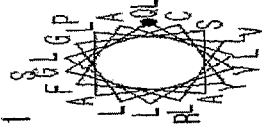
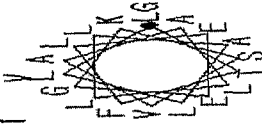
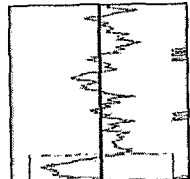
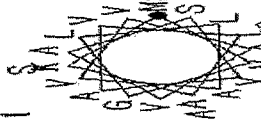
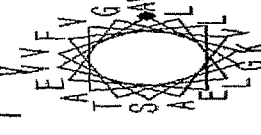
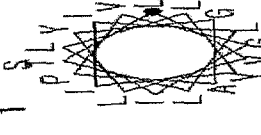
10/134

[Fig. 1c] continued

|        |                                                                              |                                                                                      |                                    |                                                                                     |                        |                                                                                    |                |                                                                                   |
|--------|------------------------------------------------------------------------------|--------------------------------------------------------------------------------------|------------------------------------|-------------------------------------------------------------------------------------|------------------------|------------------------------------------------------------------------------------|----------------|-----------------------------------------------------------------------------------|
| JO-029 | NP_631289 secreted Protein<br>(Streptomyces coelicolor A3(2)).               |   | 7<br>LWAAAAAALTLVGA<br>APAAQAAP 29 |  | IVALLVPLVLAIAV<br>L 17 |   | AAAAAALLVP 10  | No                                                                                |
| JO-030 | CA851015 Putative secreted Protein<br>(Streptomyces coelicolor A3(2)).       |   | 9<br>QFLAVCGLPVALLA<br>TALFAPLP 31 |  | LAVLPWALLAALFA<br>P 16 |   | LPVVALLA 8     |  |
| JO-031 | NP_629515 chitinase C (secreted Protein)<br>(Streptomyces coelicolor A3(2)). |  | 7<br>AAALAATLALPLAGL<br>VGLA 25    | No                                                                                  | PPLVLAALVACIVLI<br>16  |  | AAALAAPLALP 11 | No                                                                                |

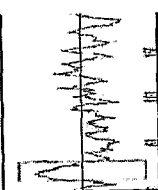
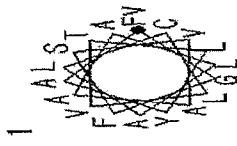
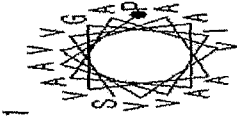
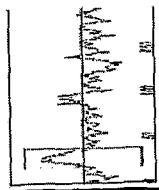
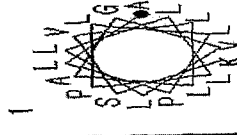
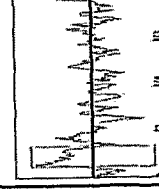
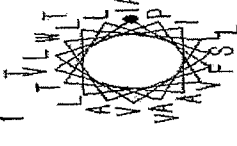
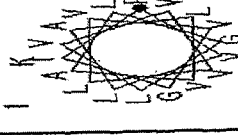
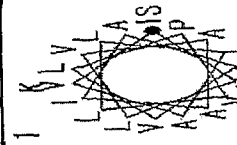
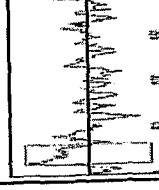
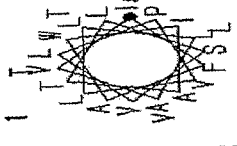
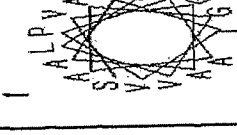
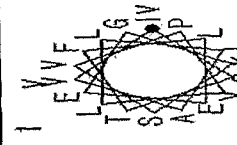
11/134

[Fig. 1c] continued

|        |                                                                                     |                                                                                      |                                     |                                                                                      |                         |    |                 |                                                                                   |
|--------|-------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------|-------------------------------------|--------------------------------------------------------------------------------------|-------------------------|----|-----------------|-----------------------------------------------------------------------------------|
| JO-032 | NP_940995 C1q and tumor necrosis factor related Protein 1 isoform 1 (Homo sapiens). |   | 6<br>QGLLAYCILLAFAS<br>GLVLSRVP 28  |   | LALAVLAAALLAVAP<br>A 16 | No | LLALLLAA 9      |  |
| JO-033 | NP_854150 POSSIBLE CONSERVED SECRETED PROTEIN (Mycobacterium bovis AF2122/97).      |   | 1<br>MKALVAVSAVAVWA<br>LLGVSSAQ 22  |   | LGLPPLLALLAGGA<br>A 16  | No | AVAVVALL 8      |  |
| JO-034 | NP_630361 Probable secreted Protein (Streptomyces coelicolor A3(2)).                |  | 7<br>LTLIIIVLLIGVPAGYL<br>VISANQ 29 |  | LVPLGALALVWAFPA<br>A 16 | No | LLLIIVLLIIVP 11 | No                                                                                |

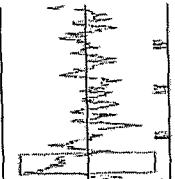
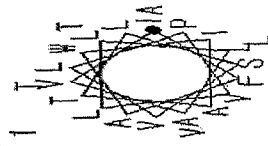
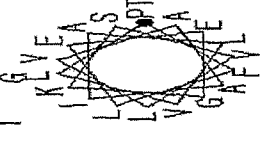
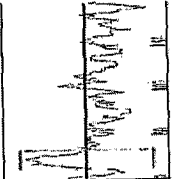
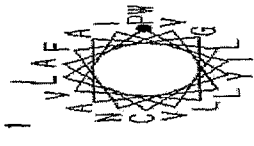
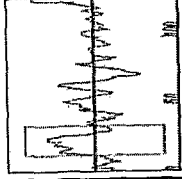
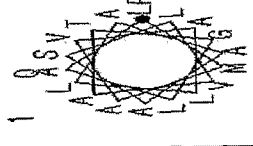
12/134

[Fig. 1c] continued

|        |                                                                            |                                                                                       |                                     |                                                                                       |                         |                                                                                     |                         |                                                                                     |
|--------|----------------------------------------------------------------------------|---------------------------------------------------------------------------------------|-------------------------------------|---------------------------------------------------------------------------------------|-------------------------|-------------------------------------------------------------------------------------|-------------------------|-------------------------------------------------------------------------------------|
| J0-035 | P39790<br>Extracellular<br>metallic Protease<br>Precursor.                 |    | 12<br>FAYLTVLCALAAV<br>SFGVPAK 33   |    | PAVIGVAANVWAAAA<br>15   |    | LALAAVVP 9              | No                                                                                  |
| J0-036 | CAA19252 Putative<br>[Pc] Protein<br>[Streptomyces<br>coelicolor A3(2)].   |    | 13<br>AAPRSVPPARALAG<br>LLLVLTAL 35 | No                                                                                    | ALPILPLLVLGALL<br>15    |    | PAALALLVA 10            | No                                                                                  |
| J0-037 | NP_625685 large<br>secreted Protein<br>[Streptomyces<br>coelicolor A3(2)]. |    | 12<br>IVALLVPLVSLTAIW<br>AFATVLT 34 |    | PIGVVLWVLGLAVL<br>AL 17 |    | IVALLVPLVLAIAAV<br>L 17 |    |
| J0-038 | NP_625685 large<br>secreted Protein<br>[Streptomyces<br>coelicolor A3(2)]. |  | 12<br>IVALLVPLVSLTAIW<br>AFATVLT 34 |  | VLVGAAAAPVILVAA<br>G 15 |  | IVALLVP 8               |  |

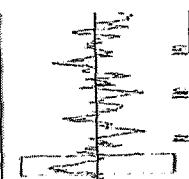
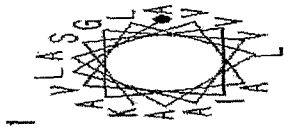
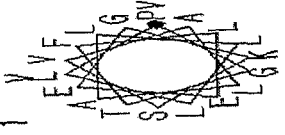
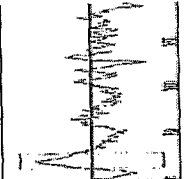
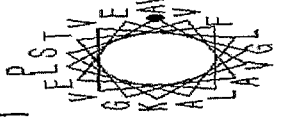
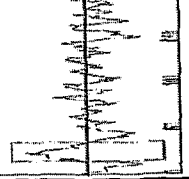
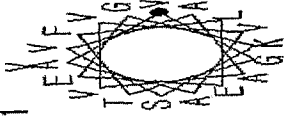
13/134

[Fig. 1d]

| #      | Origin                                                                    | Hydropathy                                                                            | HRSS<br>(Candidate Domain)            | Helix                                                                                | Preliminarily<br>modified Sequences | Helix                                                                             | finally modified<br>Sequences | Helix                                                                             |
|--------|---------------------------------------------------------------------------|---------------------------------------------------------------------------------------|---------------------------------------|--------------------------------------------------------------------------------------|-------------------------------------|-----------------------------------------------------------------------------------|-------------------------------|-----------------------------------------------------------------------------------|
| JO-039 | NP_625685large<br>secreted Protein<br>[Streptomyces<br>coelicolor A3(2)]. |    | 12<br>IVALLVPLVSLTAIW<br>AFATVLT 34   |   | LAVLVLLVLLPLVPA<br>15               |  | PLVLAIAAVL 10                 |  |
| JO-040 | NP_808800 golgi<br>PhosphoProtein 2<br>[Homo sapiens].                    |    | 12<br>KSPPLVLAALVACIIV<br>LGFNYWII 34 |   | WVAVLAPVLFAL 12                     | No                                                                                | PLVLAALVA 9                   | No                                                                                |
| JO-041 | NP_625693secreted<br>Protein<br>[Streptomyces<br>coelicolor A3(2)].       |  | 12<br>LALGTAVLSAAALLA<br>VAMPQEAQ 34  |  | AALLPIILLIPL 13                     | No                                                                                | AAALLAVA 8                    | No                                                                                |

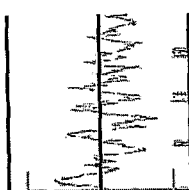
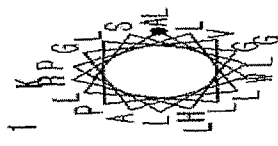
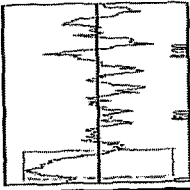
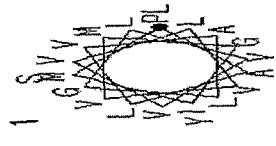
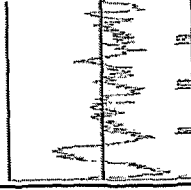
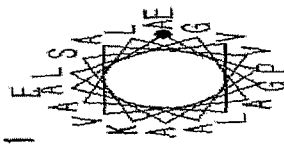
14/134

[Fig. 1d] continued

|        |                                                                                          |                                                                                      |                                      |                                                                                      |                      |                                                                                   |           |                                                                                    |
|--------|------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------|--------------------------------------|--------------------------------------------------------------------------------------|----------------------|-----------------------------------------------------------------------------------|-----------|------------------------------------------------------------------------------------|
| JO-042 | NP_004863 thymic dendritic cell-derived factor 1 [Homo sapiens].                         |   | 13<br>QLGLPPLLTLTALA<br>GGSGTASA 35  | No                                                                                   | ALAVGAAVAALLVI<br>14 |  | PLLLALA 8 |   |
| JO-043 | NP_631398 secreted Protein [Streptomyces coelicolor A3(2)].                              |   | 14<br>LTVPLGALALVAFP<br>ATA31        | No                                                                                   | LLIAALLPL 9          |  | ALALVVA 7 |   |
| JO-044 | NP_627373 Penicillin-binding Protein (secreted Protein) [Streptomyces coelicolor A3(2)]. |  | 15<br>KPAVIGSVAANVVAG<br>AGFGAYAM 37 |  | LAAVLLGLAAAPAA<br>15 | No                                                                                | VAAVWAA 8 |  |

15/134

[Fig. 1d] continued

|        |                                                                                                  |                                                                                      |                                    |                                                                                     |                        |                                                                                    |               |                |    |
|--------|--------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------|------------------------------------|-------------------------------------------------------------------------------------|------------------------|------------------------------------------------------------------------------------|---------------|----------------|----|
| JO-045 | NP_056226<br>sulfatase modifying<br>factor 2 [Homo<br>sapiens].                                  |   | 2<br>ARHGLPLLISLV<br>GAWLKLG 23    |  | WAAAAAALLVGAAP<br>A 15 | No                                                                                 | No            | PIPLLLLV 9     | No |
| JO-046 | NP_854998<br>Conserve<br>d hypothelial<br>secreted protein<br>[Mycobacterium<br>bovis AF212/97]. |   | 4<br>PMIGMVLVVVLGLA<br>VLALSYRL 28 |  | LAVCGLPWALLA 13        | No                                                                                 | No            | WLVWLPLAVLA 13 | No |
| JO-047 | NP_627512<br>secreted<br>Protein<br>[Streptomyces<br>coelicolor A3(2)].                          |  | 23<br>VLYGAAAVPVMILVA<br>AGC 39    | No                                                                                  | AAAVAVAGLAPLAL<br>14   |  | AAAVPVLVAA 10 | No             | No |

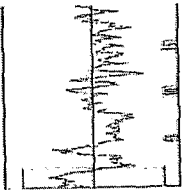
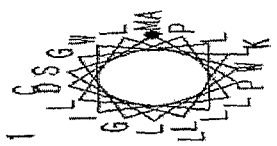
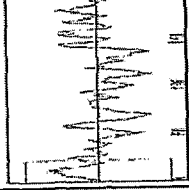
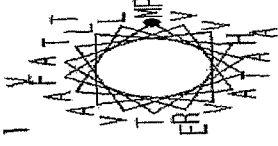
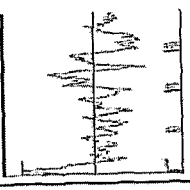
16/134

[Fig. 1d] continued

|                                                                                                |  |                                    |    |                      |  |                 |    |
|------------------------------------------------------------------------------------------------|--|------------------------------------|----|----------------------|--|-----------------|----|
| NP_110448<br>phospholipase A2,<br>group X11A [Homo<br>sapiens].                                |  | 1<br>MALLSRPALTLLLL<br>MAAVVRCQ 23 |    | ALLPALLLLLAHV<br>15  |  | PALLLLLAHV 12   | No |
| NP_003245<br>inhibitor of<br>metalloproteinase 1<br>precursor [Homo<br>sapiens].               |  | 1<br>MAPFEPLASGILL<br>WLIAPSR 23   |    | APPLAILLLLIAPA15     |  | PLAILLLLIAP 12  | No |
| NP_002978<br>small<br>inducible cytokine<br>A17 precursor [Homo<br>sapiens].                   |  | 1<br>MAPLKMALVTLILG<br>ASLQIH 23   | No | APLLALVLLALIA14      |  | PLLALVLLALIA 13 | No |
| NP_001012495<br>stromal cell derived<br>factor 1 isoform<br>gamma precursor<br>[Mus musculus]. |  | 1<br>MDAKVVAVLALVLA<br>ALCISDGK 22 |    | AVVAVLALVLAALI<br>14 |  | VVAVLALVLAAL 12 | No |

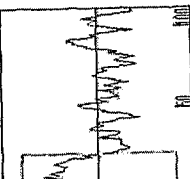
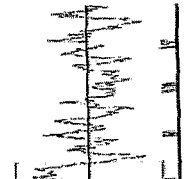
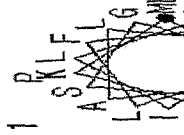
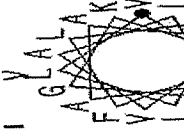
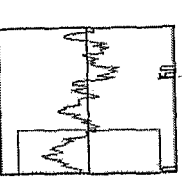
17/134

[Fig. 1e]

| #      | Origin                                                                             | Hydropathy                                                                           | HRSS<br>(Candidate Domain)          | Helix                                                                               | Preliminarily<br>modified Sequences | Helix | finally modified<br>Sequences | Helix |
|--------|------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------|-------------------------------------|-------------------------------------------------------------------------------------|-------------------------------------|-------|-------------------------------|-------|
| JO-052 | NP_775628 ficolin 3<br>isoform 2 precursor<br>[Homo sapiens].                      |   | 1<br>MDLLWILPSLWILL<br>GGPACK 22    |  | LLILPLLILPAL 13                     | No    | PLLLLPAL 9                    | No    |
| JO-053 | NP_624483 secreted<br>protein<br>[Streptomyces<br>coelicolor A3(2)].               |   | 1<br>MFRHLAAVATALAVV<br>TVTPVEAT 23 |  | LAAVAALAVVVPVA<br>14                | No    | LAAVAALAVVVP 12               | No    |
| JO-054 | NP_997465 HERV-<br>FRD provirus<br>ancestral Env<br>polyprotein [Homo<br>sapiens]. |  | 1<br>MGLLLVLITPSLAA<br>Y 17         | No                                                                                  | ILLVLVLPLAA12                       | No    | ILLVLVLPLAA 12                | No    |

18/134

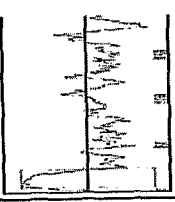
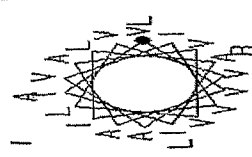
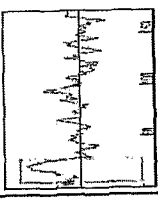
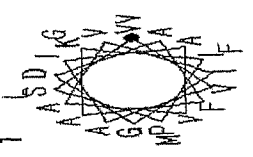
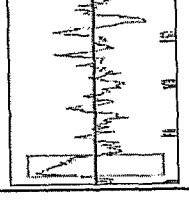
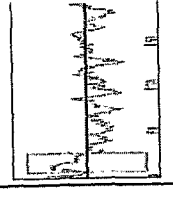
[Fig. 1e] continued

|        |                                                                                             |                                                                                      |                                     |                                                                                      |                      |    |               |                                                                                   |
|--------|---------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------|-------------------------------------|--------------------------------------------------------------------------------------|----------------------|----|---------------|-----------------------------------------------------------------------------------|
| JO-055 | NP_854234 possible conserved secreted protein (Mycobacterium bovis AF2122/97).              |   | 1<br>IMKGTKLAVVGMTV<br>AAVSLAAP 22  |   | LAWVWAIVLAAP 13      | No | LAVVWAAV 9    |  |
| JO-056 | P23284 Peptidyl-prolyl cis-trans isomerase B precursor (PPIase) (Rotamase) (Cyclophilin B). |   | 1<br>IMKVLLAAALIAGSVF<br>FLLPGP 22  |   | VLLAAALIAVLLPP1<br>5 | No | VLLAAALIA 9   |  |
| JO-057 | CAD05047 hypothetical secreted protein (Salmonella enterica subsp. Enterica serovar Typhi). |  | 1<br>MLKVTTLIASLLAAPL<br>AFSASAQ 23 |  | LVLIALLAAPLAAAI4     | No | LIALLAAPLA 10 | No                                                                                |



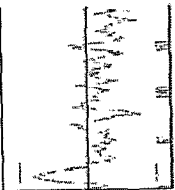
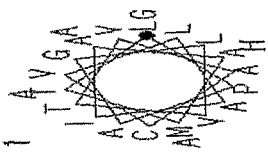
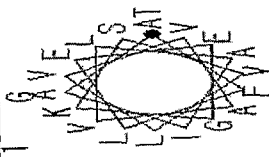
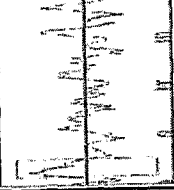
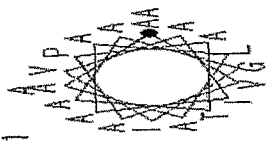
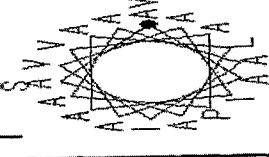
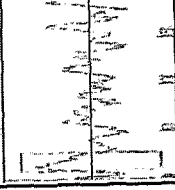
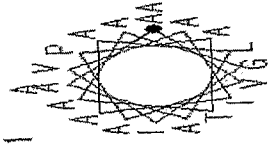
20/134

[Fig. 1e] continued

|        |                                                                                         |                                                                                       |                                     |                                                                                       |                              |                  |               |    |
|--------|-----------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------|-------------------------------------|---------------------------------------------------------------------------------------|------------------------------|------------------|---------------|----|
| JO-061 | NP_626583secreted<br>protein<br>[Streptomyces<br>coelicolor A3(2)]                      |    | 5<br>VIALIVIVAVLVVAA<br>VLALR 28    |    | VIALIVIVAVLVVAA<br>VLAL 21   | No               | WLVVAAVLAL 11 | No |
| JO-062 | NP_856548<br>SOLUBLE<br>SECRETED<br>ANTIGEN MPB53<br>[Mycobacterium<br>bovis AF212297]. |    | 6<br>VSPKAFADGIVAVAI<br>AVWL MFG 28 |    | VPIAAVAIAVWL 14              | VAVAI AVL 9      | No            | No |
| JO-063 | NP_629854secreted<br>protein<br>[Streptomyces<br>coelicolor A3(2)].                     |    | 9<br>PLIVVAAAVVAVGA<br>GLAVWATA 31  |   | AAALLALLAALPAAA<br>LLLLAA 15 | PLIVVAAA VAV 13  | No            | No |
| JO-064 | AAB59058<br>lambda receptor<br>protein [Escherichia<br>coli].                           |  | 7<br>KLPLAVAVAAGVMS<br>AQAMAVDF 28  |  | ILPLAVAVA AAVAAV<br>14       | PLAVAVA AAVAA 11 | No            | No |

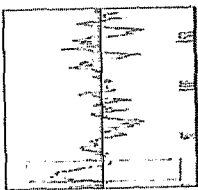
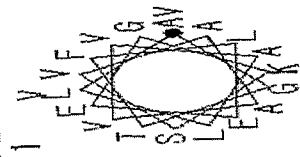
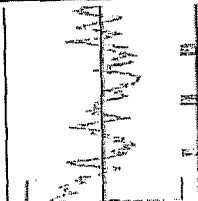
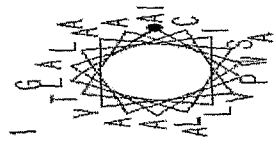
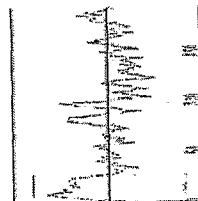
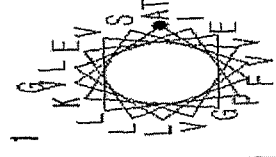
21/134

[Fig. 1f]

| #      | Origin                                                                                 | Hydropathy                                                                            | HRSS<br>(Candidate Domain)        | Helix                                                                                | Preliminarily<br>modified<br>Sequences | Helix                                                                             | finally modified<br>Sequences | Helix                                                                             |
|--------|----------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------|-----------------------------------|--------------------------------------------------------------------------------------|----------------------------------------|-----------------------------------------------------------------------------------|-------------------------------|-----------------------------------------------------------------------------------|
| JO-065 | NP_825185<br>NLP160 family<br>secreted protein<br>[Streptomyces<br>avermilis MA-4680]. |    | 8<br>LTMAIALVCAVTVL<br>GAPGAHA 30 |   | LAAIALVAVVLAPAA<br>A 16                | No                                                                                | AAIALVAVVL 10                 |  |
| JO-066 | NP_626568 secreted<br>protein<br>[Streptomyces<br>coelicolor A3(2)].                   |    | 9<br>AAALAAIIVIGATAP<br>AVAA 28   |   | AAALAAIIVIAAPAV<br>AA 18               |  | AAALAAIIVI 10                 | No                                                                                |
| JO-067 | NP_626568 secreted<br>protein<br>[Streptomyces<br>coelicolor A3(2)].                   |  | 9<br>AAALAAIIVIGATAP<br>AVAA 28   |  | AAALAAIIVIAAPAV<br>AA 18               | No                                                                                | AAAPAVAA 8                    | No                                                                                |

22/134

[Fig. 1f] continued

|        |                                                               |                                                                                      |                                     |                                                                                      |                                |                                                                                   |              |    |                                                                                    |
|--------|---------------------------------------------------------------|--------------------------------------------------------------------------------------|-------------------------------------|--------------------------------------------------------------------------------------|--------------------------------|-----------------------------------------------------------------------------------|--------------|----|------------------------------------------------------------------------------------|
| JO-068 | NP_625633 secreted protein<br>[Streptomyces coelicolor A3(2)] |   | 7<br>GAGALLASILLALP<br>FTAEAES 29   | No                                                                                   | PLIVVAAAVVAVAL<br>A<br>VAAV 19 |  | ILLALP 7     | No |   |
| JO-069 | CAC32053 putative secreted protein<br>[Mycobacterium leprae]  |   | 6<br>ALLSAVVCANWATL<br>ILAPIGAAA 28 |   | AVVAAAILLAPI 12                | No                                                                                | ALLAVVAA 8   |    |                                                                                    |
| JO-070 | NP_630954 secreted protein<br>[Streptomyces coelicolor A3(2)] |  | 6<br>AWGVVFLSPILLAG<br>AGMVLVSS 28  |  | AVVVLPILLAA 12                 | No                                                                                | AVVVLPILL 10 |    |  |

23/134

[Fig. 1f] continued

|                                                                                         |                                        |    |                 |    |                |    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
|-----------------------------------------------------------------------------------------|----------------------------------------|----|-----------------|----|----------------|----|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 997300<br>Neuroplastin precursor (Stomal cell-derived receptor 1) (SDR-1).<br>JO-071    | <br>6<br>LPGALALSLLVSGS<br>LLPGGA 27   | No | ALAALLVLLPPA 13 | No | ALAALLVP 9     | No | <br>1<br>GAVE<br>KMAEA<br>AALAA<br>LXXAA<br>VXXAA<br>AAGAA<br>GLAAE<br>LFAVE                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| NP_624483 secreted protein (Streptomyces coelicolor A3(2)).<br>JO-072                   | <br>1<br>MERHLAAVATALAVI<br>TVPVEAT 23 | No | LAAVAALAVVP 12  | No | LAAVAALAVV 10  | No | <br>1<br>YFATT<br>AALLL<br>VTTAA<br>ERVAA<br>ATAAA                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| AAA41949 Rat parathyroid gland acidic proline-rich protein mRNA, complete CDS<br>JO-073 | <br>1<br>MLVLLTAALLVLSS<br>AHGSDEEV 23 | No | LVVLLAALLVL 11  | No | LVVLLAALLVL 11 | No | <br>1<br>LAALL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL<br>LAAAL |

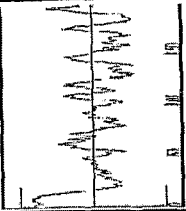
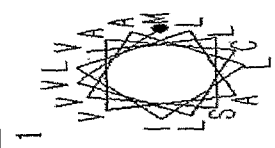
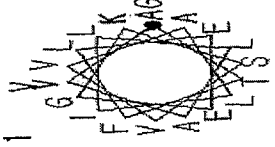
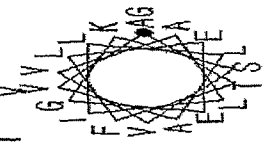
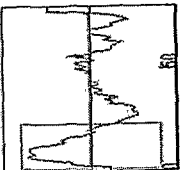
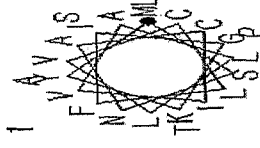
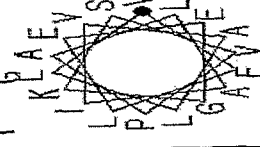
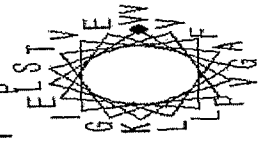
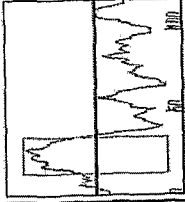
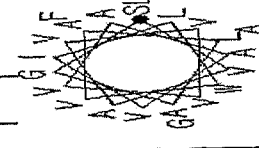
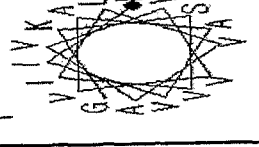
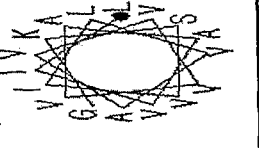
24/134

[Fig. 1f] continued

|        |                                                                                           |  |                                     |                  |    |            |    |
|--------|-------------------------------------------------------------------------------------------|--|-------------------------------------|------------------|----|------------|----|
| JO-074 | AAAT7887<br>Drosophila<br>melanogaster spatzie<br>(spz) gene                              |  | 1<br>MMTPMWWISLFKVL<br>LFAFFAT 22   | PILVLLLA 9       |    | PVLLLLAP 8 |    |
| JO-075 | NP_627867 conserved<br>secreted protein<br>[Streptomyces<br>coelicolor A3(2)]             |  | 8<br>AVSALAGLVLAGSAL<br>AVWNAAPA 30 | LVLAALAVVAAPA 13 | No | ALAVVAAP 8 | No |
| JO-076 | NP_631283 secreted<br>protein<br>[Streptomyces<br>coelicolor A3(2)]                       |  | 8<br>FLIAGVIVALLAVFTV<br>VRAVRIV 30 | VIVALLAVVVAV 12  | No | VVALLAV 8  |    |
| JO-077 | NP_003231<br>endometrial bleeding<br>associated factor<br>preproprotein [Homo<br>sapiens] |  | 1<br>MMPLWLCWALWVL<br>PLAGPGAALT 23 | PLLALVLPAP 11    | No | ALVLPAP 8  | No |

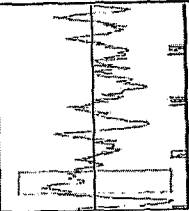
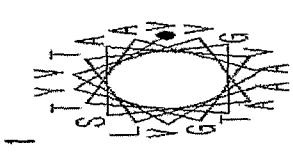
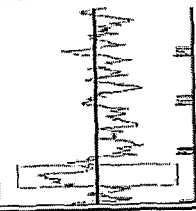
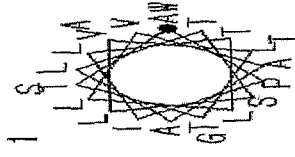
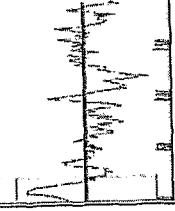
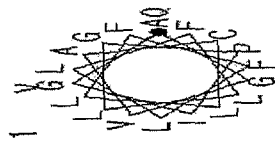
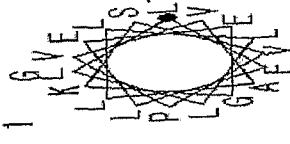
25/134

[Fig. 1g]

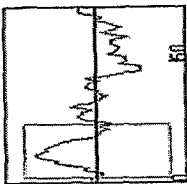
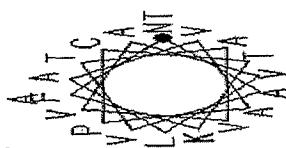
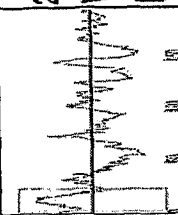
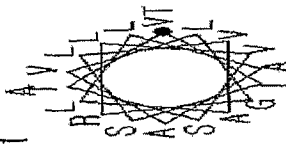
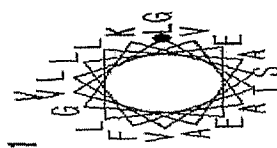
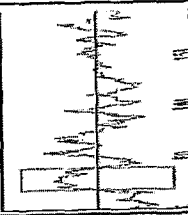
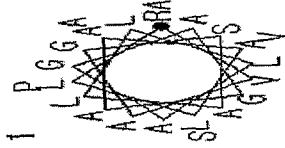
| #      | Origin                                                                                         | Hydropathy                                                                            | HRSS<br>(Candidate Domain)           | Helix                                                                                | Preliminary<br>modified<br>Sequences | Helix                                                                              | finally modified<br>Sequences | Helix                                                                              |
|--------|------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------|--------------------------------------|--------------------------------------------------------------------------------------|--------------------------------------|------------------------------------------------------------------------------------|-------------------------------|------------------------------------------------------------------------------------|
| JO-078 | CAB76313 putative<br>secreted protein<br>[Streptomyces<br>coelicolor A3(2)].                   |    | 1<br>MYLCAVALLILAVSL<br>VGGND 20     |   | AVALLILAV 9                          |   | AVALLILAV 9                   |   |
| JO-079 | P07198 Xenopsin<br>precursor [Contains:<br>Xenopsin precursor<br>fragment (XPF);<br>Xenopsin]. |    | 1<br>MYKGFILCVLLAVICA<br>NSLATPS 23  |   | VLLAVIALAP 10                        |   | VLLAVIP 7                     |   |
| JO-080 | NP_631293 secreted<br>protein<br>[Streptomyces<br>coelicolor A3(2)].                           |  | 13<br>SGALAVWLIVAAVVV<br>VAVLIGAF 35 |  | LIVAAVWVAVLI 13                      |  | LIVAAVWVAVLI 13               |  |

26/134

[Fig. 1g] continued

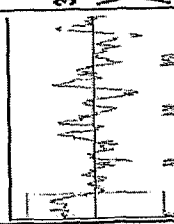
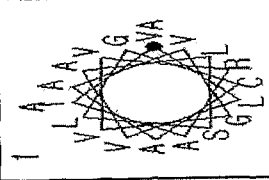
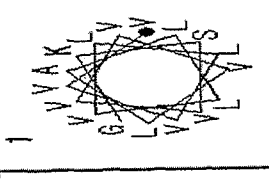
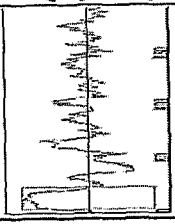
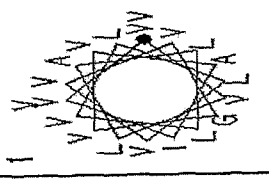
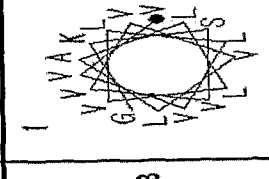
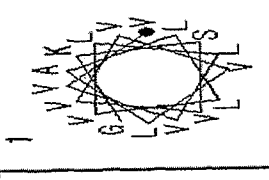
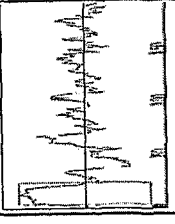
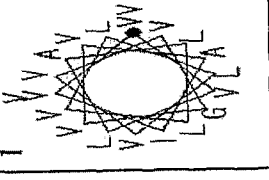
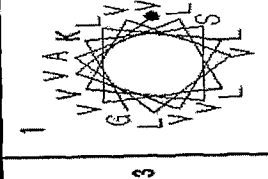
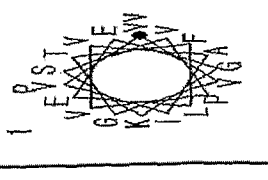
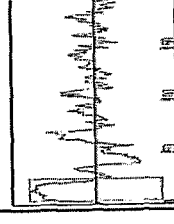
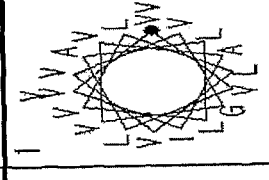
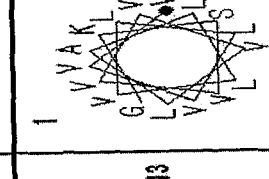
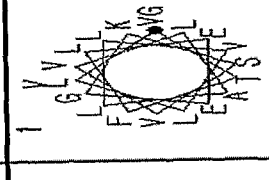
|        |                                                                                         |                                                                                      |                                    |                                                                                      |                |    |               |                                                                                    |
|--------|-----------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------|------------------------------------|--------------------------------------------------------------------------------------|----------------|----|---------------|------------------------------------------------------------------------------------|
| JO-081 | NP_626373secreted<br>protein<br>[Streptomyces<br>coelicolor A3(2)].                     |   | 16<br>VYGVASAVVAAIT<br>GTLALSPG 38 |   | AVVVAALALAP 11 | No | AVVVAAP 7     | No                                                                                 |
| JO-082 | NP_624952secreted<br>cellulose-binding<br>protein<br>[Streptomyces<br>coelicolor A3(2)] |   | 18<br>ATTLVSLAIVLLTL<br>IPWSGTA 40 |   | LAIVLLIPA 10   | No | LAIVLLIP 9    | No                                                                                 |
| JO-083 | NP_009104<br>protease, serine, 23<br>precursor [Homo<br>sapiens].                       |  | 2<br>AGIPGLLFFLLCA<br>VGQVSPY 24   |  | IPLLLLLAVVP 12 | No | LLLLLLAVVP 10 |  |

[Fig. 1g] continued

|                                                                                                                                 |                                                                                                                                        |                                                                                      |                                   |                                  |                                                                                   |
|---------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------|-----------------------------------|----------------------------------|-----------------------------------------------------------------------------------|
| <p>IAAK63068</p> <p>phytoalexin protein</p> <p>PcF precursor</p> <p>[Phytophthora</p> <p>cactorum].</p> <p><b>JO-084</b></p>    |  <p>2</p> <p>NFKTCPAVALVAVVA</p> <p>TVATAEDP 24</p> |   | <p>PAVALVAVVAVAP 13</p> <p>No</p> | <p>AVALVAVVAVVA 11</p> <p>No</p> | <p>No</p>                                                                         |
| <p>NC_003903Streptomyces coelicolor A3(2) plasmid SCP1, complete sequence.</p> <p><b>JO-085</b></p>                             |  <p>2</p> <p>VTSVLRGLVAALLAV</p> <p>LSITASTP 24</p> |   | <p>LVAALLAVLIAP 12</p> <p>No</p>  | <p>ILVAALLAVL 9</p> <p>No</p>    |  |
| <p>NP_629942peptide transport system secreted peptide binding protein [Streptomyces coelicolor A3(2)].</p> <p><b>JO-086</b></p> |  <p>20</p> <p>RLLAAGAGALLAS</p> <p>GAVAPSV 42</p>  |  | <p>ALLLAAPV 9</p> <p>No</p>       | <p>LLAAAAALLLA 11</p> <p>No</p>  | <p>No</p>                                                                         |

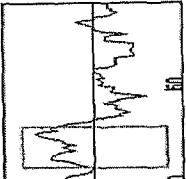
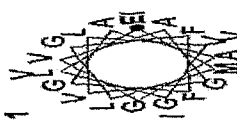
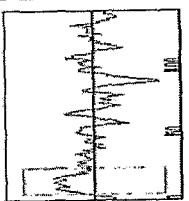
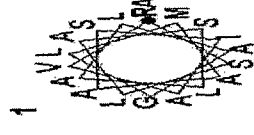
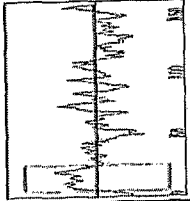
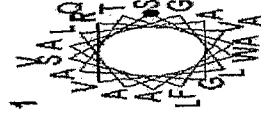
28/134

[Fig. 1g] continued

|        |                                                                             |                                                                                       |                                     |                                                                                       |                 |                                                                                     |                 |                                                                                     |                                                                                   |
|--------|-----------------------------------------------------------------------------|---------------------------------------------------------------------------------------|-------------------------------------|---------------------------------------------------------------------------------------|-----------------|-------------------------------------------------------------------------------------|-----------------|-------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------|
| JO-087 | NP_854067Possible<br>secreted protein<br>[Mycobacterium<br>bovis AF212297]. |    | 3<br>VIARWGVAAACGLSL<br>AVLAAAPT 25 |    | VAALLAVLAAAP 12 | No                                                                                  | LAVLAAAP 8      | No                                                                                  |  |
| JO-088 | NP_627802secreted<br>protein<br>[Streptomyces<br>coelicolor A3(2)].         |    | 4<br>VVIWVGWVLLVLL<br>ALVWVSR 25    |    | WVLLVLLALVWV 13 |    | WVLLVLLALVWV 13 |    |                                                                                   |
| JO-089 | NP_627802secreted<br>protein<br>[Streptomyces<br>coelicolor A3(2)].         |    | 4<br>VVIWVGWVLLVLL<br>ALVWVSR 25    |   | WVLLVLLALVWV 13 |   | WVLLVLLALVWV 13 |   |                                                                                   |
| JO-090 | NP_627802secreted<br>protein<br>[Streptomyces<br>coelicolor A3(2)].         |  | 4<br>VVIWVGWVLLVLL<br>ALVWVSR 25    |  | WVLLVLLALVWV 13 |  | VLLVLLALV 9     |  |                                                                                   |

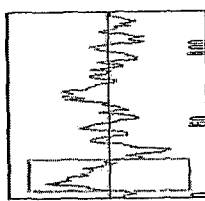
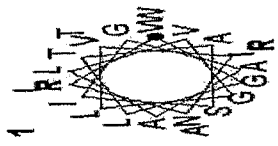
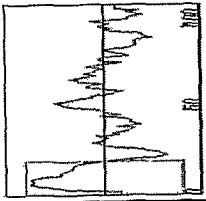
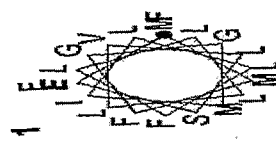
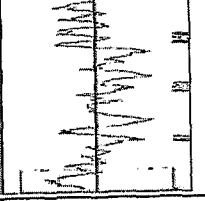
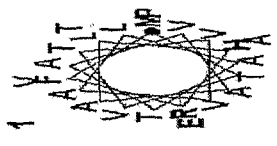
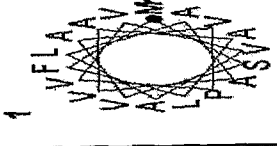
29/134

[Fig. 1h]

| #      | Origin                                                             | Hydropathy                                                                           | HRSS<br>(Candidate Domain)           | Helix                                                                                | Preliminarily<br>modified Sequences | Helix                                                                              | finally modified<br>Sequences | Helix |
|--------|--------------------------------------------------------------------|--------------------------------------------------------------------------------------|--------------------------------------|--------------------------------------------------------------------------------------|-------------------------------------|------------------------------------------------------------------------------------|-------------------------------|-------|
| JO-091 | NP_625203secreted<br>protein<br>[Streptomyces<br>coelicolor A3(2)] |   | 8<br>EELGLVLCGAVGAAG<br>FFGLMIVIV 30 |   | LGVLVGAVGAA 12                      | No                                                                                 | PVLVPVP 8                     | No    |
| JO-092 | NP_630960secreted<br>protein<br>[Streptomyces<br>coelicolor A3(2)] |   | 5<br>RVAISAAILGALALS<br>ALSATP 25    |   | VAIAAAILGALALALA<br>P 16            |   | PALALALA 8                    | No    |
| JO-093 | NP_630670secreted<br>protein<br>[Streptomyces<br>coelicolor A3(2)] |  | 5<br>SSFIVRLGAAATAG<br>ALAWAVLAQ 27  |  | FVILGAAAAGALAW<br>AVLA 18           |  | AAAAPALA 8                    | No    |

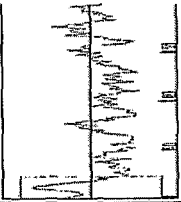
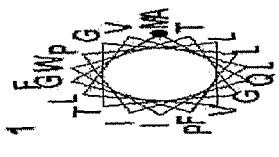
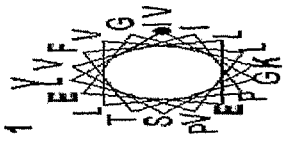
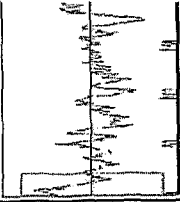
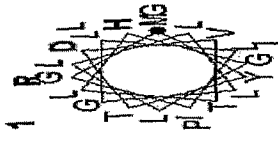
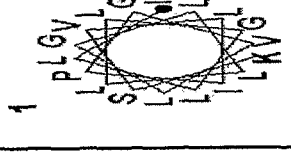
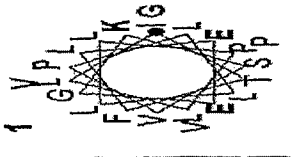
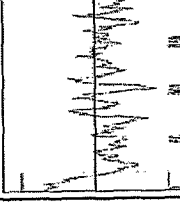
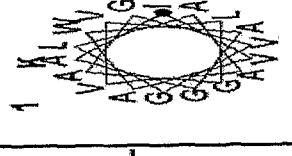
30/134

[Fig. 1h] continued

|        |                                                                     |                                                                                      |                                     |                                                                                      |                          |                                                                                    |                |    |
|--------|---------------------------------------------------------------------|--------------------------------------------------------------------------------------|-------------------------------------|--------------------------------------------------------------------------------------|--------------------------|------------------------------------------------------------------------------------|----------------|----|
| JO-094 | NP_630493secreted<br>protein<br>(Streptomyces<br>coelicolor A3(2)). |   | 5<br>VRNIVLGVLAAGISAT<br>LGWLART 27 |   | VNVLGVLAAGIALGW<br>LA 17 |   | IVLPVLAAP 9    | No |
| JO-095 | CAC2896putative<br>secreted protein<br>(Mycobacterium<br>leprae).   |   | 1<br>MESVLVLLLFLLMG<br>GFNIFF 20    |   | MLVLLLFLLMGGF<br>MFF 18  |   | ILVLLLPILLI 10 |    |
| JO-096 | NP_624483secreted<br>protein<br>(Streptomyces<br>coelicolor A3(2)). |  | 1<br>MFRHLAAVATALAVV<br>TVTPVEAT 23 |  | IMFLAAVAALAVVVPV<br>A 16 |  | LAAVAPALAVV 11 | No |

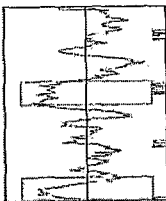
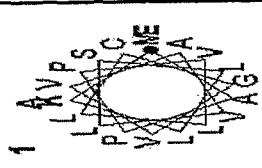
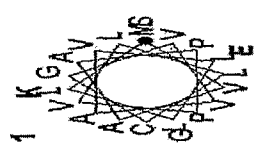
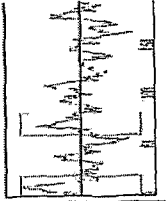
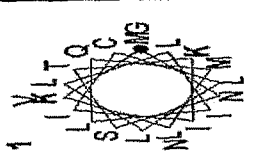
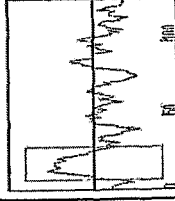
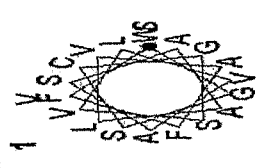
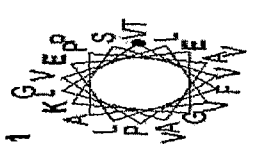
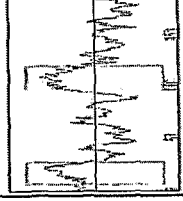
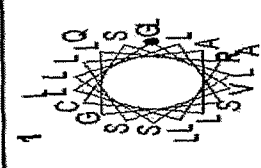
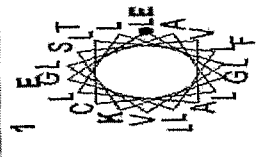
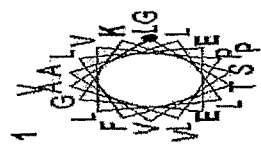
31/134

[Fig. 1h] continued

|        |                                                                                 |                                                                                      |                                   |                                                                                      |                          |                                                                                    |               |                                                                                   |
|--------|---------------------------------------------------------------------------------|--------------------------------------------------------------------------------------|-----------------------------------|--------------------------------------------------------------------------------------|--------------------------|------------------------------------------------------------------------------------|---------------|-----------------------------------------------------------------------------------|
| JO-097 | NP_037375<br>secretogranin III<br>[Homo sapiens].                               |   | 1<br>MGFLGTGTWILVVL<br>PIQAFKP 23 |   | LGGWILVVLPIAFP<br>P 16   | No                                                                                 | ILVVLPI 8     |  |
| JO-098 | NP_039199 V-set<br>and immunoglobulin<br>domain containing 4<br>[Homo sapiens]. |   | 1<br>MGILLGILLGHLTV<br>DTYGRPI 23 |   | ILLGILLGLVGPIL<br>15     |   | ILLPIILLP 9   |  |
| JO-099 | NP_733660<br>secreted hydrolase<br>[Streptomyces<br>coelicolor A3(2)].          |  | 1<br>MIAGAVVAALGVGA<br>GLWAT 19   |  | IAGAVVAALGVGAGL<br>WA 17 |  | IAPAVVAALP 10 | No                                                                                |

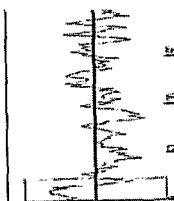
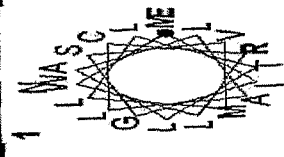
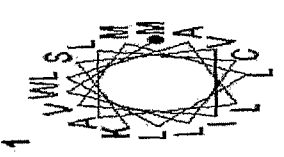
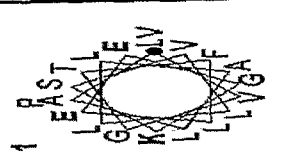
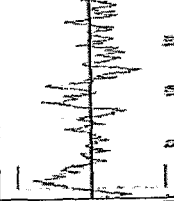
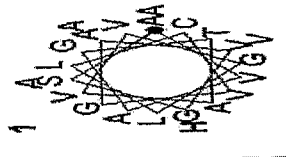
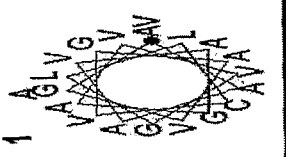
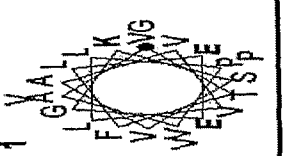
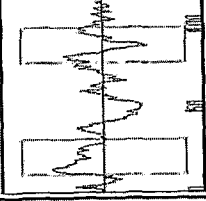
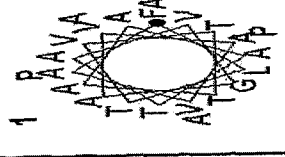
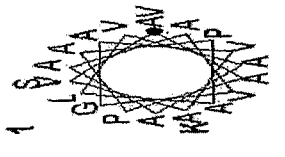
32/134

[Fig. 1h] continued

|        |                                                                      |                                                                                       |                                     |                                                                                       |                          |                                                                                     |                          |                                                                                     |  |
|--------|----------------------------------------------------------------------|---------------------------------------------------------------------------------------|-------------------------------------|---------------------------------------------------------------------------------------|--------------------------|-------------------------------------------------------------------------------------|--------------------------|-------------------------------------------------------------------------------------|--|
| JO-100 | NP_057540<br>transmembrane<br>protein 9 (Homo<br>sapiens).           |    | 1<br>MKLLSLVAVVGCLLV<br>PPAEANK 22  |    | MILLVAVVGCLLVPP<br>AA 16 |    | ILLVAVVPLLV 12<br><br>No | No                                                                                  |  |
| JO-101 | CAI74362<br>hypothetical protein<br>[Theileria annulata].            |    | 1<br>MKLMQLILLLLCIKT<br>SNGVN 21    |    | MILMILLILLCIGV 14        | No                                                                                  | LILLPLI 9<br><br>No      | No                                                                                  |  |
| JO-102 | NP_630671 secreted<br>protein<br>[Streptomyces<br>coelicolor A3(2)]. |    | 7<br>GGRWFFAVLAASAV<br>LYSGCSGSV 29 |   | FAVLAAAVLVGCGV<br>14     | No                                                                                  | AVLAAPAVLV 10<br><br>No  |   |  |
| JO-103 | NP_065695 TMEM9<br>domain family,<br>member B (Homo<br>sapiens).     |  | 7<br>GLRLGSLLSCLA<br>LSVLLAQ 29     |  | ILLGLILCLALVILL<br>A 16  |  | LALPVLLA 9<br><br>No     |  |  |

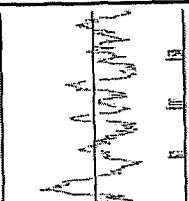
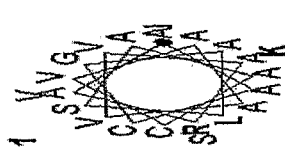
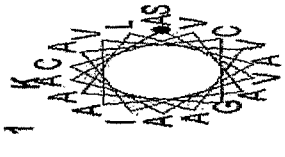
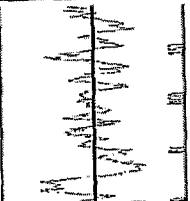
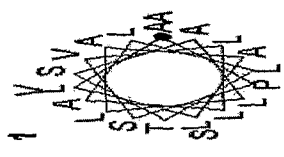
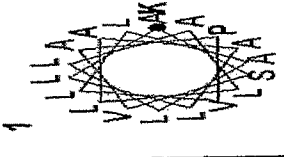
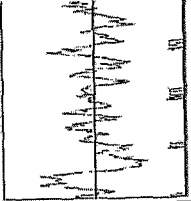
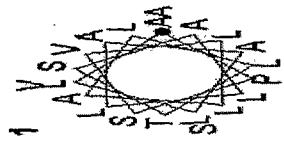
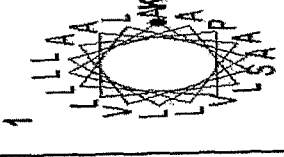
33/134

[Fig. 1i]

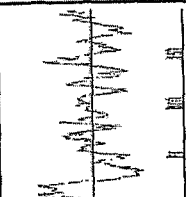
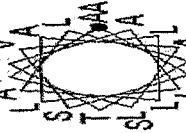
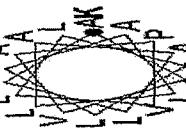
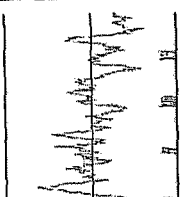
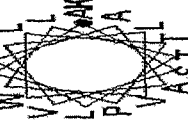
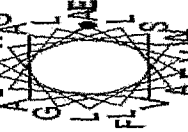
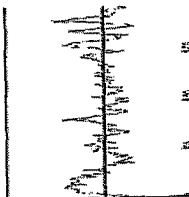
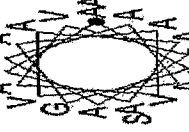
| #      | Origin                                                                                        | Hydropathy                                                                           | HRSS<br>(Candidate Domain)         | Helix                                                                                | Preliminarily<br>modified<br>Sequences | Helix                                                                              | finally modified<br>Sequences | Helix                                                                             |
|--------|-----------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------|------------------------------------|--------------------------------------------------------------------------------------|----------------------------------------|------------------------------------------------------------------------------------|-------------------------------|-----------------------------------------------------------------------------------|
| JO-104 | P06908 Pulmonary<br>surfactant-associated<br>protein A precursor<br>(SP-A) (PSP-A)<br>(PSP)   |   | 1<br>MWLRCLALTLTM<br>VSGIENNTK 23  |   | MWLRCLALTLTMVI<br>14                   |   | LALALL 7                      |  |
| JO-105 | NP_63972 putative<br>secreted protein<br>(Streptomyces<br>coelicolor A3(2)).                  |   | 10<br>ASGVAGVCLLGWA<br>TGAAAHVA 32 |   | AGVAGVCLLGWAG<br>AAVA 20               |   | VAVPLLVA 9                    |  |
| JO-106 | NP_854954 CONSER<br>VED PROBABLE<br>SECRETED<br>PROTEIN<br>(Mycobacterium<br>bovis AF212297). |  | 11<br>FAVAVAGVATAATT<br>VTLPAPA 33 |  | AVAVAGVAAAAVLA<br>PAPA 18              |  | AVAVAPVAAAA 11                | No                                                                                |

34/134

[Fig. 1i] continued

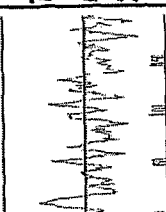
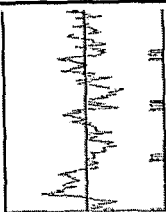
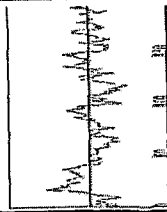
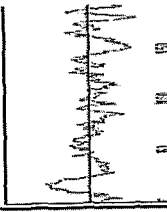
|        |                                                                        |                                                                                      |                                  |                                                                                      |                         |                                                                                    |                     |    |
|--------|------------------------------------------------------------------------|--------------------------------------------------------------------------------------|----------------------------------|--------------------------------------------------------------------------------------|-------------------------|------------------------------------------------------------------------------------|---------------------|----|
| JO-107 | NP_627759 secreted protein<br>[Streptomyces coelicolor A3(2)]          |   | 13<br>AARAVVAIVCAASL<br>AGCAI 31 |   | AAVVAIVCAALAG<br>CAI 17 |   | AAVVAIVPAA 11<br>No | No |
| JO-108 | NP_003842 cellular repressor of E1A-stimulated genes<br>[Homo sapiens] |   | 11<br>ALLAALLASTILL<br>VSPA 29   |   | ALLAALLALLLVP<br>A 16   |   | ALLAALLAP 9<br>No   | No |
| JO-109 | NP_003842 cellular repressor of E1A-stimulated genes<br>[Homo sapiens] |  | 11<br>ALLAALLASTILL<br>VSPA 29   |  | ALLAALLALLLVP<br>A 16   |  | LLALLVP 7<br>No     | No |

[Fig. 1i] continued

|        |                                                                               |                                                                                      |                                     |                                                                                      |                           |                                                                                    |                 |    |  |
|--------|-------------------------------------------------------------------------------|--------------------------------------------------------------------------------------|-------------------------------------|--------------------------------------------------------------------------------------|---------------------------|------------------------------------------------------------------------------------|-----------------|----|--|
| JO-110 | NP_003842 cellular repressor of E1A-stimulated genes [Homo sapiens].          |   | ALLAALLASTLLALL<br>VSPA 29<br>11    |   | ALLAALLALLLVP<br>A 16     |   | ALLAALLALLLV 14 |    |  |
| JO-111 | NP_000589 Homo sapiens insulin-like growth factor binding protein 3 (IGFBP3). |   | ARPTLWAAALTLLVL<br>IL 19<br>4       |   | APLWAAAALLLVLL<br>13      |   | AAALPLLVL 11    | No |  |
| JO-112 | CAB50459 putative secreted protein [Streptomyces coelicolor A3(2)].           |  | ATAAAVTAALATGVA<br>SVAAGRLA 28<br>6 |  | AAAAVAALAGVAVA<br>AGLA 18 |  | AAAVPAALAP 10   | No |  |

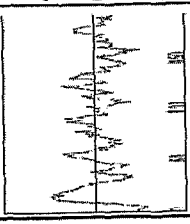
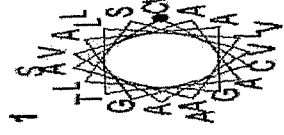
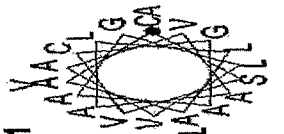
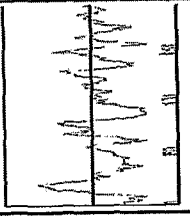
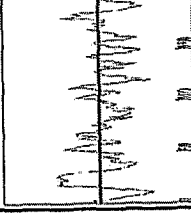
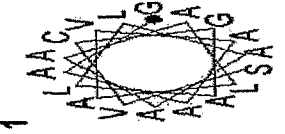
36/134

[Fig. 1i] continued

|        |                                                                     |                                                                                       |                                                 |                                                            |                               |                                                              |                |    |
|--------|---------------------------------------------------------------------|---------------------------------------------------------------------------------------|-------------------------------------------------|------------------------------------------------------------|-------------------------------|--------------------------------------------------------------|----------------|----|
| JO-113 | NP_628917secreted<br>protein<br>[Streptomyces<br>coelicolor A3(2)]. |    | 7<br>ATTALAAALAAAGCA<br>21                      | No                                                         | AALAAALAAAGCA<br>13           | No                                                           | AALAAALAA 10   | No |
| JO-114 | NP_624695secreted<br>protein<br>[Streptomyces<br>coelicolor A3(2)]. |    | 15<br>AVLAAAVTAGVTATA<br>VTASPGVAALPAGP<br>A 44 | 1<br>A T S P<br>A A A A<br>P A A A<br>A A A A<br>G G T T   | AVLAAAVAGVAAVA<br>PGVAALPA 22 | 1<br>V G G A<br>A A A A<br>V V V V<br>P L A A<br>A A A A     | AVLAAAVP 8     | No |
| JO-115 | NP_624695secreted<br>protein<br>[Streptomyces<br>coelicolor A3(2)]. |    | 15<br>AVLAAAVTAGVTATA<br>VTASPGVAALPAGP<br>A 44 | 1<br>A T S P<br>A A A A<br>P A A A<br>A A A A<br>G G T T   | AVLAAAVAGVAAVA<br>PGVAALPA 22 | 1<br>V G G A<br>A A A A<br>V V V V<br>P L A A<br>A A A A     | VAALPAPA 8     | No |
| JO-116 | NP_624791secreted<br>protein<br>[Streptomyces<br>coelicolor A3(2)]. |  | 9<br>AVTGTALAAVSAVL<br>TACGG 28                 | 1<br>G I L T K<br>A A A A<br>P A A A<br>A A A A<br>L C V S | AVGALAAVAVLAC<br>GG 16        | 1<br>C G A V C<br>L A A A<br>S V V V<br>T A G A<br>G A A L A | ALALAVPAVLP 11 | No |

37/134

[Fig. 1]]

| #      | Origin                                                                                   | Hydropathy                                                                           | HRSS<br>(Candidate Domain)        | Helix                                                                                | Preliminarily<br>modified Sequences | Helix                                                                              | finally modified<br>Sequences | Helix |
|--------|------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------|-----------------------------------|--------------------------------------------------------------------------------------|-------------------------------------|------------------------------------------------------------------------------------|-------------------------------|-------|
| JQ-117 | CAB4559putative<br>secreted protein<br>[Streptomyces<br>coelicolor A3(2)].               |   | 10<br>CAALLTAAVAVSLGA<br>AGC 27   |   | CAALLAAVAVLGAA<br>GC 16             |   | AALLPAAVAVP 11<br>No          | No    |
| JQ-118 | NP_627066secreted<br>protein<br>[Streptomyces<br>coelicolor A3(2)].                      |   | 13<br>FAIGTAVVVALAGMIN<br>GPWL 31 | No                                                                                   | FAIGAVVVALAGMGP<br>WIL 17           | No                                                                                 | AVVVALAP 8<br>No              | No    |
| JQ-119 | NP_630174secreted<br>substrate-binding<br>protein<br>[Streptomyces<br>coelicolor A3(2)]. |  | 16<br>GAAAVALTAAALAA<br>GC 32     |  | GAAAVALAAAALLA<br>GC 16             |  | AAAVALPAAAALLA 14<br>No       | No    |

38/134

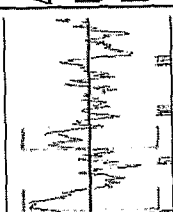
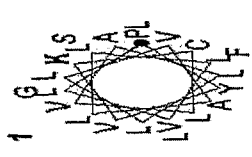
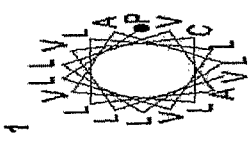
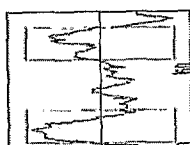
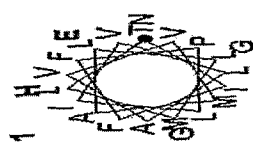
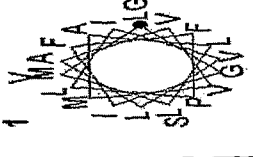
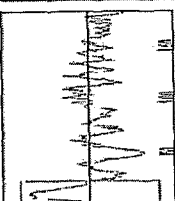
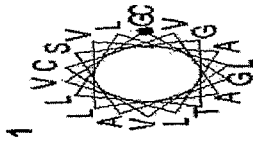
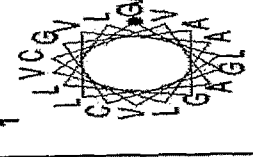
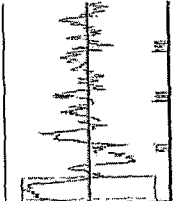
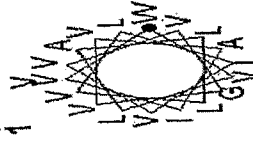
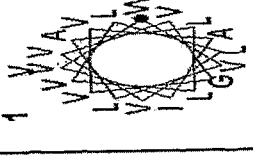
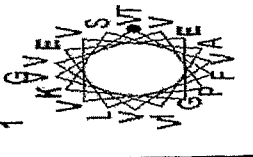
[Fig. 1]] continued

|        |                                                                                                   |  |                                     |    |                          |  |                                          |                  |    |                                   |
|--------|---------------------------------------------------------------------------------------------------|--|-------------------------------------|----|--------------------------|--|------------------------------------------|------------------|----|-----------------------------------|
| JO-120 | P05727<br>Apolipoprotein A-IV<br>precursor (Apo-AIV)<br>(ApoA-IV).<br>Homo sapiens                |  | 3<br>LKAVLTALVAVAG<br>ARAESAD 25    | No | LAVLLALVAVAGAA<br>VA 17  |  | 1<br>GAVV<br>LAVLLALVAVAGAA<br>VA 17     | AVWLPLALVAVAP 13 |    | 1<br>P V V K G<br>LAVLLALVAVAP 13 |
| JO-121 | Q62067 Serum<br>paraoxonase/lectina<br>se 3.<br>Mus musculus                                      |  | 2<br>GKLVALTLLGACLAL<br>IGERLLNF 24 |    | GLVALLGACLALIG<br>LLF 18 |  | 1<br>L A L L<br>GLVALLGACLALIG<br>LLF 18 | LVALPLP 8        | No |                                   |
| JO-122 | NP_627123probable<br>secreted penicillin-<br>binding<br>protein[Streptomyces<br>coelicolor A3(2)] |  | 3<br>STVWGLLLVGGFF<br>LGYHLVQI 25   |    | WWGLLLVGGFFLG<br>LVI 18  |  | 1<br>S V L L<br>WWGLLLVGGFFLG<br>LVI 18  | WVPLLNP 10       |    | 1<br>G V E P<br>WVPLLNP 10        |



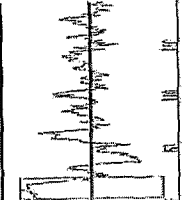
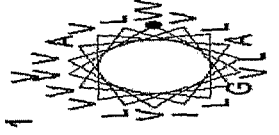
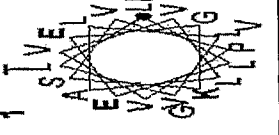
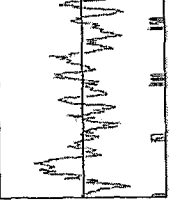
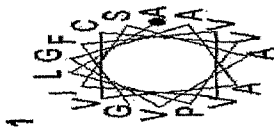
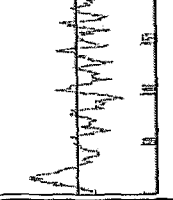
40/134

[Fig. 1]] continued

|        |                                                                                        |                                                                                       |                                    |                                                                                       |                          |                                                                                     |                 |                                                                                     |
|--------|----------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------|------------------------------------|---------------------------------------------------------------------------------------|--------------------------|-------------------------------------------------------------------------------------|-----------------|-------------------------------------------------------------------------------------|
| JO-126 | NP_060220 all-trans-13,14-dihydroretinol saturase [Homo sapiens].                      |    | 4<br>PLVLLAVLLAVLC<br>KVYGLFS 26   |    | PLVLLAVLLAVLC<br>VL 17   |    | LVLLAVLLAVLP 14 | No                                                                                  |
| JO-127 | AL627273<br>Salmonella enterica<br>serovar Typhi<br>(Salmonella typhi)<br>strain CT18, |    | 4<br>TLMLAMVVALVILP<br>FFINHGGE 26 |    | LMLAMVVALVILP<br>FIGG 19 |    | LLAPVVALVILP 12 | No                                                                                  |
| JO-128 | NP_625987 secreted<br>protein<br>[Streptomyces<br>coelicolor A3(2)]                    |    | 4<br>TRGVAVLAVCVLL<br>TGSAGCG 25   |   | GVAVLAVCVLLGA<br>GCG 18  |   | VLAVLAVPVLLP 13 | No                                                                                  |
| JO-129 | CAB4547 putative<br>secreted protein<br>[Streptomyces<br>coelicolor A3(2)]             |  | 4<br>WIAVGVVLLVLL<br>ALVVSR 25     |  | WIAVGVVLLVLL<br>ALVV 20  |  | WIAVVPVV 10     |  |

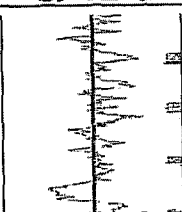
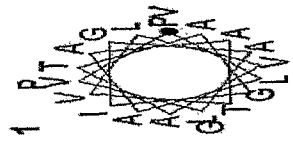
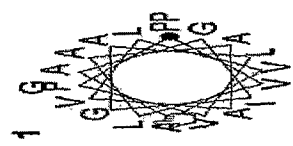
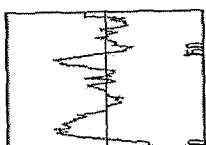
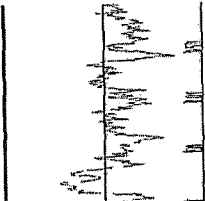
41/134

[Fig. 1k]

| #      | Origin                                                                               | Hydropathy                                                                            | HRSS<br>(Candidate Domain)  | Helix                                                                                | Preliminarily<br>modified Sequences | Helix                                                                             | finally modified<br>Sequences | Helix                                                                             |
|--------|--------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------|-----------------------------|--------------------------------------------------------------------------------------|-------------------------------------|-----------------------------------------------------------------------------------|-------------------------------|-----------------------------------------------------------------------------------|
| JO-130 | CAB45474putative<br>secreted protein<br>[Streptomyces<br>coelicolor A3(2)]           |    | WIAWGVVLLVLL<br>ALVVSR 25   |   | WIAWGVVLLVLL<br>ALVW 20             |  | LLVLLALVWP 11                 |  |
| JO-131 | CAB36805putative<br>secreted protein<br>[Streptomyces<br>coelicolor A3(2)].          |    | VLLRRALPVCVAGV<br>ASIVFG 36 |   | VLLALPVCVAGVAI<br>VWFG 19           |  | VLLALPWVAAAP 11               | No                                                                                |
| JO-132 | NP_628377NLP60-<br>family secreted<br>protein<br>[Streptomyces<br>coelicolor A3(2)]. |  | VLTTAVTVCAITVL<br>AAPG 26   |  | VLAVVVAIVLAAP 13                    | No                                                                                | AVVWPVIVLAAP 12               | No                                                                                |

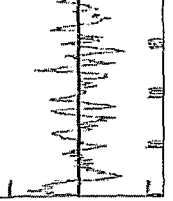
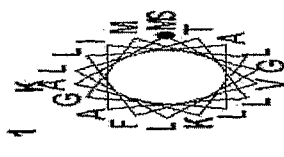
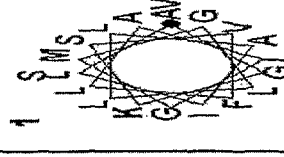
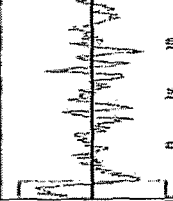
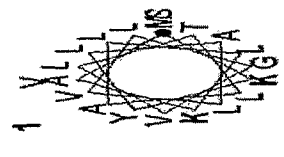
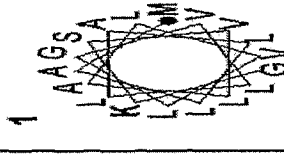
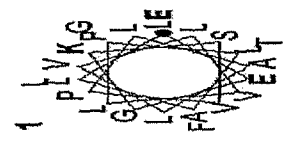
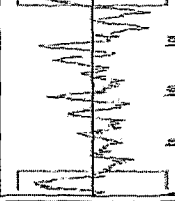
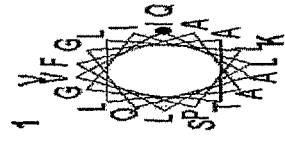
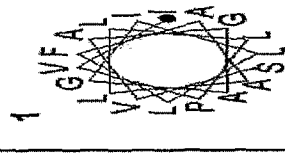
42/134

[Fig. 1k] continued

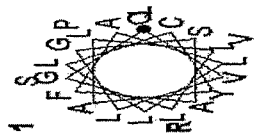
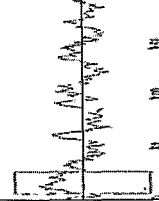
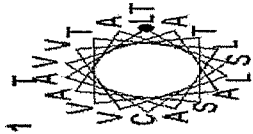
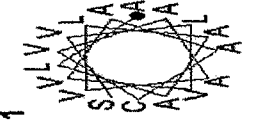
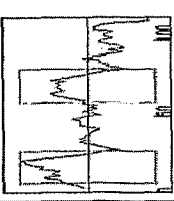
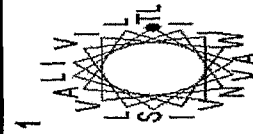
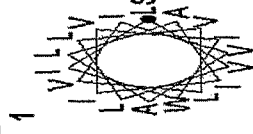
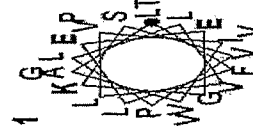
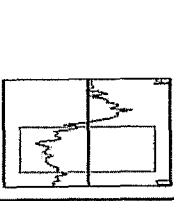
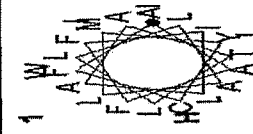
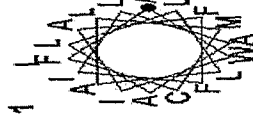
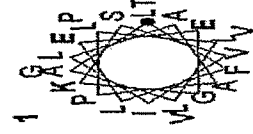
|        |                                                                                       |                                                                                      |                                                |                                                                                     |                             |                                                                                   |               |    |
|--------|---------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------|------------------------------------------------|-------------------------------------------------------------------------------------|-----------------------------|-----------------------------------------------------------------------------------|---------------|----|
| JO-133 | CAB59594 putative<br>secreted protein<br>[Streptomyces<br>coelicolor A3(2)].          |   | 9<br>PRKPVLAGIGATAVL<br>VTAAALVPG 32           |  | PPVLAGIGAAVLVAA<br>ALVPG 20 |  | AVLVPAALVP 11 | No |
| JO-134 | NP_624974secreted<br>protein<br>[Streptomyces<br>coelicolor A3(2)]                    |   | 8<br>VVAALTLVLGATGA<br>22                      | No                                                                                  | VVAALLVLGAGA 12             | No                                                                                | VVAALPLVLP 10 | No |
| JO-135 | NP_733682secreted<br>ATP/GTP binding<br>protein<br>[Streptomyces<br>coelicolor A3(2)] |  | 9<br>WGPVAVMTTAFALA<br>VGSQGAVALPGAP<br>AKA 39 | No                                                                                  | AFALAVGGAAVALP<br>GAPAA 19  | No                                                                                | AAVALPAAAP 10 | No |

43/134

[Fig. 1k] continued

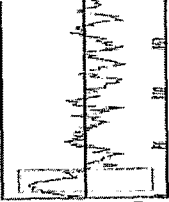
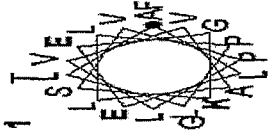
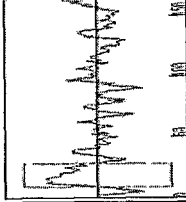
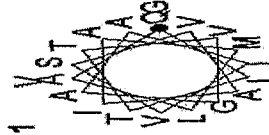
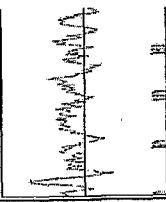
|        |                                                                                                                          |                                                                                      |                                |                                                                                      |                      |                                                                                    |                |                                                                                   |
|--------|--------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------|--------------------------------|--------------------------------------------------------------------------------------|----------------------|------------------------------------------------------------------------------------|----------------|-----------------------------------------------------------------------------------|
| JO-136 | P27169 Serum<br>paraoxonase/arylesterase 1 (PON 1)<br>(Serum)<br>aryldialkylphosphatase 1 (A-esterase 1)<br>Homo sapiens |   | 1<br>MAKLIALTLLGMGLA<br>LF 17  |   | ALIALLLGMGLALF<br>14 |   | LIALPLLP 8     | No                                                                                |
| JO-137 | P52430 Serum<br>paraoxonase/arylesterase 1 (PON 1)<br>(Serum)<br>aryldialkylphosphatase 1 (A-esterase 1)<br>Homo sapiens |   | 1<br>MAKLLALTIVGLVLA<br>LYK 18 |   | MALLALLVGLVLA<br>14  |   | LLALPLVLVLP 13 |  |
| JO-138 | NP_626569 secreted<br>protein<br>(Streptomyces<br>coelicolor A3(2))                                                      |  | 6<br>IVPLLAAFLIGTAG<br>QAQ 24  |  | IVPLLAAFLIGAGA<br>16 |  | IVPLLAAP 9     | No                                                                                |

[Fig. 1k] continued

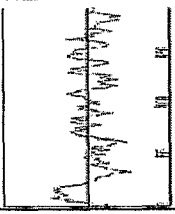
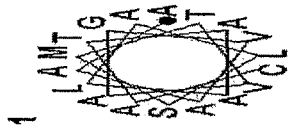
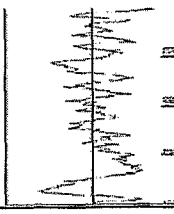
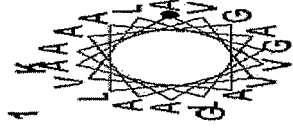
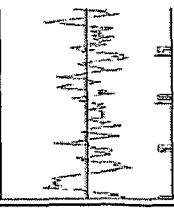
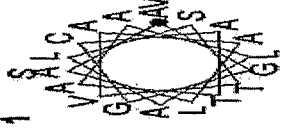
|        |                                                                                            |                                                                                       |                                     |                                                                                       |                            |                                                                                     |               |                                                                                     |
|--------|--------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------|-------------------------------------|---------------------------------------------------------------------------------------|----------------------------|-------------------------------------------------------------------------------------|---------------|-------------------------------------------------------------------------------------|
| JO-139 | NP_940995 C1q and tumor necrosis factor related protein 1 isoform 1 [Homo sapiens].        |    | 6<br>QGLLAYCILLAFAS<br>GLVLSRVP 28  |    | GILLACILLAFAGLV<br>LVP 18  |    | LLLAPLLAP 10  | No                                                                                  |
| JO-140 | NP_628174 large secreted protein [Streptomyces coelicolor A3(2)]                           |    | 6<br>SARLAALTVAAVCSA<br>ASTVLT 28   |    | ALAALVAAYCAAVLV<br>15      |    | LAAALPVAAP 10 | No                                                                                  |
| JO-141 | CAB83860 putative protein-export integral membrane protein [Neisseria meningitidis Z2491]. |    | 6<br>TLWVNIISALAVVL<br>VLLQH 27     |   | LWVNIISALAVVL<br>17        |   | ALAVVLVLL 10  |   |
| JO-142 | NP_001009551 constrictin-like isoform 2 [Homo sapiens].                                    |  | 7<br>AFCYMLALLTTAALIF<br>FAIWHII 29 |  | AFCMLALLAALIFF<br>AIWII 20 |  | LALLPAALI 10  |  |

45/134

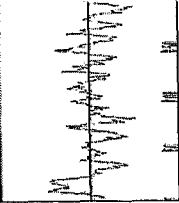
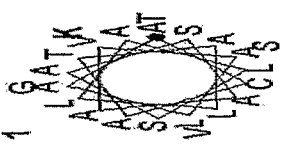
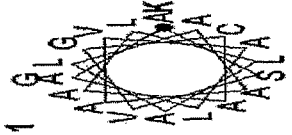
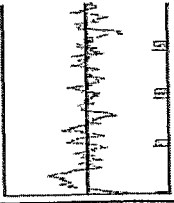
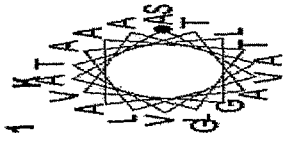
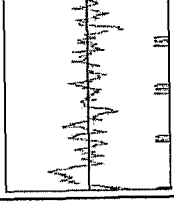
[Fig. 11]

| #      | Origin                                                                       | Hydropathy                                                                            | HRSS<br>(Candidate Domain)         | Helix                                                                               | Prelimarily<br>modified Sequences | Helix                                                                             | finally modified<br>Sequences | Helix                                                                              |
|--------|------------------------------------------------------------------------------|---------------------------------------------------------------------------------------|------------------------------------|-------------------------------------------------------------------------------------|-----------------------------------|-----------------------------------------------------------------------------------|-------------------------------|------------------------------------------------------------------------------------|
| JO-143 | NP_626808 secreted<br>protein<br>[Streptomyces<br>coelicolor A3(2)]          |    | 7<br>VFALLFLAWLGVY<br>ATVKA VGA 29 |  | VFALLFLAWLGV<br>VA VGA 20         |  | ALLPLLAVLP 11                 |   |
| JO-144 | NP_639798 putative<br>secreted protein<br>[Streptomyces<br>coelicolor A3(2)] |    | 9<br>QALMAIVSVLAAGV<br>ITLGV 28    |  | ALMAIVSVLAAGV<br>V 16             |  | AIAPVLAAP 10                  | No                                                                                 |
| JO-145 | NP_000492 Homo<br>sapiens elastin<br>(supravalvular aortic<br>stenosis,      |  | 6<br>AAAPRPGVLLLL 18               | No                                                                                  | AAAPPGVLLLL 12                    | No                                                                                | AAAPVLLLLL 10                 |  |

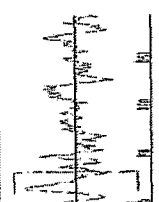
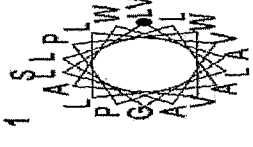
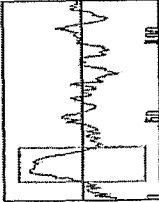
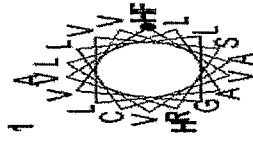
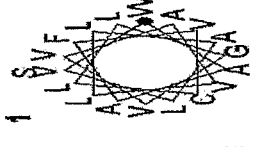
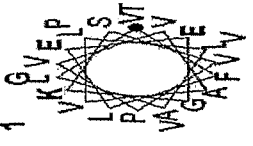
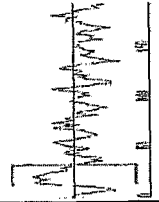
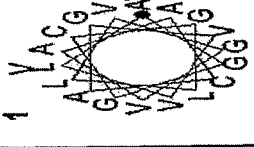
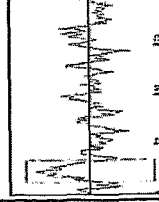
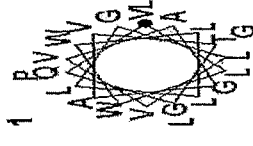
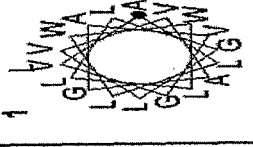
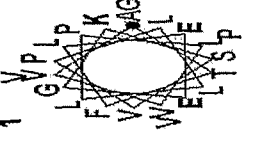
[Fig. 1l] continued

|                                                                                             |                                                                                      |                                  |                                                                                      |                           |                                                                                   |                 |    |    |
|---------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------|----------------------------------|--------------------------------------------------------------------------------------|---------------------------|-----------------------------------------------------------------------------------|-----------------|----|----|
| NP_630680 secreted<br>sugar binding protein<br>[Streptomyces<br>coelicolor A3(2)].          |   | 6<br>AAVGAVTMSLALA<br>ATACGGG 26 |   | AAVGAVMILALAAA<br>CGGG 18 |  | AAVAVLALAP 11   | No | No |
| CAB56129 putative<br>secreted protein<br>[Streptomyces<br>coelicolor A3(2)].                |   | 10<br>AALAAVVAAGSLVT<br>AGAA 28  |   | AALAAVVAAGLVA<br>GAA 17   |  | AALAAVVAAP 11   | No | No |
| NP_625109<br>secreted solute-<br>binding lipoprotein<br>[Streptomyces<br>coelicolor A3(2)]. |  | 11<br>AALAAVTSALAATA<br>CGG 28   |  | AALAAVLAACGG<br>G 15      | No                                                                                | AALAAVPLALAP 12 | No | No |

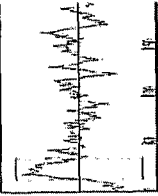
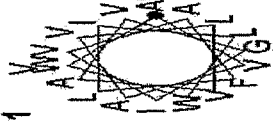
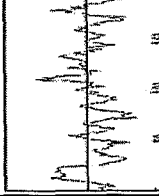
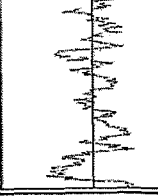
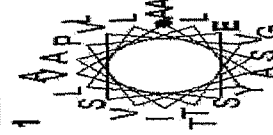
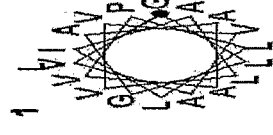
[Fig. 1l] continued

|        |                                                                                                 |                                                                                      |                          |                                                                                      |                        |                                                                                   |                        |    |
|--------|-------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------|--------------------------|--------------------------------------------------------------------------------------|------------------------|-----------------------------------------------------------------------------------|------------------------|----|
| JO-149 | NP_73357[secreted,<br>sugar-binding protein<br>(Streptomyces<br>coelicolor A3(2))]              |   | AALVAASASLALLA<br>TACTG  |   | AALVAAALALLAAC<br>G 16 |  | ALAVAAPALALLP 13<br>No | No |
| JO-150 | NP_630126<br>secreted chitinase<br>(secreted protein)<br>(Streptomyces<br>coelicolor<br>A3(2)). |   | AALTAATTTVAAVGL<br>AL 25 |   | AALAAVAAVGLAL<br>14    | No                                                                                | AALPAAAP 8<br>No       | No |
| JO-151 | NP_630126<br>secreted chitinase<br>(secreted protein)<br>(Streptomyces<br>coelicolor<br>A3(2)). |  | AALTAATTTVAAVGL<br>AL 25 |  | AALAAVAAVGLAL<br>14    | No                                                                                | AAAPVAAVP 9<br>No      | No |

[Fig. 1l] continued

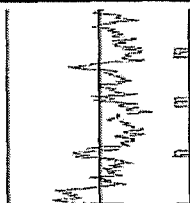
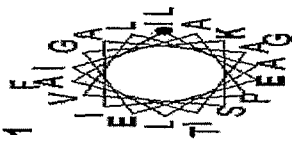
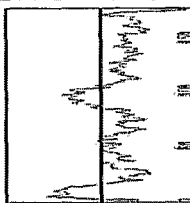
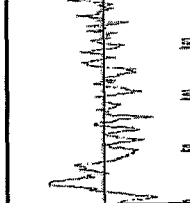
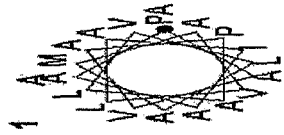
|        |                                                                                           |                                                                                       |                                      |                                                                                       |                            |                                                                                     |               |                                                                                     |  |
|--------|-------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------|--------------------------------------|---------------------------------------------------------------------------------------|----------------------------|-------------------------------------------------------------------------------------|---------------|-------------------------------------------------------------------------------------|--|
| JO-152 | NP_872425<br>secretory protein<br>LOC348174 [Homo<br>sapiens].                            |    | 12<br>HLLAVLLALLGTAWA<br>EVWPPQLQ 34 | No                                                                                    | LLAVLLALLGAWAV<br>WPPL 18  |    | LLAVLLALLP 10 | No                                                                                  |  |
| JO-153 | NP_630107 secreted<br>protein<br>[Streptomyces<br>coelicolor A3(2)].                      |    | 12<br>HVRSLALLVAVVGL<br>LCVFAHAE 34  |    | WLALLVAVWGLLCV<br>FAA 18   |    | VLALLVAVWP 10 |    |  |
| JO-154 | NP_733688 peptide-<br>binding transport<br>protein<br>[Streptomyces<br>coelicolor A3(2)]. |    | 13<br>ALVVGACAAVGVLL<br>SGCTGGVS 34  |   | ALVVGACAAVGVLL<br>GCGGV 19 |   | ALVWPAAVP 9   | No                                                                                  |  |
| JO-155 | NP_629904 secreted<br>protein<br>[Streptomyces<br>coelicolor A3(2)].                      |  | 13<br>AVGGTVAGAVVGL<br>LLWWLLPLG 35  |  | AVGVAGAVVGLLL<br>WWLLPL 20 |  | AVVLPPLL 9    |  |  |

[Fig. 1m]

| #      | Origin                                                                                         | Hydropathy                                                                            | HRSS<br>(Candidate Domain)                                                                                                                                                                                                                                                                             | Helix                                                                                | Preliminarily<br>modified<br>Sequences                                                                                                                                                                                                        | Helix                                                                              | finally modified<br>Sequences                                                     | Helix                                                                              |
|--------|------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------|------------------------------------------------------------------------------------|
| JO-156 | YP_177862 MCE-<br>FAMILY PROTEIN<br>MCE3A<br>[Mycobacterium<br>tuberculosis H37Rv].            |    | 13<br>A <sup>13</sup> W <sup>14</sup> T <sup>15</sup> L <sup>16</sup> I <sup>17</sup> L <sup>18</sup> F <sup>19</sup> A <sup>20</sup> V <sup>21</sup> G <sup>22</sup> V <sup>23</sup> A <sup>24</sup> V <sup>25</sup><br>LVTA <sup>26</sup> SFT <sup>35</sup>                                          |   | AWW <sup>1</sup> L <sup>2</sup> IL <sup>3</sup> FA <sup>4</sup> VG <sup>5</sup> AV <sup>6</sup> L <sup>7</sup><br>VAV <sup>8</sup> F <sup>9</sup> 19                                                                                          |   | AVIP <sup>1</sup> VAV <sup>2</sup> LVP <sup>10</sup>                              |   |
| JO-157 | CAA19627 putative<br>secreted solute<br>binding protein<br>[Streptomyces<br>coelicolor A3(2)]. |    | 9<br>G <sup>9</sup> M <sup>10</sup> L <sup>11</sup> F <sup>12</sup> G <sup>13</sup> A <sup>14</sup> A <sup>15</sup> V <sup>16</sup> S <sup>17</sup> A <sup>18</sup> G <sup>19</sup> A <sup>20</sup> V <sup>21</sup><br>LAGC <sup>26</sup>                                                              |   | G <sup>1</sup> M <sup>2</sup> L <sup>3</sup> F <sup>4</sup> G <sup>5</sup> A <sup>6</sup> A <sup>7</sup> A <sup>8</sup> V <sup>9</sup> A <sup>10</sup> G <sup>11</sup> A <sup>12</sup> V <sup>13</sup> L <sup>14</sup><br>AG <sup>16</sup> 16 |   | AAV <sup>1</sup> PA <sup>2</sup> AV <sup>3</sup> L <sup>4</sup> AP <sup>11</sup>  | No                                                                                 |
| JO-158 | NP_639884 putative<br>large secreted<br>protein<br>[Streptomyces<br>coelicolor A3(2)].         |  | 10<br>G <sup>10</sup> V <sup>11</sup> A <sup>12</sup> TV <sup>13</sup> WS <sup>14</sup> Y <sup>15</sup> L <sup>16</sup> AI <sup>17</sup> S <sup>18</sup> L <sup>19</sup> L <sup>20</sup><br>SEP <sup>21</sup> V <sup>22</sup> AA <sup>23</sup> T <sup>24</sup> GL <sup>25</sup> AL <sup>26</sup><br>36 |  | GV <sup>1</sup> AV <sup>2</sup> VL <sup>3</sup> AIL <sup>4</sup> LP <sup>5</sup> V <sup>6</sup> AA <sup>7</sup><br>AGL <sup>20</sup> 20                                                                                                       |  | VAV <sup>1</sup> PV <sup>2</sup> VL <sup>3</sup> AIL <sup>4</sup> P <sup>11</sup> |  |

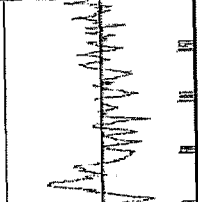
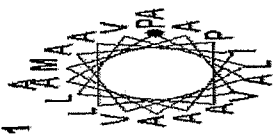
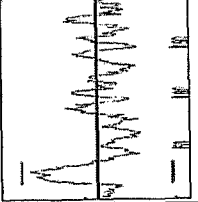
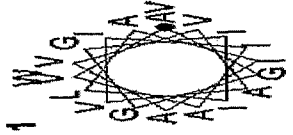
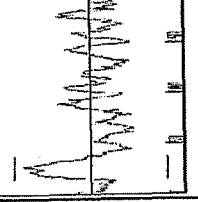
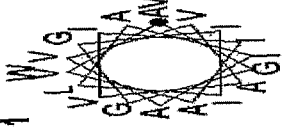
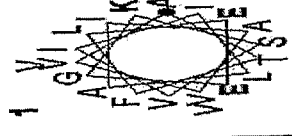
50/134

[Fig. 1m] continued

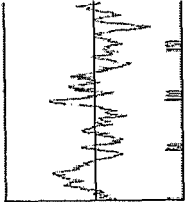
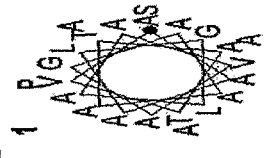
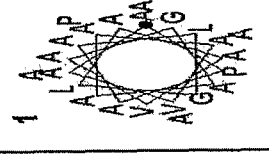
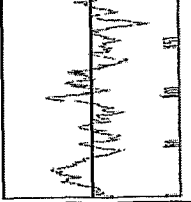
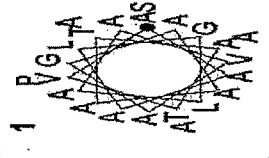
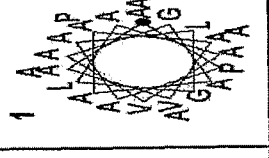
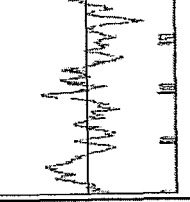
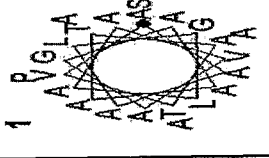
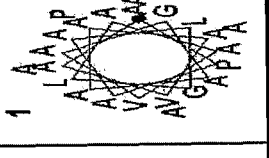
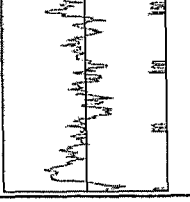
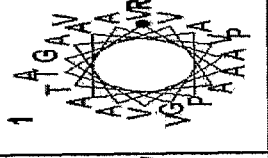
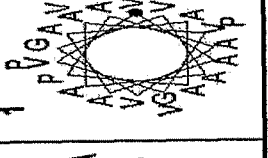
|        |                                                                                        |                                                                                      |                                      |                                                                                      |                          |                                                                                    |                 |                                                                                   |
|--------|----------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------|--------------------------------------|--------------------------------------------------------------------------------------|--------------------------|------------------------------------------------------------------------------------|-----------------|-----------------------------------------------------------------------------------|
| JO-159 | P24327 Foldase<br>protein prsA<br>precursor.                                           |   | IIAIAITATSILAL 17                    | No                                                                                   | IIAIAAILAL 11            | No                                                                                 | IIAIAIPAILAL 12 |  |
| JO-160 | CAB84808<br>putative membrane<br>lipoprotein [Neisseria<br>meningitidis Z2491].        |   | IIASALITFALAAC 18                    | No                                                                                   | IIAALIAFALAAC 13         | No                                                                                 | ALIAPALAAP 10   | No                                                                                |
| JO-161 | NP_639883 putative<br>large secreted<br>protein<br>[Streptomyces<br>coelicolor A3(2)]. |  | IIIRAGPAIALVAMA<br>LTQVELAPHAVAAA 39 |  | PAALVAMALVLP<br>AVAAA 20 |  | AAIALVAPAL 10   | No                                                                                |

51/134

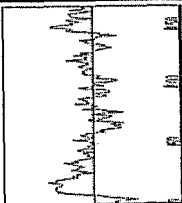
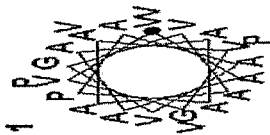
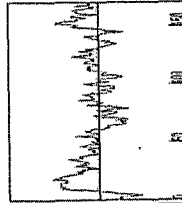
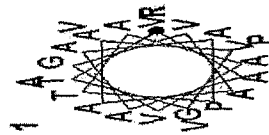
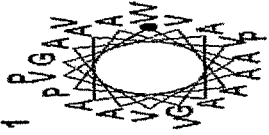
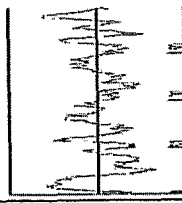
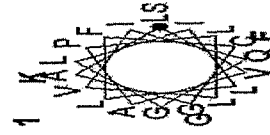
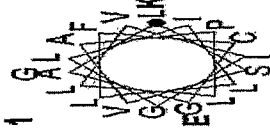
[Fig. 1m] continued

|        |                                                                                        |                                                                                      |                                         |                                                                                      |                            |                                                                                    |                   |    |                                                                                    |
|--------|----------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------|-----------------------------------------|--------------------------------------------------------------------------------------|----------------------------|------------------------------------------------------------------------------------|-------------------|----|------------------------------------------------------------------------------------|
| JO-162 | NP_639883 putative<br>large secreted<br>protein<br>[Streptomyces<br>coelicolor A3(2)]. |   | 11LRAGPAIALVAMA<br>LTQVELAPHAVAAA<br>39 |   | PAIALVAMALVLAP<br>AVAAA 20 |   | LAPAVAAAP 9       | No |                                                                                    |
| JO-163 | NP_627362 secreted<br>protein<br>[Streptomyces<br>coelicolor A3(2)].                   |   | 15NAQVAIIVSAVVAIALI<br>IGGGVW 37        |   | AVAIIVAVVAIALIGG<br>GVW 20 |   | VAIIVPAVVAIALI 15 |    |   |
| JO-164 | NP_627362 secreted<br>protein<br>[Streptomyces<br>coelicolor A3(2)].                   |  | 15NAQVAIIVSAVVAIALI<br>IGGGVW 37        |  | AVAIIVAVVAIALIGG<br>GVW 20 |  | AVAIAIALI 9       |    |  |

[Fig. 1m] continued

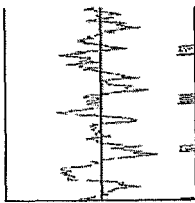
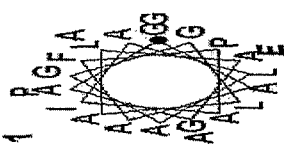
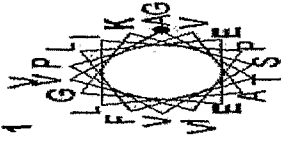
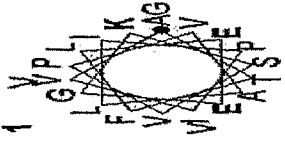
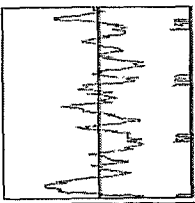
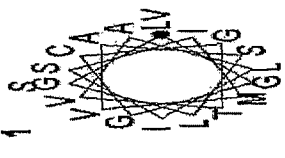
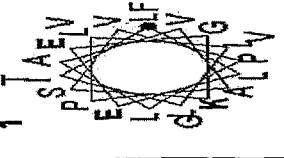
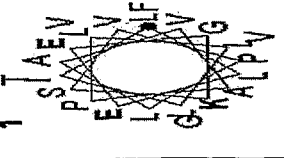
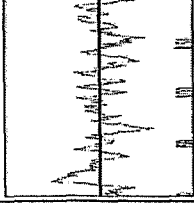
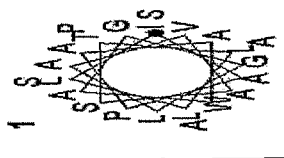
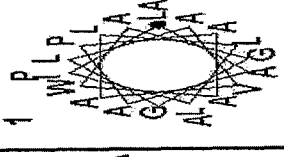
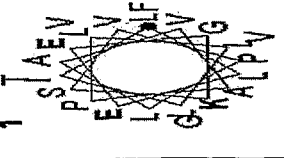
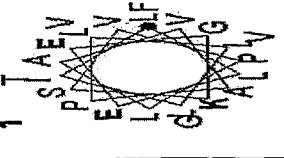
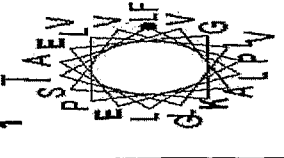
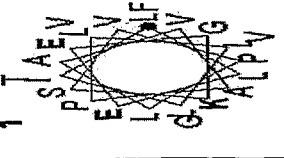
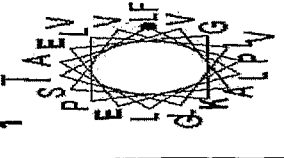
|        |                                                                    |                                                                                       |                                            |                                                                                       |                                       |                                                                                     |               |    |
|--------|--------------------------------------------------------------------|---------------------------------------------------------------------------------------|--------------------------------------------|---------------------------------------------------------------------------------------|---------------------------------------|-------------------------------------------------------------------------------------|---------------|----|
| JO-165 | NP_624625secreted<br>protein<br>[Streptomyces<br>coelicolor A3(2)] |    | 14<br>LAAVTATAAGAVAA<br>LGLAASPAAP 40      |    | LAAVAAAAGAVAAAL<br>GLAAPAAAAP 24      |    | LAAVPAAAP 9   | No |
| JO-166 | NP_624625secreted<br>protein<br>[Streptomyces<br>coelicolor A3(2)] |    | 14<br>LAAVTATAAGAVAA<br>LGLAASPAAP 40      |    | LAAVAAAAGAVAAAL<br>GLAAPAAAAP 24      |    | AVAALPLAAP 10 | No |
| JO-167 | NP_624625secreted<br>protein<br>[Streptomyces<br>coelicolor A3(2)] |   | 14<br>LAAVTATAAGAVAA<br>LGLAASPAAP 40      |   | LAAVAAAAGAVAAAL<br>GLAAPAAAAP 24      |   | LAAPAAAAP 9   | No |
| JO-168 | NP_626936secreted<br>protein<br>[Streptomyces<br>coelicolor A3(2)] |  | 9<br>LAAWTGVAAGV<br>AATPAAARAVVPV<br>PL 39 |  | LAAVGVAAAVGVA<br>APAAAAAVPVPVPL<br>28 |  | LAAVPVAAP 12  | No |

[Fig. 1n]

| #      | Origin                                                             | Hydropathy                                                                            | HRSS<br>(Candidate Domain)                  | Helix                                                                                | Preliminarily<br>modified Sequences | Helix                                                                              | finally modified<br>Sequences | Helix                                                                              |
|--------|--------------------------------------------------------------------|---------------------------------------------------------------------------------------|---------------------------------------------|--------------------------------------------------------------------------------------|-------------------------------------|------------------------------------------------------------------------------------|-------------------------------|------------------------------------------------------------------------------------|
| JO-169 | NP_626936secreted<br>protein<br>[Streptomyces<br>coelicolor A3(2)] |    | 9<br>LAAVVTGVAAGV<br>AATPAAARAVVPV<br>PL 39 |   | LAAVVGVAAGV<br>APAAAAAVPVPPL<br>28  |   | VAAPAAAP 9<br>No              |                                                                                    |
| JO-170 | NP_626936secreted<br>protein<br>[Streptomyces<br>coelicolor A3(2)] |    | 9<br>LAAVVTGVAAGV<br>AATPAAARAVVPV<br>PL 39 |   | LAAVVGVAAGV<br>APAAAAAVPVPPL<br>28  |   | AVPVPVPL 8<br>No              |                                                                                    |
| JO-171 | NP_085072 matrilin<br>2 isoform b precursor<br>[Homo sapiens]      |  | 5<br>LAGCFLLILGQIVLP<br>A 21                |  | LAGCFLLILGIVLLPA<br>16              |  | ILLIPVILLP 10                 |  |

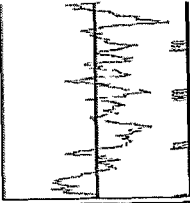
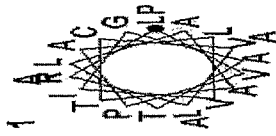
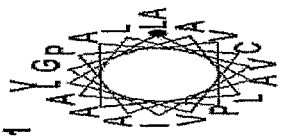
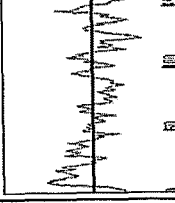
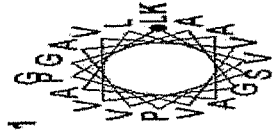
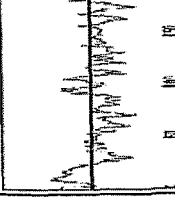
54/134

[Fig. 1n] continued

|        |                                                                                                |                                                                                      |                                        |                                                                                      |                              |                                                                                    |                |                                                                                    |                                                                                    |                                                                                    |                                                                                    |                                                                                    |                                                                                    |
|--------|------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------|----------------------------------------|--------------------------------------------------------------------------------------|------------------------------|------------------------------------------------------------------------------------|----------------|------------------------------------------------------------------------------------|------------------------------------------------------------------------------------|------------------------------------------------------------------------------------|------------------------------------------------------------------------------------|------------------------------------------------------------------------------------|------------------------------------------------------------------------------------|
| JO-172 | CAB94057 putative<br>secreted protein<br>[Streptomyces<br>coelicolor A3(2)].                   |   | 20<br>LGAGALALGGALIA<br>PFAAGPAEAV 44  |   | ALAIAFPFAAGPAAV<br>14        | No                                                                                 | ALALPALAIAP 11 | No                                                                                 |   | No                                                                                 | ALALPALAIAP 11                                                                     | No                                                                                 |   |
| JO-173 | NP_624384 secreted<br>protein<br>[Streptomyces<br>coelicolor A3(2)]                            |   | 4<br>LGLSVMISILAVTGC<br>GG 21          |   | LGLAVMILLAVGCGG              | No                                                                                 | AVIPILAVP 9    | No                                                                                 |  | No                                                                                 | AVIPILAVP 9                                                                        | No                                                                                 |  |
| JO-174 | NP_733505 large,<br>multifunctional<br>secreted protein<br>[Streptomyces<br>coelicolor A3(2)]. |  | 13<br>LILLITSAVALGGAW<br>AAPASSAAPA 37 |  | LILLLAVALGGAWAA<br>PAAAPA 21 |  | LILLPAVALP 11  |  |  |  |  |  |                                                                                    |

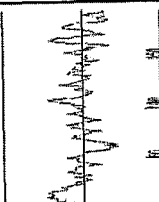
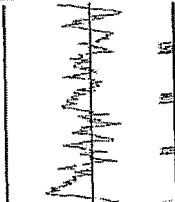
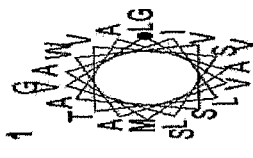
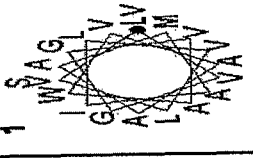
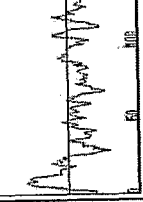
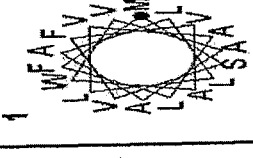
55/134

[Fig. 1n] continued

|        |                                                                             |                                                                                      |                                    |                                                                                      |                            |                                                                                    |                |    |
|--------|-----------------------------------------------------------------------------|--------------------------------------------------------------------------------------|------------------------------------|--------------------------------------------------------------------------------------|----------------------------|------------------------------------------------------------------------------------|----------------|----|
| JO-175 | CAB45630 putative<br>secreted protein<br>[Streptomyces<br>coelicolor A3(2)] |   | 6<br>LRLVCTAALTAGIVL<br>APVPAAA 27 |   | LLVCAALAGIVLAPV<br>PAAA 19 |   | IVLAPVAAAA 10  | No |
| JO-176 | NP_627887 secreted<br>protein<br>[Streptomyces<br>coelicolor A3(2)].        |   | 6<br>LPVWVGAGPVGLA<br>AAA 22       |   | LPVWVGAGPVLAA<br>AA 16     |   | WVVPVLA AAA 11 | No |
| JO-177 | P06832<br>Bacilloolysin<br>precursor                                        |  | 7<br>LSVAVAASFMSLTI<br>20          |  | LVAVAAMLI 9                |  | LVAVAAP 7      | No |

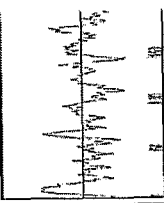
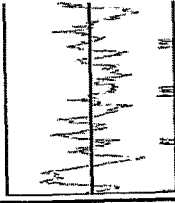
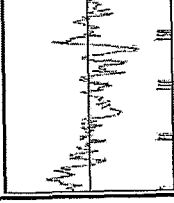
56/134

[Fig. 1n] continued

|        |                                                                             |                                                                                       |                                     |                                                                                                                                           |                           |                                                                                                                                                |                |    |    |
|--------|-----------------------------------------------------------------------------|---------------------------------------------------------------------------------------|-------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------|---------------------------|------------------------------------------------------------------------------------------------------------------------------------------------|----------------|----|----|
| JO-178 | NP_625998<br>secreted hydrolase<br>[Streptomyces<br>coelicolor A3(2)].      |    | 13<br>LYLAAGAALTGAATA<br>27         | No                                                                                                                                        | LVL AAGAALGAAA<br>13      | No                                                                                                                                             | LVL AAPAALP 10 | No | No |
| JO-179 | NP_625057 secreted<br>protein<br>[Streptomyces<br>coelicolor A3(2)]         |    | 10<br>LVLSVTLIAMAAASV<br>WAVGGSV 31 | <br>1<br>GAWV<br>TAAWV<br>AAGV<br>MAGV<br>SLGV<br>LVAV | LVLVLIAMAAAVWAV<br>GGV 18 | <br>1<br>SAGL<br>WAGL<br>LVLV<br>GAVL<br>AVAV<br>LAVL<br>AVAV | LIAPAAAVP 9    | No | No |
| JO-180 | NP_443750 ADP-<br>ribosyltransferase 5<br>precursor [Homo<br>sapiens].      |    | 1<br>MALAALMIALGSLGL<br>15          | No                                                                                                                                        | MALAALMIALGGL             | No                                                                                                                                             | ALAALPIALP 10  | No | No |
| JO-181 | CAB84257 putative<br>secreted protein<br>[Neisseria<br>meningitidis Z2491]. |  | 4<br>MFLSAVLLSAAQQT<br>YVWADTVF 25  | No                                                                                                                                        | MFLAVLLAAAVWA<br>VF 16    | <br>1<br>WFAF<br>LVLV<br>VAVL<br>AVAV<br>LAVL<br>AVAV       | AVLLLPAAA 9    | No | No |

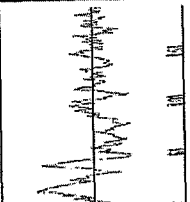
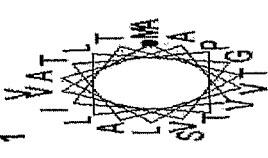
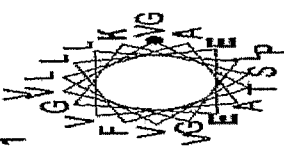
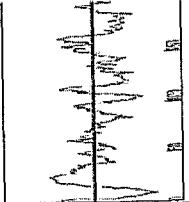
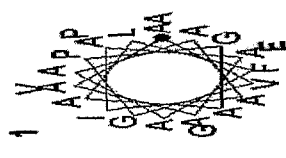
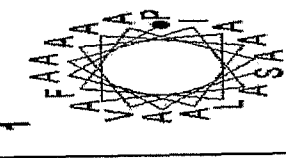
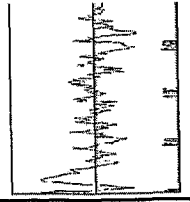
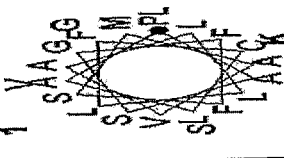
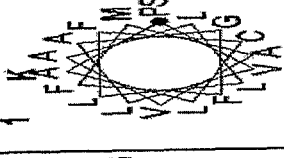
57/134

[Fig. 1o]

| #      | Origin                                                                       | Hydropathy                                                                            | HRSS<br>(Candidate Domain)           | Helix                                          | Preliminarily<br>modified Sequences | Helix                                          | finally modified<br>Sequences | Helix                                                     |
|--------|------------------------------------------------------------------------------|---------------------------------------------------------------------------------------|--------------------------------------|------------------------------------------------|-------------------------------------|------------------------------------------------|-------------------------------|-----------------------------------------------------------|
| JO-182 | P00634 Alkaline<br>phosphatase<br>precursor (APase).                         |    | 6<br>IALALLPLLPVTK<br>20             | No                                             | IALALLPLFPV 12                      | No                                             | IALALLPLL 9                   | 1<br>V GALL K G<br>L F V L E<br>L E P T S P               |
| JO-183 | NP_000933<br>peptidylprolyl<br>isomerase B<br>precursor [Homo<br>sapiens].   |    | 9<br>MKVLLAAALIAGSVF<br>FLLPG 29     | 1<br>P S K L F<br>A S L L G<br>L I G V A L A L | MVLLAAALIAGVFFL<br>LLPG 19          | 1<br>V L L A V<br>F A L L A<br>L A A L L L     | VLLAAALIAP 10                 | 1<br>G A E P<br>K A A A P<br>L P V L T<br>V L G L F V A V |
| JO-184 | CAB7125 putative<br>secreted protein.<br>[Streptomyces<br>coelicolor A3(2)]. |  | 1<br>MTAPAVLTPPVVIG<br>AGPIGLAAAA 25 | 1<br>A V P A<br>T G A L L I<br>V L A L P G V A | MAPAVLPPVVVIGAG<br>PIGLAAAA 23      | 1<br>A V P A<br>L G A L L I<br>V L A L P G V A | APAVLPPVVV 11                 | No                                                        |

58/134

[Fig. 1o] continued

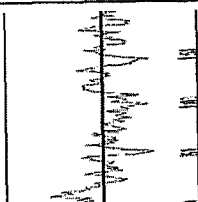
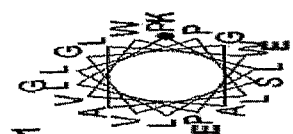
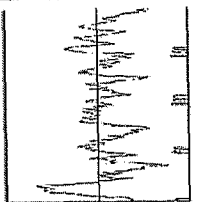
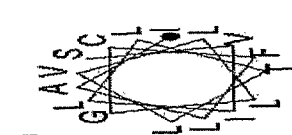
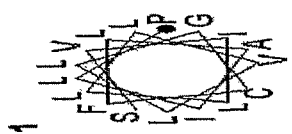
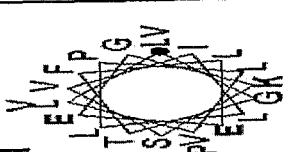
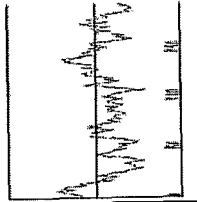
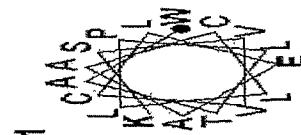
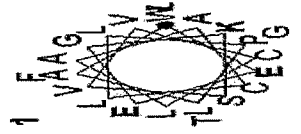
|        |                                                                                                        |                                                                                      |                                          |                                                                                      |                       |                                                                                    |               |    |
|--------|--------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------|------------------------------------------|--------------------------------------------------------------------------------------|-----------------------|------------------------------------------------------------------------------------|---------------|----|
| JO-185 | CAC31847 possible<br>secreted protein<br>[Mycobacterium<br>leprae].                                    |   | 13<br>MVVGLLVAALTTTP<br>TAVA 31          |   | 15<br>MVVGLLVAALIPAVA |   | WVGLLVAAL 9   |    |
| JO-186 | NP_626948 secreted<br>protein<br>[Streptomyces<br>coelicolor A3(2)].                                   |   | 14<br>PAHAAAAIAAAFL<br>AAGPGVAVGEPAAP 42 |   | 16<br>PAAAAAIAAAFLA   |   | AIAAAAPLAA 11 | No |
| JO-187 | NP_059120 cat eye<br>syndrome critical<br>region protein 1<br>(isoform a precursor<br>[Homo sapiens]). |  | 10<br>PALCELLAVANISFF<br>GSAL 28         |  | 17<br>PALCFLLAVAMFFG  |  | ILLAVAP 7     | No |

[Fig. 1o] continued

|                                                             |  |                                     |  |                                                                  |                              |  |                                                      |                      |  |                                                                |
|-------------------------------------------------------------|--|-------------------------------------|--|------------------------------------------------------------------|------------------------------|--|------------------------------------------------------|----------------------|--|----------------------------------------------------------------|
| NP_006519 tissue factor pathway inhibitor 2 (Homo sapiens). |  | <p>3 PARPLGLSILLFLTE<br/>AAL 21</p> |  | <p>1 ELLLS<br/>V F G A L S L<br/>A L S L</p>                     | <p>PARPLGILLFLAAL<br/>15</p> |  | <p>1 VGL<br/>L L L L L<br/>A L S L</p>               | <p>LIILLPLAAL 10</p> |  | <p>1 G A E P<br/>K L P L L<br/>L L L L L<br/>V L G L F V V</p> |
| NP_071447 tubulointerstitial nephritis antigen-like         |  | <p>5 PASLLLLVLASHCCL<br/>GSA 22</p> |  | <p>1 S A L G<br/>L L L L L<br/>S A L G<br/>K S C L A S L</p>     | <p>PALLLVACCLGA<br/>14</p>   |  | <p>1 V V F<br/>L L L L L<br/>T S P L<br/>V L G L</p> | <p>ALLLVLA 8</p>     |  | <p>1 V V F<br/>L L L L L<br/>T S P L<br/>V L G L</p>           |
| NP_071447 tubulointerstitial nephritis antigen-like         |  | <p>5 PLGLLLLLPLAGHLA<br/>L 20</p>   |  | <p>1 G P L L L<br/>L L L L L<br/>V L L L L<br/>E G L L L L L</p> | <p>PLLLLLPLAGLAL<br/>14</p>  |  | <p>1 VGL<br/>L L L L L<br/>A L S L</p>               | <p>LIILLPLA 9</p>    |  | <p>1 V V F<br/>L L L L L<br/>T S P L<br/>V L G L</p>           |

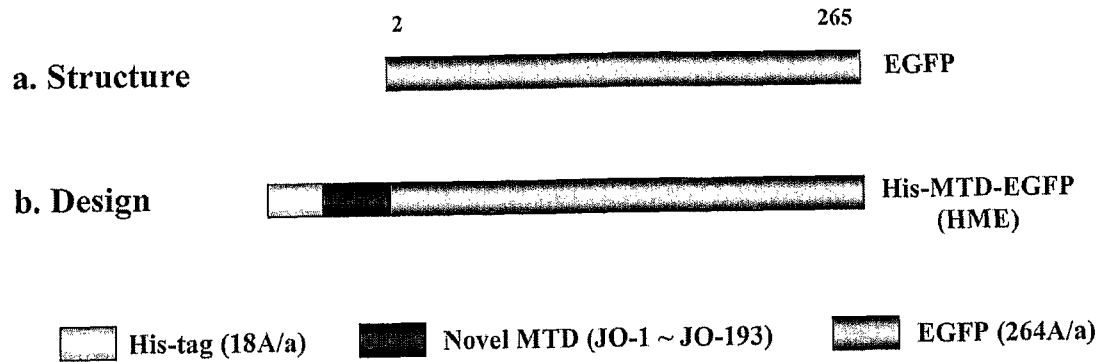
60/134

[Fig. 1o] continued

|        |                                                                                           |                                                                                      |                                |                                                                                      |                         |                                                                                    |             |                                                                                   |    |
|--------|-------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------|--------------------------------|--------------------------------------------------------------------------------------|-------------------------|------------------------------------------------------------------------------------|-------------|-----------------------------------------------------------------------------------|----|
| JO-191 | NP_056322<br>epidermal growth<br>factor-like protein 6<br>precursor [Homo<br>sapiens].    |   | 2<br>PLPWSLALPILLSW<br>VAGG 19 | No                                                                                   | PLPWIALPILLWVA<br>GG 16 |   | IALPILLP 8  | No                                                                                |    |
| JO-192 | NP_628035secreted<br>penicillin-binding<br>protein<br>[Streptomyces<br>coelicolor A3(2)]. |   | 4<br>PLRRIALFCGLLVLT<br>LI 21  |   | PLIALFCGLLVLLI<br>15    |   | ILLVPLLI 8  |  | No |
| JO-193 | NP_663380<br>cathexin H isoform<br>b precursor [Homo<br>sapiens].                         |  | 2<br>WATPLLCAELCV<br>15        |  | WALPLLCAALCV 12         |  | LPLLPAALV 9 | No                                                                                |    |

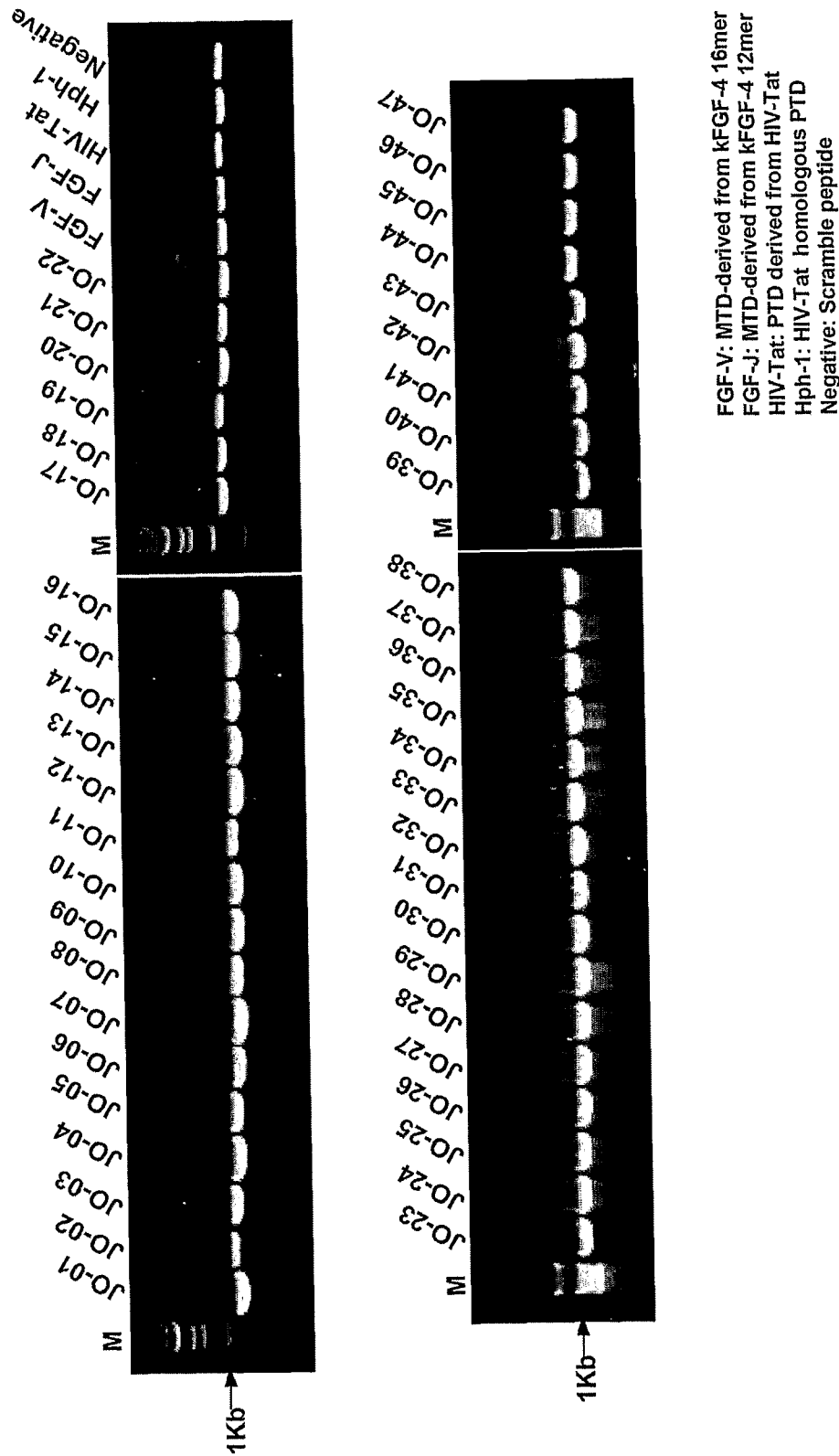
61/134

[Fig. 2]



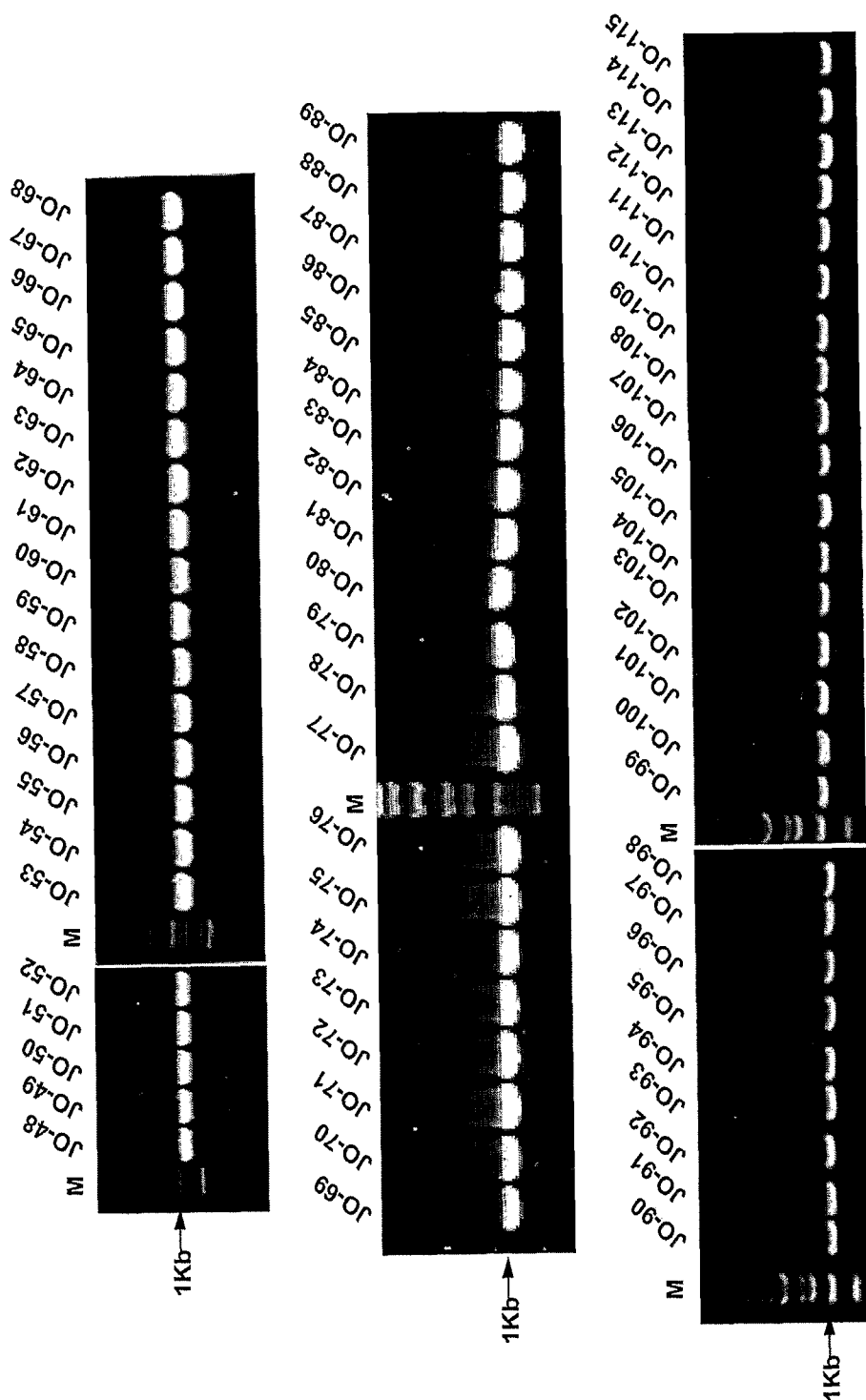
62/134

[Fig. 3a]

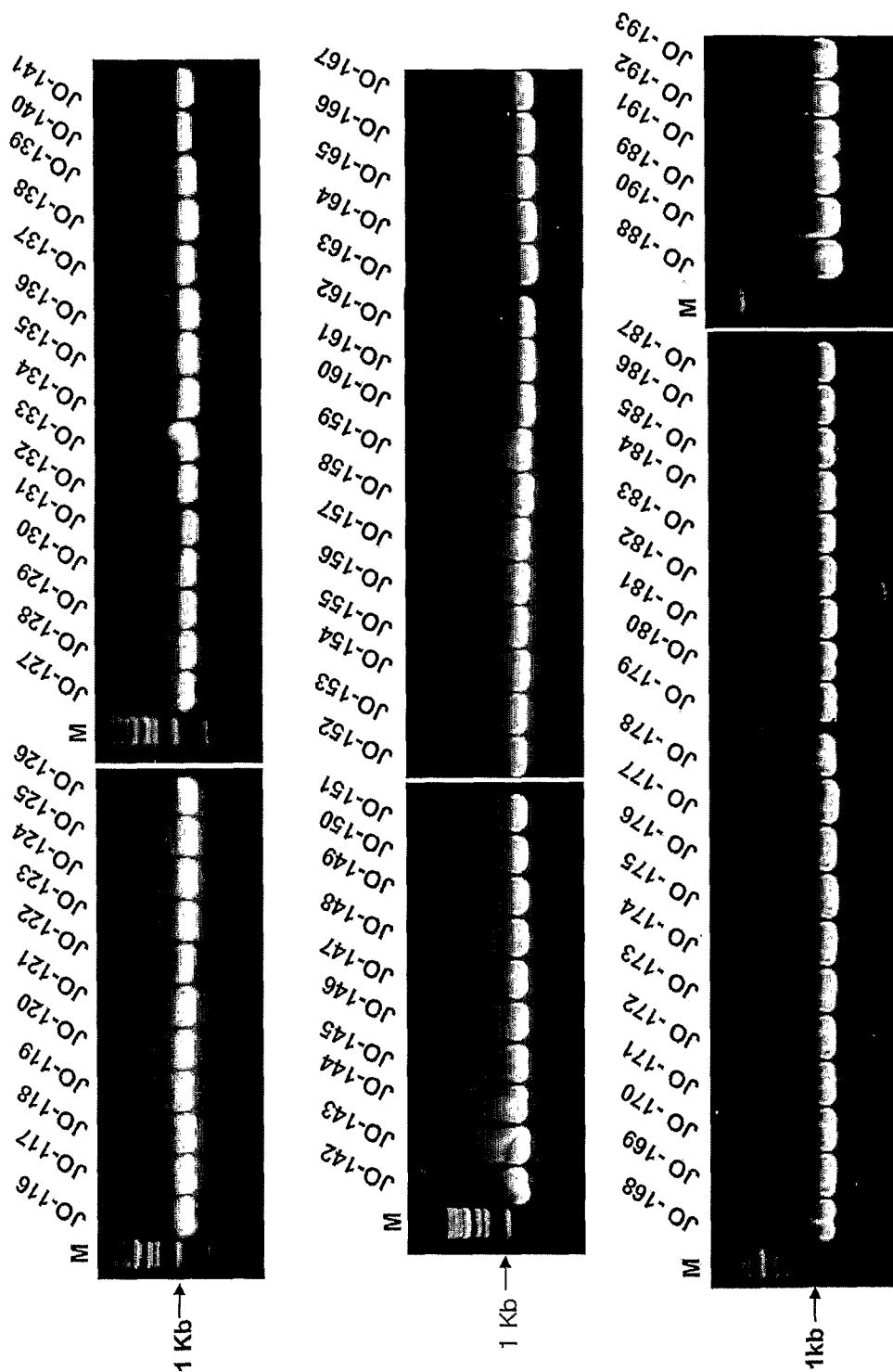


63/134

[Fig. 3b]

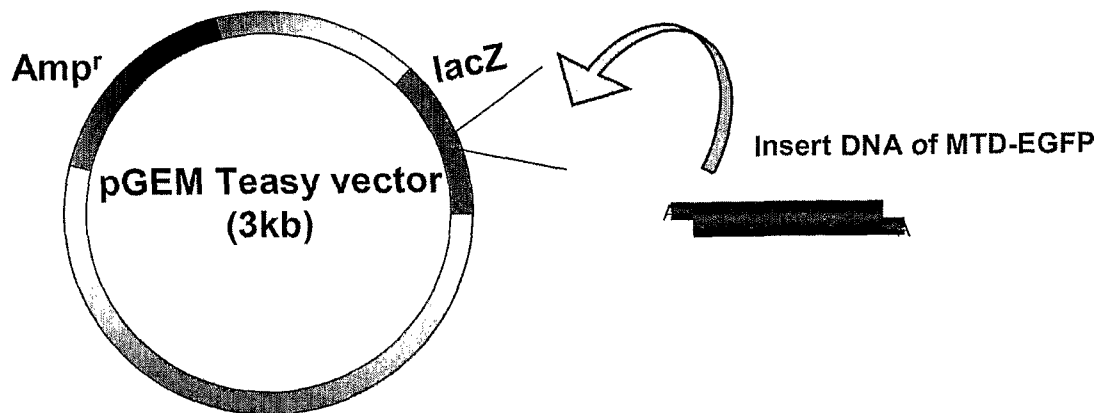


[Fig. 3c]

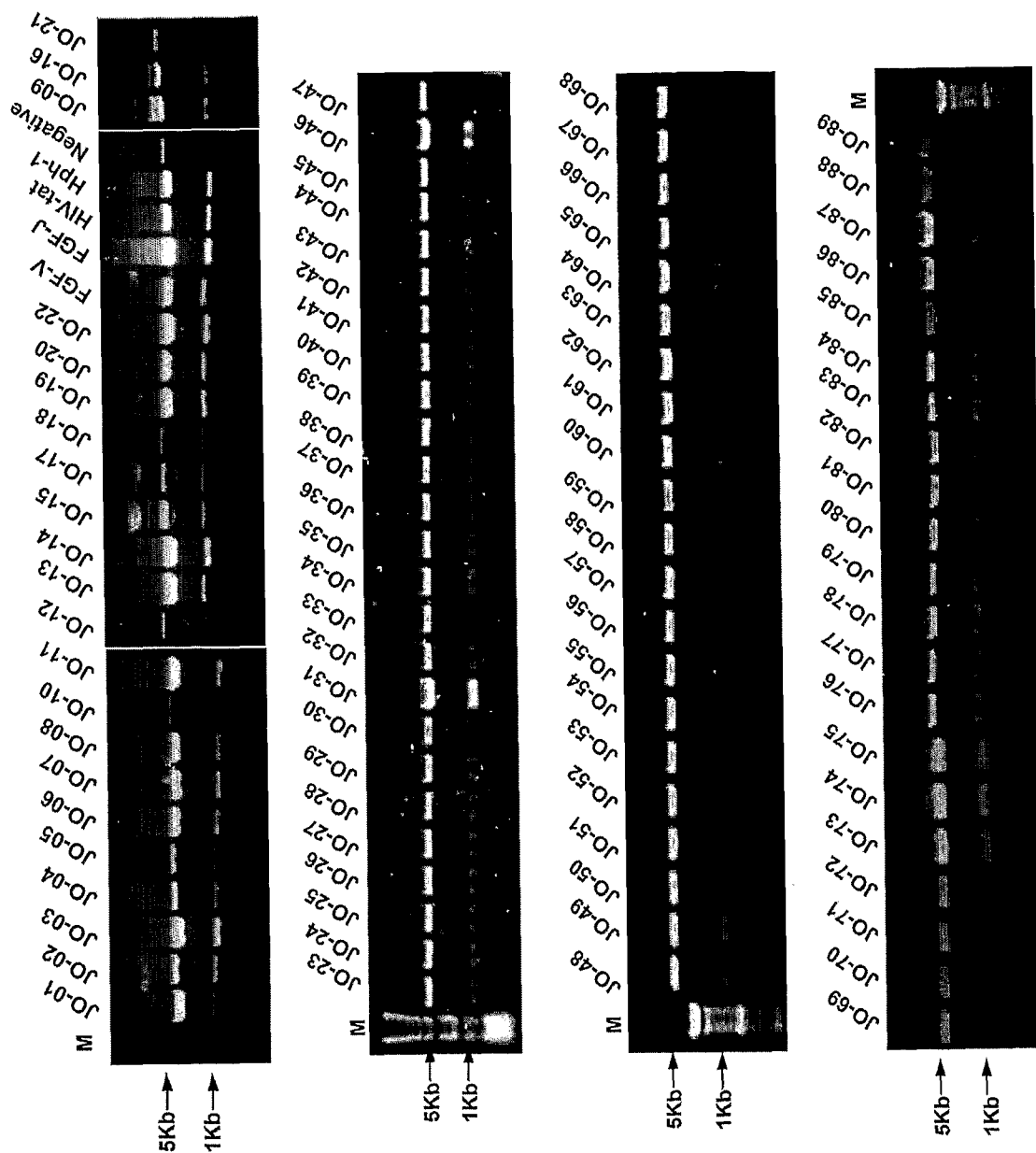


65/134

[Fig. 4a]

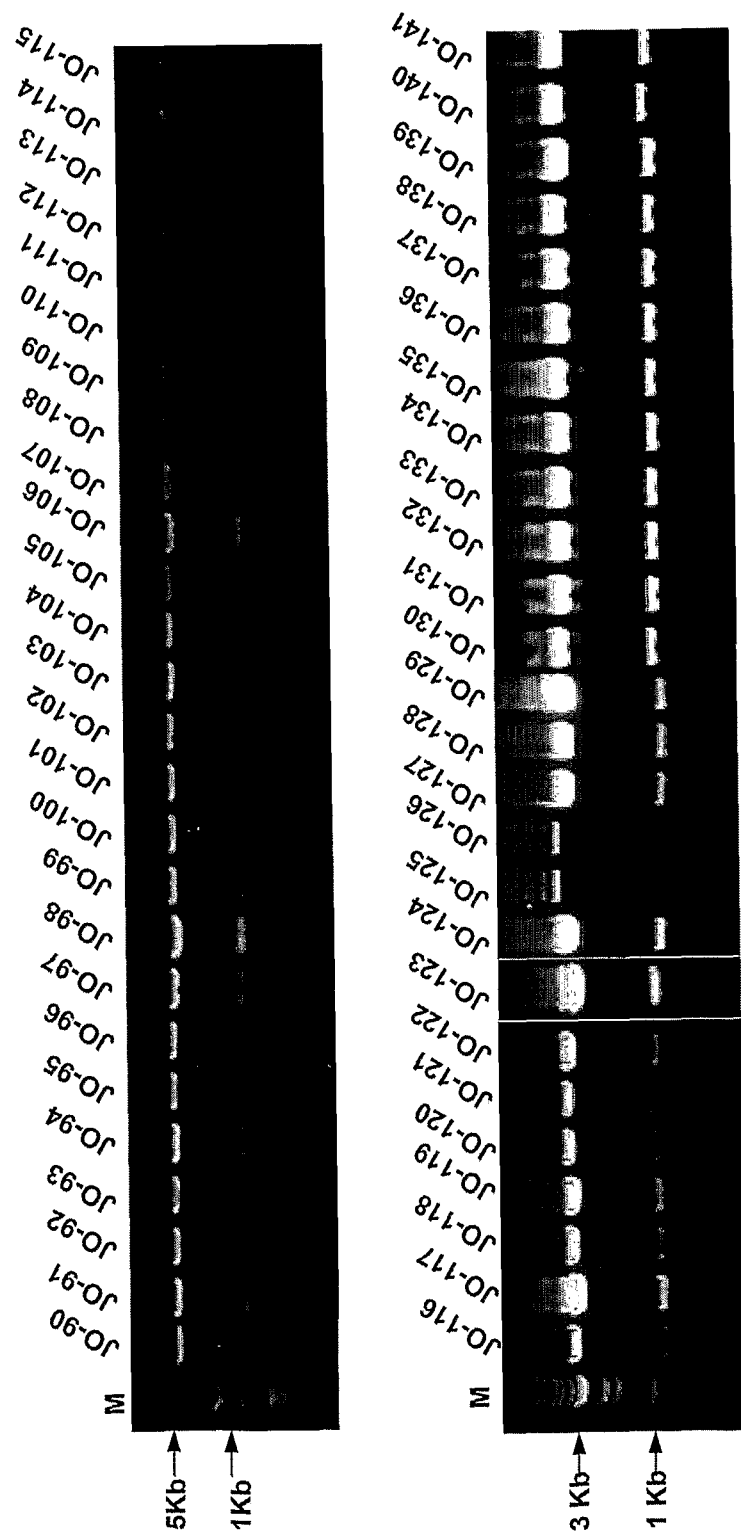


[Fig. 4b]

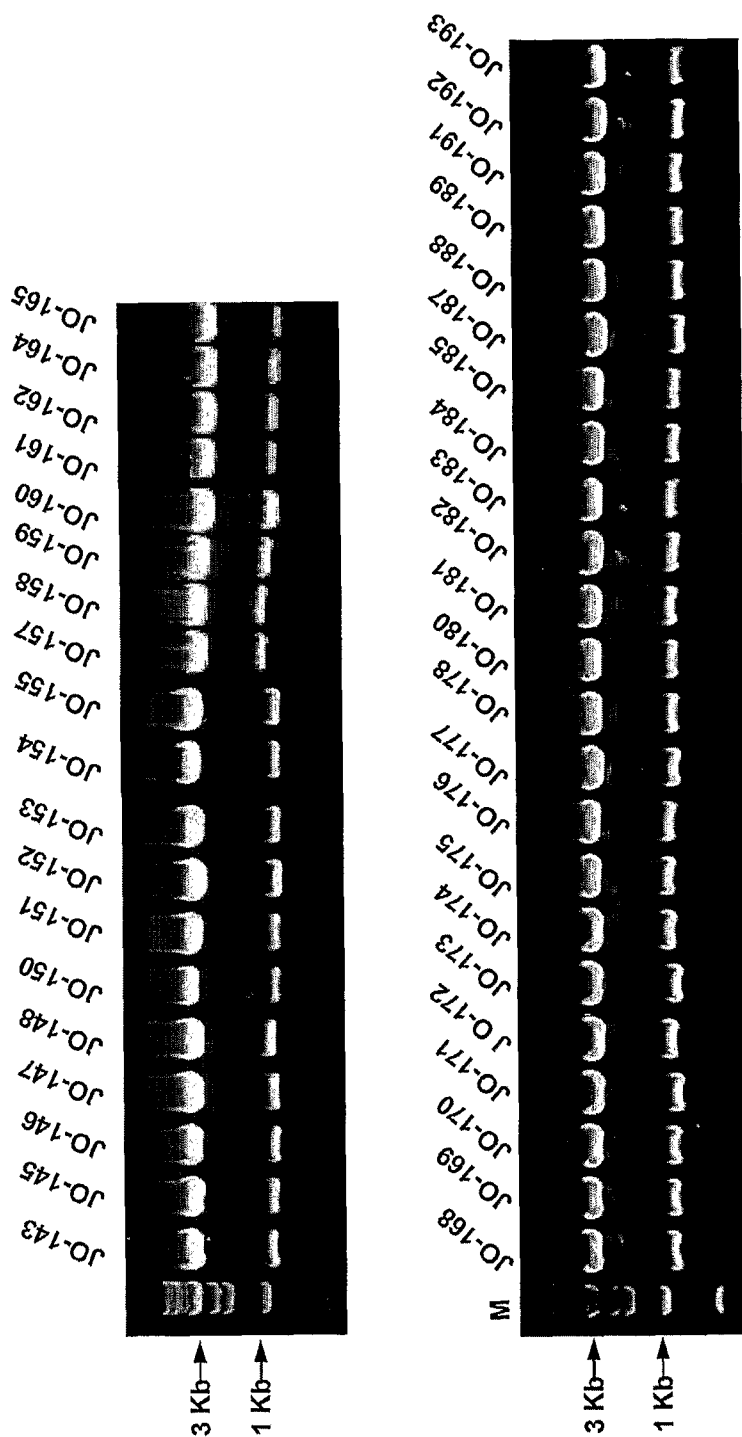


67/134

[Fig. 4c]

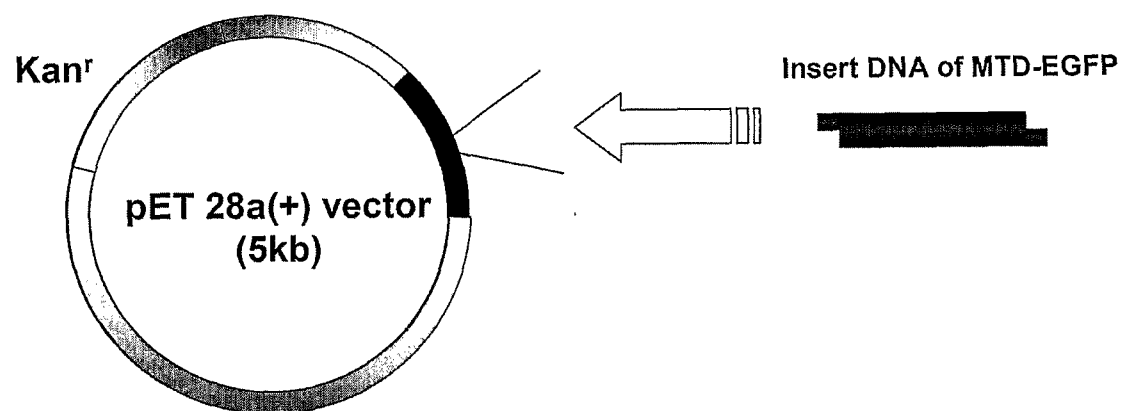


[Fig. 4d]

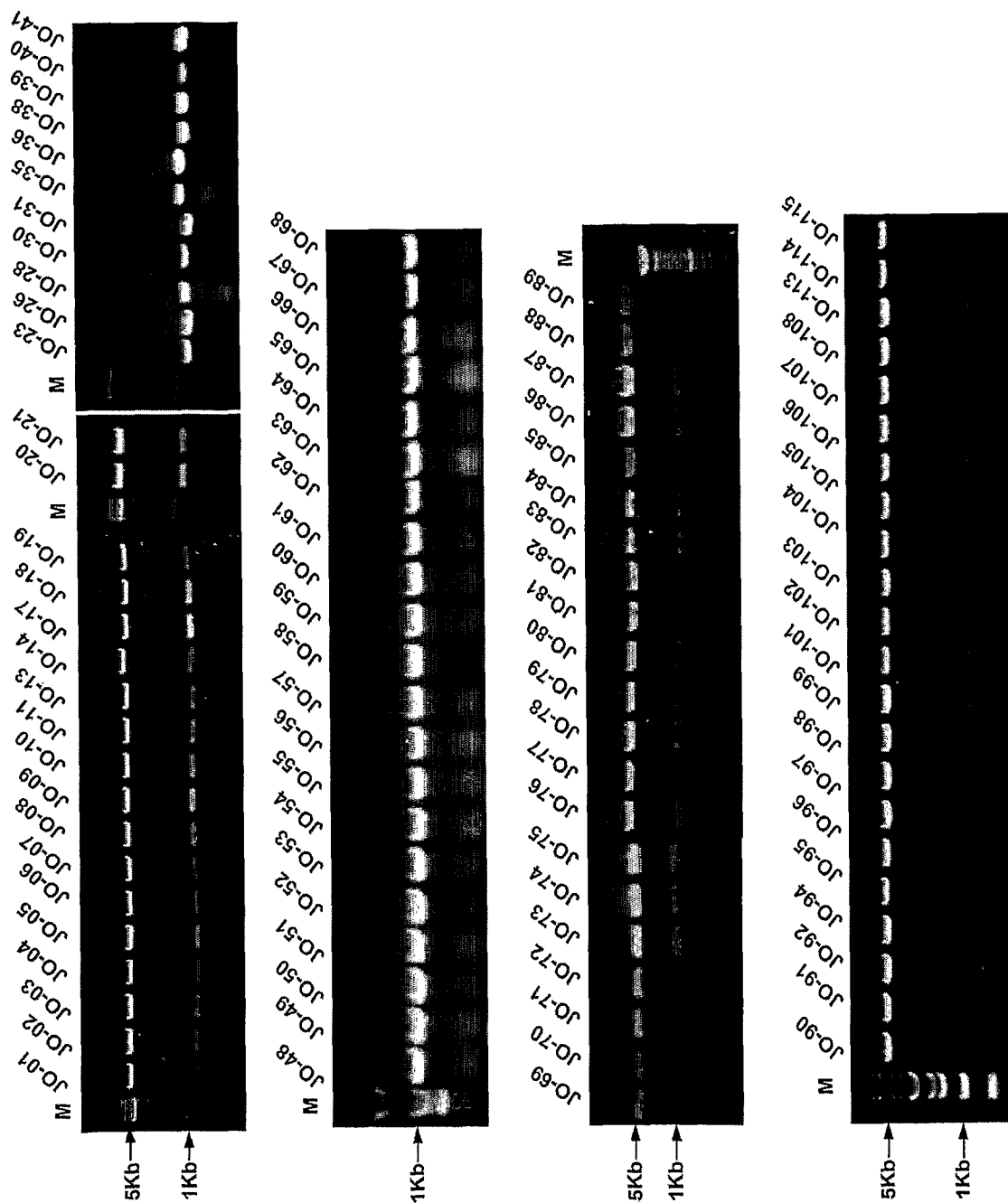


69/134

[Fig. 5a]

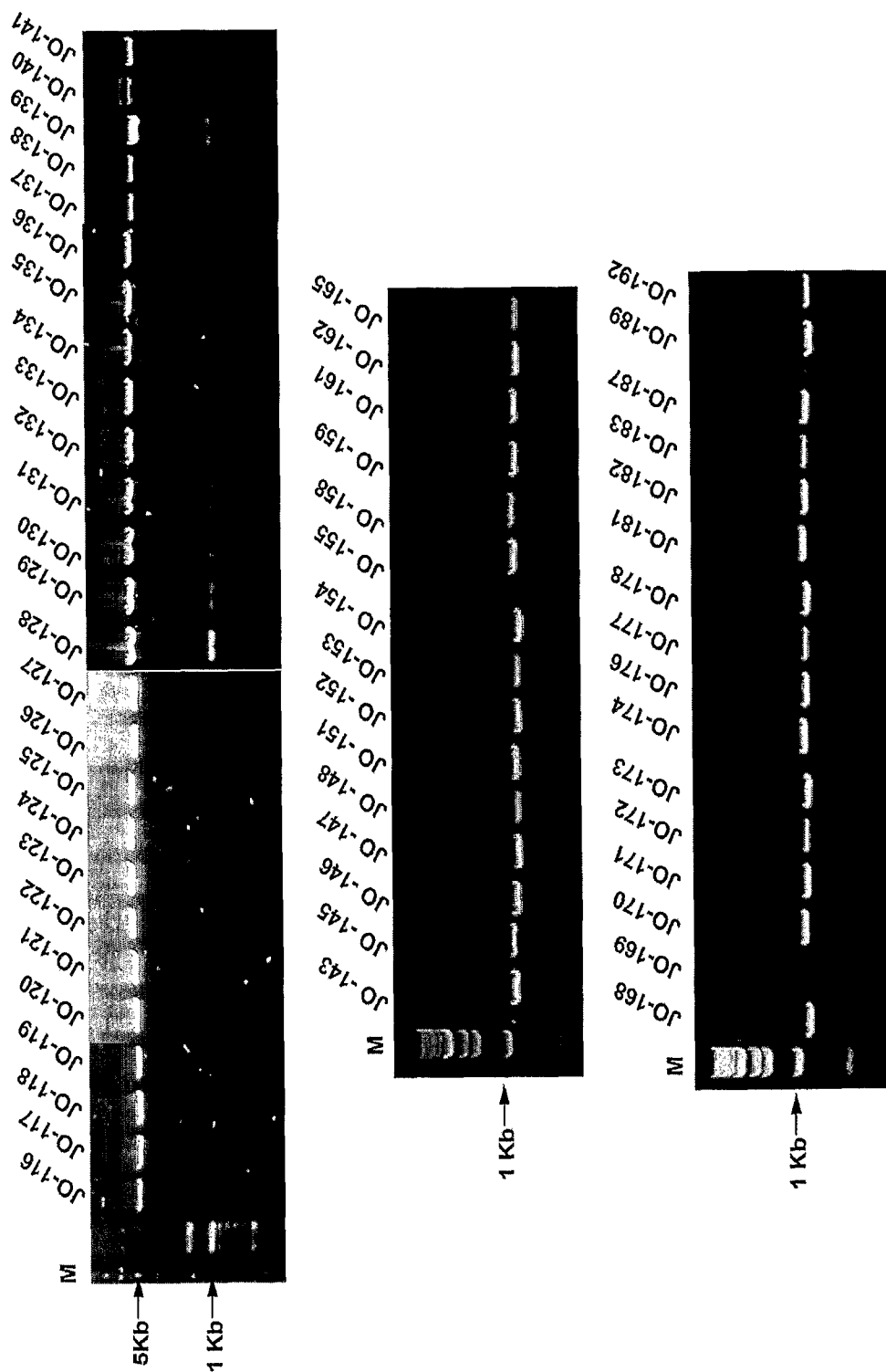


[Fig. 5b]



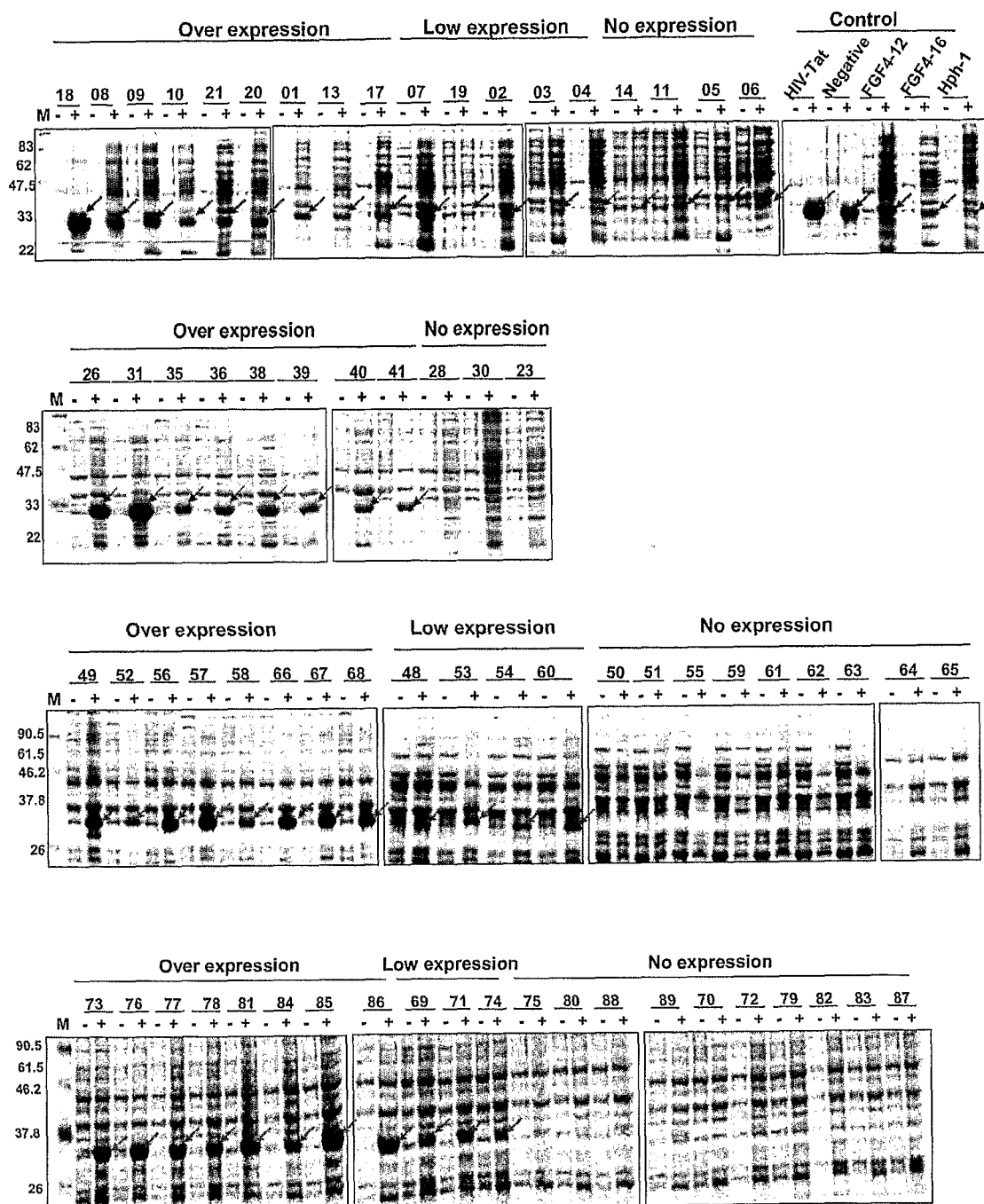
71/134

[Fig. 5c]



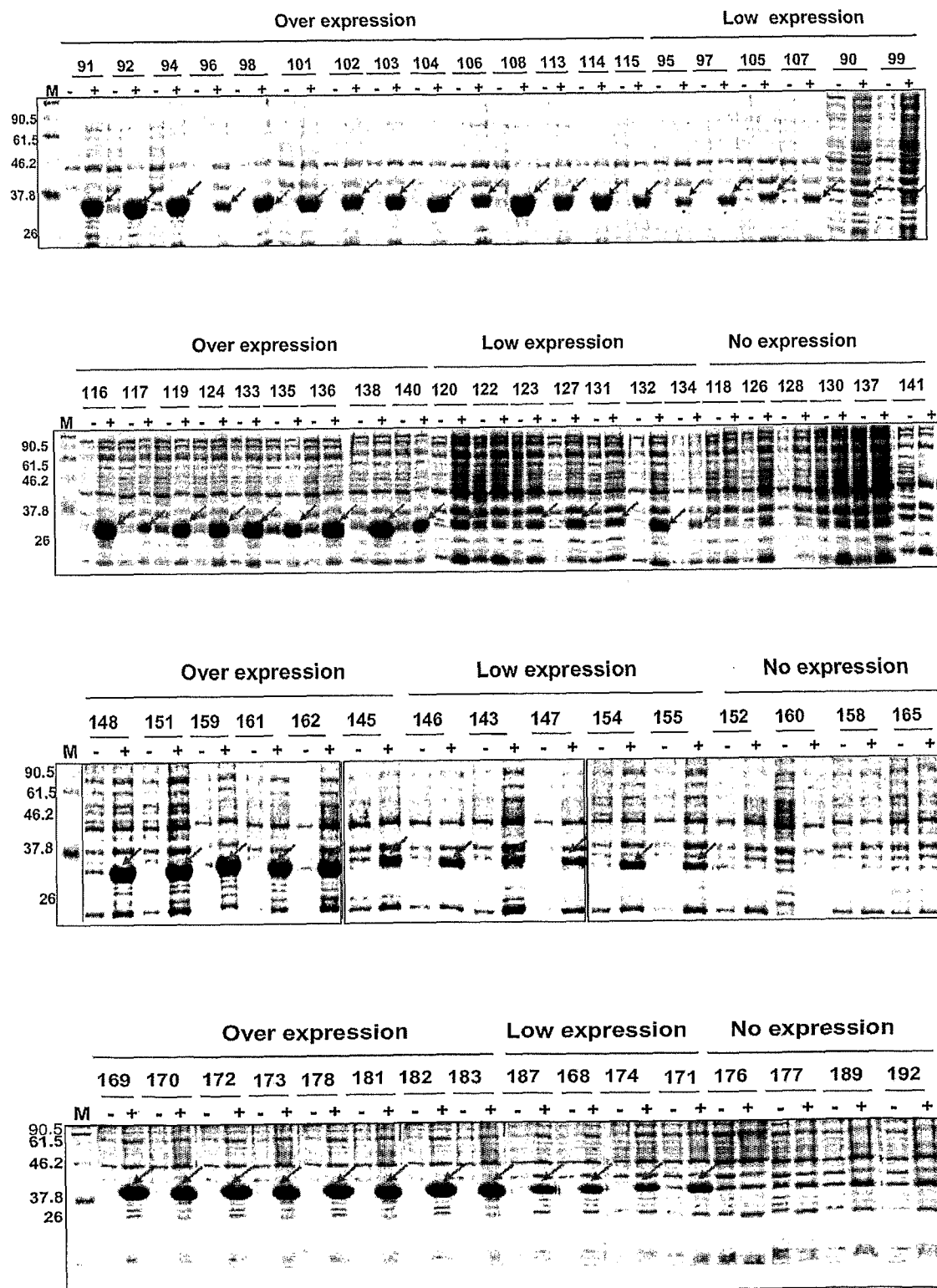
72/134

[Fig. 6a]



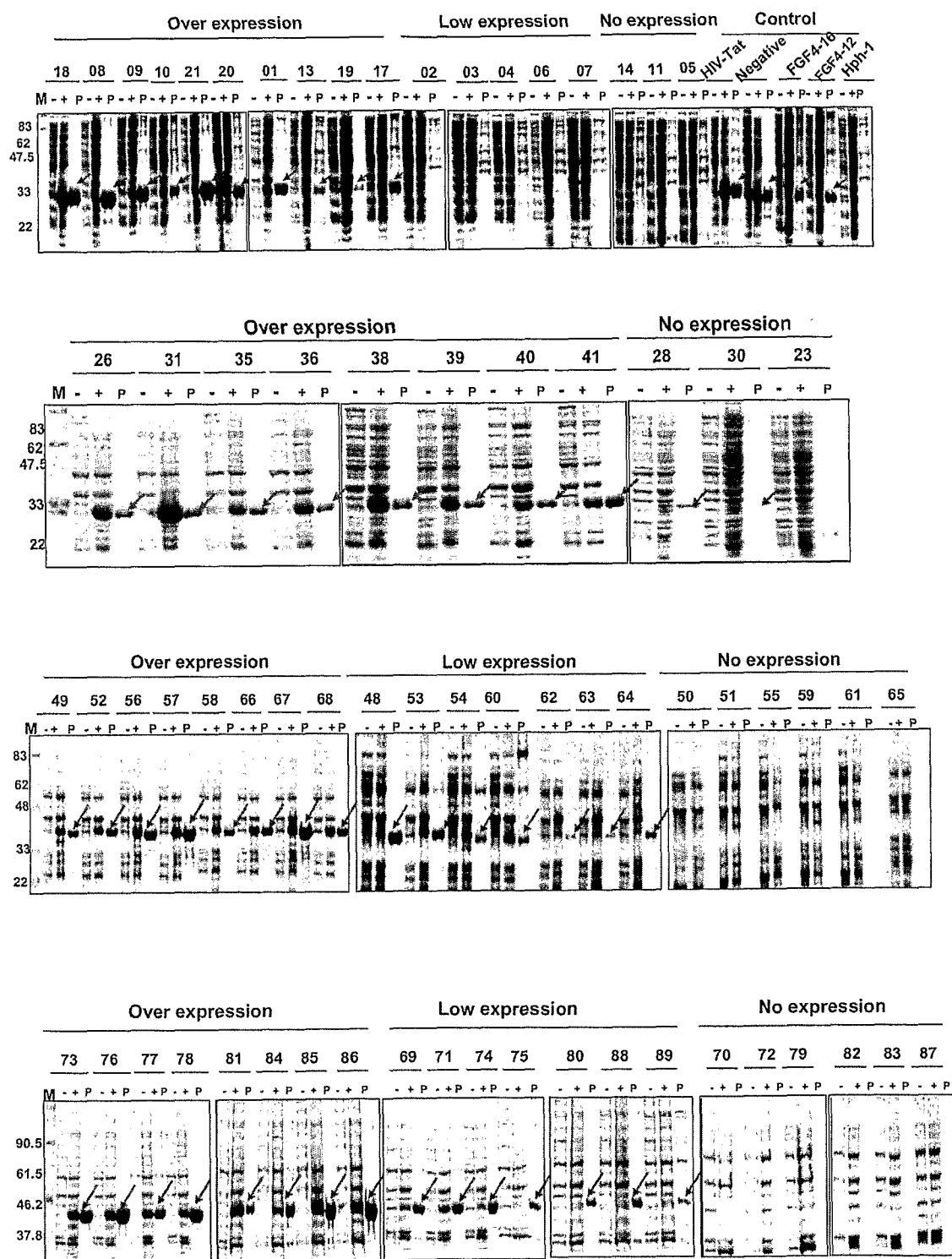
73/134

[Fig. 6b]



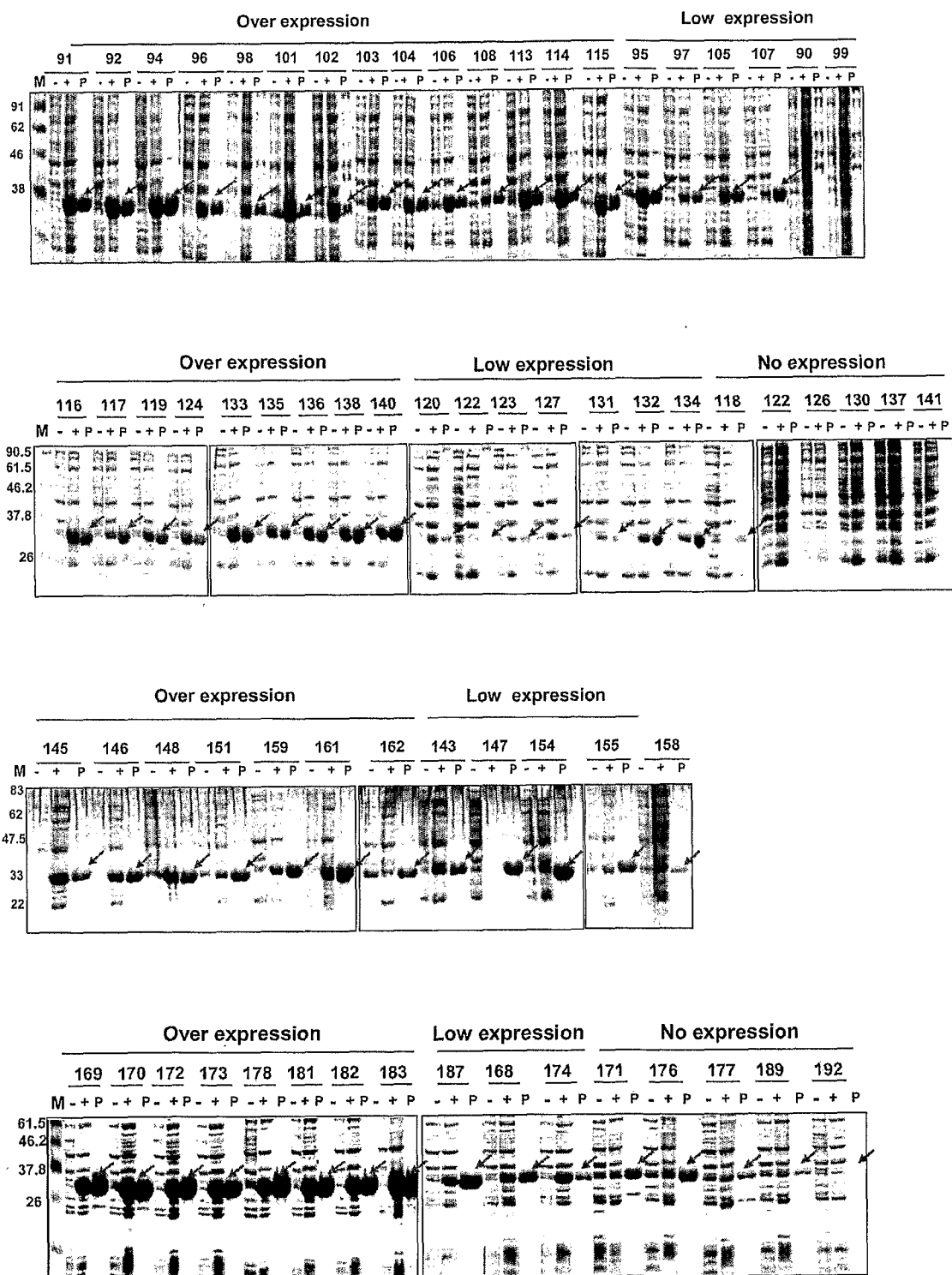
74/134

[Fig. 7a]



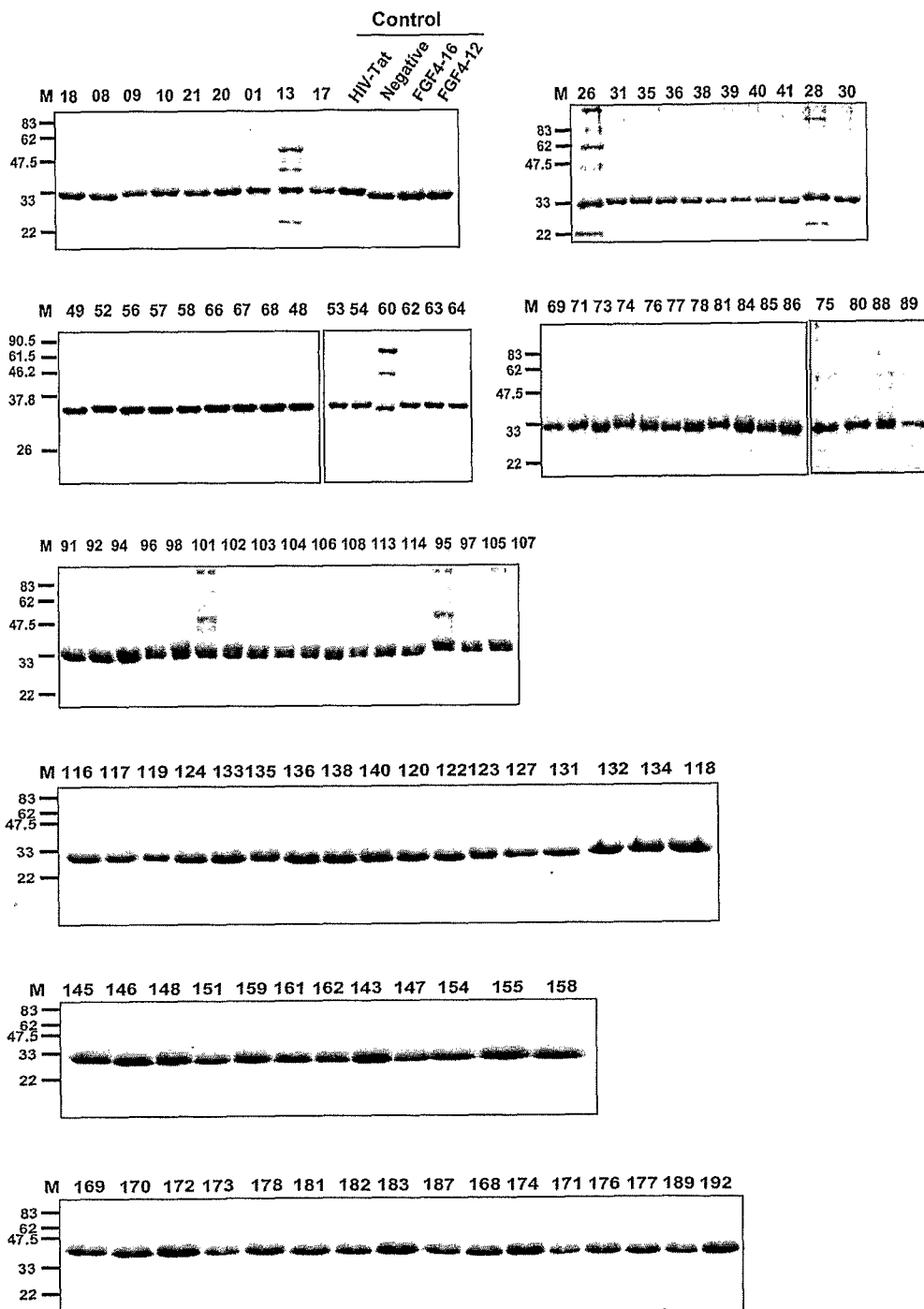
75/134

[Fig. 7b]



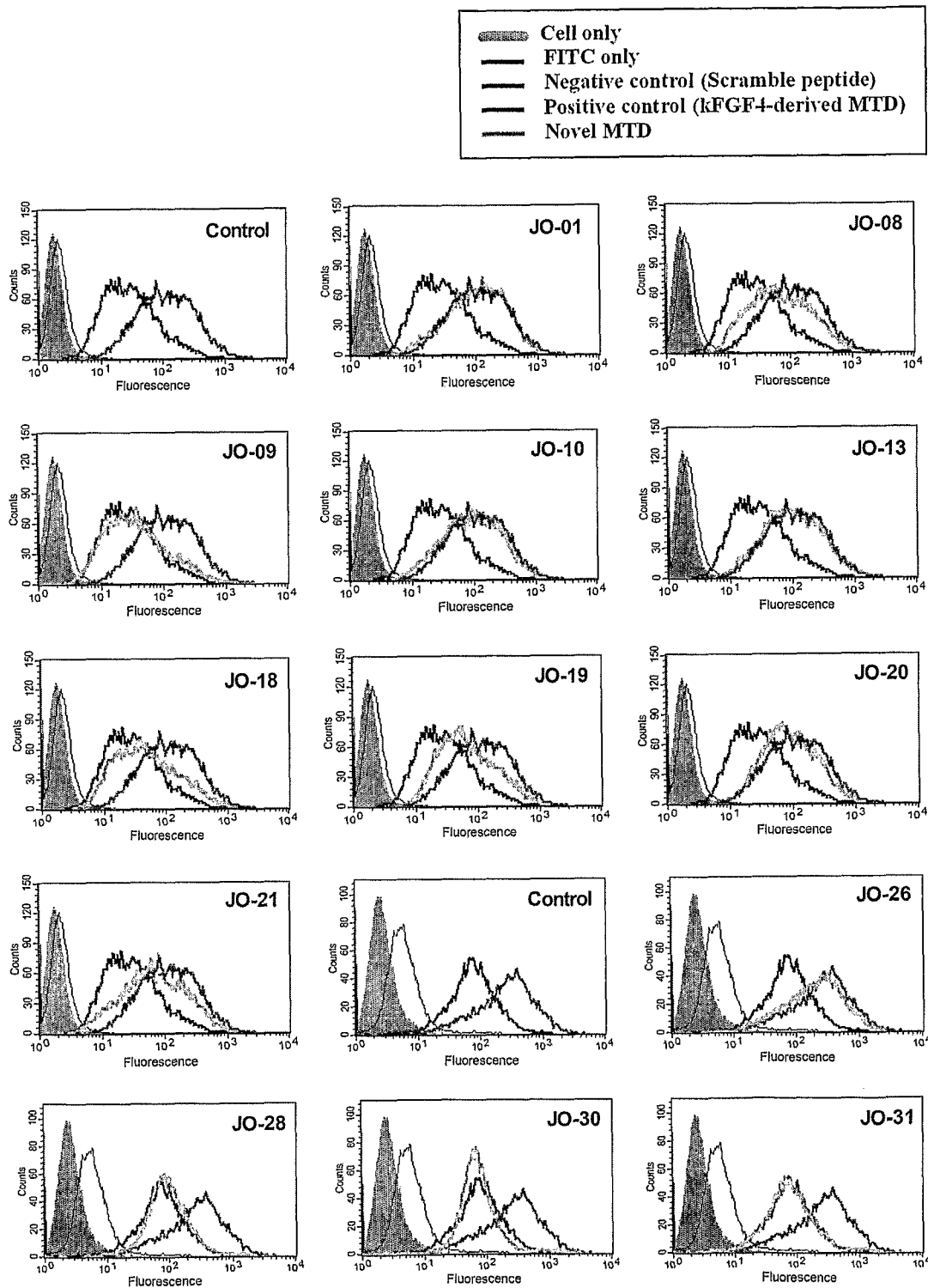
76/134

[Fig. 8]



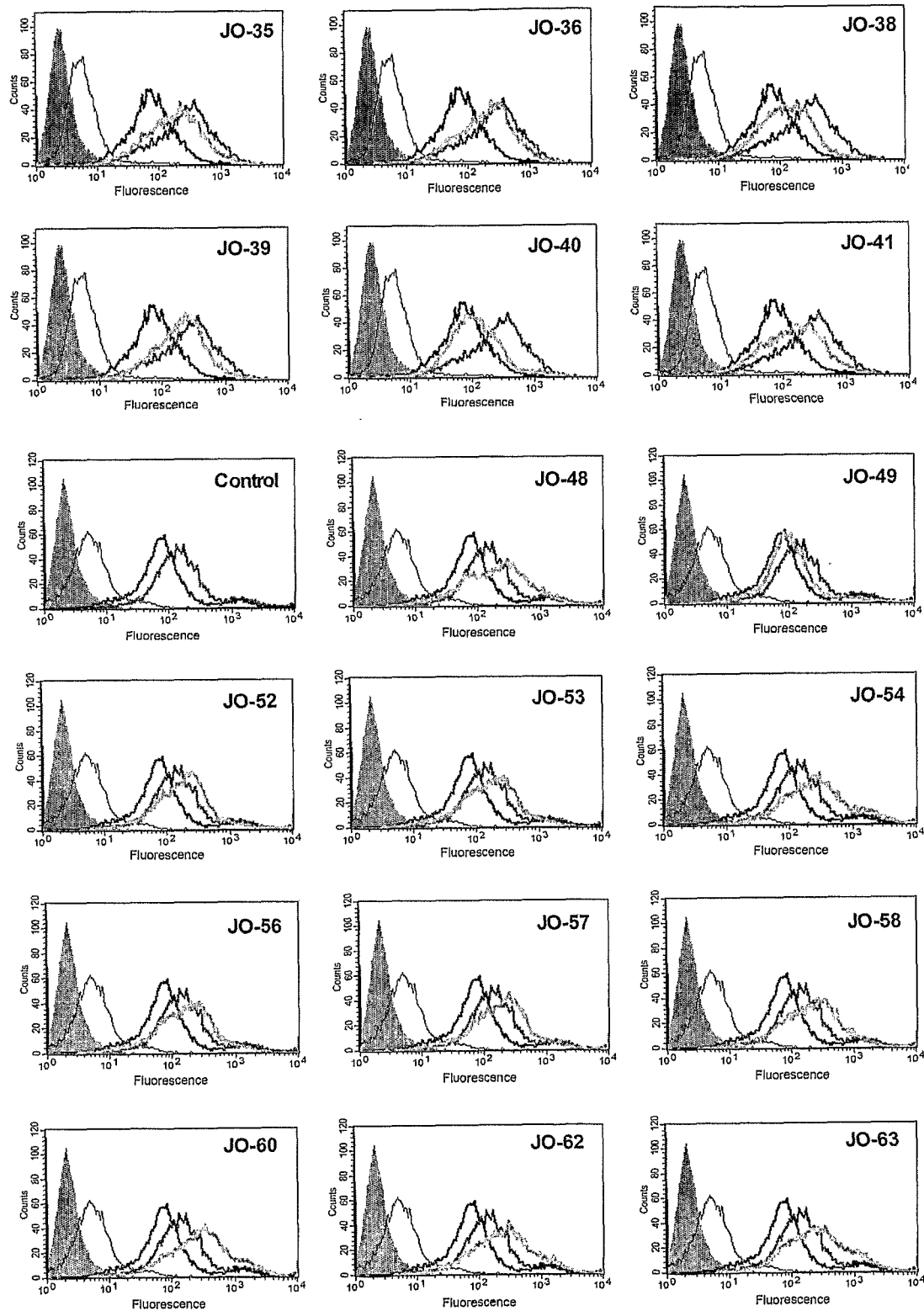
77/134

[Fig. 9a]



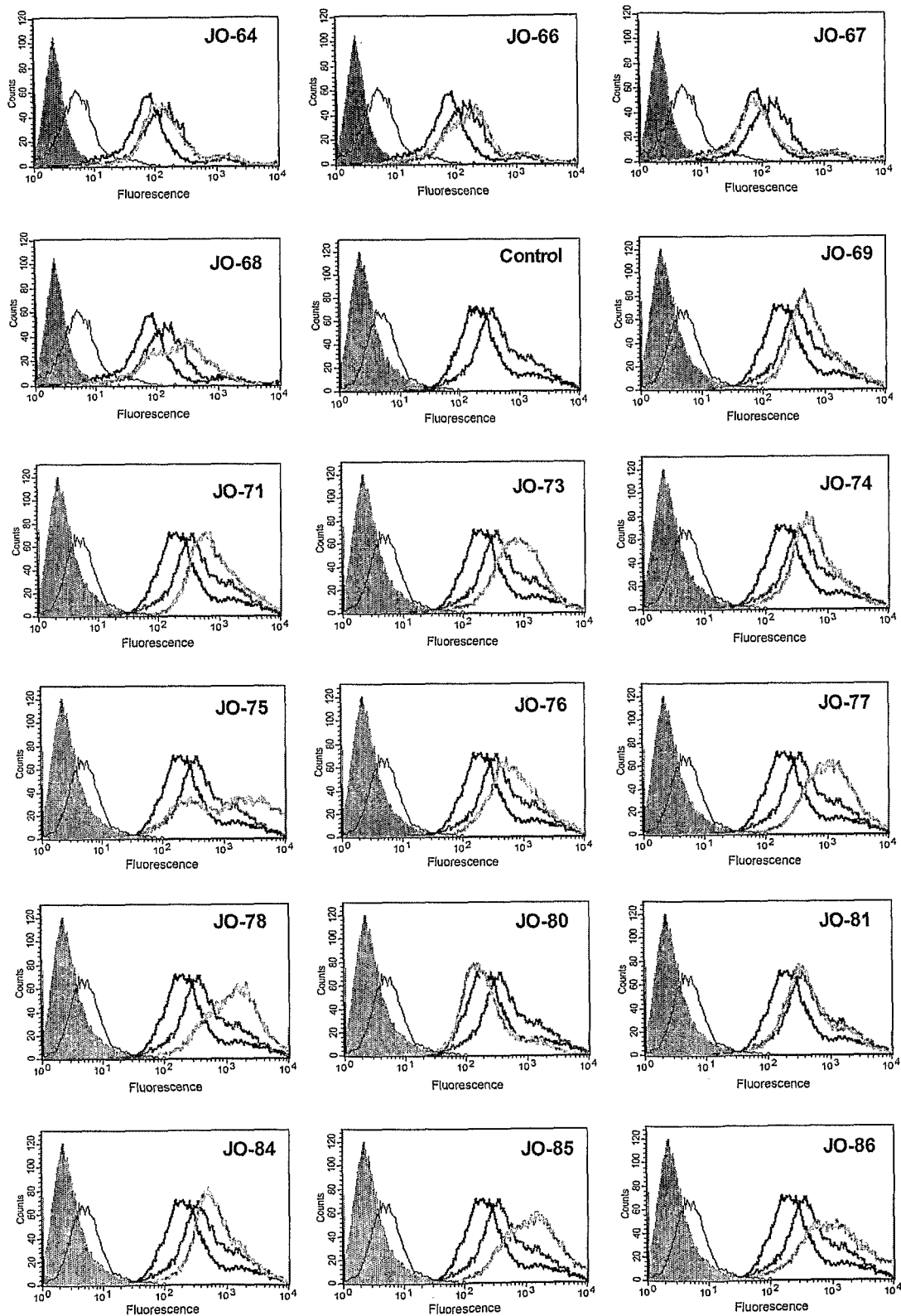
78/134

[Fig. 9b]



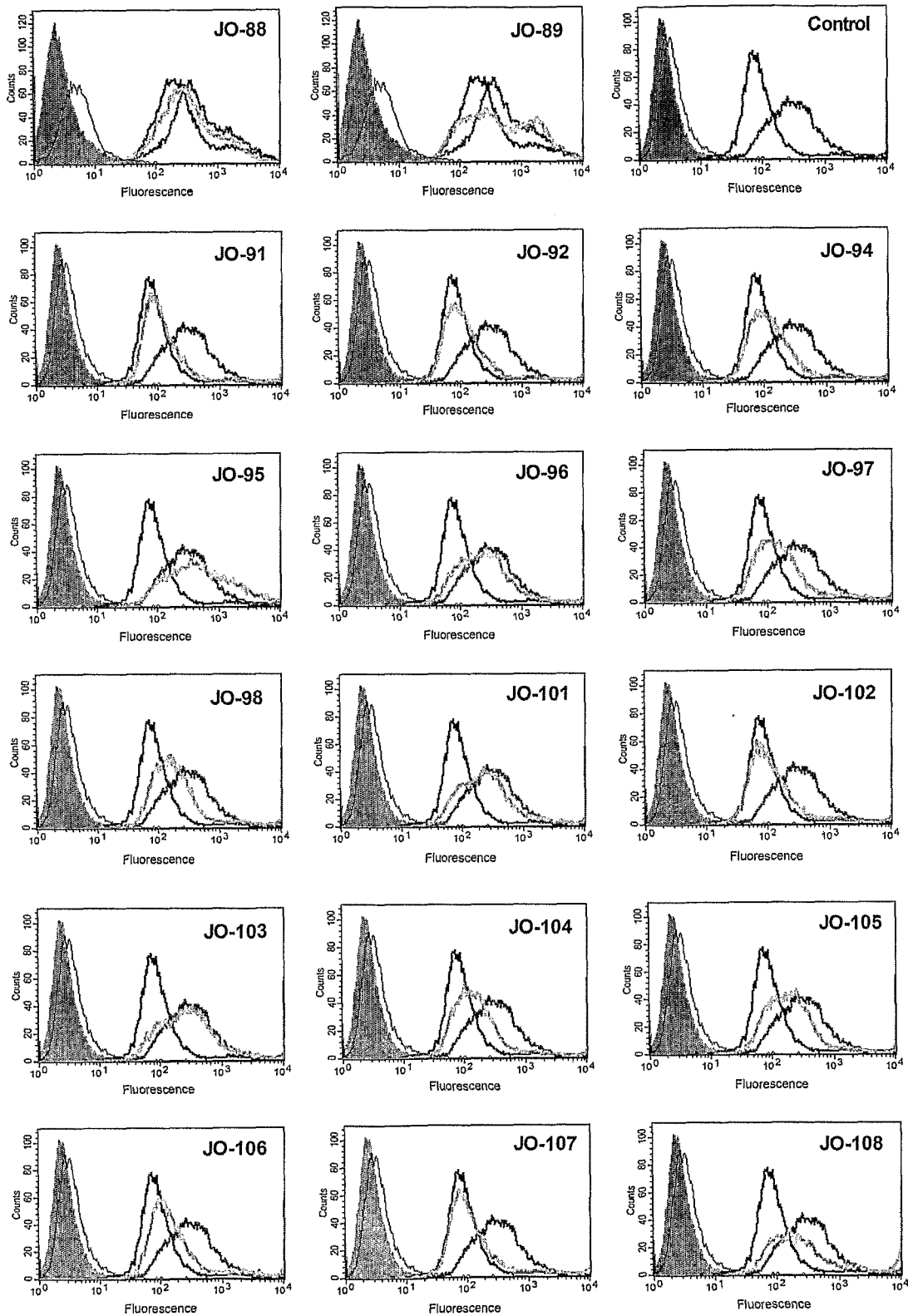
79/134

[Fig. 9c]



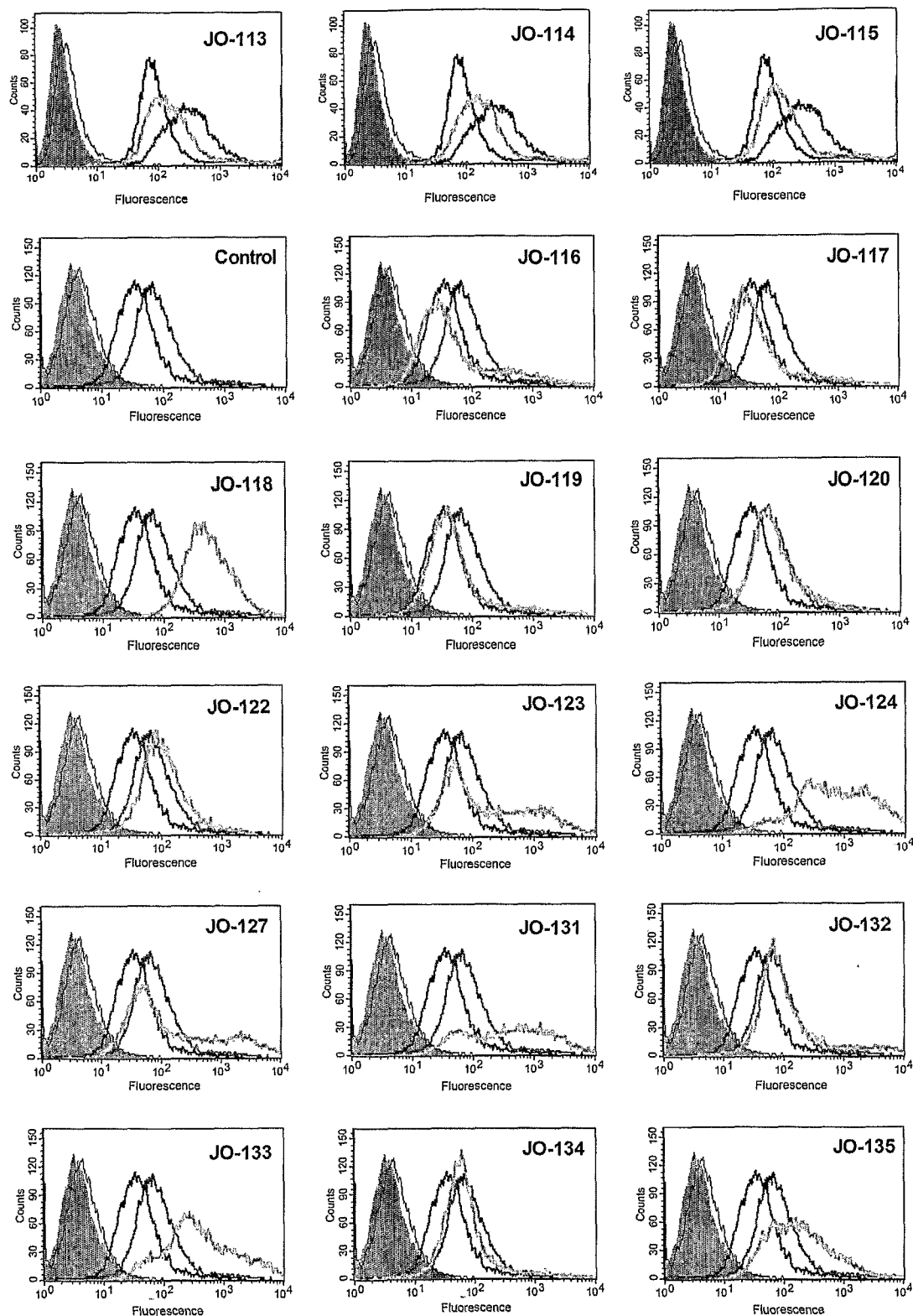
80/134

[Fig. 9d]



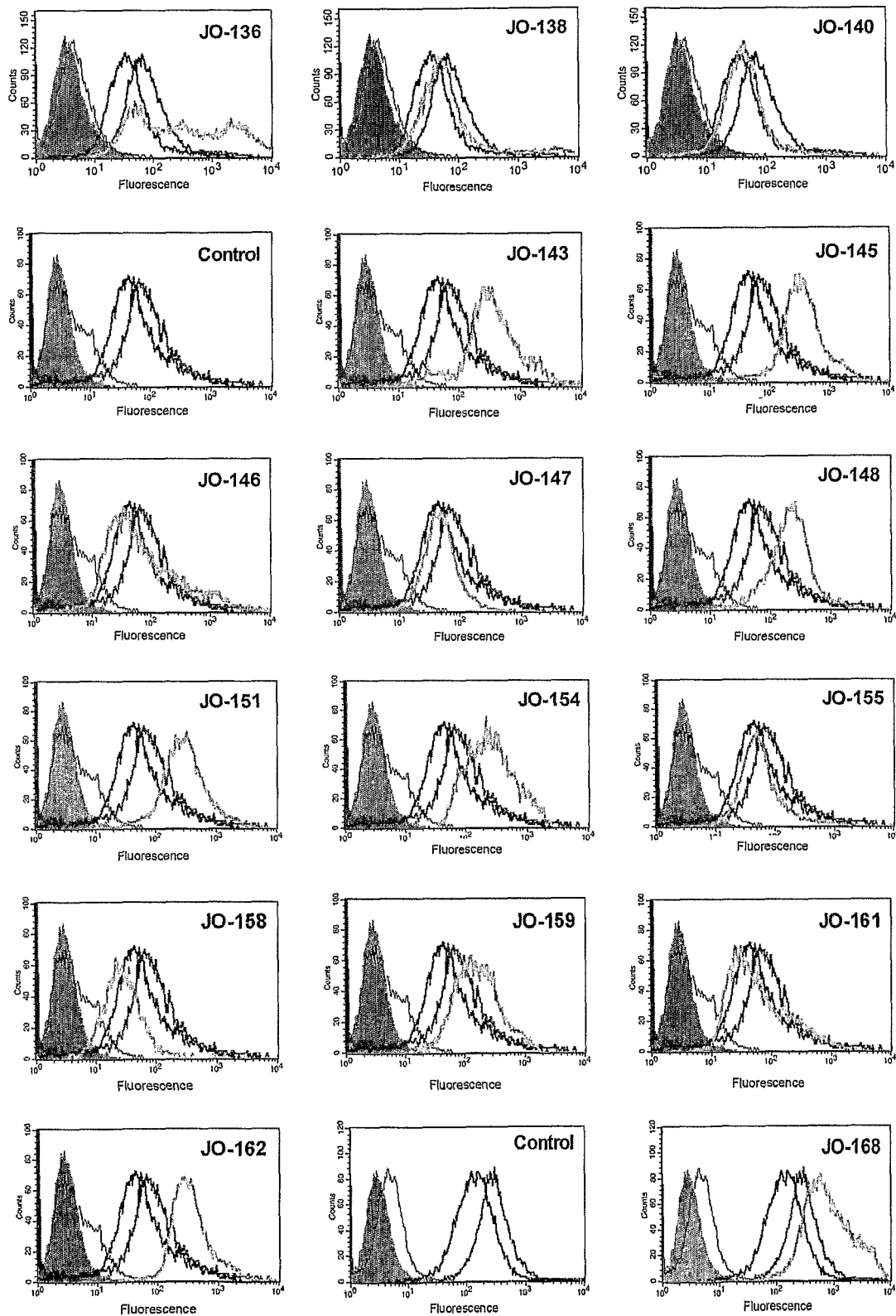
81/134

[Fig. 9e]



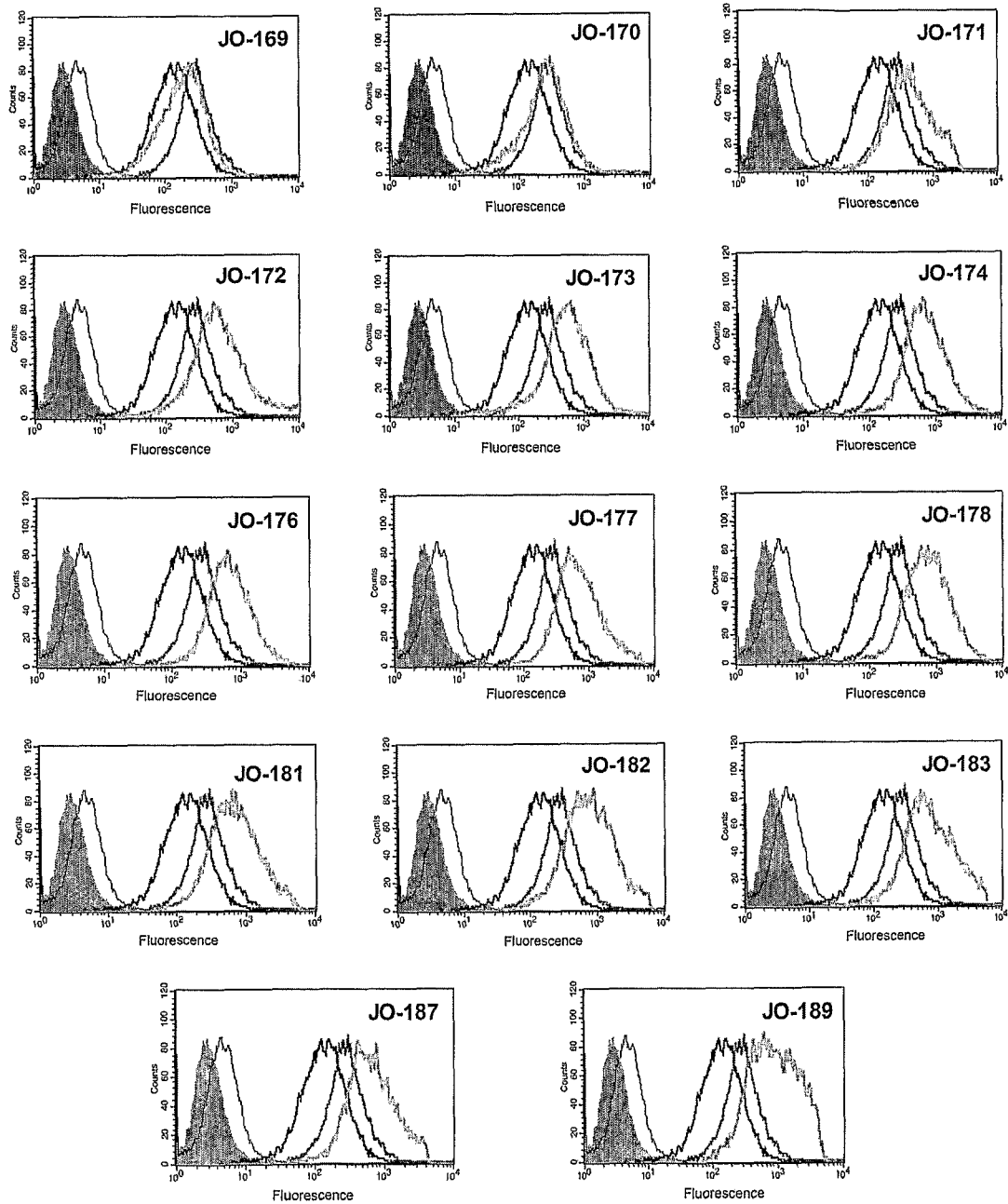
82/134

[Fig. 9f]

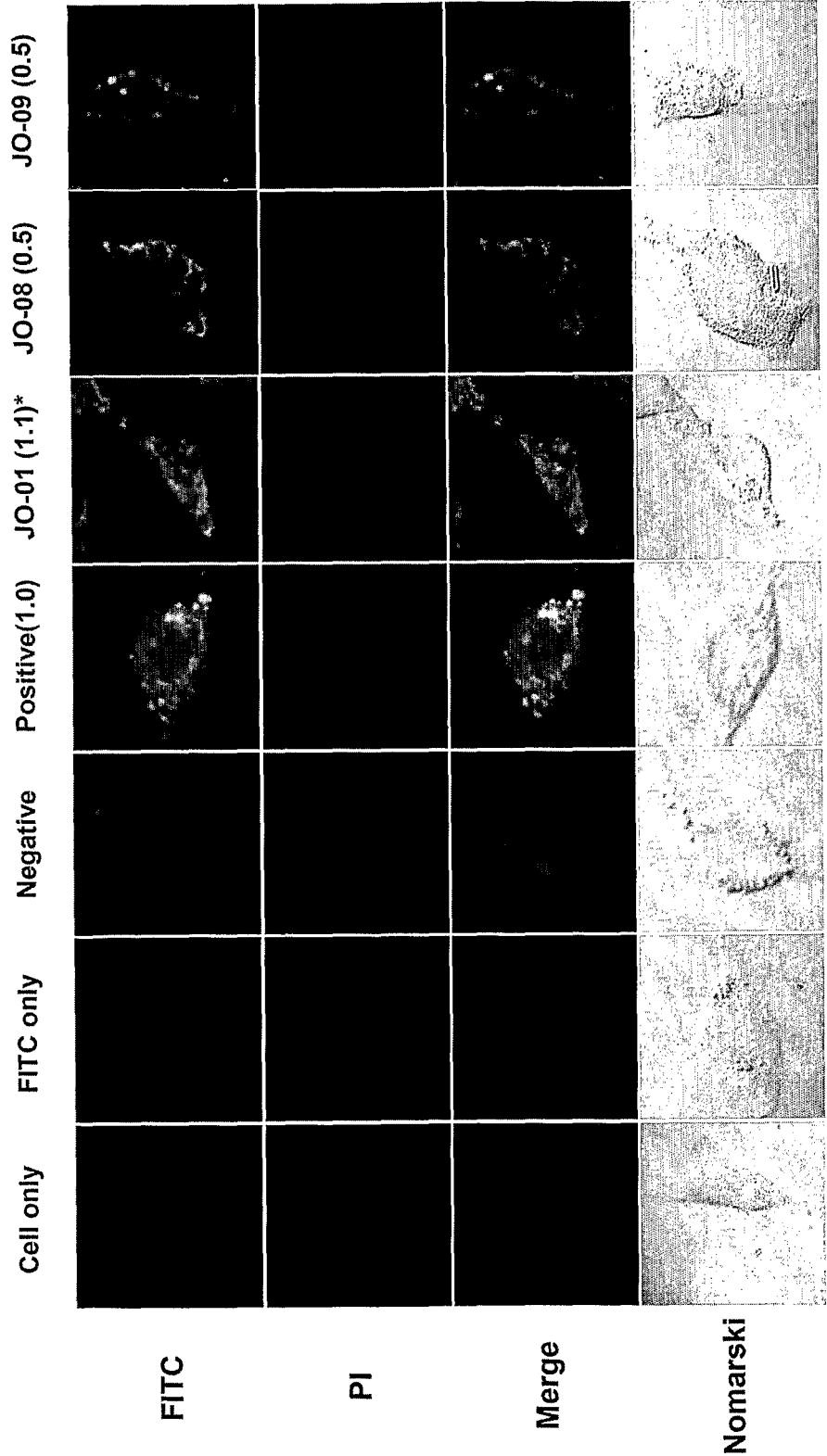


83/134

[Fig. 9g]

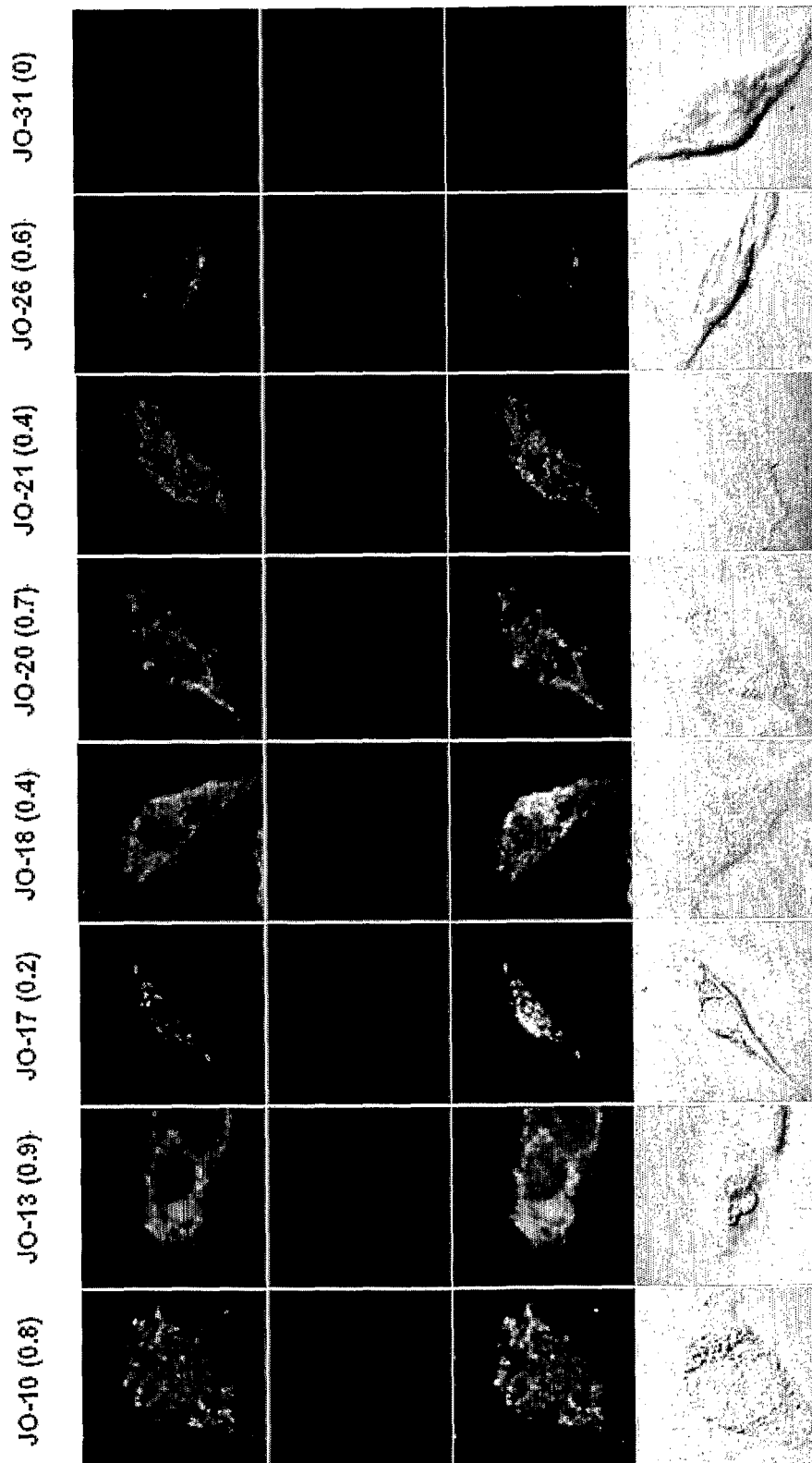


[Fig. 10a]

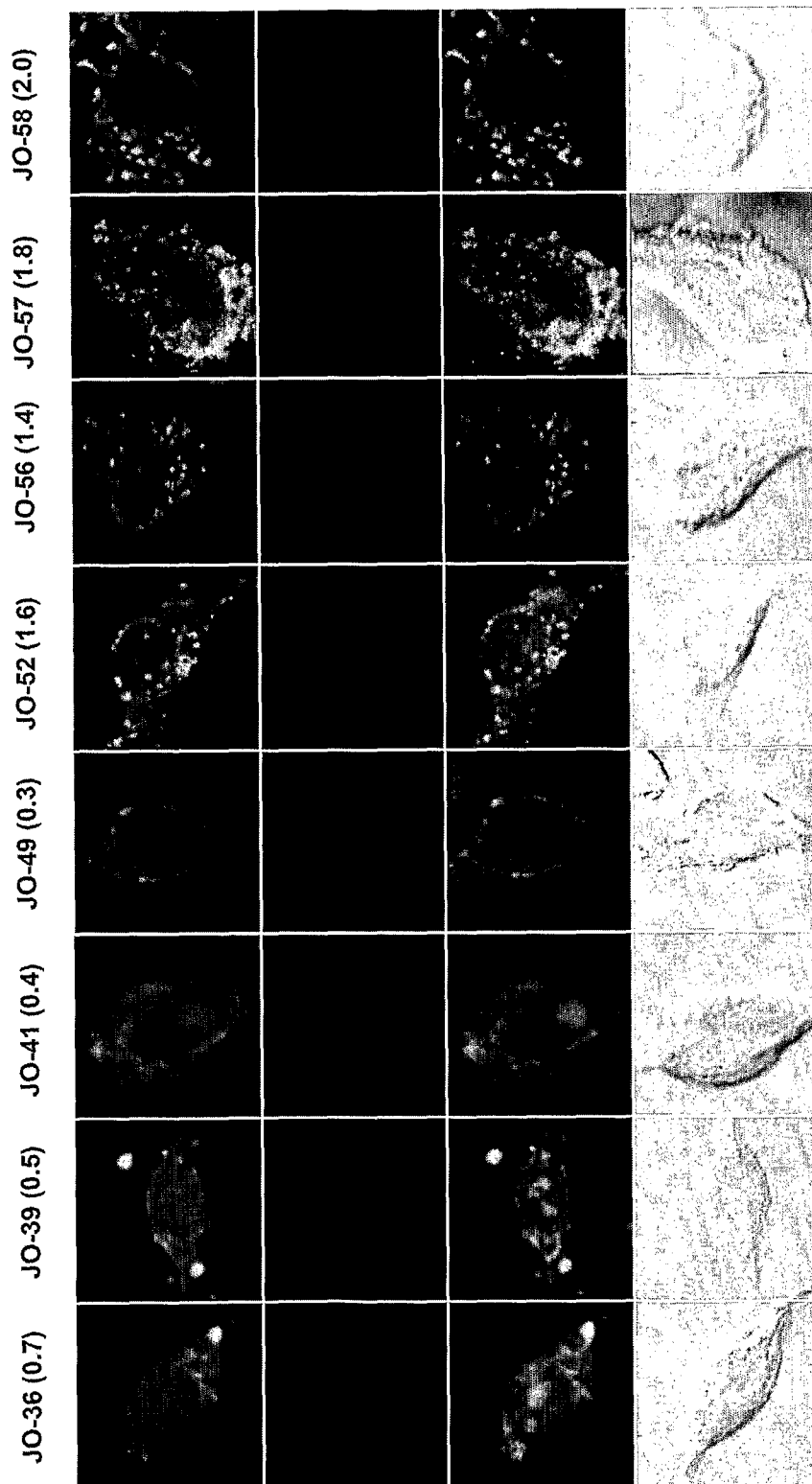


\* : Relative fold of cell permeability compared to the positive (MTD-derived from kEGF4)

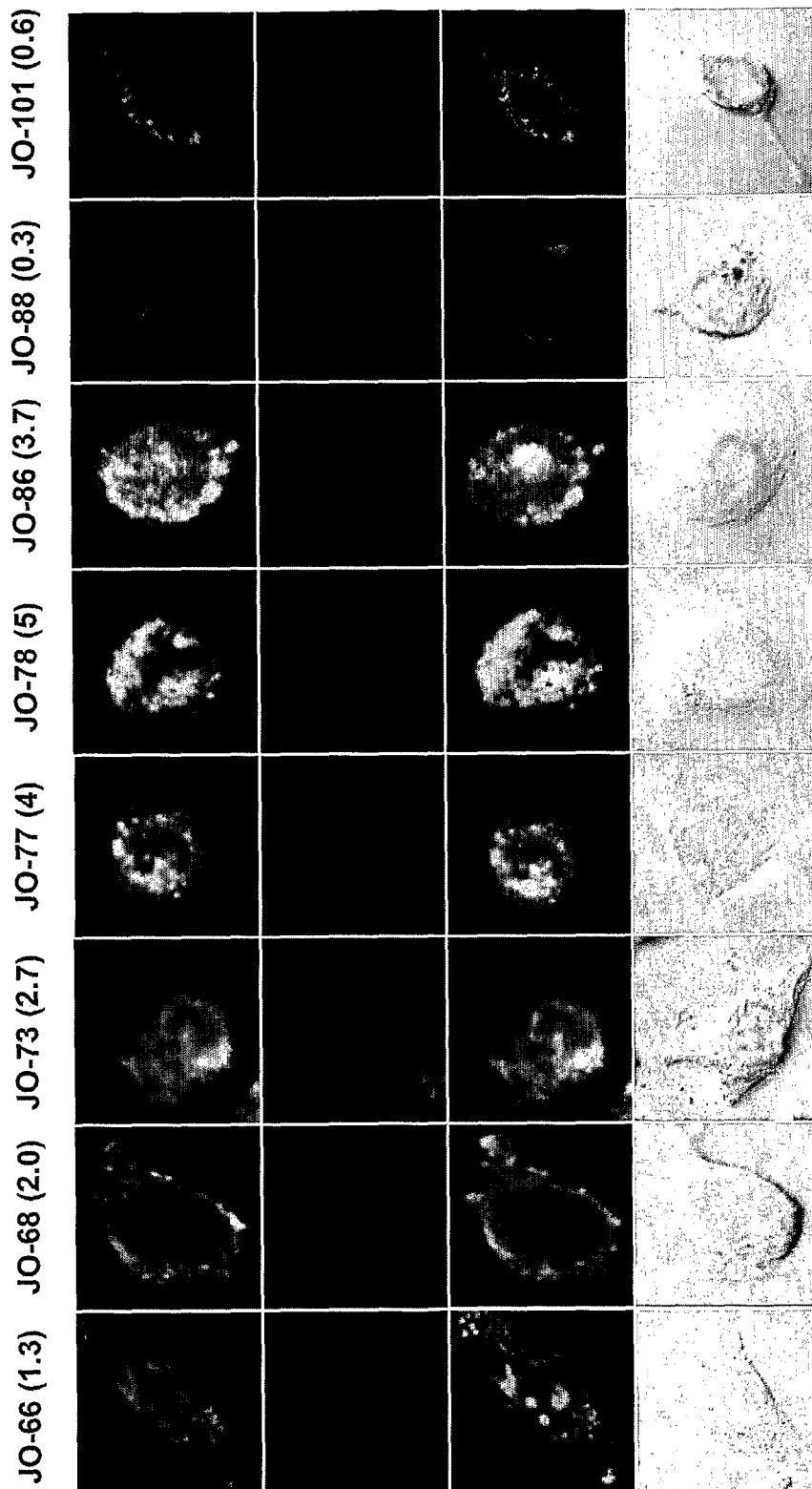
[Fig. 10b]



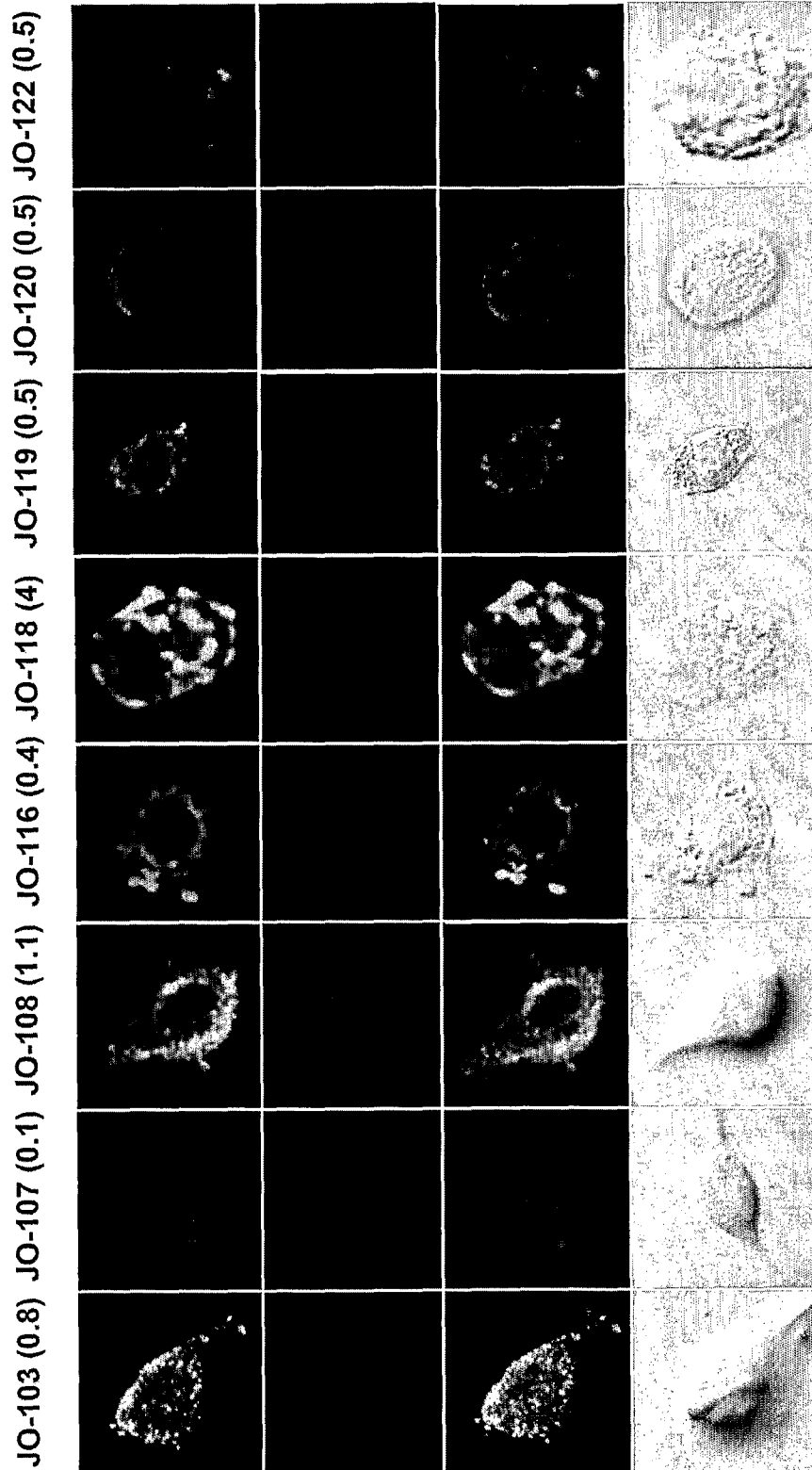
[Fig. 10c]



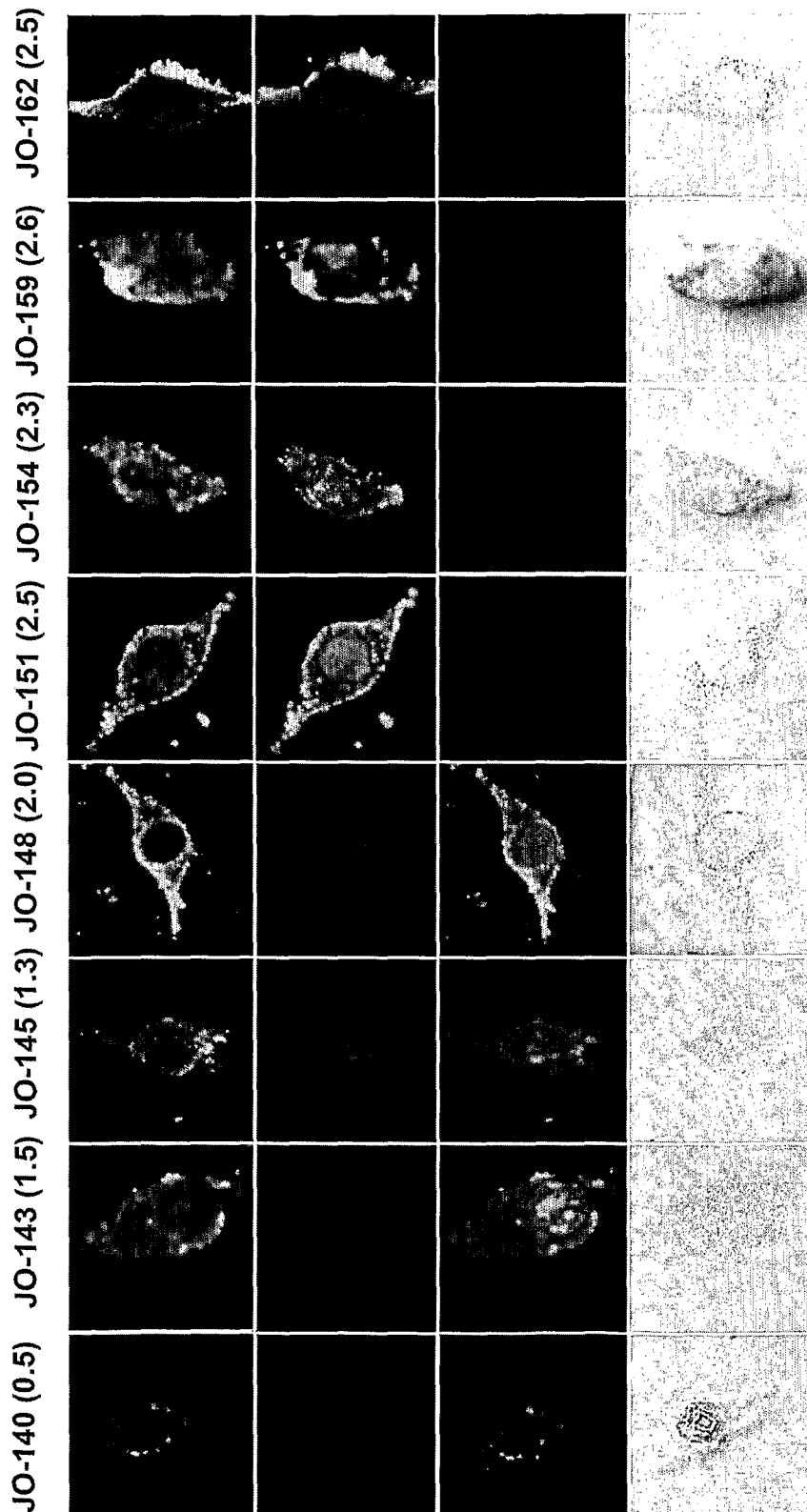
[Fig. 10d]



[Fig. 10e]

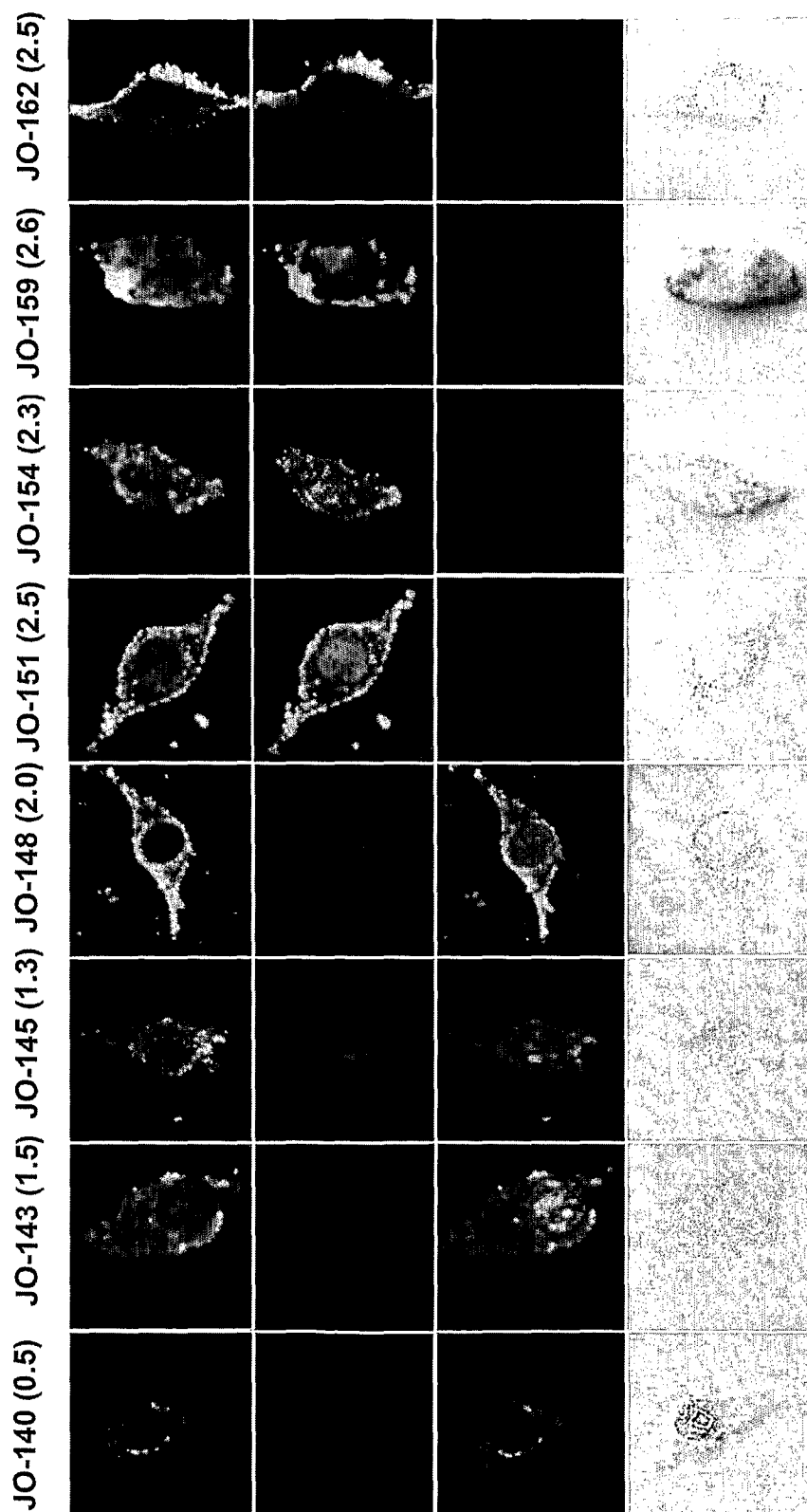


[Fig. 10f]

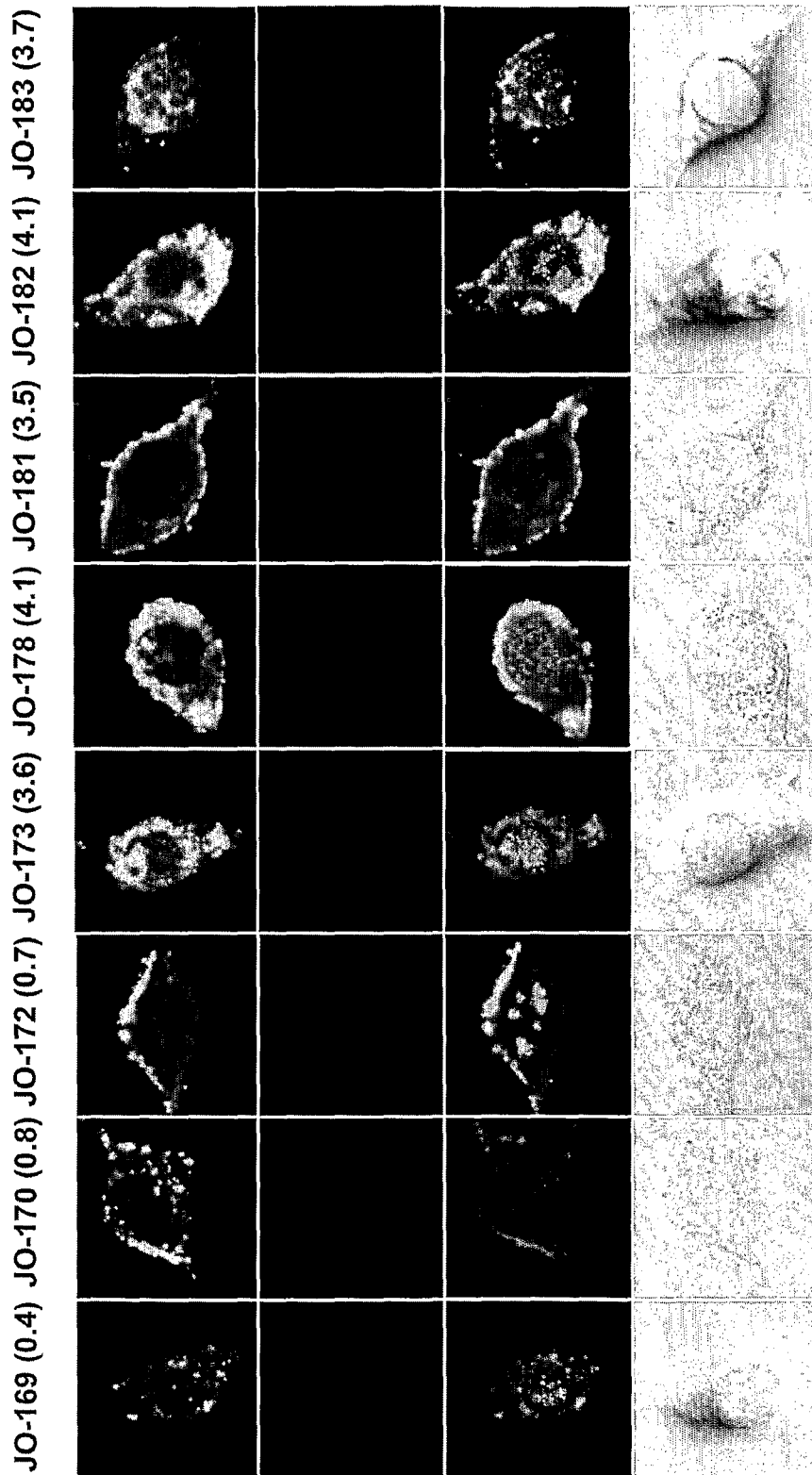


90/134

[Fig. 10g]



[Fig. 10h]



92/134

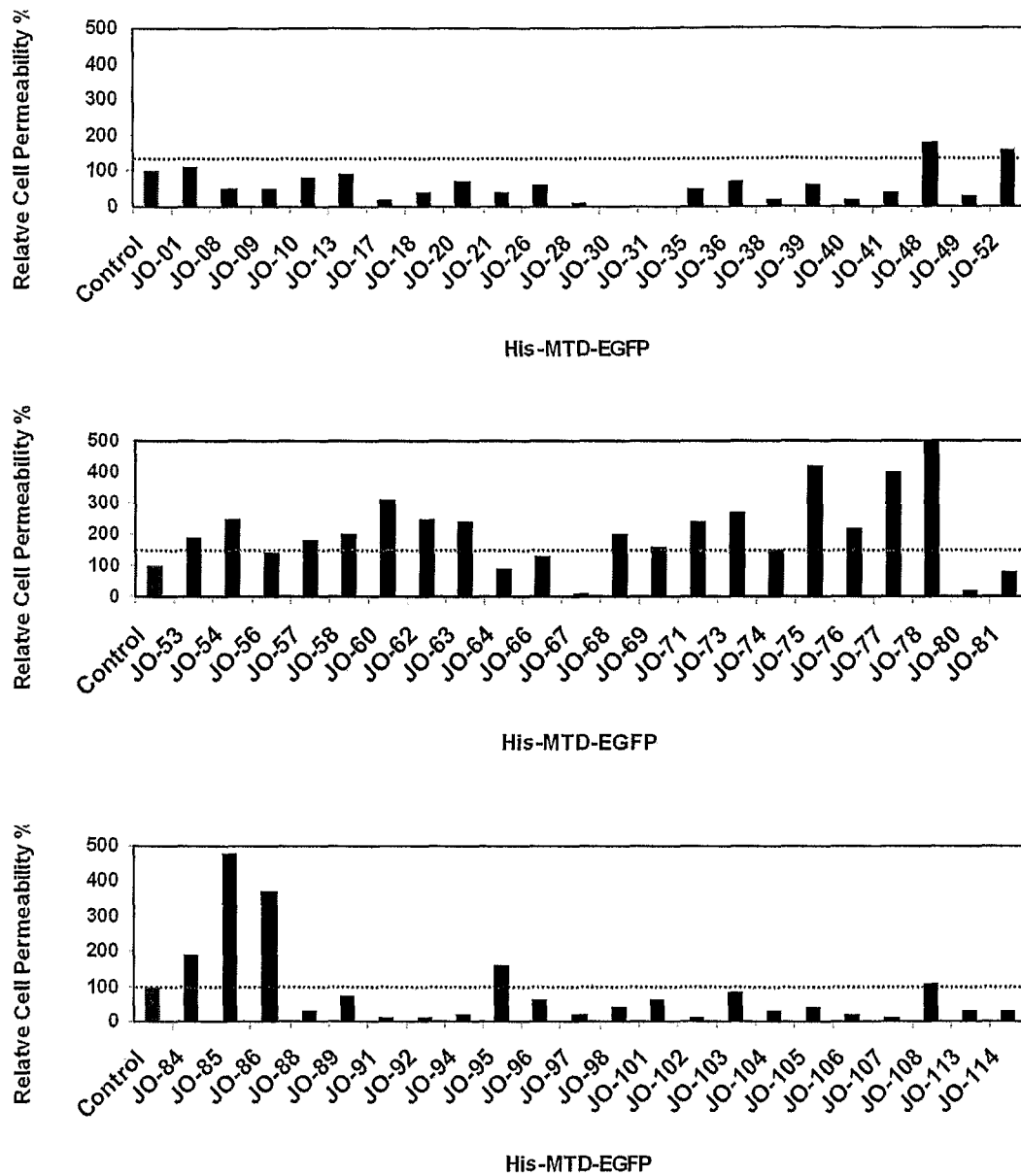
[Fig. 10i]

JO-187 (3.2)



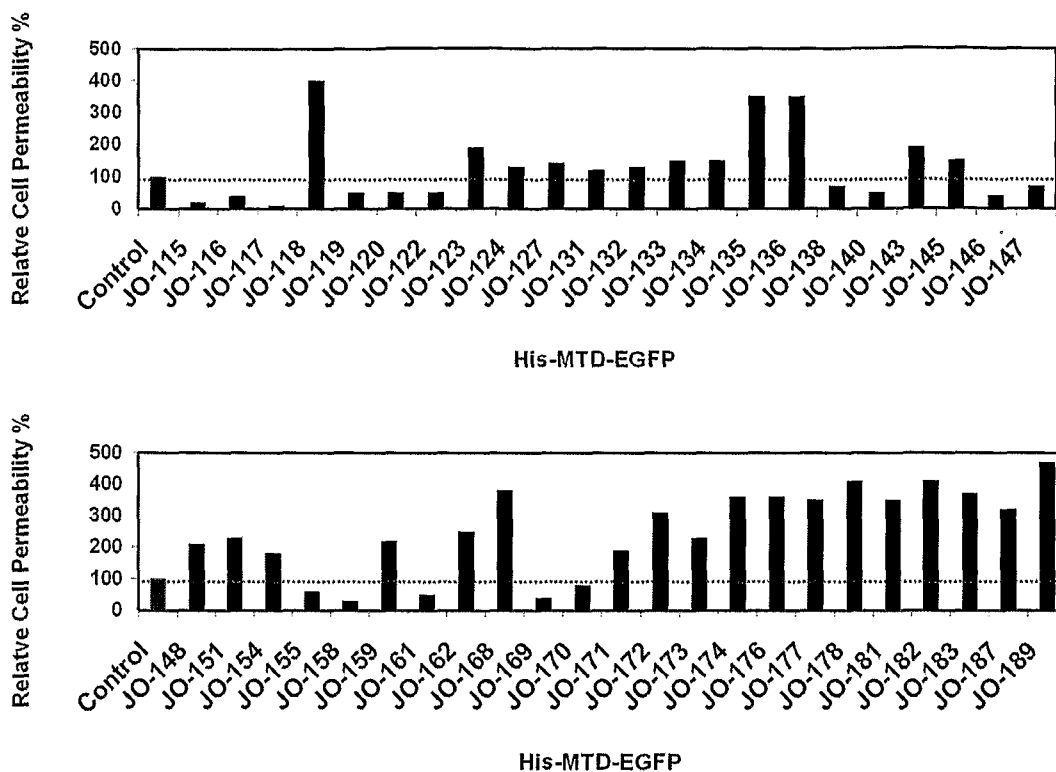
93/134

[Fig. 11a]



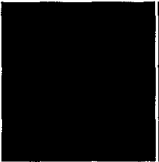
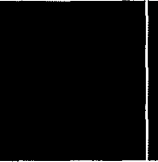
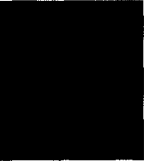
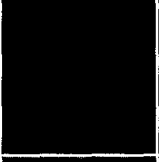
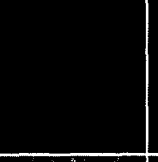
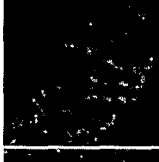
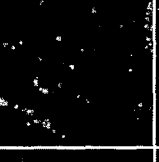
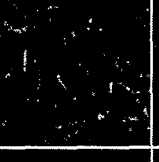
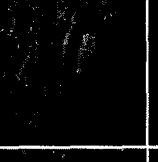
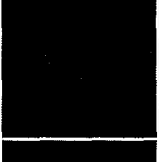
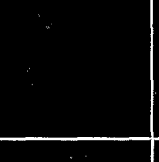
94/134

[Fig. 11b]



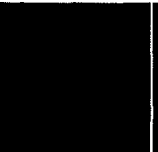
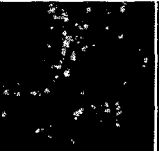
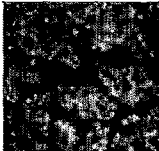
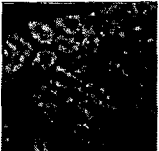
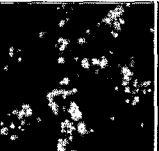
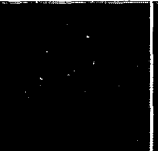
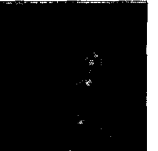
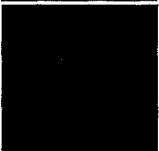
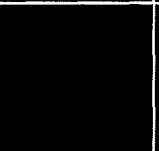
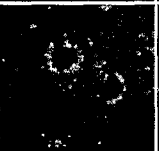
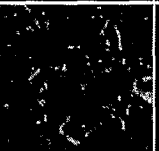
95/134

[Fig. 12a]

|                  | Liver                                                                               | Kidney                                                                              | Spleen                                                                              | Lung                                                                                | Heart                                                                                | Brain                                                                                 |
|------------------|-------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------|
| FITC only        |    |    |    |    |    |    |
| Negative control |    |    |    |    |    |    |
| Positive control |    |    |    |    |    |    |
| JO-01 (1.1)      |   |   |   |   |   |   |
| JO-08 (0.5)      |  |  |  |  |  |  |

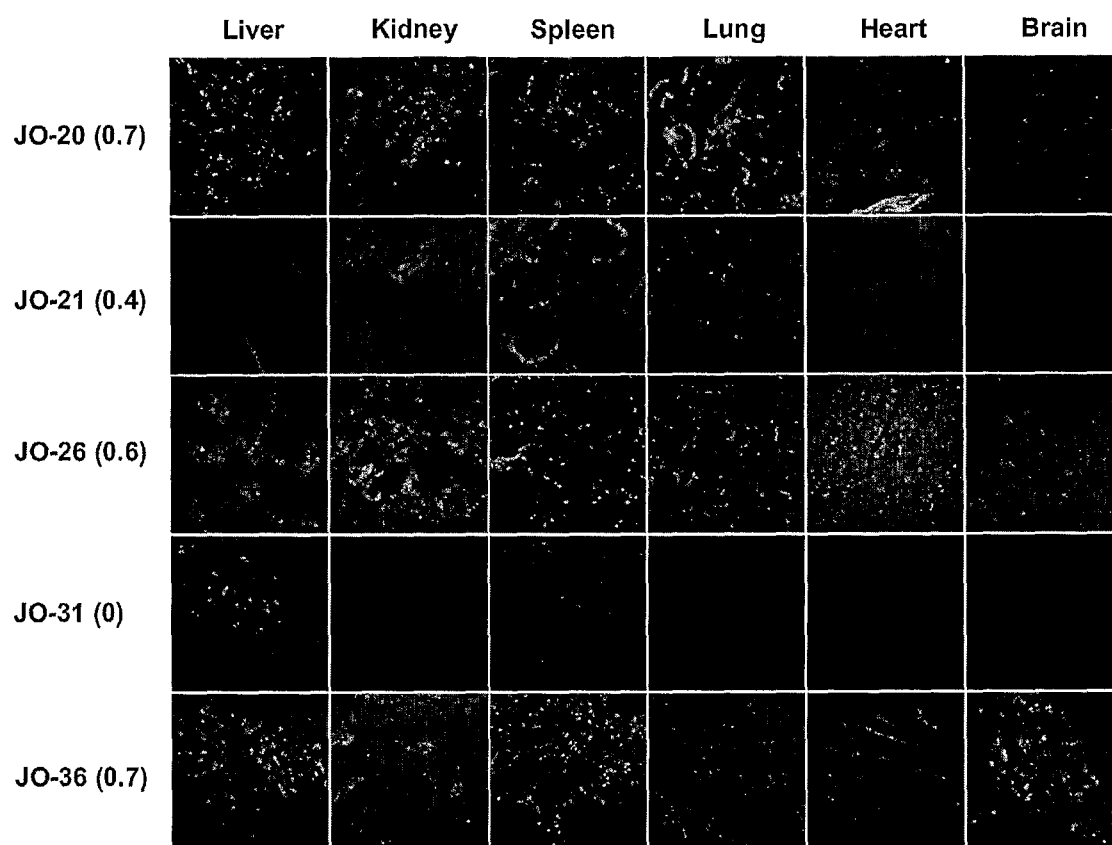
96/134

[Fig. 12b]

|             | Liver                                                                               | Kidney                                                                              | Spleen                                                                              | Lung                                                                                | Heart                                                                                | Brain                                                                                 |
|-------------|-------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------|
| JO-09 (0.5) |    |    |    |    |    |    |
| JO-10 (0.8) |    |    |    |    |    |    |
| JO-13 (0.9) |    |    |    |    |    |    |
| JO-17 (0.2) |   |   |   |   |   |   |
| JO-18 (0.4) |  |  |  |  |  |  |

97/134

[Fig. 12c]



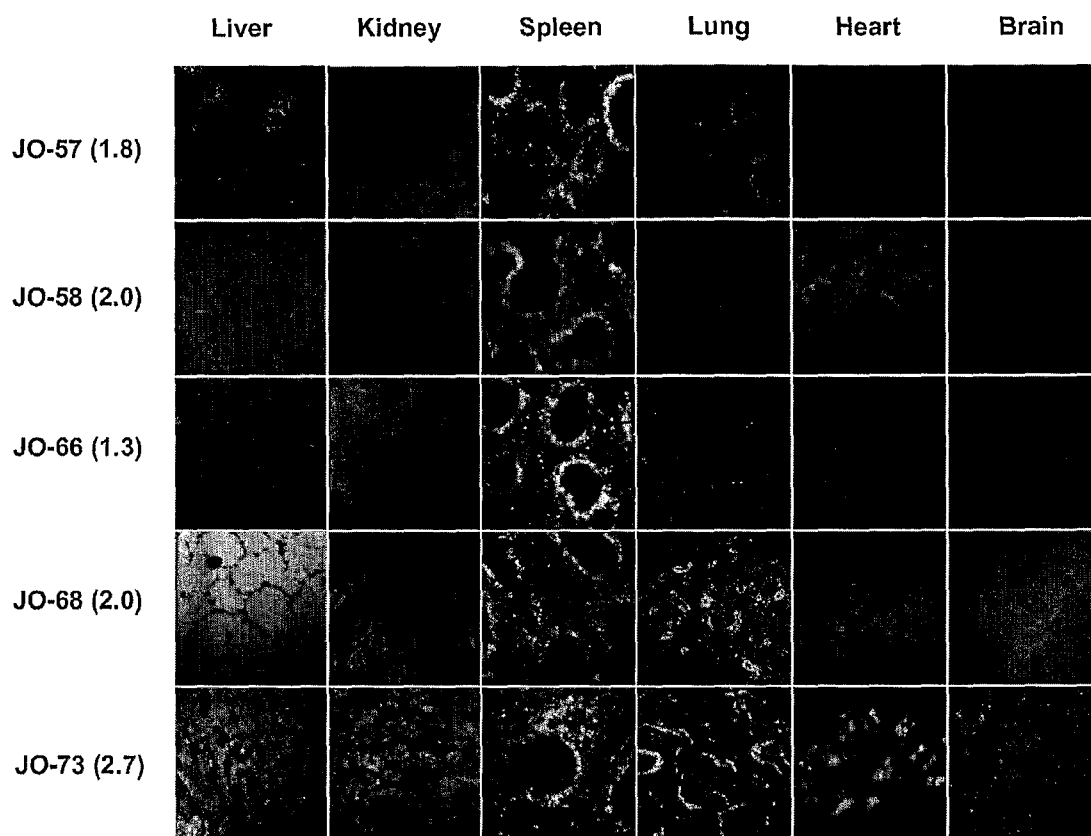
98/134

[Fig. 12d]

|             | Liver | Kidney | Spleen | Lung | Heart | Brain |
|-------------|-------|--------|--------|------|-------|-------|
| JO-39 (0.5) |       |        |        |      |       |       |
| JO-41 (0.4) |       |        |        |      |       |       |
| JO-49 (0.3) |       |        |        |      |       |       |
| JO-52 (1.6) |       |        |        |      |       |       |
| JO-56 (1.4) |       |        |        |      |       |       |

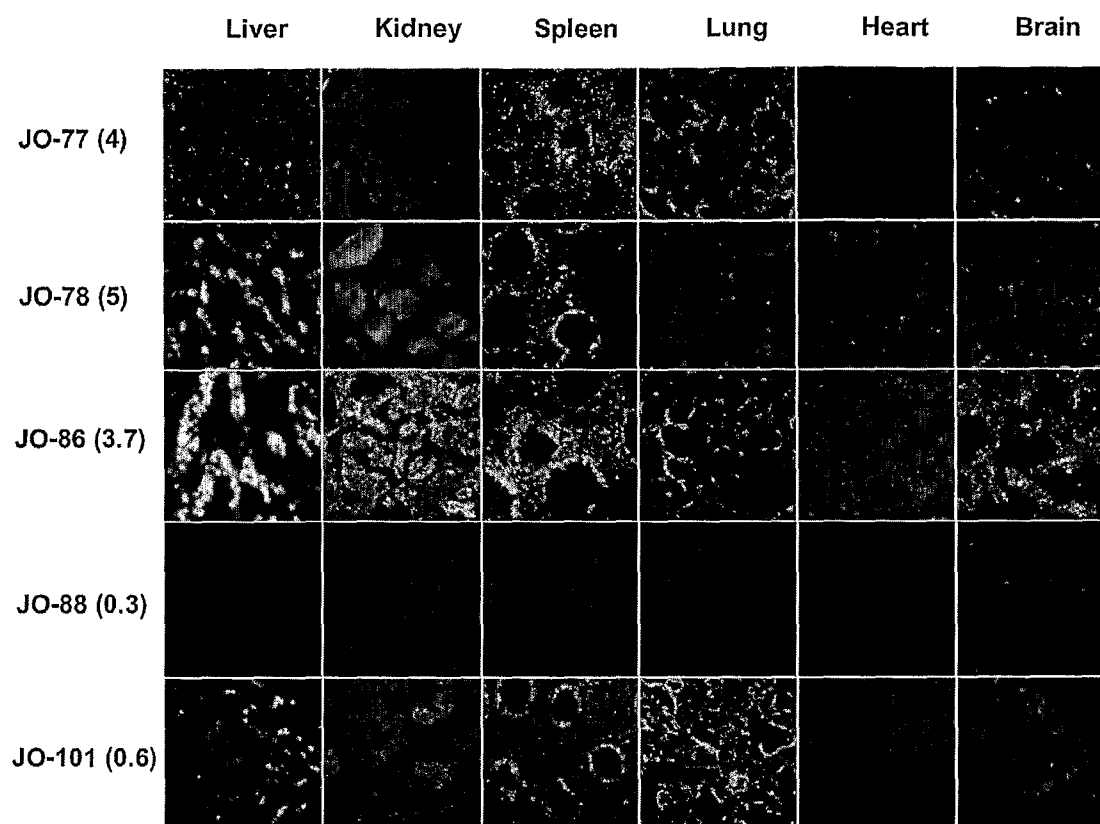
99/134

[Fig. 12e]



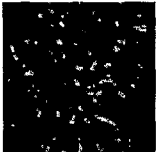
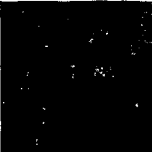
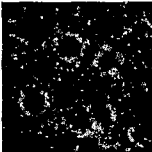
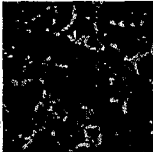
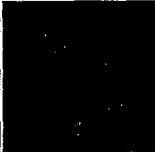
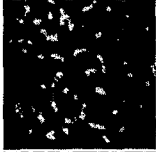
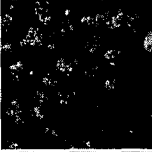
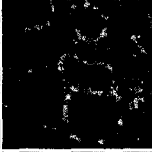
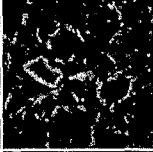
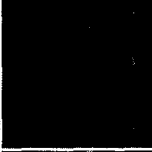
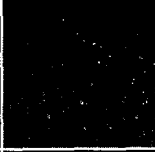
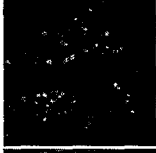
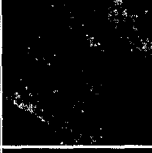
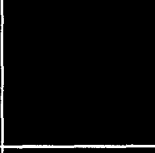
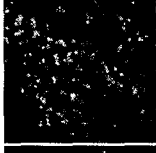
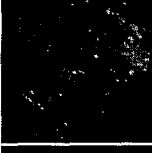
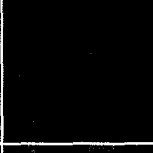
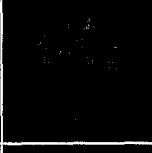
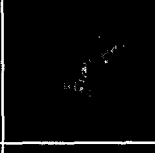
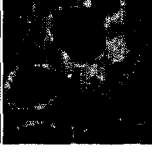
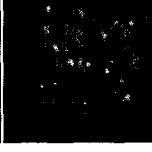
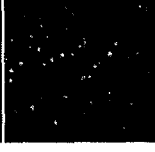
100/134

[Fig. 12f]



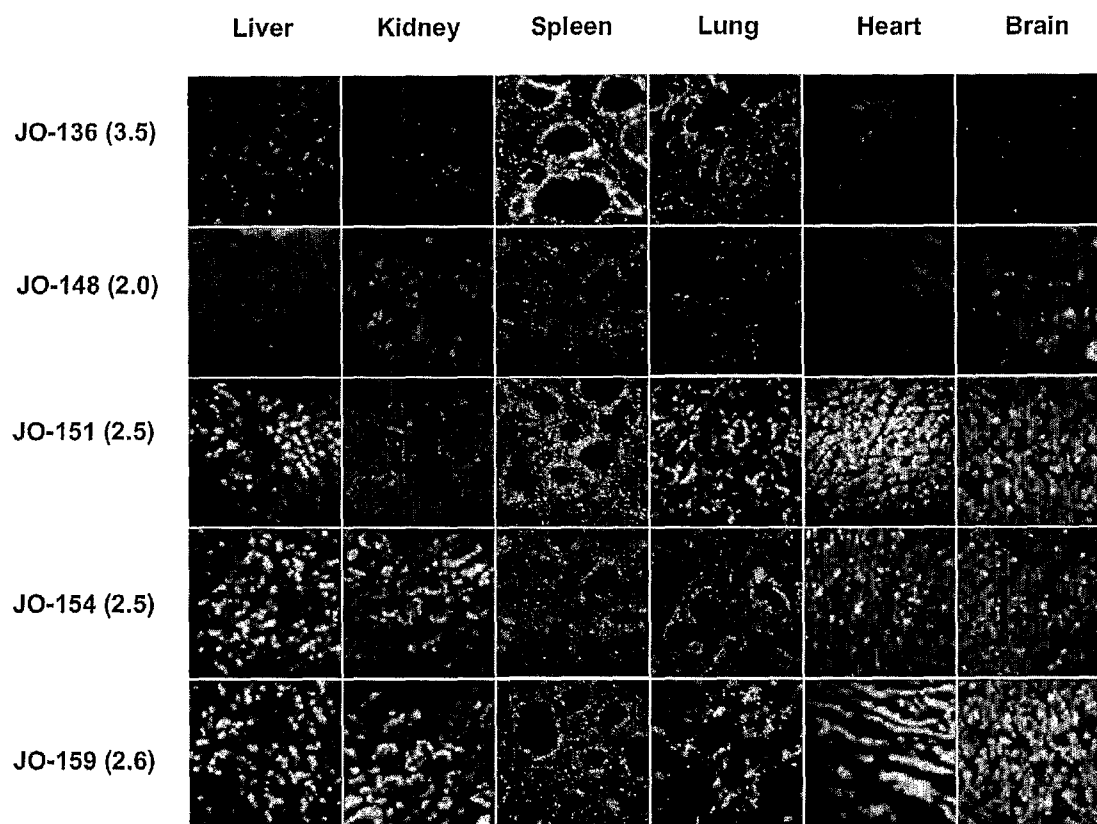
101/134

[Fig. 12g]

|              | Liver                                                                               | Kidney                                                                              | Spleen                                                                              | Lung                                                                                | Heart                                                                                | Brain                                                                                 |
|--------------|-------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------|
| JO-103 (0.8) |    |    |    |    |    |    |
| JO-108 (1.1) |    |    |    |    |    |    |
| JO-132 (1.3) |    |    |    |    |    |    |
| JO-133 (1.5) |   |   |   |   |   |   |
| JO-134 (1.5) |  |  |  |  |  |  |

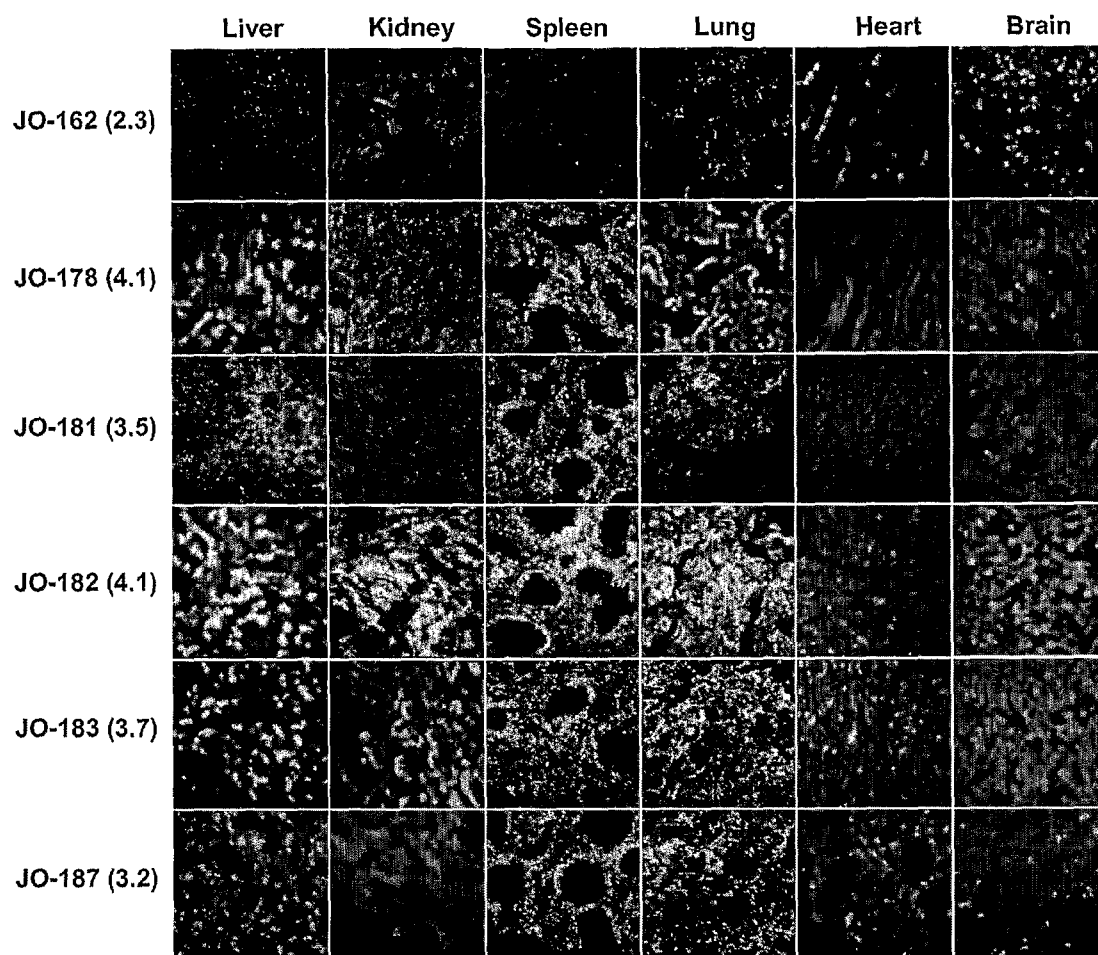
102/134

[Fig. 12h]



103/134

[Fig. 12i]



[Fig. 13a]

| ID    | Origin                         | Protein                                                         | MTD sequence          | A/a # | Induction | Purification | Cell permeability |
|-------|--------------------------------|-----------------------------------------------------------------|-----------------------|-------|-----------|--------------|-------------------|
| JO-01 | <i>Streptomyces coelicolor</i> | CAC04038 putative NLP/P60-family secreted protein               | AVVVC AIVLAA<br>P     | 12    | ++++      | ++++         | 1.1               |
| JO-02 | <i>Homosapiens</i>             | NP_057021 phosphatidylinositol glycan, class T precursor        | PLALLVLLLLG<br>P      | 12    | -         | -            | ND                |
| JO-03 | <i>Homo sapiens</i>            | NP_072171 chorionic somatomammotropin hormone 2 isoform 3       | LLLAFAALLCLP          | 11    | -         | -            | ND                |
| JO-04 | <i>Homo sapiens</i>            | NP_932156 nudix -type motif 9 isoform a                         | LLGALAAVLLA<br>LA     | 13    | -         | -            | ND                |
| JO-05 | <i>Homo sapiens</i>            | NP_057327 NAD(P)H:quinone oxidoreductase type 3, polypeptide A2 | PVLLALGVGLV<br>LLGLAV | 17    | -         | -            | ND                |
| JO-06 | <i>Streptomyces coelicolor</i> | CAD55300 putative secreted protein.                             | AAAAVLLAA             | 9     | -         | -            | ND                |
| JO-07 | <i>Streptomyces coelicolor</i> | NP_629514 secreted protein                                      | IVVAVVVI              | 8     | -         | -            | ND                |
| JO-08 | <i>Streptomyces coelicolor</i> | CAB57190 putative secreted chitin binding protein               | AVLAPVVAV             | 9     | ++++      | ++++         | 0.5               |

[Fig. 13a] continued

|       |                                |                                                         |                                    |    |      |      |     |
|-------|--------------------------------|---------------------------------------------------------|------------------------------------|----|------|------|-----|
| JO-09 | <i>Streptomyces coelicolor</i> | CAB51015 putative secreted protein                      | LAVCGLP <sup>+</sup> VVAL<br>LA    | 13 | ++++ | ++++ | 0.5 |
| JO-10 | <i>Streptomyces coelicolor</i> | NP_625021 glycosyl hydrolase (secreted protein)         | LGGA <sup>+</sup> VVAAPV<br>AAAVAP | 17 | ++++ | ++++ | 0.8 |
| JO-11 | <i>Streptomyces coelicolor</i> | NP_630686 secreted protein                              | LLVLAVLLAV<br>LP                   | 13 | -    | -    | ND  |
| JO-12 | <i>Homo sapiens</i>            | NP_057329 dehydrogenase/reductase (SDR family) member 8 | LLLLLLPLLIV                        | 12 | -    | -    | ND  |
| JO-13 | <i>Streptomyces coelicolor</i> | NP_639877 putative secreted protein                     | LAAAALAVLPL                        | 11 | ++++ | ++++ | 0.9 |
| JO-14 | <i>Homo sapiens</i>            | NP_699201 protease inhibitor 16 precursor               | FLMLLLPLL<br>LVA                   | 14 | -    | -    | ND  |
| JO-15 | <i>Streptomyces coelicolor</i> | NP_639871 putative secreted protein                     | AAAAAALGLAA<br>AVPA                | 15 | -    | -    | ND  |
| JO-16 | <i>Neisseria meningitidis</i>  | CAB85250 putative secreted protein                      | LLLAALLLAF<br>AV                   | 14 | -    | -    | ND  |
| JO-17 | <i>Streptomyces coelicolor</i> | NP_626397 small secreted hydrophilic protein            | ALAAVVLPLGI<br>AA                  | 14 | ++++ | ++++ | 0.2 |

106/134

[Fig. 13a] continued

|       |                                |                                                   |                     |    |      |      |     |
|-------|--------------------------------|---------------------------------------------------|---------------------|----|------|------|-----|
| JO-18 | <i>Streptomyces coelicolor</i> | CAB57190 putative secreted chitin binding protein | AALIGAVLAPV<br>VAV  | 14 | ++++ | ++++ | 0.4 |
| JO-19 | <i>Streptomyces coelicolor</i> | NP_626007 secreted cellulose-binding protein      | AAGIAVAIAAIV<br>PLA | 15 | -    | -    | ND  |

107/134

[Fig. 13b]

| ID    | Origin                         | Protein                            | MTD sequence         | A/a # | Induction | Purification | Cell permeability |
|-------|--------------------------------|------------------------------------|----------------------|-------|-----------|--------------|-------------------|
| JO-20 | <i>Streptomyces coelicolor</i> | NP_625632 secreted protein         | IAVAIAAIVPLA         | 12    | ++++      | ++++         | 0.7               |
| JO-21 | <i>Mycobacterium leprae</i>    | CAC31790 putative secreted protein | VAMAAAAVLA<br>APALA  | 15    | ++++      | ++++         | 0.4               |
| JO-22 | <i>Streptomyces coelicolor</i> | CAB38593 putative secreted protein | AALALGVAAAP<br>AAAPA | 16    | -         | -            | ND                |
| JO-23 | <i>Streptomyces coelicolor</i> | NP_630266 secreted Protein         | LAVLVLLVLLP          | 11    | -         | -            | ND                |
| JO-24 | <i>Streptomyces coelicolor</i> | NP_630165 secreted Protein         | VVAVLAPVL            | 9     | -         | -            | ND                |
| JO-25 | <i>Streptomyces coelicolor</i> | NC_003888 secreted Protein         | AALLPLLLLLP          | 12    | -         | -            | ND                |
| JO-26 | <i>Streptomyces coelicolor</i> | NP_627363 secreted Protein         | PAAVAALLVI           | 10    | ++++      | ++++         | 0.6               |
| JO-27 | <i>Streptomyces coelicolor</i> | NP_631288 secreted Protein         | LLIAALLP             | 8     | -         | -            | ND                |

[Fig. 13b] continued

|       |                                |                                                                   |                  |    |      |      |     |
|-------|--------------------------------|-------------------------------------------------------------------|------------------|----|------|------|-----|
| JO-28 | <i>Streptomyces coelicolor</i> | NP_630325 secreted Protein                                        | AAVVLPLAAA<br>IP | 12 | -    | ++   | 0.1 |
| JO-29 | <i>Streptomyces coelicolor</i> | NP_631289 secreted Protein                                        | AAAAAALLVP       | 10 | -    | -    | ND  |
| JO-30 | <i>Streptomyces coelicolor</i> | CAB51015 Putative secreted Protein                                | LPVVALLA         | 8  | ++++ | ++   | ND  |
| JO-31 | <i>Streptomyces coelicolor</i> | NP_629515 chitinase C (secreted Protein)                          | AAALAAPLALP      | 11 | -    | ++++ | ND  |
| JO-32 | <i>Homo sapiens</i>            | NP_940995 C1q and tumor necrosis factor related Protein1 isoform1 | LLLALLAA         | 9  | -    | -    | ND  |
| JO-33 | <i>Mycobacterium bovis</i>     | NP_854150 POSSIBLE CONSERVED SECRETED PROTEIN                     | AVAVVALL         | 8  | -    | -    | ND  |
| JO-34 | <i>Streptomyces coelicolor</i> | NP_630361 Probable secreted Protein                               | LLLIVLLIVP       | 11 | -    | -    | ND  |
| JO-35 | <i>Bacillus subtilis</i>       | P39790 Extracellular metalloProtease Precursor                    | LALAAAVVP        | 9  | ++++ | ++++ | 0.5 |
| JO-36 | <i>Streptomyces coelicolor</i> | CAA19252 Putative lipoprotein                                     | PAALALLVA        | 10 | ++++ | ++++ | 0.7 |

109/134

[Fig. 13b] continued

|       |                                |                                  |                      |    |     |     |     |
|-------|--------------------------------|----------------------------------|----------------------|----|-----|-----|-----|
| JO-37 | <i>Streptomyces coelicolor</i> | NP_625685 large secreted Protein | IVALLVPLVLA<br>IAAVL | 17 | -   | -   | ND  |
| JO-38 | <i>Streptomyces coelicolor</i> | NP_625685 large secreted Protein | IVALLVPT             | 8  | +++ | +++ | 0.2 |
| JO-39 | <i>Streptomyces coelicolor</i> | NP_625685 large secreted Protein | PLVLAIAAVL           | 10 | +++ | +++ | 0.6 |

110/134

[Fig. 13c]

| ID    | Origin                         | Protein                                                 | MTD sequence      | A/a # | Induction | Purification | Cell permeability |
|-------|--------------------------------|---------------------------------------------------------|-------------------|-------|-----------|--------------|-------------------|
| JO-40 | <i>Homo sapiens</i>            | NP_808800 golgi PhosPhoProtein 2                        | PLVLAALVA         | 9     | ++++      | +++          | 0.2               |
| JO-41 | <i>Streptomyces coelicolor</i> | NP_626993 secreted Protein                              | AAALLAVA          | 8     | ++++      | +++          | 0.4               |
| JO-42 | <i>Homo sapiens</i>            | NP_004863 thymic dendritic cell-derived factor 1        | PLLLALA           | 8     | -         | -            | ND                |
| JO-43 | <i>Streptomyces coelicolor</i> | NP_631398 secreted Protein                              | ALALVVA           | 7     | -         | -            | ND                |
| JO-44 | <i>Streptomyces coelicolor</i> | NP_627373 Penicillin-binding Protein (secreted Protein) | VAAVVVAA          | 8     | -         | -            | ND                |
| JO-45 | <i>Homo sapiens</i>            | NP_056226 sulfatase modifying factor 2                  | PLLPLLLV          | 9     | -         | -            | ND                |
| JO-46 | <i>Mycobacterium bovis</i>     | NP_854998 Conserved hypothetical secreted protein       | VVLVVVLPLAV<br>LA | 13    | -         | -            | ND                |
| JO-47 | <i>Streptomyces coelicolor</i> | NP_627512 secreted Protein                              | AAAVPVLVAA        | 10    | -         | -            | ND                |

[Fig. 13c] continued

|       |                                |                                                                                            |                  |    |      |      |      |     |
|-------|--------------------------------|--------------------------------------------------------------------------------------------|------------------|----|------|------|------|-----|
| JO-48 | <i>Homo sapiens</i>            | NP_110448 phospholipase A2, group XIA                                                      | PALLLLAAV<br>V   | 12 | ++   | ++   | ++   | 1.8 |
| JO-49 | <i>Homo sapiens</i>            | NP_003245 tissue inhibitor of metalloproteinase 1 precursor                                | PLAIIIIIIAP      | 12 | ++++ | ++++ | ++++ | 0.3 |
| JO-50 | <i>Homo sapiens</i>            | NP_002978 small inducible cytokine A17 precursor                                           | PLLALVLLALI<br>A | 12 | -    | -    | -    | ND  |
| JO-51 | <i>Mus musculus</i>            | NP_001012495 stromal cell derived factor 1 isoform gamma precursor                         | VVAVLALVLAA<br>L | 9  | -    | -    | -    | ND  |
| JO-52 | <i>Homo sapiens</i>            | NP_775628 ficolin 3 isoform 2 precursor                                                    | PLIIIIIPAL       | 9  | ++++ | ++++ | ++++ | 1.6 |
| JO-53 | <i>Streptomyces coelicolor</i> | NP_624483 secreted protein                                                                 | LAAVAALAVV<br>VP | 12 | ++   | ++   | ++   | 1.9 |
| JO-54 | <i>Homo sapiens</i>            | NP_997465 HERV-FRD provirus ancestral Env polyprotein                                      | LLLLVLIIPLAA     | 12 | ++   | ++   | ++   | 2.5 |
| JO-55 | <i>Mycobacterium bovis</i>     | NP_854234 possible conserved secreted protein                                              | LAVVVAAV         | 9  | -    | -    | -    | ND  |
| JO-56 | <i>Homo sapiens</i>            | p23284 Peptidyl-prolyl cis-trans isomerase B precursor (PPIase) (Rotamase) (Cyclophilin B) | VLLAAALIA        | 9  | ++++ | ++++ | ++++ | 1.4 |

112/134

[Fig. 13c] continued

|       |                            |                                                                                           |            |    |      |     |     |
|-------|----------------------------|-------------------------------------------------------------------------------------------|------------|----|------|-----|-----|
| JO-57 | <i>Salmonella enterica</i> | CAD05047 hypothetical secreted protein                                                    | LIALLAAPLA | 10 | ++++ | +++ | 1.8 |
| JO-58 | <i>Homo sapiens</i>        | P05067 Amyloid beta A4 protein precursor (APP) (ABPP) (Alzheimer disease amyloid protein) | LALLLLAA   | 8  | ++++ | +++ | 2.0 |

113/134

[Fig. 13d]

| ID    | Origin                          | Protein                                          | MTD sequence    | A/a # | Induction | Purification | Cell permeability |
|-------|---------------------------------|--------------------------------------------------|-----------------|-------|-----------|--------------|-------------------|
| JO-59 | <i>Homo sapiens</i>             | NP_004878 small inducible cytokine B14 precursor | LLAAALLLLL<br>A | 12    | -         | -            | ND                |
| JO-60 | <i>Streptomyces coelicolor</i>  | NP_626589 secreted protein                       | VIALIVVA        | 10    | ++        | ++           | 3.1               |
| JO-61 | <i>Streptomyces coelicolor</i>  | NP_626589 secreted protein                       | VVLVVA AVLAL    | 11    | -         | -            | ND                |
| JO-62 | <i>Mycobacterium bovis</i>      | NP_856548 SOLUBLE SECRETED ANTIGEN MPB53         | VAVAI VVL       | 9     | -         | ++           | 2.5               |
| JO-63 | <i>Streptomyces coelicolor</i>  | NP_629854 secreted protein                       | PLIVVVA AVV     | 13    | -         | ++           | 2.4               |
| JO-64 | <i>Escherichia coli</i>         | AAB59058 lambda receptor protein                 | PLAVAVA A       | 11    | -         | ++           | 0.9               |
| JO-65 | <i>Streptomyces avermitilis</i> | NP_825185 NLP/T60-family secreted protein        | AAIALVAVVL      | 10    | -         | -            | ND                |
| JO-66 | <i>Streptomyces coelicolor</i>  | NP_626568 secreted protein                       | AAALAAIAVI      | 10    | ++++      | ++++         | 1.3               |

114/134

[Fig. 13d] continued

|       |                                |                                                                           |             |    |      |      |     |
|-------|--------------------------------|---------------------------------------------------------------------------|-------------|----|------|------|-----|
| JO-67 | <i>Streptomyces coelicolor</i> | NP_626568 secreted protein                                                | AAAPAVAA    | 8  | ++++ | ++++ | 0.1 |
| JO-68 | <i>Streptomyces coelicolor</i> | NP_625639 secreted protein                                                | LLAALP      | 7  | ++++ | ++++ | 2.0 |
| JO-69 | <i>Mycobacterium leprae</i>    | CAC32053 putative secreted protein                                        | ALLAVVAA    | 8  | ++   | ++++ | 1.6 |
| JO-70 | <i>Streptomyces coelicolor</i> | NP_630954 secreted protein                                                | AVVVVLPILL  | 10 | -    | -    | ND  |
| JO-71 | <i>Mus musculus</i>            | P97300 Neuroligin precursor (Stromal cell-derived receptor 1) (SDR-1)     | ALALLLVLP   | 9  | ++   | ++++ | 2.4 |
| JO-72 | <i>Streptomyces coelicolor</i> | NP_624483 secreted protein                                                | LVVLLAALLVL | 11 | -    | -    | ND  |
| JO-73 | <i>Drosophila melanogaster</i> | AAA41949 Rat parotid gland acidic proline-rich protein mRNA, complete CDS | PVLLLLA     | 8  | ++++ | ++++ | 2.7 |
| JO-74 | <i>Streptomyces coelicolor</i> | AAA17887 Drosophila melanogaster spatula (spz) gene                       | ALAVVAAP    | 8  | ++   | ++++ | 1.5 |
| JO-75 | <i>Streptomyces coelicolor</i> | NP_627867 conserved secreted protein                                      | VIVALLAV    | 8  | -    | ++   | 4.2 |

115/134

[Fig. 13d] continued

|       |                                |                                                                |            |   |                |                |     |
|-------|--------------------------------|----------------------------------------------------------------|------------|---|----------------|----------------|-----|
| JO-76 | <i>Homo sapiens</i>            | NP_631283 secreted protein                                     | ALVLLPLAP  | 8 | ++<br>++<br>++ | ++<br>++<br>++ | 2.2 |
| JO-77 | <i>Streptomyces coelicolor</i> | NP_003231 endometrial bleeding associated factor preproprotein | AVALLLILAV | 9 | ++<br>++<br>++ | ++<br>++<br>++ | 4.0 |

116/134

[Fig. 13e]

| ID    | Origin                         | Protein                                      | MTD sequence  | A/a # | Induction | Purification | Cell permeability |
|-------|--------------------------------|----------------------------------------------|---------------|-------|-----------|--------------|-------------------|
| JO-78 | <i>Xenopus laevis</i>          | CAB76313 putative secreted protein           | VLLAVIP       | 7     | +++       | +++          | 5.0               |
| JO-79 | <i>Xenopus laevis</i>          | P07198 Xenopsin precursor                    | LIVAAVVVVAVLI | 13    | -         | -            | ND                |
| JO-80 | <i>Streptomyces coelicolor</i> | NP_631293 secreted protein                   | AVVVAAP       | 7     | -         | ++           | 0.2               |
| JO-81 | <i>Streptomyces coelicolor</i> | NP_626373 secreted protein                   | LAAVLLIP      | 9     | +++       | +++          | 0.8               |
| JO-82 | <i>Streptomyces coelicolor</i> | NP_624952 secreted cellulose-binding protein | LLLLLLAVVP    | 10    | -         | -            | ND                |
| JO-83 | <i>Homo sapiens</i>            | NP_009104 protease, serine, 23 precursor     | AVAVVVAV<br>A | 11    | -         | -            | ND                |
| JO-84 | <i>Streptomyces coelicolor</i> | AAK63068 phytoalexin protein PcoF precursor  | LVAALLAVL     | 9     | +++       | +++          | 1.9               |
| JO-85 | <i>Streptomyces coelicolor</i> | NC_003903 plasmid SCP1, complete sequence,   | LIAAAAALLLA   | 11    | +++       | +++          | 4.8               |

117/134

[Fig. 13e] continued

|       |                                |                                                                     |                   |    |      |      |     |
|-------|--------------------------------|---------------------------------------------------------------------|-------------------|----|------|------|-----|
| JO-86 | <i>Mycobacterium bovis</i>     | NP_629842 peptide transport system secreted peptide binding protein | LAVLAAAP          | 8  | ++++ | +++  | 3.7 |
| JO-87 | <i>Mycobacterium bovis</i>     | NP_854067 Possible secreted protein.                                | VVVLLVLLALV<br>VV | 13 | -    | -    | ND  |
| JO-88 | <i>Streptomyces coelicolor</i> | NP_627802 secreted protein                                          | VVIAVVP           | 7  | -    | ++   | 0.3 |
| JO-89 | <i>Streptomyces coelicolor</i> | NP_627802 secreted protein                                          | VLLVLLALV         | 9  | -    | ++   | 0.7 |
| JO-90 | <i>Streptomyces coelicolor</i> | NP_627802 secreted protein                                          | VLLVLLALV         | 9  | ++   | -    | ND  |
| JO-91 | <i>Streptomyces coelicolor</i> | NP_625203 secreted protein                                          | PVLPVPAVP         | 8  | ++++ | ++++ | 0.1 |
| JO-92 | <i>Streptomyces coelicolor</i> | NP_630960 secreted protein                                          | PALALALA          | 8  | ++++ | ++++ | 0.1 |
| JO-93 | <i>Streptomyces coelicolor</i> | NP_630670 secreted protein                                          | AAAAPALA          | 8  | -    | -    | ND  |
| JO-94 | <i>Streptomyces coelicolor</i> | NP_630493 secreted protein                                          | IVLPVLAAP         | 9  | ++++ | ++++ | 0.2 |

118/134

[Fig. 13e] continued

|       |                                |                                    |             |    |      |      |     |
|-------|--------------------------------|------------------------------------|-------------|----|------|------|-----|
| JO-95 | <i>Mycobacterium leprae</i>    | CAC29994 putative secreted protein | LVLLLLPLLI  | 10 | ++   | ++   | 1.6 |
| JO-96 | <i>Streptomyces coelicolor</i> | NP_624483 secreted protein         | LAAVAPALAVV | 11 | ++++ | ++++ | 0.6 |

[Fig. 13f]

| ID     | Origin                         | Protein                                                                          | MTD sequence     | A/a # | Induction | Purification | Cell permeability |
|--------|--------------------------------|----------------------------------------------------------------------------------|------------------|-------|-----------|--------------|-------------------|
| JO-97  | <i>Homo sapiens</i>            | NP_037375 secretogranin III                                                      | ILVLVLPI         | 8     | ++        | ++           | 0.2               |
| JO-98  | <i>Homo sapiens</i>            | NP_009199 V-set and immunoglobulin domain containing 4                           | ILLPLLLP         | 9     | ++++      | ++++         | 0.4               |
| JO-99  | <i>Streptomyces coelicolor</i> | NP_733650 secreted hydrolase                                                     | IAPAVVAALP       | 10    | ++        | -            | ND                |
| JO-100 | <i>Homo sapiens</i>            | NP_057540 transmembrane protein 9                                                | LLLVAVVPLLV<br>P | 12    | -         | -            | ND                |
| JO-101 | <i>Theileria annulata</i>      | CAI74362 hypothetical protein                                                    | LILLPLPI         | 9     | ++++      | ++           | 0.6               |
| JO-102 | <i>Streptomyces coelicolor</i> | NP_630671 secreted protein                                                       | AVLAAPAVLV       | 10    | ++++      | ++++         | 0.1               |
| JO-103 | <i>Homo sapiens</i>            | NP_065695 TMEM9 domain family, member B                                          | LALPVLLA         | 9     | ++++      | ++++         | 0.8               |
| JO-104 | <i>Canis lupus familiaris</i>  | P06908 Pulmonary surfactant-associated protein A precursor (SP-A) (PSP-A) (PSAP) | LALALL           | 7     | ++++      | ++++         | 0.3               |

[Fig. 13f] continued

|        |                                |                                                                 |                    |    |      |      |     |
|--------|--------------------------------|-----------------------------------------------------------------|--------------------|----|------|------|-----|
| JO-105 | <i>Streptomyces coelicolor</i> | NP_639721 putative secreted protein                             | VAVPLL VVA         | 9  | ++   | ++++ | 0.4 |
| JO-106 | <i>Mycobacterium bovis</i>     | NP_854954 CONSERVED PROBABLE SECRETED PROTEIN                   | AVAVAPVAAA<br>A    | 11 | ++++ | ++++ | 0.2 |
| JO-107 | <i>Streptomyces coelicolor</i> | NP_627759 secreted protein                                      | AAAVVAAVPA<br>A    | 11 | ++   | ++++ | 0.1 |
| JO-108 | <i>Homo sapiens</i>            | NP_003842 cellular repressor of E1A-stimulated genes            | ALLAALLAP          | 9  | ++++ | ++++ | 1.1 |
| JO-109 | <i>Homo sapiens</i>            | NP_003842 cellular repressor of E1A-stimulated genes            | LLALLVP            | 7  | -    | -    | ND  |
| JO-110 | <i>Homo sapiens</i>            | NP_003842 cellular repressor of E1A-stimulated genes            | ALLAALLALLV<br>LLV | 14 | -    | -    | ND  |
| JO-111 | <i>Homo sapiens</i>            | NP_000589 insulin-like growth factor binding protein 3 (IGFBP3) | AAALPLLVLPP        | 11 | -    | -    | ND  |
| JO-112 | <i>Streptomyces coelicolor</i> | CAB59459 putative secreted protein                              | AAAVPAALAP         | 10 | -    | -    | ND  |
| JO-113 | <i>Streptomyces coelicolor</i> | NP_628917 secreted protein                                      | AALAVAALAA         | 10 | +++  | ++++ | 0.3 |

121/134

[Fig. 13f] continued

|        |                                |                            |          |   |      |     |
|--------|--------------------------------|----------------------------|----------|---|------|-----|
| JO-114 | <i>Streptomyces coelicolor</i> | NP_624695 secreted protein | AVLAAAVP | 8 | ++++ | 0.3 |
| JO-115 | <i>Streptomyces coelicolor</i> | NP_624695 secreted protein | VAALPAPA | 8 | ++++ | 0.2 |

122/134

[Fig. 13g]

| ID     | Origin                         | Protein                                                  | MTD sequence      | A/a # | Induction | Purification | Cell permeability |
|--------|--------------------------------|----------------------------------------------------------|-------------------|-------|-----------|--------------|-------------------|
| JO-116 | <i>Streptomyces coelicolor</i> | NP_624791 secreted protein                               | ALALAVPAVLP       | 11    | +++       | +++          | 0.4               |
| JO-117 | <i>Streptomyces coelicolor</i> | CAB45579 putative secreted protein                       | AALLPAAVAVP       | 11    | +++       | +++          | 0.1               |
| JO-118 | <i>Streptomyces coelicolor</i> | NP_627066 secreted protein                               | AVVVALAP          | 8     | -         | -            | 4.0               |
| JO-119 | <i>Streptomyces coelicolor</i> | NP_630174 secreted substrate-binding protein             | AAVALPAAAA<br>LLA | 14    | +++       | +++          | 0.5               |
| JO-120 | <i>Homo sapiens</i>            | P06727 Apolipoprotein A-IV precursor (Apo-AIV) (ApoA-IV) | AVVLPLALVAV<br>AP | 13    | ++        | ++           | 0.5               |
| JO-121 | <i>Mus musculus</i>            | Q62087 Serum paraoxonase/lactonase 3                     | LVALPLLP          | 8     | -         | -            | ND                |
| JO-122 | <i>Streptomyces coelicolor</i> | NP_627123 probable secreted penicillin-binding protein   | VVVPLLLIVP        | 10    | -         | ++           | 0.5               |
| JO-123 | <i>Mycobacterium leprae</i>    | CAC30224 putative secreted hydrolase                     | LAVVLAVP          | 8     | ++        | ++           | 1.9               |

123/134

[Fig. 13g] continued

|        |                                |                                                                           |                |    |     |     |     |
|--------|--------------------------------|---------------------------------------------------------------------------|----------------|----|-----|-----|-----|
| JO-124 | <i>Plasmodium cynomolgi</i>    | OZZQAM circumsporozoite protein precursor                                 | LLAVPILLVP     | 10 | +++ | +++ | 1.8 |
| JO-125 | <i>Homo sapiens</i>            | Q15166 Serum paraoxonase/lactonase 3                                      | LVALVLLP       | 8  | -   | -   | ND  |
| JO-126 | <i>Homo sapiens</i>            | NP_060220 all-trans-13,14-dihydroretinol saturase                         | LVLALLVLLA VLP | 14 | -   | -   | ND  |
| JO-127 | <i>Salmonella typhi</i>        | AL627273 Salmonella enterica serovar Typhi (Salmonella typhi) strain CT18 | LLAPVVVALVILP  | 12 | ++  | -   | 1.4 |
| JO-128 | <i>Streptomyces coelicolor</i> | NP_625987 secreted protein                                                | VLAVALVPVLL LP | 13 | ++  | -   | ND  |
| JO-129 | <i>Streptomyces coelicolor</i> | CAB45474 putative secreted protein                                        | VVIAVVPVVV     | 10 | -   | -   | ND  |
| JO-130 | <i>Streptomyces coelicolor</i> | CAB45474 putative secreted protein                                        | LLVLLALVVVP    | 11 | -   | -   | ND  |
| JO-131 | <i>Streptomyces coelicolor</i> | CAB36605 putative secreted protein                                        | VLLALPVVAAP    | 11 | ++  | ++  | 1.2 |
| JO-132 | <i>Streptomyces coelicolor</i> | NP_628377 NLP/P60-family secreted protein                                 | AVVVPAILVLAAP  | 12 | ++  | ++  | 1.3 |

124/134

[Fig. 13g] continued

|        |                                |                                    |                         |    |      |     |     |
|--------|--------------------------------|------------------------------------|-------------------------|----|------|-----|-----|
| JO-133 | <i>Streptomyces coelicolor</i> | CAB59594 putative secreted protein | AVLVPAALVP <sup>a</sup> | 11 | ++++ | +++ | 1.5 |
| JO-134 | <i>Streptomyces coelicolor</i> | NP_624974 secreted protein         | VVAALPLVLP              | 10 | ++   | ++  | 1.5 |

[Fig. 13h]

| ID     | Origin                         | Protein                                                                                       | MTD sequence   | A/a # | Induction | Purification | Cell permeability |
|--------|--------------------------------|-----------------------------------------------------------------------------------------------|----------------|-------|-----------|--------------|-------------------|
| JO-135 | <i>Streptomyces coelicolor</i> | NP_733682 secreted ATP/GTP binding protein                                                    | AAVALFAAAP     | 10    | ++++      | ++++         | 3.5               |
| JO-136 | <i>Homo sapiens</i>            | P27169 Serum paraoxonase/arylesterase 1 (PON 1) (Serum arylalkylphosphatase 1) (A-esterase 1) | LIALPLLP       | 8     | ++++      | ++++         | 3.5               |
| JO-137 | <i>Homo sapiens</i>            | P52430 Serum paraoxonase/arylesterase 1 (PON 1) (Serum arylalkylphosphatase 1) (A-esterase 1) | LLALPLVLVLA LP | 13    | -         | -            | ND                |
| JO-138 | <i>Streptomyces coelicolor</i> | NP_626569 secreted protein                                                                    | IVPLLLAAP      | 9     | ++++      | ++++         | 0.7               |
| JO-139 | <i>Homo sapiens</i>            | NP_940995 C1q and tumor necrosis factor related protein 1 isoform 1                           | LLALPLLLAP     | 10    | -         | -            | ND                |
| JO-140 | <i>Streptomyces coelicolor</i> | NP_626174 large secreted protein                                                              | LAALPVAAVP     | 10    | ++++      | ++++         | 0.5               |
| JO-141 | <i>Neisseria meningitidis</i>  | CAB38860 putative protein-export integral membrane protein                                    | ALAVIVLVLL     | 10    | -         | -            | ND                |

126/134

[Fig. 13h] continued

|        |                                |                                                     |                   |    |      |      |     |
|--------|--------------------------------|-----------------------------------------------------|-------------------|----|------|------|-----|
| JO-142 | <i>Homo sapiens</i>            | NP_001009551<br>cornichon-like isoform 2            | LALLPAALI         | 10 | -    | -    | ND  |
| JO-143 | <i>Streptomyces coelicolor</i> | NP_626808 secreted protein                          | ALLPLAVVLP        | 11 | ++   | ++   | 1.9 |
| JO-144 | <i>Streptomyces coelicolor</i> | NP_639798 putative secreted protein                 | AIAVPVLAAP        | 10 | -    | -    | ND  |
| JO-145 | <i>Homo sapiens</i>            | NP_000492<br>elastin(supravalvular aortic stenosis) | AAAPVLLLLL        | 10 | ++++ | ++++ | 1.5 |
| JO-146 | <i>Streptomyces coelicolor</i> | NP_630680 secreted sugar binding protein            | AAAVAVLALAP       | 11 | ++++ | ++++ | 0.4 |
| JO-147 | <i>Streptomyces coelicolor</i> | CAB56129 putative secreted protein                  | AALAAALVVAAP      | 11 | ++   | ++   | 0.7 |
| JO-148 | <i>Streptomyces coelicolor</i> | NP_625109 secreted solute-binding lipoprotein       | AALAAVPLALA<br>P  | 12 | +++  | +++  | 2.1 |
| JO-149 | <i>Streptomyces coelicolor</i> | NP_733579 secreted sugar binding protein            | ALAVAAPALAL<br>LP | 13 | -    | -    | ND  |
| JO-150 | <i>Streptomyces coelicolor</i> | NP_630126 secreted chitinase (secreted protein)     | AALPAAAP          | 8  | -    | -    | ND  |

127/134

[Fig. 13h] continued

|        |                                |                                                 |            |    |     |     |     |
|--------|--------------------------------|-------------------------------------------------|------------|----|-----|-----|-----|
| JO-151 | <i>Streptomyces coelicolor</i> | NP_630126 secreted chitinase (secreted protein) | AAAPVAAVP  | 9  | +++ | +++ | 2.3 |
| JO-152 | <i>Homo sapiens</i>            | NP_872425 secretory protein LOC348174           | LLAVLLALLP | 10 | -   | -   | ND  |

[Fig. 13]

| ID     | Origin                            | Protein                                           | MTD sequence | A/a # | Induction | Purification | Cell permeability |
|--------|-----------------------------------|---------------------------------------------------|--------------|-------|-----------|--------------|-------------------|
| JO-153 | <i>Streptomyces coelicolor</i>    | NP_630107 secreted protein                        | VLALLVAVVP   | 10    | -         | -            | ND                |
| JO-154 | <i>Streptomyces coelicolor</i>    | NP_733688 peptide-binding transport protein       | ALVVPAAVP    | 9     | ++        | ++           | 1.8               |
| JO-155 | <i>Streptomyces coelicolor</i>    | NP_629904 secreted protein                        | AVVLPPLL     | 9     | ++        | ++           | 0.6               |
| JO-156 | <i>Mycobacterium tuberculosis</i> | YP_177852 MCE-FAMILY PROTEIN MCE3A                | AVIPVAVLVP   | 10    | -         | -            | ND                |
| JO-157 | <i>Streptomyces coelicolor</i>    | CAA19627 putative secreted solute binding protein | AAAVPAAVLAP  | 11    | -         | -            | ND                |
| JO-158 | <i>Streptomyces coelicolor</i>    | NP_639884 putative large secreted protein         | VAVPVVLAILP  | 11    | -         | ++           | 0.3               |
| JO-159 | <i>Bacillus subtilis</i>          | P24327 Foldase protein prsA precursor             | IAIAAIPAILAL | 12    | ++++      | ++++         | 2.2               |
| JO-160 | <i>Neisseria meningitidis</i>     | CAB84808 putative membrane lipoprotein            | ALAAPALAAP   | 10    | -         | -            | ND                |
| JO-161 | <i>Streptomyces coelicolor</i>    | NP_639883 putative large secreted protein         | AAIALVAPAL   | 10    | ++++      | ++++         | 0.5               |

[Fig. 13i] continued

|        |                                |                                           |                    |    |     |     |     |
|--------|--------------------------------|-------------------------------------------|--------------------|----|-----|-----|-----|
| JO-162 | <i>Streptomyces coelicolor</i> | NP_639883 putative large secreted protein | LAPAVAAAP          | 9  | +++ | +++ | 2.5 |
| JO-163 | <i>Streptomyces coelicolor</i> | NP_627362 secreted protein                | VAIVPAVVAIA<br>LII | 15 | -   | -   | ND  |
| JO-164 | <i>Streptomyces coelicolor</i> | NP_627362 secreted protein                | AVVAIALII          | 9  | -   | -   | ND  |
| JO-165 | <i>Streptomyces coelicolor</i> | NP_624625 secreted protein                | LAAVPAAP           | 9  | -   | -   | ND  |
| JO-166 | <i>Streptomyces coelicolor</i> | NP_624625 secreted protein                | AVAALPLAAP         | 10 | -   | -   | ND  |
| JO-167 | <i>Streptomyces coelicolor</i> | NP_624625 secreted protein                | LAAPAAAAP          | 9  | -   | -   | ND  |
| JO-168 | <i>Streptomyces coelicolor</i> | NP_626936 secreted protein                | LAAVVPVAAA<br>VP   | 12 | ++  | ++  | 3.8 |
| JO-169 | <i>Streptomyces coelicolor</i> | NP_626936 secreted protein                | VAAPAAAAP          | 9  | +++ | +++ | 0.4 |

130/134

[Fig. 13i] continued

|        |                                |                                          |             |    |      |     |     |
|--------|--------------------------------|------------------------------------------|-------------|----|------|-----|-----|
| JO-170 | <i>Streptomyces coelicolor</i> | NP_626936 secreted protein               | AVPVVPVL    | 8  | ++++ | +++ | 0.8 |
| JO-171 | <i>Homo sapiens</i>            | NP_085072 matrilin 2 isoform b precursor | LLLPVLLP    | 10 | ++   | -   | 1.9 |
| JO-172 | <i>Streptomyces coelicolor</i> | CAB94057 putative secreted protein       | ALALPALALAP | 11 | ++++ | +++ | 3.1 |

[Fig. 13]]

| ID     | Origin                         | Protein                                           | MTD sequence    | A/a # | Induction | Purification | Cell permeability |
|--------|--------------------------------|---------------------------------------------------|-----------------|-------|-----------|--------------|-------------------|
| JO-173 | <i>Streptomyces coelicolor</i> | NP_624384 secreted protein                        | AVIPILAVP       | 9     | +++       | +++          | 2.3               |
| JO-174 | <i>Streptomyces coelicolor</i> | NP_733505 large, multifunctional secreted protein | LILLPVAVLP      | 11    | ++        | ++           | 3.6               |
| JO-175 | <i>Streptomyces coelicolor</i> | CAB45630 putative secreted protein                | IVLAPVPAAA      | 10    | -         | -            | ND                |
| JO-176 | <i>Streptomyces coelicolor</i> | NP_627887 secreted protein                        | VVVVPVLAAA<br>A | 11    | -         | -            | 3.6               |
| JO-177 | <i>Bacillus</i>                | P06832 Bacillolysin precursor                     | LVAVAAP         | 7     | -         | -            | 3.5               |
| JO-178 | <i>Streptomyces coelicolor</i> | NP_625998 secreted hydrolase                      | LVLAAAPAAALP    | 10    | ++        | +++          | 4.7               |
| JO-179 | <i>Streptomyces coelicolor</i> | NP_625057 secreted protein                        | LIAPAAAVP       | 9     | -         | -            | ND                |
| JO-180 | <i>Homo sapiens</i>            | NP_443750 ADP-ribosyltransferase 5 precursor      | ALAALPIALP      | 10    | -         | -            | ND                |

[Fig. 13]] continued

|        |                                |                                                                          |             |    |      |      |     |
|--------|--------------------------------|--------------------------------------------------------------------------|-------------|----|------|------|-----|
| JO-181 | <i>Neisseria meningitidis</i>  | CAB84257 putative secreted protein                                       | AVLLLPAAA   | 9  | ++++ | ++++ | 3,5 |
| JO-182 | <i>Escherichia coli</i>        | P00634 Alkaline phosphatase precursor (APase)                            | IALALLPLL   | 9  | ++++ | ++++ | 4,1 |
| JO-183 | <i>Homo sapiens</i>            | NP000933 peptidylprolyl isomerase B precursor                            | VLLAAALIAP  | 10 | ++++ | ++++ | 3,7 |
| JO-184 | <i>Streptomyces coelicolor</i> | CAB71258 putative secreted protein,                                      | APAVLPVVVI  | 11 | -    | -    | ND  |
| JO-185 | <i>Mycobacterium leprae</i>    | CAC31847 possible secreted protein                                       | VVGLLVAAL   | 9  | -    | -    | ND  |
| JO-186 | <i>Streptomyces coelicolor</i> | NP_626948 secreted protein                                               | AAIAAAAPLAA | 11 | -    | -    | ND  |
| JO-187 | <i>Homo sapiens</i>            | NP_059120 cat eye syndrome critical region protein 1 isoform a precursor | LLLA VAP    | 7  | ++   | ++   | 3,2 |
| JO-188 | <i>Homo sapiens</i>            | NP_006519 tissue factor pathway inhibitor 2                              | LILLPLAAL   | 10 | -    | -    | 4,7 |

133/134

[Fig. 13]] continued

|        |                     |                                                                                                      |           |   |   |   |    |
|--------|---------------------|------------------------------------------------------------------------------------------------------|-----------|---|---|---|----|
| JO-189 | <i>Mus musculus</i> | P97299 Secreted frizzled-related protein 2 precursor (sFRP-2) (Secreted apoptosis-related protein 1) | ALLLVLA   | 8 | - | - | ND |
| JO-190 | <i>Homo sapiens</i> | NP_071447 tubulointerstitial nephritis antigen-like 1                                                | LLLLLLPLA | 9 | - | - | ND |
| JO-191 | <i>Homo sapiens</i> | NP_056322 epidermal growth factor-like protein 6 precursor                                           | LALPLLLP  | 8 | - | - | ND |

[Fig. 13k]

| ID     | Origin                         | Protein                                       | MTD sequence | A/a # | Induction | Purification | Cell permeability |
|--------|--------------------------------|-----------------------------------------------|--------------|-------|-----------|--------------|-------------------|
| JO-192 | <i>Streptomyces coelicolor</i> | NP_628035 secreted penicillin-binding protein | ELVLPILI     | 8     | -         | -            | ND                |
| JO-193 | <i>Homo sapiens</i>            | NP_683880 cathepsin H isoform b precursor     | LPLLLPAALV   | 9     | -         | -            | ND                |

\* ND: Recombinant His-MTD-EGFP proteins of novel MTDs could not be prepared overall due to failure either on construction of expression vector, inducible expression, purification or preparation in soluble form.

## INTERNATIONAL SEARCH REPORT

International application No.

**PCT/KR2008/000525**

### Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.b of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of :

a. type of material



a sequence listing



table(s) related to the sequence listing

b. format of material



on paper



in electronic form

c. time of filing/furnishing



contained in the international application as filed



filed together with the international application in electronic form



furnished subsequently to this Authority for the purposes of search

2. ☐ In addition, in the case that more than one version or copy of a sequence listing and/or table relating thereto has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that in the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

3. Additional comments:

**INTERNATIONAL SEARCH REPORT**

International application No.

**PCT/KR2008/000525****Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)**

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.: 21-25  
because they relate to subject matter not required to be searched by this Authority, namely:  
Claims 21-25 pertain to methods for treatment of the human or animal body by therapy as well as diagnostic methods, and thus relate to a subject matter which this International Searching Authority is not required, under Article 17(2)(a)(i) of the PCT and Rule 39.1(iv) of the Regulations under the PCT, to search.
2. ☒ Claims Nos.: 4-9  
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:  
A method of claims 4-9 is unclear because the meaning of wording 'identifying' and 'single hydrophobic region' are vague, and technical feature for 'optimizing and minimizing' step is not clearly defined.
3. ☐ Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

**Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)**

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

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

**Remark on Protest**

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.

## INTERNATIONAL SEARCH REPORT

International application No.  
**PCT/KR2008/000525****A. CLASSIFICATION OF SUBJECT MATTER****C07K 7/04(2006.01)i, C07K 7/06(2006.01)i, C07K 7/08(2006.01)i**

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

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)

IPC 8 C07K 7/04, C07K 7/06, C07K 7/08

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

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

NCBI database (SEQ ID NOS:1-193), PubMed (Key words: PTD, transduction, delivery)

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

| Category* | Citation of document, with indication, where appropriate, of the relevant passages                                                                                               | Relevant to claim No. |
|-----------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------|
| A         | WO 03/097671 A1 (CREAGENE INC.) 27 Nov. 2003<br>See abstract, claims and pages 7 lines 11-16.                                                                                    | 1-3,10-20             |
| A         | US 7,166,692 B2 (CANBREX BIO SCIENCE WALKERSVILLE INC.) 23 Jan. 2007<br>See abstract and examples.                                                                               | 1-3,10-20             |
| A         | Ho, A. et al. 'Synthetic protein transduction domains: Enhanced transduction potential in vitro and in vivo', Jan. 15, 2001, Vol. 61, pages 474-477.<br>See abstract and Fig. 2. | 1-3,10-20             |



Further documents are listed in the continuation of Box C.



See patent family annex.

\* Special categories of cited documents:

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

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

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

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

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

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

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

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

"&amp;" document member of the same patent family

Date of the actual completion of the international search

17 APRIL 2008 (17.04.2008)

Date of mailing of the international search report

**17 APRIL 2008 (17.04.2008)**

Name and mailing address of the ISA/KR

Korean Intellectual Property Office  
Government Complex-Daejeon, 139 Seonsa-ro, Seo-  
gu, Daejeon 302-701, Republic of Korea

Facsimile No. 82-42-472-7140

Authorized officer

KIM, KYOUNG MI

Telephone No. 82-42-481-8161



**INTERNATIONAL SEARCH REPORT**

Information on patent family members

International application No.

**PCT/KR2008/000525**

| Patent document<br>cited in search report | Publication<br>date | Patent family<br>member(s) | Publication<br>date |
|-------------------------------------------|---------------------|----------------------------|---------------------|
| W003097671A1                              | 27. 11. 2003        | AU2003218815A1             | 02. 12. 2003        |
|                                           |                     | EP1495045A1                | 12. 01. 2005        |
|                                           |                     | EP1495045A4                | 17. 08. 2005        |
|                                           |                     | JP2006502703T2             | 26. 01. 2006        |
|                                           |                     | KR1020050025143            | 11. 03. 2005        |
|                                           |                     | US20050154188A1            | 14. 07. 2005        |
|                                           |                     | US2005154188A1             | 14. 07. 2005        |
|                                           |                     | US2005154188AA             | 14. 07. 2005        |
|                                           |                     | US7101844BB                | 05. 09. 2006        |
| US07166692B2                              | 23. 01. 2007        | CA2516677AA                | 16. 09. 2004        |
|                                           |                     | EP1601373A2                | 07. 12. 2005        |
|                                           |                     | EP1601373A4                | 11. 07. 2007        |
|                                           |                     | JP2006521819T2             | 28. 09. 2006        |
|                                           |                     | US20040209797A1            | 21. 10. 2004        |
|                                           |                     | US2006106197AA             | 18. 05. 2006        |
|                                           |                     | W02004078933A2             | 16. 09. 2004        |
|                                           |                     | W02004078933A3             | 06. 05. 2005        |
|                                           |                     |                            |                     |