



Office de la Propriété

Intellectuelle
du Canada

Un organisme
d'Industrie Canada

Canadian
Intellectual Property
Office

An agency of
Industry Canada

CA 2319084 C 2011/02/01

(11)(21) **2 319 084**

(12) **BREVET CANADIEN**
CANADIAN PATENT

(13) **C**

(86) Date de dépôt PCT/PCT Filing Date: 1998/07/24
(87) Date publication PCT/PCT Publication Date: 1999/07/29
(45) Date de délivrance/Issue Date: 2011/02/01
(85) Entrée phase nationale/National Entry: 2000/07/24
(86) N° demande PCT/PCT Application No.: JP 1998/003328
(87) N° publication PCT/PCT Publication No.: 1999/037767
(30) Priorité/Priority: 1998/01/23 (JP10/11281)

(51) Cl.Int./Int.Cl. *C12N 15/12*(2006.01),
A01N 37/46(2006.01), *A01N 63/00*(2006.01),
C07K 14/47(2006.01), *C07K 14/78*(2006.01)

(72) Inventeur/Inventor:
WAKAMIYA, NOBUTAKA, JP

(73) Propriétaire/Owner:
FUSO PHARMACEUTICAL INDUSTRIES, LTD., JP

(74) Agent: KIRBY EADES GALE BAKER

(54) Titre : NOUVELLE COLLECTINE

(54) Title: NOVEL COLLECTIN

(57) Abrégé/Abstract:

A gene encoding a novel collectin protein which is expected to exhibit an antibacterial activity, an antiviral activity, etc. particularly in the human body, and its amino acid sequence.



PCT

特許協力条約に基づいて公開された国際出願

(51) 国際特許分類6 C12N 15/12, C07K 14/47, A01N 63/00	A1	(11) 国際公開番号 (43) 国際公開日	WO99/37767 1999年7月29日(29.07.99)
--	----	-------------------------------	------------------------------------

(21) 国際出願番号 (22) 国際出願日 (30) 優先権データ 特願平10/11281	PCT/JP98/03328 1998年7月24日(24.07.98) 1998年1月23日(23.01.98) JP	(81) 指定国 CA, US 添付公開書類 国際調査報告書
(71) 出願人 (米国を除くすべての指定国について) 扶桑薬品工業株式会社 (FUSO PHARMACEUTICAL INDUSTRIES, LTD.)[JP/JP] 〒541-0045 大阪府大阪市中央区道修町1丁目7番10号 Osaka, (JP) (72) 発明者 ; および (75) 発明者 / 出願人 (米国についてのみ) 若宮伸隆(WAKAMIYA, Nobutaka)[JP/JP] 〒567-0826 大阪府茨木市大池1丁目9-20 Osaka, (JP) (74) 代理人 弁理士 角田嘉宏, 外(SUMIDA, Yoshihiro et al.) 〒650-0031 兵庫県神戸市中央区東町123番地の1 貿易ビル3階 Hyogo, (JP)		

(54)Title: NOVEL COLLECTIN

(54)発明の名称 新規コレクチン

(57) Abstract

A gene encoding a novel collectin protein which is expected to exhibit an antibacterial activity, an antiviral activity, etc. particularly in the human body, and its amino acid sequence.

SPECIFICATION

Novel Collectin

5 [Field of the Invention]

The present invention relates to a novel collectin which is useful for investigating mechanisms of biological defense, and is expected to be applied for utilizing as materials for medicines because it may have physiological activities including anti-viral activities and the like.

10

[Background Art]

Collectin is a generic name of proteins having calcium-dependent carbohydrate recognition domain (CRD) and collagen-like region, and the member of these proteins is conceived to involve in basic immunity systems 15 against a wide spectrum of microorganisms such as bacteria and viruses.

20

The collectins which have been identified heretofore include mannan-binding protein (MBP), surfactant protein A (SP-A), surfactant protein D (SP-D) and conglutinin. These collectins are known to be constituted from basic structures comprising four unique regions of: (1) calcium-dependent carbohydrate recognition domain (CRD), (2) neck region, (3) collagen-like region and (4) N-terminal region containing cysteine [Malhortra *et al.*, *Eur.J.Immunol.* Vol.22, 1437-1445, 1992] (see, Fig. 1(a)). A subunit can be formed from the three basic structures through making a triple helix in the collagen-like region, and such subunit constitutes an oligomer, e.g., trimer, 25 tetramer and hexamer.

25

In vertebrates, mechanisms involving cellular immune responses and specific antibody reactions are considered as dominant host-defense systems 30 against invasion of the pathogenic bacteria, viruses and the like. Recently,

involvement in nonspecific immune responses of the lectins such as conglutinin has been suggested, for example, it was reported that the lectins may play important roles in neutralizing and removing the various microorganisms in infants having insufficient maternal antibodies and undeveloped specific defense systems [Super *et al.*, *Lancet*, Vol.2, 1236-1239, 1989].

Moreover, with respect to the roles of the lectins in the biological host-defense systems, it was reported that the host becomes susceptible to infection by, for example, a reduction of the MBP concentration in blood due to genetic mutation of the MBP gene [Sumiya *et al.*, *Lancet*, Vol.337, 1569-1570, 1991].

The present inventors have found that the conglutinin and the mannabinding protein can inhibit infection and hemagglutination activity of H1 and H3 Type Influenza A viruses [Wakamiya *et al.*, *Glycoconjugate J.*, Vol.8, 235, 1991; Wakamiya *et al.*, *Biochem. Biophys. Res. Comm.*, Vol.187, 1270-1278, 1992].

Thereafter, the present inventors isolated a cDNA clone encoding the conglutinin, and found that closer correlation may exist between the conglutinin gene and various surfactant protein D gene [Suzuki *et al.*, *Biochem. Biophys. Res. Comm.*, Vol.191, 335-342, 1993].

As described above, the collectin has been expected to be useful in investigating mechanisms of biological defense, and be applicable for utilizing as materials for medicines, however, the presence of any other molecular species belonging to this protein family has not been elucidated.

[Disclosure of the Invention]

The present invention was accomplished in consideration of the aforementioned state of art, and is directed to provide a novel collectin which can be expected to exhibit physiological activities such as anti-bacterial, anti-viral activity, especially in human body.

5

Accordingly, to provide the following polynucleotide and protein which share characteristic structures of those belonging to the collectin family, and which are distinct from the collectins reported heretofore is intended by the present invention:

10

[1] A polynucleotide comprising the nucleotide sequence which encodes a protein having the amino acid sequence set out in SEQ ID NO: 2;

[2] A polynucleotide comprising the nucleotide sequence set out in SEQ ID NO: 1;

15

[3] A polynucleotide which encodes a collectin protein, wherein said polynucleotide can hybridize under a stringent condition with a probe produced from a genomic clone which shares high homology to a consensus collectin amino acid sequence set out in SEQ ID NO:3

20

Glu-Lys-Cys-Val-Glu-Met-Tyr-Thr-Asp-Gly-Lys-Trp-Asn-Asp-Arg-Asn-Cys-Leu-Gln-Ser-Arg-Leu-Ala-Ile-Cys-Glu-Phe;

25

[4] A polynucleotide which can hybridize with any of the polynucleotide according to any of [1] to [3], wherein the protein encoded by said polynucleotide comprises: (1) Ca^{2+} -dependent carbohydrate recognition domain (CRD), (2) neck region, (3) collagen-like region, and (4) N-terminal region containing cysteine;

[5] A collectin protein encoded by the polynucleotide according to any of [3] or [4];

[6] A collectin protein comprising the amino acid sequence set out in SEQ ID NO:2;

30

[7] A collectin protein comprising the amino acid sequence encoded by

the polynucleotide comprising the nucleotide sequence set out in SEQ ID NO: 1;

[8] The collectin protein according to any of [5] to [7], wherein the amino acid sequence of the protein comprises deletion, substitution and/or 5 addition of one or more amino acids, and wherein the protein comprises: (1) Ca^{2+} -dependent carbohydrate recognition domain (CRD), (2) neck region, (3) collagen-like region, and (4) N-terminal region containing cysteine.

In one aspect there is provided an isolated polynucleotide comprising a nucleotide sequence encoding a protein comprising the following amino 10 acid sequence as set out in SEQ ID NO: 2:

Met-Asn-Gly-Phe-Ala-Ser-Leu-Leu-Arg-Arg-Asn-Gln-Phe-Ile-Leu-Leu-
Val-Leu-Phe-Leu-Leu-Gln-Ile-Gln-Ser-Leu-Gly-Leu-Asp-Ile-Asp-Ser-
Arg-Pro-Thr-Ala-Glu-Val-Cys-Ala-Thr-His-Thr-Ile-Ser-Pro-Gly-Pro-
Lys-Gly-Asp-Asp-Gly-Glu-Lys-Gly-Asp-Pro-Gly-Glu-Gly-Lys-His-
15 Gly-Lys-Val-Gly-Arg-Met-Gly-Pro-Lys-Gly-Ile-Lys-Gly-Glu-Leu-Gly-
Asp-Met-Gly-Asp-Arg-Gly-Asn-Ile-Gly-Lys-Thr-Gly-Pro-Ile-Gly-Lys-
Lys-Gly-Asp-Lys-Gly-Glu-Lys-Gly-Leu-Leu-Gly-Ile-Pro-Gly-Glu-Lys-
Gly-Lys-Ala-Gly-Thr-Val-Cys-Asp-Cys-Gly-Arg-Tyr-Arg-Lys-Phe-Val-
Gly-Gln-Leu-Asp-Ile-Ser-Ile-Ala-Arg-Leu-Lys-Thr-Ser-Met-Lys-Phe-
20 Val-Lys-Asn-Val-Ile-Ala-Gly-Ile-Arg-Glu-Thr-Glu-Gly-Lys-Phe-Tyr-
Tyr-Ile-Val-Gln-Glu-Gly-Lys-Asn-Tyr-Arg-Glu-Ser-Leu-Thr-His-Cys-
Arg-Ile-Arg-Gly-Gly-Met-Leu-Ala-Met-Pro-Lys-Asp-Glu-Ala-Ala-Asn-
Thr-Leu-Ile-Ala-Asp-Tyr-Val-Ala-Lys-Ser-Gly-Phe-Phe-Arg-Val-Phe-
Ile-Gly-Val-Asn-Asp-Leu-Glu-Arg-Glu-Gly-Gln-Tyr-Met-Phe-Thr-Asp-
25 Asn-Thr-Pro-Leu-Gln-Asn-Tyr-Ser-Asn-Trp-Asn-Glu-Gly-Glu-Pro-Ser-
Asp-Pro-Tyr-Gly-His-Glu-Asp-Cys-Val-Glu-Met-Leu-Ser-Ser-Gly-Arg-
Trp-Asn-Asp-Thr-Glu-Cys-His-Leu-Thr-Met-Tyr-Phe-Val-Cys-Glu-Phe-
Ile-Lys-Lys-Lys.

In another aspect there is provided an isolated polynucleotide 30 comprising the following nucleotide sequence and encoding the amino acid sequence set out in SEQ ID NO: 2:

4a

atgaatggct ttgcacccctt gcttcgaaga aaccaattta tcctcctgg
 actatttctt ttgcaaattc agagtctggg tctggatatt gatagccgtc
 ctaccgctga agtctgtgcc acacacacaa tttcaccagg acccaaagga
 gatgatggtg aaaaaggaga tccaggagaa gagggaaagc atggcaaagt
 5 gggacgcattt gggccgaaag gaattaaagg agaactgggt gatatggag
 atcggggcaa tattggcaag actgggcca ttggaaagaa gggtgacaaa
 gggaaaaaaag gtttgcttgg aataccttgg aaaaaaggca aagcaggtac
 tgtctgtat tgtgaaagat accggaaatt ttttggacaa ctggatatta
 gtattgcccgg 10 gctcaagaca tctatgaagt ttgtcaagaa tgtgatagca
 gggatttaggg aaactgaaga gaaattctac tacatcgatc aggaagagaa
 gaactacagg gaatccctaa cccactgcag gattcggggt ggaatgctag
 ccatgcccaa gatgaaagct gccaacacac tcattcgatc ctatgttgc
 aagagtggct tcttcgggt gttcattggc gtgaatgacc ttgaaaggaa
 15 gggacagtac atgttcacag acaacactcc actgcagaac tatagcaact
 ggaatgaggg ggaacccagc gaccctatg gtcattgagga ctgtgtggag
 atgctgatc ctggcagatg gaatgacaca gagtgccatc ttaccatgt
 ctttgtctgt gagttcatca agaagaaaaa g

In yet another aspect there is provided a probe that is capable of hybridizing under conditions of high stringency to a polynucleotide or 20 complement thereof that encodes a peptide that comprises a region substantially identical to the following [a] consensus collectin amino acid sequence as set out in SEQ ID NO: 3

Glu-Lys-Cys-Val-Glu-Met-Tyr-Thr-Asp-Gly-Lys-Trp-Asn-Asp-Arg-Asn-Cys-Leu-Gln-Ser-Arg-Leu-Ala-Ile-Cys-Glu-Phe.

25 In still yet another aspect there is provided the probe according to Claim 3 which comprises an amplification product of a PCR reaction performed using one or more primers selected from SEQ ID NO: 7 and SEQ ID NO: 8:

ttttgatgga ggctccatac c (SEQ ID NO: 7)
 30 ctgccaacac actcatcgct g (SEQ ID NO: 8).

4b

In still yet another aspect there is provided a collectin protein comprising the following amino acid sequence as set out in SEQ ID NO: 2:

5 Met-Asn-Gly-Phe-Ala-Ser-Leu-Leu-Arg-Arg-Asn-Gln-Phe-Ile-Leu-Leu-
 Val-Leu-Phe-Leu-Leu-Gln-Ile-Gln-Ser-Leu-Gly-Leu-Asp-Ile-Asp-Ser-
 Arg-Pro-Thr-Ala-Glu-Val-Cys-Ala-Thr-His-Thr-Ile-Ser-Pro-Gly-Pro-
 Lys-Gly-Asp-Asp-Gly-Glu-Lys-Gly-Asp-Pro-Gly-Glu-Glu-Gly-Lys-His-
 Gly-Lys-Val-Gly-Arg-Met-Gly-Pro-Lys-Gly-Ile-Lys-Gly-Glu-Leu-Gly-
 Asp-Met-Gly-Asp-Arg-Gly-Asn-Ile-Gly-Lys-Thr-Gly-Pro-Ile-Gly-Lys-
 Lys-Gly-Asp-Lys-Gly-Glu-Lys-Gly-Leu-Leu-Gly-Ile-Pro-Gly-Glu-Lys-
 10 Gly-Lys-Ala-Gly-Thr-Val-Cys-Asp-Cys-Gly-Arg-Tyr-Arg-Lys-Phe-Val-
 Gly-Gln-Leu-Asp-Ile-Ser-Ile-Ala-Arg-Leu-Lys-Thr-Ser-Met-Lys-Phe-
 Val-Lys-Asn-Val-Ile-Ala-Gly-Ile-Arg-Glu-Thr-Glu-Glu-Lys-Phe-Tyr-
 Tyr-Ile-Val-Gln-Glu-Glu-Lys-Asn-Tyr-Arg-Glu-Ser-Leu-Thr-His-Cys-
 Arg-Ile-Arg-Gly-Gly-Met-Leu-Ala-Met-Pro-Lys-Asp-Glu-Ala-Ala-Asn-
 15 Thr-Leu-Ile-Ala-Asp-Tyr-Val-Ala-Lys-Ser-Gly-Phe-Phe-Arg-Val-Phe-
 Ile-Gly-Val-Asn-Asp-Leu-Glu-Arg-Glu-Gly-Gln-Tyr-Met-Phe-Thr-Asp-
 Asn-Thr-Pro-Leu-Gln-Asn-Tyr-Ser-Asn-Trp-Asn-Glu-Gly-Glu-Pro-Ser-
 Asp-Pro-Tyr-Gly-His-Glu-Asp-Cys-Val-Glu-Met-Leu-Ser-Ser-Gly-Arg-
 Trp-Asn-Asp-Thr-Glu-Cys-His-Leu-Thr-Met-Tyr-Phe-Val-Cys-Glu-Phe-
 20 Ile-Lys-Lys-Lys-Lys.

In still yet another aspect there is provided a collectin protein comprising the amino acid sequence as set out in SEQ ID NO: 2 encoded by the following nucleotide sequence:

25 atgaatggct ttgcattcctt gcttcgaaga aaccaattta tcctcctgg
 actatttctt ttgcaaattc agagtctggg tctggatatt gatagccgtc
 ctaccgctga agtctgtgcc acacacacaa tttcaccagg acccaaagga
 gatgatggtg aaaaaggaga tccaggagaa gagggaaagc atggcaaagt
 gggacgcattt gggccgaaag gaattaaagg agaactgggt gatatggag
 atcggggcaa tattggcaag actgggcca ttgggaagaa ggtgacaaa
 gggaaaaag gtttgcttgg aatacctgga gaaaaaggca aagcaggtac
 30 tgtctgtat tgtggaaagat accggaaatt ttttggacaa ctggatatta
 gtattgccccg gctcaagaca tctatgaagt ttgtcaagaa tgtgatagca
 gggatttaggg aaactgaaga gaaattctac tacatcggtc aggaagagaa
 gaactacagg gaatccctaa cccactgcag gattcgggtt ggaatgctag
 35 ccatgccccaa ggtatggact gccaacacac tcattcgctga ctatgttgc
 aagagtggct tctttcggtt gttcattggc gtgaatgacc ttgaaaggaa
 gggacagttac atgttccacag acaacactcc actgcagaac tatacaact
 ggaatgaggg ggaacccagc gacccctatg gtcattgagga ctgtgtggag
 40 atgctgagct ctggcagatg gaatgacaca gagtgcacatc ttaccatgt
 ctttgtctgt gagttcatca agaagaaaaa g

In still yet another aspect there is provided an isolated polynucleotide consisting of the 6th to the 836th contiguous nucleotides in the nucleotide sequence as set out in SEQ ID NO:1.

5 In still yet another aspect there is provided a polypeptide consisting of the amino acid sequence as set out in SEQ ID NO:2.

[Brief Description of the Drawings]

Figure 1 is a profile showing basic structures and overviews of the principal collectins reported heretofore;

10 Figure 2 shows the alignment of the preceding half portions of amino acid sequences of three collectins reported heretofore;

Figure 3 shows the alignment of the latter half portions of the amino acid sequences in Figure 2;

15 Figure 4 shows each of the primers used for sequencing the novel collectin polynucleotide of the present invention, and maps of the nucleotide sequence which were read from the sequencer (a, jb); and an ORF of the obtained collectin (a);

Figure 5 shows the alignment of the preceding half portions of amino acid sequences of the three collectins reported heretofore and the novel collectin of the present invention;

20 Figure 6 shows the alignment of the latter half portions of the amino acid sequences in Figure 5;

Figure 7 illustrates a comparison of the basic structures of the three collectins reported heretofore and the novel collectin of the present invention comprising (1) Ca²⁺-dependent carbohydrate recognition domain (CRD), (2) neck region, (3) collagen-like region, and (4) N-terminal region containing cysteine;

Figure 8 shows a result of genomic Southern analysis with the novel

collectin of the present invention;

Figure 9 shows a result of genomic Northern analysis of various human tissues, i.e., (a) heart, (b) brain, (c) placenta, (d) lung, (e) liver, (f) skeletal muscle, (g) kidney and (h) pancreas with the novel collectin of the present invention to clarify the tissue distribution of the collectin;

Figure 10 shows a result of genomic Southern analysis of genes from various vertebrates, i.e., (a) human, (b) monkey, (c) rat, (d) mouse, (e) dog, (f) cow, (g) rabbit and (h) chicken with the novel collectin of the present invention to elucidate the conservation of the collectin during the species; and

Figure 11 shows a phylogenetic tree of various collectins.

[Best Mode for Carrying Out the Invention]

In the preferred embodiment of the present invention, the probe in the above [3] may be amplification products by PCR which was performed using the primers which have the following sequences:

TTTGATGGAGGCTCCATACC (SEQ ID NO: 7) ; and
CTGCCAACACACTCATCGCTG (SEQ ID NO: 8).

Thus, desired polynucleotide encoding the collectin protein can be suitably obtained.

Moreover, in the preferred embodiment, the polynucleotide may be cDNA.

Further, the protein of the present invention may preferably consist of the amino acid sequence which is identical to that derived from human, because it can be expected to exhibit physiological activities such as anti-bacterial, anti-viral activity in human body, thereby useful physiologically active material for medicines may be resulted. Therefore, it is intended that the protein of the present invention may be the collectin protein derived from

human. We examined various human tissues, and expression of the collectin protein in human liver which seems to be useful could be revealed.

The stringent hybridization condition in the inventions of the above [3] and [4] may include for example, a series of the following steps for the hybridization: prehybridization in a solution of 5 x SSC (prepared by diluting 20 x SSC (3 M NaCl, 0.3 M sodium citrate)), 1% blocking agent (Boehringer Mannheim), 0.1% N-lauroyl sarcosine, and 0.02% SDS, at 68°C for an hour; and hybridization in a solution of 5 x SSC, 1% blocking agent, 0.1% N-lauroyl sarcosine, and 0.02% SDS containing cDNA probes (10 ng/ml), at 55°C for 16 hours; washing in a solution of 2 x SSC/0.1% SDS for 5 minutes, 2 times; and washing in a solution of 0.5 x SSC/0.1% SDS at 55°C for 15 minutes, 2 times. However, several modifications/alterations of these conditions may be made, based on the knowledge of the skilled art, such as the concentration of the solution, incubation temperature and time.

In addition, the N-terminal region containing cysteine in the above [8] may contain at least one cysteine residue, preferably one cysteine residue.

Further, deletion, substitution and/or addition of one or more amino acids in the above [8] may be those which result in less changes of hydrophilic/hydrophobic nature, acidic/basic nature, and constitutional residues of the collectin proteins, without bringing much alterations of the properties in the above-described four regions, specifically in both of the regions: (1) Ca^{2+} -dependent carbohydrate recognition domain (CRD) and (3) collagen-like region. Taking the structural sequences of the proteins belonging to the collectin family reported heretofore into account, for example, deletion, substitution and/or addition of 1-10 amino acid residues in (1) Ca^{2+} -dependent carbohydrate recognition domain (CRD) and in (2) neck region, 1-100, preferably 1-15 amino acid residues in (3) collagen-like region, and 1-20 amino acid residues in (4) N-terminal region containing cysteine and

the signal sequence may be allowed.

The present invention will be described in more detail by the non-limiting illustrative examples. It is intended that the present invention encompasses all modifications and variations which occur to those skilled in the art upon consideration of the disclosures herein, and in particular those embodiments which are within the broadest proper interpretation of the claims and their requirements.

10 The Examples illustrate: the search on EST data base (Example 1); preparation of the probes for screening (Example 2); screening of cDNA library derived from human liver (Example 3); sequencing of nucleotide sequence of the novel collectin (Example 4); genomic Southern analysis of the novel collectin (Example 5); Northern analysis of the novel collectin with various human tissues (Example 6); genomic Southern analysis of the novel collectin with tissues from various species of animals (Example 7); and genetic study of the novel collectin (Example 8).

Example 1: Search on EST Data Base

20 Highly conserved regions between molecules of known collectin proteins, i.e., MBP, SP-A and SP-D were searched by comparing the amino acid sequences thereof (see Figures 2 and 3, in which amino acid residues which were recognized to be homologous between those proteins were boxed). As a result, it was suggested that the region consisting of 27 amino acids, namely from amino acid 220 to 246 of human MBP sequence (shown in Figure 3, reversed characters), was highly homologous. Therefore, some consensus sequences corresponding to this region were prepared, and conducted searches on EST (Expressed Sequence Tags) data base with such sequences. For this search, the EST data base published on October 11, 1996 was employed, which included 676750 sequences.

Consequently, some data comprising homologous amino acid sequences were obtained. Searches on GenBank/EST data base with thus obtained data of the amino acid sequences were further conducted, and deduced whether they were derived from known, or unknown substances. Thereby, it was confirmed that data including highly homologous but unknown nucleotide sequence (registered as: R29493) could be identified when the following amino acid sequence was used as a consensus sequence:

5 Glu-Lys-Cys-Val-Glu-Met-Tyr-Thr-Asp-Gly-Lys-Trp-Asn-Asp-Arg-Asn-Cys-Leu-Gln-Ser-Arg-Leu-Ala-Ile-Cys-Glu-Phe (SEQ ID NO: 3).

10 The data contained the sequence of 5'-terminal 326 nucleotides of a clone F1-1006D from human embryonic (22 weeks old) liver cDNA library.

15 Thereafter, the clone was kindly provided from the owner, Mr. Hee-Sup Shin (Pohang Institute of Science & Technology (Pohang, Korea)). The insert size of the clone was about 600 bp, with the 5'-terminal end being incorporated adjacent to the nucleotide sequence set out in SEQ ID NO: 4, while 3'-terminal end being incorporated into plasmid pSK(-) (pBluescriptTMIISK(-)) at XhoI restriction site.

20 Example 2: Preparation of the Probes for Screening

The insert of the clone described above in Example 1 was excised using EcoRI and XhoI, then incorporated into pUC18, and sequenced using a primer (Pharmacia, M13 Universal Primer (SEQ ID NO: 5, 5'-fluorescein-CGACGTTGAAAACGACGGCCAGT-3')) and M13 Reverse Primer (SEQ ID NO: 6, 5'-fluorescein-CAGGAAACAGCTATGAC-3').

25 In the nucleotide sequence obtained, an open reading flame was selected through matching it to the collectin amino acid sequence. The nucleotide sequence corresponding to the amino acid sequence which could be read from the above open reading flame was picked out, and primers for

digoxigenin (DIG) labeled cDNA probes (Reverse Primer (SEQ ID NO: 7) and Forward Primer (SEQ ID NO: 8)) corresponding to the parts of the nucleotide sequences were produced using DNA/RNA Synthesizer (Applied Biosystems, 392A). DIG labeling was achieved using PCR DIG Probe Synthesis Kit (Boehringer Mannheim). The reaction mixture contained: DNA fragments which were the excised inserts from the clone F1-1006D with EcoRI and XhoI (4.4 ng/ μ l, 12 μ l: 52.8 ng), 10 x buffer: 5 μ l, 25 mM MgCl₂ : 5 μ l, dNTP (PCR Labeling Mix): 2.5 μ l, 20 μ M Reverse Primer : 2.5 μ l, 20 μ M Forward Primer : 5 μ l, H₂O : 18 μ l, Taq Polymerase : 0.5 μ l. PCR reaction was carried out using ZymoreactorTM (Atto Corp.,) through 35 cycles of: 1 minute at 92°C, 1 minute at 55°C, and 2 minutes at 72°C.

Example 3: Screening of cDNA Library Derived from Human Liver

First, phage cDNA library was titrated as follows. Escherichia coli Y1090r⁻ which had been cultured at 37°C for 16 hours in mLB medium (LB medium (1 g tryptone, 0.5g yeast extract and 0.5 g NaCl in total volume of 100 ml) containing 10 mM MgSO₄ and 0.2% maltose), 0.2 ml, and 0.1 ml of cDNA library serially diluted with SM buffer (5.8 g NaCl, 2 g MgSO₄·7H₂O, 2 M Tris-HCl (pH 7.5) 25 ml, and 2% gelatin 5 ml in total volume of 1L) were incubated at 37°C for 15 minutes, then the mixtures were added to 2.5 ml of LB-TOP agar (0.75% agar/LB medium) to make homogenous solutions, and plated onto LB Plates (Iwaki Glass, 90 mm ϕ , 1.5% agar/LB medium). The added solutions were hardened at a room temperature for 15 minutes, then incubated for 5 hours at 42°C. The plaques on each of the plates were counted, and the titer of the phage was calculated. The titer calculated hereby was 2.3×10^{10} pfu/ml.

The screening with thus titrated cDNA library using the probes prepared in Example 2 was conducted as follows. Escherichia coli Y1090r⁻ which had been cultured at 37°C for 16 hours in mLB medium, 0.6 ml and cDNA library diluted with SM buffer to 1×10^5 pfu were incubated at 37°C

for 15 minutes, then the mixture was added to 7.5 ml of LB-TOP agar (0.75% agarose) to make a homogenous solution. The solution was plated onto ten LB square plates of 140 mm² (Nissui Seiyaku), hardened at a room temperature for 15 minutes, then the plates were incubated for 5 hours at 5 42°C. After plaque formation of each of the plates was determined, the transfer to the nylon membranes, using NytranTM 13N (Schleicher and Schuell Co.) as a filter was performed. The filters (12.5 cm x 9.0 cm in size) were immersed in distilled water for 10 minutes to be wet, then the excess water was removed on Whatmann 3MM Paper. The filters were placed on the plates 10 having the plaques formed thereon. After standing for two minutes, the filters were recovered and air dried for 10 minutes. The phage DNA on the filters was denatured for 2 minutes with 0.2 M NaOH/1.5 M NaCl, followed by neutralization with 0.4 M Tris-HCl (pH7.6) / 2 x SSC for 2 minutes and washing with 2 x SSC for 2 minutes. Thereafter, the DNA was fixed on the 15 membrane by UV irradiation with GS GENE LINKER (BioRad).

Hybridization and detection of the signals were conducted as follows. The filters were soaked in 2 x SSC, and the excess water was removed using Whatmann 3MM Paper. Then, the filters were placed in a hybridization bag and prehybridization at 68°C for one hour in a hybridization solution (5 x 20 SSC, 1% blocking agent, 0.1% N-lauroyl sarcosine and 0.02% SDS) was performed. Subsequently, the hybridization solution was removed from the bag, and the hybridization solution containing DIG labeled cDNA probe at a concentration of 10 ng/ml was added thereto, and hybridization was proceeded at 55°C for 16 hours. After the hybridization was completed, the filters were washed in a solution of 2 x SSC/0.1% SDS at a room temperature 25 for 5 minutes, 2 times; and further washed in a solution of 0.5 x SSC/0.1% SDS for 15 minutes at 55°C, 2 times. Then, SDS was removed using DIG buffer I (100 mM Tris-HCl, 150 mM NaCl (pH7.5)) for 1 minute, and the filters were blocked with DIG buffer II (1% blocking agent in DIG buffer I) 30

for 30 minutes. After washing the filters with DIG buffer I for one minute, a solution of alkaline phosphatase labeled anti-DIG antibody (Boehringer Mannheim) which was diluted to 5000-fold in DIG buffer II was added, and the reactions between antigen and antibody were allowed for 30 minutes. The filters were then washed twice with DIG buffer I for 15 minutes at a room temperature. Through the subsequent treatment of the filters with DIG buffer III (100 mM Tris-HCl, 100 mM NaCl (pH 9.5), 50 mM MgCl₂) for 3 minutes, the concentration of Mg²⁺ was elevated. Finally, a solution of NBT/BCIP (WAKO Chem., Co.) in DIG buffer III was added for color development, thereby 13 positive clones were identified.

The plaques corresponding to these clones were excised from the plates and placed in the tubes containing 1 ml of SM buffer. After stirring for 10 minutes, each of the buffer solution was serially diluted with SM buffer, and 0.1 ml of the diluted solution was mixed with 0.2 ml cultures of Escherichia coli Y1090r which had been cultured in mLB medium for 16 hours at 37°C. The mixture was incubated for 15 minutes at 37°C, and added to 2.5 ml of LB-TOP agarose to make a homogenous solution. The solution was plated onto ten LB plates (90mm ϕ), hardened at a room temperature for 15 minutes, then the plates were incubated for 5 hours at 42°C. With respect to several plaques thus obtained, the secondary screening was performed essentially in accordance with the procedures of the primary screening as described above.

Example 4: Sequencing of the Novel Collectin Nucleotide

The plaques from suitable two clones (HLI1-3M and HLI1-9) which were selected from the positive clones obtained in the above secondary screening were excised from the plates, and placed respectively in the tubes containing 1 ml of SM buffer. After stirring, 50 μ l of each of the solution was added to 4.95 ml of mLB medium with 50 μ l cultures of Escherichia

coli Y1090r which had been cultured in mLB medium for 16 hours at 37°C. The mixture was cultured for 16 hours at 37°C, then one drop of chloroform was added thereto. After stirring for 3 minutes, the mixture was subjected to centrifuge at 10,000 rpm for 5 minutes to obtain a supernatant therefrom.

5

The insert DNA was amplified by PCR with TaKaRa LA PCR KitTM Ver.2 (TAKARA Syuzo, Co.) using the resulting supernatant as a template. PCR reactions contained the supernatant: 11 μ l, 10 x LA PCR Buffer II (Mg²⁺free): 2.5 μ l, 25 mM MgCl₂: 5 μ l, dNTP Mix: 8 μ l, 20 μ M λ gt11 Reverse Primer (SEQ ID NO: 9: 5'-TTGACACCAGACCAACTGGTAATG-3'): 2.5 μ l, 20 μ M λ gt11 Forward Primer (SEQ ID NO: 10: 5'-GGTGGCGACGACTCCTGGAGCCCCG-3'): 1 μ l, LA Taq polymerase: 0.5 μ l, and H₂O : to make final volume 50 μ l. The PCR reaction was performed using Applied Biosystems Gene Amp PCR System 9600, with 30 cycles of: 10 seconds at 98°C, and 5 minutes at 68°C. The PCR product was verified by the electrophoresis with 1% agarose gel, and purified through excising from the gel. For this purification step, Sephaglas BandPrep Kit (Pharmacia) was used.

20

The excised DNA fragment was incorporated into pCR2.1 vector (Invitrogen, TA Cloning Kit). The recombinant vector was transformed into TOP10F' cell included in the Invitrogen TA Cloning Kit. The transformants were cultured in LB medium (containing 100 μ g/ml ampicillin), and two plasmid for each of the clones (HLI1-3M-1, HLI1-3M-2, HLI1-9-1 and HLI1-9-2) were extracted by alkaline SDS method followed by nucleotide sequencing with Autoread Sequencing Kit (Pharmacia) and A.L.F. Autosequencer. M13 Universal Primer (SEQ ID NO: 5) and M13 Reverse Primer (SEQ ID NO: 6) from the Autoread Sequencing Kit were used first, then full length nucleotide sequence was determined based on the resulting elucidated nucleotide sequences using the following primers (3MUO-9R3)

30

which were produced on a DNA/RNA synthesizer and labeled with FITC (Pharmacia, Fluore Prime™):

3MU0: 5'-fluorescein-TAATGGTAGCGACCGGCGCT-3' (SEQ ID NO: 11),

3MU1: 5'-fluorescein-AAACCAATTATACTCCTGG-3' (SEQ ID NO: 12),

5 3MU2: 5'-fluorescein-AATATTGGCAAGACTGGGCC-3' (SEQ ID NO: 13),

3MR1: 5'-fluorescein-GATGAGTGTGTTGGCAGCAT-3' (SEQ ID NO: 14),

3MR2: 5'-fluorescein-GTATCTTCCACAATCACAGA-3' (SEQ ID NO: 15),

3MR3: 5'-fluorescein-TTAATTCCCTTCGGCCCCAT-3' (SEQ ID NO: 16),

3MR4: 5'-fluorescein-GCAAAAGAAATAGTACCAAGG-3' (SEQ ID NO:

10 17),

3MR5: 5'-fluorescein-CATATCACCCAGTTCTCCTT-3' (SEQ ID NO: 18),

9U1 : 5'-fluorescein-AGCAGGGATTAGGGAAACTG-3' (SEQ ID NO: 19),

9U3 : 5'-fluorescein-CTGTGAGCGTCATTACAGTT-3' (SEQ ID NO: 20),

9U4 : 5'-fluorescein-GGTTGTCTATATGTCAAATG-3' (SEQ ID NO: 21),

15 9U5 : 5'-fluorescein-TATGCCATGGCTATACTTG-3' (SEQ ID NO: 22),

7U3 : 5'-fluorescein-ATCGCTGACTATGTTGCCAA-3' (SEQ ID NO: 23),

9R1 : 5'-fluorescein-CAAGTATGCCATGGCCATA-3' (SEQ ID NO: 24),

9R2 : 5'-fluorescein-AACTGTAATGACGCTCACAG-3' (SEQ ID NO: 25),

and

20 9R3 : 5'-fluorescein-CATTGACATATGAACAAACC-3' (SEQ ID NO: 26)

As a result, the obtained cDNA clone contained 1295 bases set out in SEQ ID NO: 1, which comprises ORF (open reading frame) of 831 base pairs encoding 277 amino acids as shown in SEQ ID NO: 2.

25

The outline of this nucleotide sequencing strategy is shown in Figure 4. An ORF of the obtained collectin is illustrated in Figure 4 (a), wherein G-X-Y denotes a collagen-like region. Further, in Figure 4 (b), each of the primer names and maps of the nucleotide sequence which were read from the sequencer (shown as arrows), as well as M13 Universal Primer (shown as U)

30

and M13 Reverse Primer (shown as R) are illustrated.

Figures 5 and 6 show the alignment of the amino acid sequence of the novel collectin of the present invention and those of three collectin proteins reported heretofore. Similarly to Figures 2 and 3, amino acid residues which were recognized to be homologous were boxed.

Furthermore, under structural studies of the sequence of this novel collectin protein, as shown schematically in Figure 7, it was suggested that this protein comprised: (a) N-terminal region containing cysteine, (b) collagen-like region, (c) neck region and (d) carbohydrate recognition domain, as in the case with the known collectins.

However, homology search results on GenBank data base of the DNA and amino acid sequence indicated that the sequence of the obtained protein is distinct from those of the collectins identified previously and it derives from a novel collectin.

Example 5: Genomic Southern Analysis of the Novel Collectin

Genomic Southern analysis was performed in order to clarify whether the novel collectin gene comprising the cDNA sequence shown in Example 4 was a single copy gene or a multi copy gene.

Four μ g of genomic DNA extracted from placenta was digested with a restriction enzyme, EcoRI, HindIII, BamHI, XbaI or SacI, followed by electrophoresis with 0.7% agarose gel at 100 mA, for 3 hours. After the electrophoresis was completed, the DNA was transferred to a nylon membrane (Nytran 13N) to prepare a membrane for the analysis.

For the transfer step, the electrophoresed gel was immersed in 100 ml

of 0.25 N HCl for 10 minutes, washed three times with distilled water, then immersed twice in 100 ml of a denaturalizing solution (1.5 M NaCl, 0.5 M NaOH) for 15 minutes, and immersed in 100 ml of a neutralizing solution (0.5 M Tris-HCl, 3 M NaCl (pH 6.8)) for 30 minutes so that the depurination, denaturalization and neutralization were accomplished. Thereafter, the DNA was transferred using Vacuum Blotting System (Toyobo Engineering, VB-30). In this step, the membrane which had been pretreated by immersing in 2 x SSC for 5 minutes and in 20 x SSC for 5 minutes was used, with a pad which had been soaked with 20 x SSC. After the transfer was terminated, fixation of the DNA was performed by UV irradiation.

Hybridization probe employed for the Southern analysis was the DIG labeled DNA probe corresponding to the cDNA sequence of ORF of the novel collectin as obtained in the above Example 4, which was labeled using the above-described PCR DIG Probe Synthesis Kit. Prior to hybridization, the probe was boiled for 10 minutes, and rapidly frozen with dry ice/ethanol for 5 minutes.

First, the membrane which was subjected to the transfer was immersed in 2 x SSC for 5 minutes, then prehybridization was performed in ExpressHybTM Hybridization Solution (Clonetech), 10 ml at 68°C for 30 minutes. Subsequently, the above frozen probe was diluted to 10 ng/ml in ExpressHyb Hybridization Solution and 2 ml of this solution was used for hybridization at 68°C for one hour.

Following hybridization, the membrane was washed by shaking: two times in 20 ml of 2 x SSC, 0.1%SDS at a room temperature for 5 minutes and then two times in 20 ml of 0.2 x SSC, 0.1%SDS at 68°C for 15 minutes. Next, the membrane was washed two times with 50 ml of DIG buffer I (100 mM Tris-HCl, 150 mM NaCl (pH 7.5)) at a room temperature for one minute

in order to remove SDS, and was blocked in DIG buffer II' (1.5% blocking agent in DIG buffer I), 50 ml at a room temperature for one hour. Thereafter, the membrane was treated with 10 ml of alkaline phosphatase labeled anti-DIG antibody which was diluted to 5000-fold in DIG buffer I containing 5 TM 0.2% Tween20 for 30 minutes followed by washing two times by shaking in 50 ml of DIG buffer I which contains 0.2% Tween20 at a room temperature for 20 minutes. After soaking the membrane twice in 10 ml of DIG buffer III at a room temperature for 3 minutes, it was placed in a hybridization bag, and 10 CSPD (registered trade name, Boehringer Mannheim: chemiluminescence substrate) which was diluted to 100-fold in DIG buffer III was added thereto so that the solution can spread over the membrane. Subsequently, the membrane was exposed to Instant Film 57 TM (Polaroid).

Consequently, it was speculated that the gene of the obtained novel 15 collectin has been a single copy gene, because only one or two signals could be detected from the respective genomic DNA which was digested with each of the restriction enzymes, as shown in the lanes in Figure 8.

Example 6: Northern Analysis of the Novel Collectin
with Various Human Tissues

In order to examine the expression of the mRNA of the novel collectin of the present invention in various human tissues, analysis by Northern hybridization was performed.

25 Hybridization probe employed for this analysis was the DIG labeled RNA probe corresponding to the cDNA sequence of ORF of the novel collectin (SEQ ID NO: 1), which was labeled using DIG RNA Labeling Kit (SP6/T7, Boehringer Mannheim). The analyzed membrane was Human Multiple Tissue Northern (MTN) Blot (Clonetech) containing each poly A⁺ RNA from human (a) heart, (b) brain, (c) placenta, (d) lung, (e) liver, (f) 30

skeletal muscle, (g) kidney and (h) pancreas, which was prepared by: modification of the electric charge of a nylon membrane prior to transferring the RNA, the RNA transfer from a 1.2% formaldehyde denaturalized agarose gel which had been previously loaded with 2 μ g of the above each poly A⁺ RNA and electrophoresed, and then followed by a fixation using UV irradiation.

5

10

15

20

25

30

Hybridization was carried out using the above-described probe solution and membrane, in accordance with the following procedures. First, the membrane was immersed in 2 x SSC for 5 minutes, then prehybridization was performed in 10ml of hybridization solution (5 x SSC, 10 x Denhardt's solution, 10 mM sodium phosphate buffer (pH 6.5), 50% formamide, 0.5%SDS, 0.1 mg/ml sermon sperm DNA), at 65°C for 3 hours. Subsequently, the probe which had previously been boiled for 10 minutes and rapidly frozen with dry ice/ethanol for 5 minutes was diluted in the hybridization solution to be 1 μ g/ml, and 2 ml of thus diluted probe solution was used for hybridization at 65°C for 18 hours.

Following hybridization, the membrane was washed by shaking: two times in 20 ml of 2 x SSC, 0.1%SDS at a room temperature for 5 minutes and then two times in 20 ml of 0.1 x SSC, 0.1%SDS at 68°C for 15 minutes. Next, the membrane was washed two times with 50 ml of DIG buffer I at a room temperature for one minute in order to remove SDS, and was blocked in 50 ml of DIG buffer II' at a room temperature for one hour. Thereafter, the membrane was treated with 10 ml of alkaline phosphatase labeled anti-DIG antibody which was diluted to 5000-fold in DIG buffer I which contains 0.2% Tween20 for 30 minutes, followed by washing two times by shaking in 50 ml of DIG buffer I which contains 0.2% Tween20 at a room temperature for 20 minutes. After soaking the membrane twice in 10 ml of DIG buffer III at a room temperature for 3 minutes, it was placed in a hybridization bag, and

CSPD which was diluted to 100-fold in DIG buffer III was added thereto so that the solution can spread over the membrane. Subsequently, the membrane was exposed to Instant Film 612 (Polaroid).

5 As a consequence, it was revealed that mRNA of the collectin of the present invention having 1.2 kb and 3.8 kb in size has been expressed in liver (lane e) and in placenta (lane c), with more amount being expressed in liver while less but certain amount being expressed in placenta as shown in Figure 9.

10

Example 7: Genomic Southern Analysis of the Novel Collectin in Various Animals

15 In order to elucidate conservation of the collectin gene of the present invention between other species of animals, analysis by genomic Southern hybridization was performed.

20 Hybridization probe employed for this analysis was the DIG labeled DNA probe corresponding to the cDNA sequence of ORF of the novel collectin as obtained in the above Example 4, which was labeled using the above-described PCR DIG Probe Synthesis Kit (Boehringer Mannheim), while the analyzed membrane was ZOO-BLOT (Clonetech). This membrane contains each genomic DNA obtained from (a) human placenta, (b) Rhesus monkey kidney, (c) Sprague-Dawley rat kidney, (d) Balb/c mouse kidney, (e) canine kidney, (f) bovine kidney, (g) rabbit kidney and (h) chicken liver, which was prepared by: modification of the electric charge of a nylon membrane prior to transferring the genomic DNA, the DNA transfer from an agarose gel which had been previously loaded with 4 μ g of the above each genomic DNA which were digested with restriction enzyme EcoRI, and electrophoresed, then finally followed by a fixation using UV irradiation.

25

30

Hybridization was carried out using the above-described probe and membrane, in accordance with the following procedures. First, the membrane was immersed in 2 x SSC for 5 minutes, then prehybridization was performed in 10ml of ExpressHyb Hybridization Solution at 65°C for 30 minutes. Subsequently, the probe which had previously been frozen as described above was diluted in the ExpressHyb Hybridization Solution to be 10 ng/ml, and 2 ml of thus diluted probe solution was used for hybridization at 65°C for one hour.

Following hybridization, the membrane was washed by shaking: two times in 20 ml of 2 x SSC, 0.1%SDS at a room temperature for 5 minutes and then two times in 20 ml of 0.2 x SSC, 0.1%SDS at 68°C for 15 minutes. Next, the membrane was washed two times with DIG buffer I at a room temperature for one minute in order to remove SDS, and was blocked in 50 ml of DIG buffer II' at a room temperature for one hour. Thereafter, the membrane was treated with 10 ml of alkaline phosphatase labeled anti-DIG antibody which was diluted to 5000-fold in DIG buffer I which contains 0.2% Tween20 for 30 minutes followed by washing two times with shaking in 50 ml of DIG buffer I which contains 0.2% Tween20 at a room temperature for 20 minutes. After soaking the membrane twice in 10 ml of DIG buffer III at a room temperature for 3 minutes, it was placed in a hybridization bag, and CSPD which was diluted to 100-fold in DIG buffer III was added thereto so that the solution can spread over the membrane. Subsequently, the membrane was exposed to Instant Film 57.

The result of this analysis is shown in Figure 10, wherein signals of DNA can be found in all lanes except for the lane h which was loaded with DNA from chicken. Therefore, it was illustrated that the collectin gene of the present invention has been conserved between the mammalian species.

Example 8: Genetic Analysis of the Novel Collectin

To elucidate the genetic positional relevance of the present collectin between the known collectins, analysis was performed based on the DNA sequence as obtained, and a phylogenetic tree was created.

5

The collectins selected as subjects for analysis were: human MBP (Mannan- Binding Protein), human SP-A (Surfactant Protein A), rat MBP-A, rat MBP-C, rat SP-D, mouse MBP-A, mouse MBP-C, rabbit MBP, monkey MBP-A, monkey MBP-C, bovine SP-D, bovine MBP, bovine conglutinin (bKg), and bovine collectin 43 (CL43). Each of the amino acid sequence was retrieved from GenBank data base, then using the regions containing lectin domains from the obtained data, multiple alignment was produced by clustalw method. Thereby, a phylogenetic tree was created using N-J method with Phylipl Version 3.57c package program.

10

Consequently, as shown in Figure 11, although SP-D, bovine collectin 43 and bovine conglutinin have constituted single cluster, additionally MBP and SP-A have respectively constituted separate clusters, while the collectin gene of the present invention has not belonged to any of these clusters. 20 Accordingly, it was speculated that the collectin of the present invention may constitute a distinct cluster which is genetically diverse from those of the collectins reported heretofore.

[Industrial Applicability]

As set forth above, a novel collectin gene and protein having characteristic structures of the collectins, which are different from the collectins reported so far, are provided by the present invention. Such protein is expected to exhibit physiological activities such as anti-bacterial, anti-viral activity, especially in human body, thereby medicinal applications, as well as tools for investigating mechanisms of biological defense systems may be 30

provided.

SEQUENCE LISTING

<110> FUSO PHARMACEUTICAL INDUSTRIES, LTD.

<120> Novel Collectin

<130> 45202-NP

<140> 2,319,084

<141> 1998-07-24

<150> JP 10-11281

<151> 1998-01-23

<160> 26

<210> 1

<211> 1595

<212> DNA

<213> Homo Sapiens

<220>

<221> CDS

<222> (6)..(836)

<400> 1

cagca atg aat ggc ttt gca tcc ttg ctt cga aga aac caa ttt atc 47
Met Asn Gly Phe Ala Ser Leu Leu Arg Arg Asn Gln Phe Ile
1 5 10

ctc ctg gta cta ttt ctt ttg caa att cag agt ctg ggt ctg gat att 95
Leu Leu Val Leu Phe Leu Leu Gln Ile Gln Ser Leu Gly Leu Asp Ile
15 20 25 30

gat agc cgt cct acc gct gaa gtc tgt gcc aca cac aca att tca cca 143
Asp Ser Arg Pro Thr Ala Glu Val Cys Ala Thr His Thr Ile Ser Pro
35 40 45

23

gga ccc aaa gga gat gat ggt gaa aaa gga gat cca gga gaa gag gga	191		
Gly Pro Lys Gly Asp Asp Gly Glu Lys Gly Asp Pro Gly Glu Glu Gly			
50	55	60	
aag cat ggc aaa gtg gga cgc atg ggg ccg aaa gga att aaa gga gaa	239		
Lys His Gly Lys Val Gly Arg Met Gly Pro Lys Gly Ile Lys Gly Glu			
65	70	75	
ctg ggt gat atg gga gat cgg ggc aat att ggc aag act ggg ccc att	287		
Leu Gly Asp Met Gly Asp Arg Gly Asn Ile Gly Lys Thr Gly Pro Ile			
80	85	90	
ggg aag aag ggt gac aaa ggg gaa aaa ggt ttg ctt gga ata cct gga	335		
Gly Lys Lys Gly Asp Lys Gly Glu Lys Gly Leu Leu Gly Ile Pro Gly			
95	100	105	110
gaa aaa ggc aaa gca ggt act gtc tgt gat tgt gga aga tac cgg aaa	383		
Glu Lys Gly Lys Ala Gly Thr Val Cys Asp Cys Gly Arg Tyr Arg Lys			
115	120	125	
ttt gtt gga caa ctg gat att agt att gcc cgg ctc aag aca tct atg	431		
Phe Val Gly Gln Leu Asp Ile Ser Ile Ala Arg Leu Lys Thr Ser Met			
130	135	140	
aag ttt gtc aag aat gtg ata gca ggg att agg gaa act gaa gag aaa	479		
Lys Phe Val Lys Asn Val Ile Ala Gly Ile Arg Glu Thr Glu Glu Lys			
145	150	155	
ttc tac tac atc gtg cag gaa gag aag aac tac agg gaa tcc cta acc	527		
Phe Tyr Tyr Ile Val Gln Glu Glu Lys Asn Tyr Arg Glu Ser Leu Thr			
160	165	170	
cac tgc agg att cgg ggt gga atg cta gcc atg ccc aag gat gaa gct	575		
His Cys Arg Ile Arg Gly Gly Met Leu Ala Met Pro Lys Asp Glu Ala			
175	180	185	190
gcc aac aca ctc atc gct gac tat gtt gcc aag agt ggc ttc ttt cgg	623		
Ala Asn Thr Leu Ile Ala Asp Tyr Val Ala Lys Ser Gly Phe Phe Arg			
195	200	205	

gtg ttc att ggc gtg aat gac ctt gaa agg gag gga cag tac atg ttc	671		
Val Phe Ile Gly Val Asn Asp Leu Glu Arg Glu Gly Gln Tyr Met Phe			
210	215	220	
aca gac aac act cca ctg cag aac tat agc aac tgg aat gag ggg gaa	719		
Thr Asp Asn Thr Pro Leu Gln Asn Tyr Ser Asn Trp Asn Glu Gly Glu			
225	230	235	
ccc agc gac ccc tat ggt cat gag gac tgt gtg gag atg ctg agc tct	767		
Pro Ser Asp Pro Tyr Gly His Glu Asp Cys Val Glu Met Leu Ser Ser			
240	245	250	
ggc aga tgg aat gac aca gag tgc cat ctt acc atg tac ttt gtc tgt	815		
Gly Arg Trp Asn Asp Thr Glu Cys His Leu Thr Met Tyr Phe Val Cys			
255	260	265	270
gag ttc atc aag aag aaa aag taactccct catcctacgt atttgctatt	866		
Glu Phe Ile Lys Lys Lys Lys			
275			
ttcctgtgac cgtcattaca gttattgtta tccatccttt tttcctgat tgtactacat	926		
ttgatctgag tcaacatagc tagaaaatgc taaaactgagg tatggagcct ccatcatcat	986		
gctctttgt gatgatttc atatttcac acatggatg ttattgaccc aataactcgc	1046		
caggttacat gggcttgag agagaatttt aattactaat tgtgcacgag atagttggtt	1106		
gtctatatgt caaatgagtt gttctcttg tattgctct accatctctc cctagagcac	1166		
tctgtgtcta tcccagtgga taatttcca gtttactggt gatgattagg aaggttgtt	1226		
atggtaggc taacctgccc tggcccaaag ccagacatgt acaagggtt tctgtgagca	1286		
atgataagat cttgaatcc aagatgccc gatgtttac cagtcacacc ctatggccat	1346		
ggctatactt ggaagttctc cttgttgca cagacataga aatgctttaa ccccaagcct	1406		
ttatatgggg gacttctagc tttgtgtctt gtttcagacc atgtggaatg ataaataactc	1466		
ttttgtgct tctgatctat cgatttcaact aacatatacc aagtaggtgc tttgaacccc	1526		
tttctgtagg ctcacaccc tt aatctcaggc ccctatatacg tcacacttg atttaagaaa	1586		
aacggagcc	1595		

<210> 2

<211> 277

<212> PRT

<213> Homo Sapiens

25

<400> 2

Met	Asn	Gly	Phe	Ala	Ser	Leu	Leu	Arg	Arg	Asn	Gln	Phe	Ile	Leu	Leu
1															15

Val	Leu	Phe	Leu	Leu	Gln	Ile	Gln	Ser	Leu	Gly	Leu	Asp	Ile	Asp	Ser
														30	

Arg	Pro	Thr	Ala	Glu	Val	Cys	Ala	Thr	His	Thr	Ile	Ser	Pro	Gly	Pro
															45

Lys	Gly	Asp	Asp	Gly	Glu	Lys	Gly	Asp	Pro	Gly	Glu	Gly	Lys	His
50				55					60					

Gly	Lys	Val	Gly	Arg	Met	Gly	Pro	Lys	Gly	Ile	Lys	Gly	Glu	Leu	Gly
															80
65				70					75						

Asp	Met	Gly	Asp	Arg	Gly	Asn	Ile	Gly	Lys	Thr	Gly	Pro	Ile	Gly	Lys
															95
85								90							

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

Gly	Lys	Ala	Gly	Thr	Val	Cys	Asp	Cys	Gly	Arg	Tyr	Arg	Lys	Phe	Val
															125
115							120								

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

Val	Lys	Asn	Val	Ile	Ala	Gly	Ile	Arg	Glu	Thr	Glu	Glu	Lys	Phe	Tyr
															160
145							150				155				

Tyr	Ile	Val	Gln	Glu	Glu	Lys	Asn	Tyr	Arg	Glu	Ser	Leu	Thr	His	Cys
															175
165								170							

Arg	Ile	Arg	Gly	Gly	Met	Leu	Ala	Met	Pro	Lys	Asp	Glu	Ala	Ala	Asn
															190
180								185				190			

Thr	Leu	Ile	Ala	Asp	Tyr	Val	Ala	Lys	Ser	Gly	Phe	Phe	Arg	Val	Phe
															205
195								200				205			

26

Ile Gly Val Asn Asp Leu Glu Arg Glu Gly Gln Tyr Met Phe Thr Asp
 210 215 220

Asn Thr Pro Leu Gln Asn Tyr Ser Asn Trp Asn Glu Gly Glu Pro Ser
 225 230 235 240

Asp Pro Tyr Gly His Glu Asp Cys Val Glu Met Leu Ser Ser Gly Arg
 245 250 255

Trp Asn Asp Thr Glu Cys His Leu Thr Met Tyr Phe Val Cys Glu Phe
 260 265 270

Ile Lys Lys Lys Lys

275

<210> 3

<211> 27

<212> PRT

<213> Artificial Sequence

<220>

<223> Consensus sequence of collectins which were reported
 heretofore.

<400> 3

Glu Lys Cys Val Glu Met Tyr Thr Asp Gly Lys Trp Asn Asp Arg Asn
 5 10 15

Cys Leu Gln Ser Arg Leu Ala Ile Cys Glu Phe
 20 25

<210> 4

<211> 14

<212> DNA

<213> Artificial Sequence

<220>

<223> Pre-sequence of an Insert.

<400> 4

gaattcggca cgag

14

<210> 5

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> M13 Universal Primer Sequence for Sequencing.

<400> 5

cgacgttgta aaacgacggc cagt

24

<210> 6

<211> 17

<212> DNA

<213> Artificial Sequence

<220>

<223> M13 Reverse Primer Sequence for Sequencing.

<400> 6

caggaaacag ctatgac

17

<210> 7

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Sequence of a Reverse Primer for Screening a Novel
Collectin.

<400> 7

tttgatgga ggctccatac c

21

<210> 8

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Sequence of a Forward Primer for Screening a Novel
Collectin.

<400> 8

ctgccaacac actcatcgct g

21

<210> 9

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> Sequence of a fÉgt11 Reverse Primer for Sequencing.

<400> 9

ttgacaccag accaactggtaatg

24

<210> 10

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> Sequence of a fÉgt11 Forward Primer for Sequencing.

<400> 10

ggtgtggcgacg actcctggag ccccg

24

<210> 11
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Sequence of a Synthetic Primer for Sequencing a Novel
Collectin.

<400> 11
taatggtagc gaccggcgct

20

<210> 12
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Sequence of a Synthetic Primer for Sequencing a Novel
Collectin.

<400> 12
aaacccaattt atactcctgg

20

<210> 13
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Sequence of a Synthetic Primer for Sequencing a Novel
Collectin.

<400> 13
aatattggca agactgggcc

20

30

<210> 14
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Sequence of a Synthetic Primer for Sequencing a Novel
Collectin.

<400> 14
gatgagtgtg ttggcagcat

20

<210> 15
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Sequence of a Synthetic Primer for Sequencing a Novel
Collectin.

<400> 15
gtatcttcca caatcacaga

20

<210> 16
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Sequence of a Synthetic Primer for Sequencing a Novel
Collectin.

<400> 16
ttaattcctt tcggcccat

20

<210> 17
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Sequence of a Synthetic Primer for Sequencing a Novel
Collectin.

<400> 17
gcaaaaagaaaa tagtaccagg

20

<210> 18
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Sequence of a Synthetic Primer for Sequencing a Novel
Collectin.

<400> 18
catatcaccc agttctcctt

20

<210> 19
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Sequence of a Synthetic Primer for Sequencing a Novel
Collectin.

<400> 19
agcaggatt agggaaactg

20

<210> 20
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Sequence of a Synthetic Primer for Sequencing a Novel
Collectin.

<400> 20
ctgtgagcgt cattacagtt

20

<210> 21
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Sequence of a Synthetic Primer for Sequencing a Novel
Collectin.

<400> 21
ggttgtctat atgtcaaatg

20

<210> 22
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Sequence of a Synthetic Primer for Sequencing a Novel
Collectin.

<400> 22
tatggccatg gctatacttg

20

<210> 23
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Sequence of a Synthetic Primer for Sequencing a Novel
Collectin.

<400> 23
atcgctgact atgttgccaa

20

<210> 24
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Sequence of a Synthetic Primer for Sequencing a Novel
Collectin.

<400> 24
caagtatagc catggccata

20

<210> 25
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Sequence of a Synthetic Primer for Sequencing a Novel
Collectin.

<400> 25
aactgtaatg acgctcacag

20

<210> 26

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Sequence of a Synthetic Primer for Sequencing a Novel
Collectin.

<400> 26

catttgacat atgaacaacc

20

CLAIMS:

1. An isolated polynucleotide consisting of the 6th to the 836th contiguous nucleotides in the nucleotide sequence as set out in SEQ ID NO:1.

5

2. The polynucleotide according to Claim 1 which is a cDNA.

3. A method for isolating the polynucleotide according to Claim 1, the method comprises the steps of (1) preparing a probe which is 10 complementary to the polynucleotide according to Claim 1, (2) hybridizing the probe with a candidate polynucleotide at 55°C in a hybridization solution comprising 5×SSC, 1% blocking agent, 0.1% N-Lauroyl sarcosine and 0.02% SDS, (3) removing any non-hybridized probe at 55°C with a washing solution comprising 2×SSC, and (4) isolating the hybridized polynucleotide.

15

4. A polypeptide consisting of the amino acid sequence as set out in SEQ ID NO:2.

5. The polypeptide according to Claim 4 which is encoded by the 6th 20 to the 836th contiguous nucleotides in the nucleotide sequence as set out in SEQ ID NO:1.

1/11

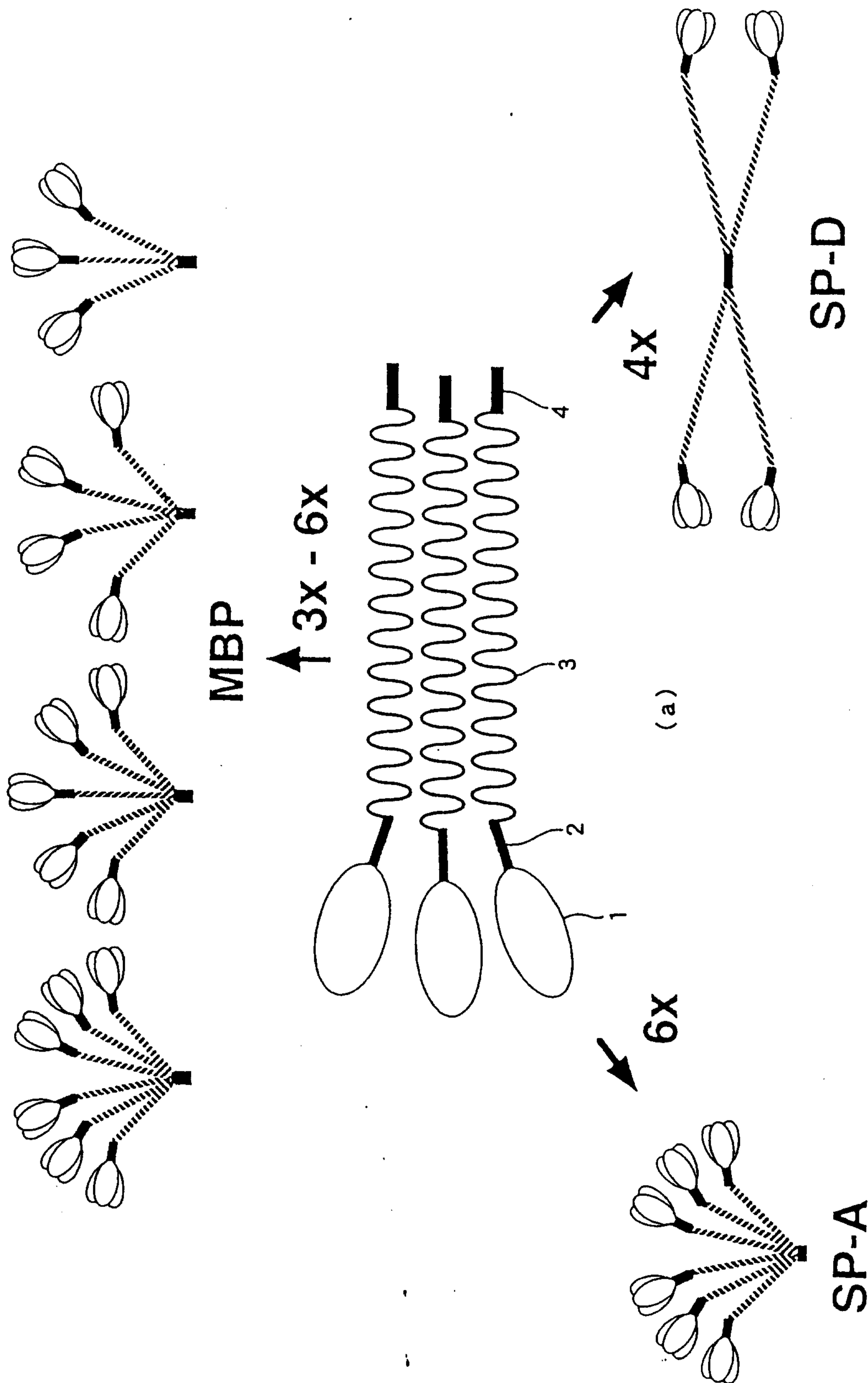


Fig. 1

2/11

MBP	MSLFPS-IPPLILSMVAASYSETVTCEDAQKI-----CPAVIACSS--PGINNGFPGKDGTDGKGEKGEPE
SP-A	MWLCPLAIIITIIMA-----ASGAACEVKDVCV-----GSPG
SP-D	MLFLH-SALVLIQ-PLCAYLEAEMKTYSHRTTPSACTIV-MCSSVESGLPGRDGRDGRCPRGEKGDPG 70

	IPCTPGSHGLPGRDGR-----DGVKGDPGPPMGP-----ETP-----
	IPGAAGQAGMPCQAGPVGPKGDNGSVGEPGPKGDTGPGVPGPAGRECP-----GPKLQGNIGPQGKPGP 140

	QGLRGHQGP-----PGKLGPPGNPGPSGSPGPKCOK
	CPPCNNGLPGAPGVPG-----RGEKGEPPGERGPPGL-----
	KGEAGPKGEVGAAPGMQGSAGARGIAAGPKGERGVPGSPGARGPPGLKGDK 210

Fig. 2

3/11

MBP	[G]PG-[K]SPD[G]DSS[AA]---	-SERKAL[QTEMARIK]KWL[TF]SI[G]KQVGNKFFLTNGEIMTF[E]KV	
SP-A	--PAHLD[EE]LQATL[H]D---	[FRH]Q[IL]QTRGALS-[L]Q[SI]---	MTVGEKVFSSNGQSITF[DAI]
SP-D	[G]IP[G]DK[G]AK[G]ES[G]LPD[V]ASL[R]QQ[VEAL]Q[GOV]Q[HL]Q[A]AFS[QY]KK[VE]LFP[N]G[OSV]GERK[IF]K[IA]GFV[K]P[TEA]	280.	
		KAL[CV]VKFQASVAT[PRNAENGAI]Q[N]I---	KEEAFLGITD[E]KTE[G]QFV[D]LITGNR[IT]YTNWNEGEPNNAGS
		[Q]EAC[AR]AG[G]RI[AV]PRN[PEE]NEAIASFV[K]YNT[AY]V[G]LIE[G]PSP[G]DFRY[SD]GTPVNYT[N]WYRGE[P]AGR[G]	[Q]LCTQAGGQI[ASPR]SAENNA[AL]Q[QL]VAKNEAAFLSMTDSKTE[G]K[FT]YPTGESI[M]YS[N]WAP[G]EP[N]D[G]GG
			350
		DED[C]V[LL]K[N]GQWN[D]VPC[ST]SH[AVCE]FPI*	
		[KE]QCVEM[Y]ID[G]QWN[D]RNCLY[S]R[IT]ICEF*---	
		SEDCVEIIFTNGKWNDR[AC]GEKRIWVCEF*---	

Fig. 3

4/11

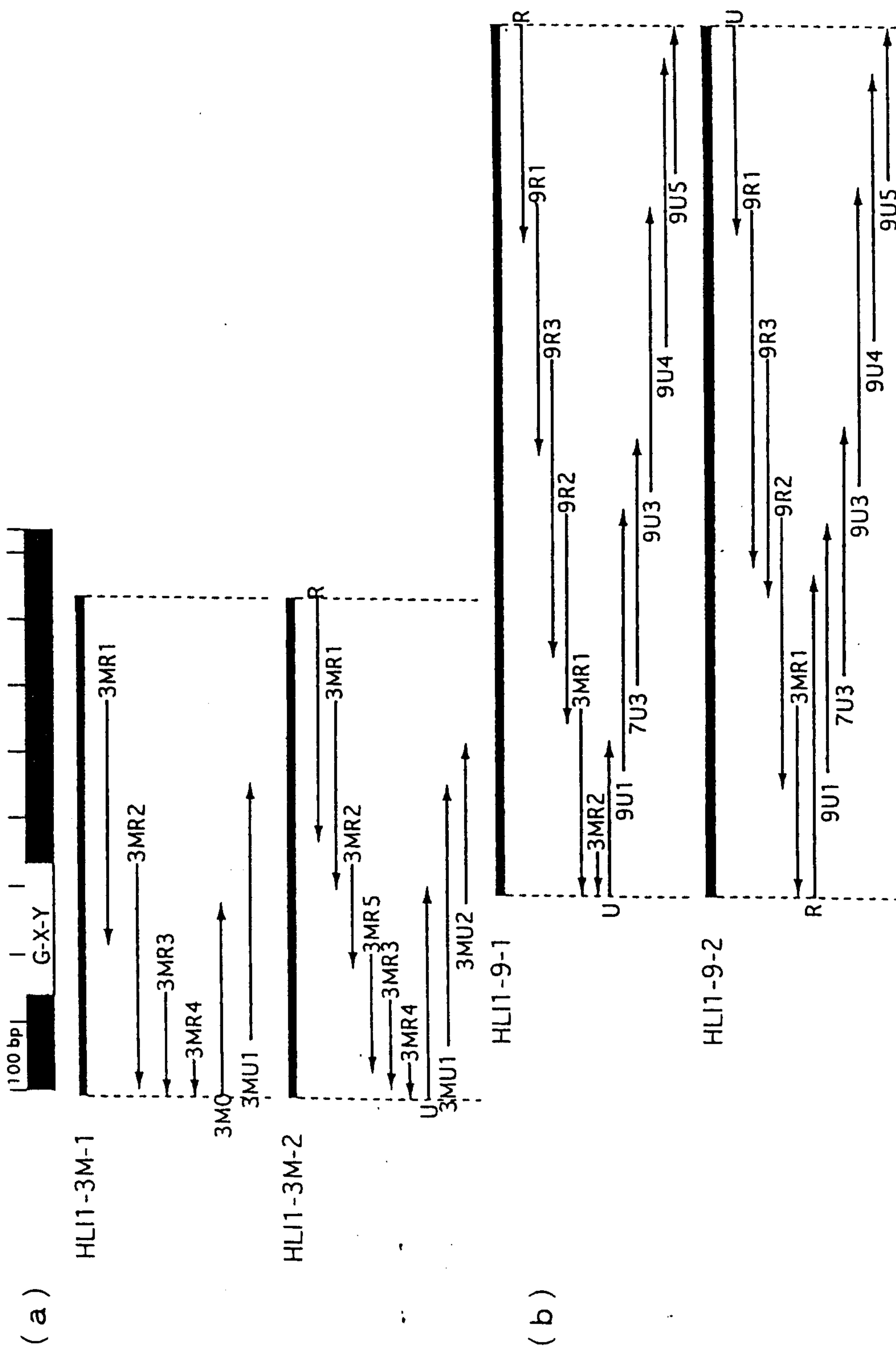


Fig. 4

5/11

MSLFPS-1---P₁LLSMVAAASYSEtvTCEDAQK1---CPAVIACSS--PGINGPGRDGT₁GEKG
 MSLIOPLA₁---T₁IMA---ASGAACEVKDV₁V---G
 SP-A MSLIOPLA₁---T₁IMA---
 SP-D MSLIOPLA₁-S---AVLITQ-P₁Cyl₁EAEMKTYSHRTTPSACTIV-MCSSVESGLPGRDGRDGT₁GEKG
 novel collectin MNGFASLLRRNQFILLVLFL₁Q-IQSLGLDIDSRP₁TAEVCAHTIISP---GPKGD₁DG 70

EPG SPGIPGTPGSHG₁PG₁GRDGR-
 DPGIPGAAGQAGMPGQAGPVGPKGDNNGS₁VGEPGPKGDI₁GPSGPPGVPGPAGREGPLGKQGNIGPQGK
 EKGDPGEE₁SKH₁---KVGRMCPKGIKGE₁LGDRGN₁IGKTGPI₁GK 140

QGIRGLQGP-
 CPPGNNGLP₁GAPGV₁GE---P₁GKLGPPGPNP₁GSP₁GPK
 PGPKGEEVGA₁GPKGEVGA₁GMQGSAGARGLAC₁GPKGERGVPGNAGAAGSAGAMGPQGSPGARGPPGLK
 -KGDKGEKGLLGI₁PEKGKAG- 210

Fig. 5

6/11

MBP	GGKGDPG-KSPDGSSTAA	SERKALQTEMARIKKWLTSLGKQVGNKFFLTNGEIMTF
SP-A	---PAHLDEELQATLHD	---FRHQIILQTRGALS-LQGSII-----MTVGEKVKFSSNGQSITF
SP-D	GDKGIPGDKGAKGESGIPDVASIRQVEALQGQVQHILQAAFSQYKKVVELFPNGQSVGEKIEFKIAGFVKPF	novel collectin --- RITSMKFVKN -- VVAGIRETEEKFFYYIVQEEFKNY 280
		EKVKAALCVKEQASVATPRNAENGATIONI --- KEE -AFLGITDEKTEGQFVDLIGNRIT-YTNWNNEGEP
		DAIQEACARAGGRIAVPRNPEENEAIASFVIKIVNTY-AYVGLITEGPSPGDFRYSDGTPVN-YTNWYRGEPEP
		TEAQCQCTQAGGQLASPRSAANAAALQQLVAKNEA-AFLSMTDSKTEGKFIIYPTGESEV-YSNWAPGEPEP
		RESLTHCRIRGGMAMPKDEAANTLIADYMAKSFFRVEIGVNDLEREGQYMFIDNTPLQNYSNWNEGEP 350
		NNAGSDEDCCVLLIKNGQWNDVPCSTSHIAVCEFPI*---
		AGRQG-KFQCVEMYIDGQWNDRNCLYSRITICEF*---
		NDDGGSEDCVEIETNCKWNDRACGEKRIWCEF*---
		SDPYGHEDDCVEMLSSGRWNDRTECHILTMYFVCEFIKKKK*

Fig. 6

7/11

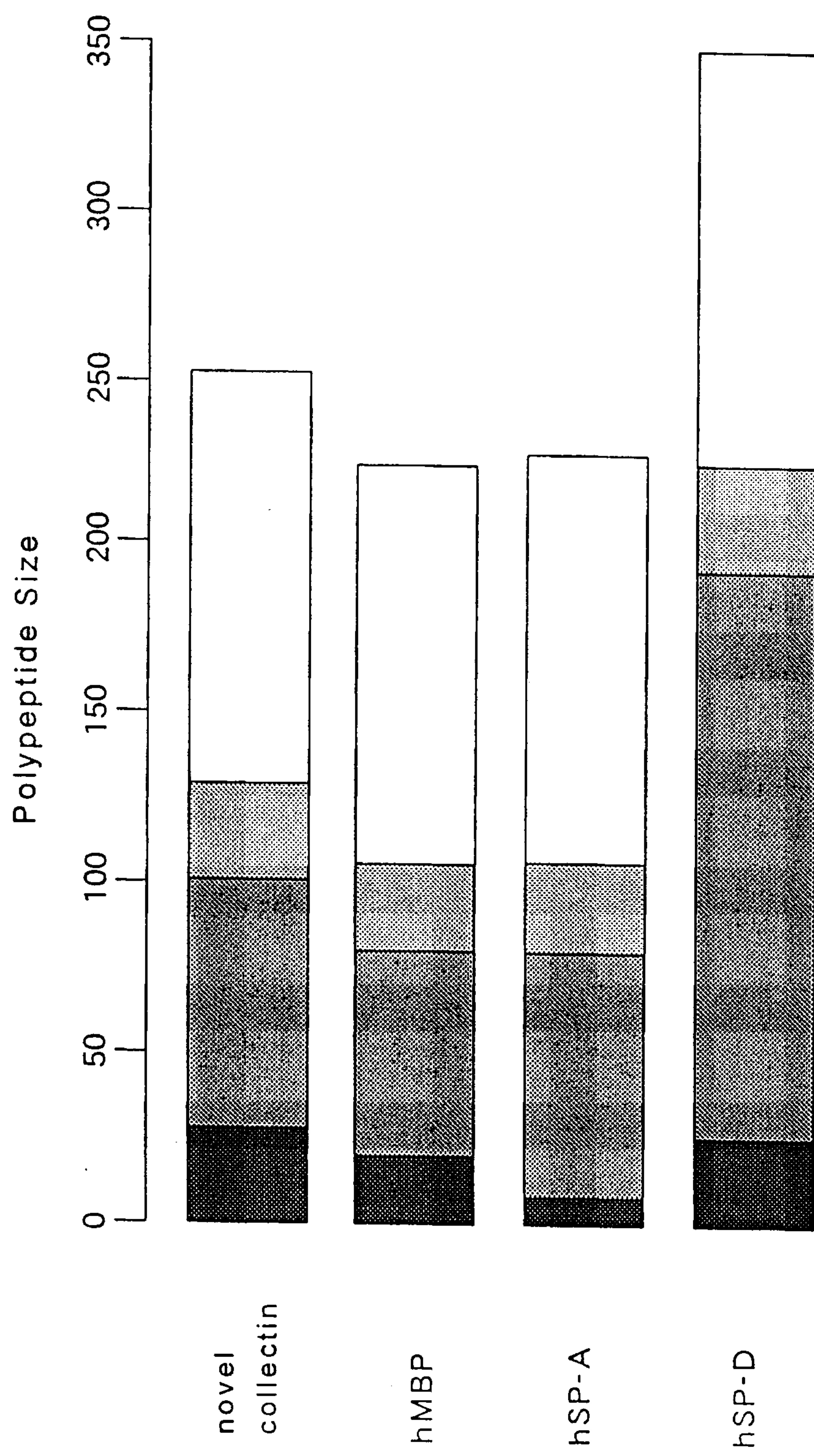


Fig. 7

8/11

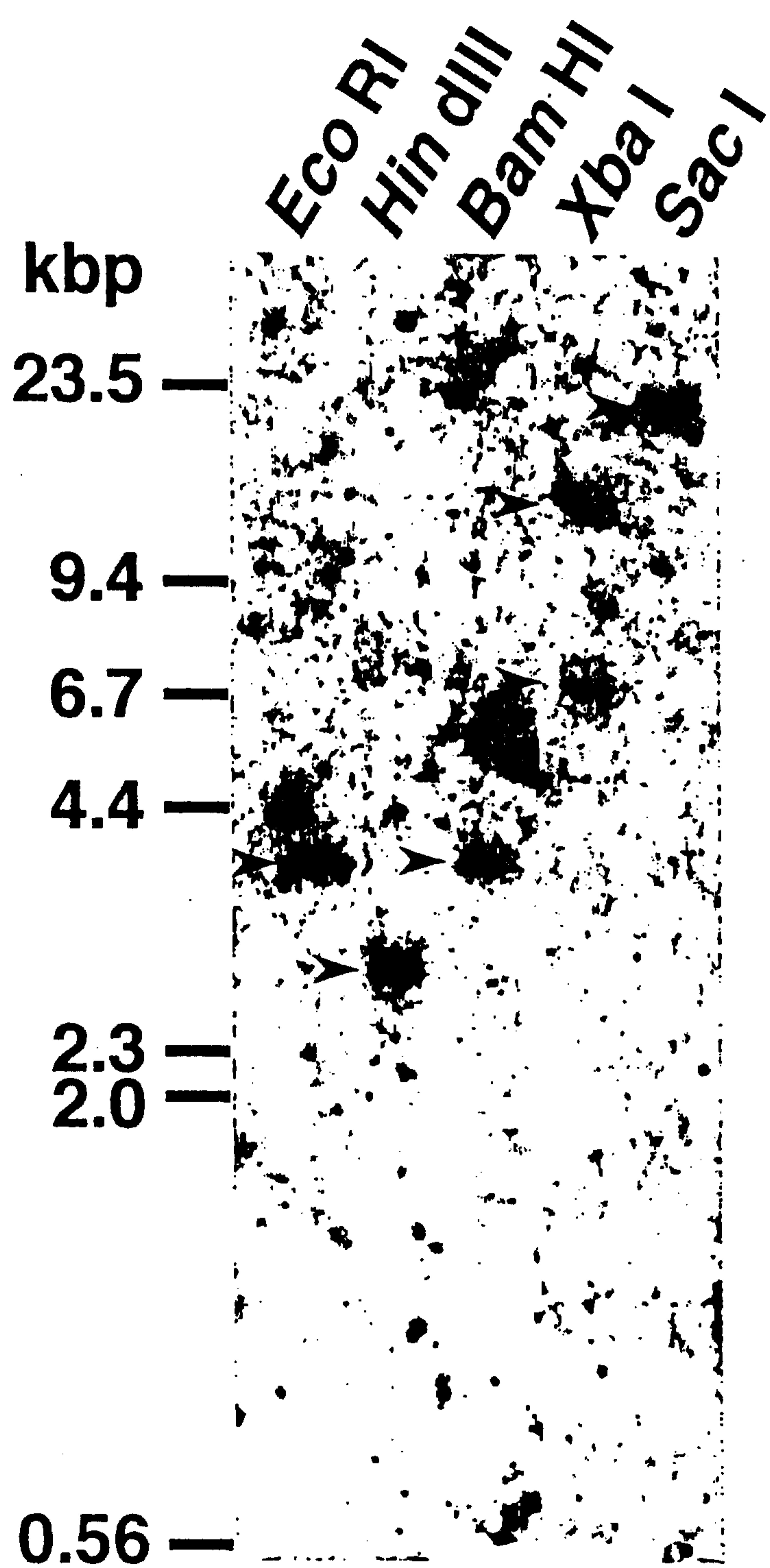


Fig.8

9/11

a b c d e f g h

kb**9.5 —****7.5 —****4.4 —****-3.8****2.4 —****1.35 —****-1.2**

Fig. 9

10/11

a b c d e f g h

kbp

23.9 —

9.4 —

6.7 —

4.4 —

2.3 **2.0** **==**

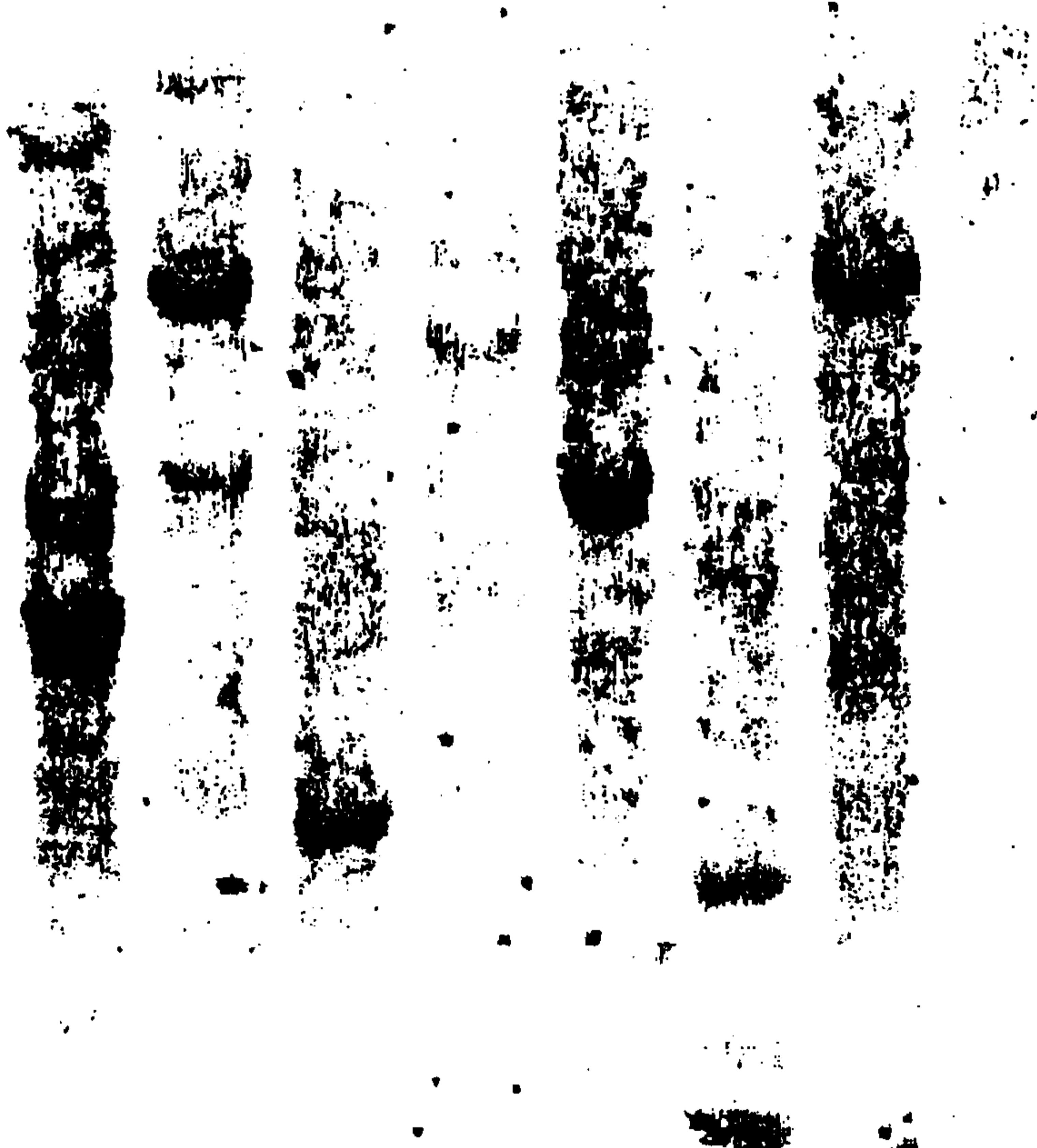


Fig.10

11/11

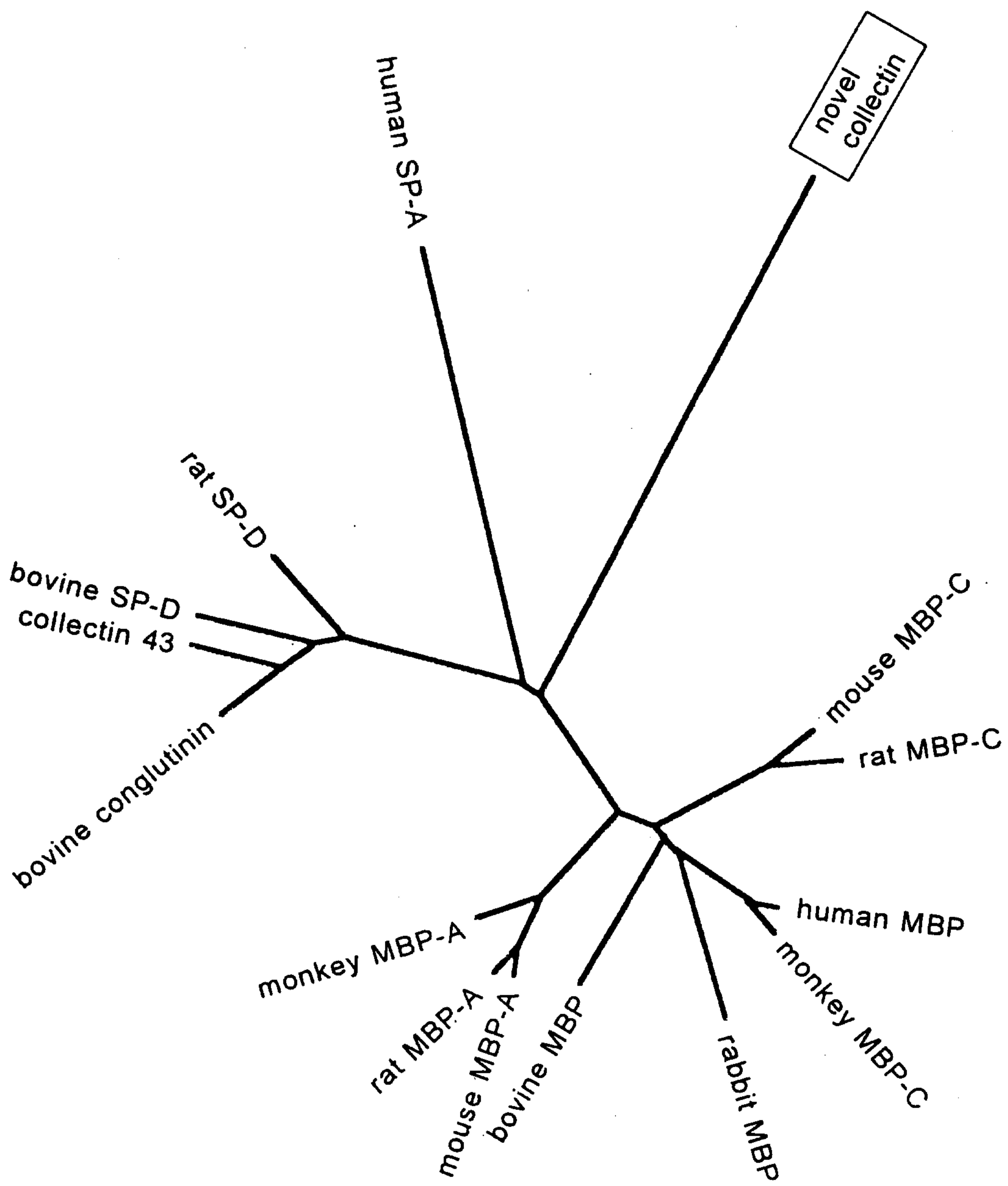


Fig. 11