

(12) PATENT
(19) AUSTRALIAN PATENT OFFICE

(11) Application No. **AU 199734598 B2**
(10) Patent No. **720930**

(54) Title
Differentiation inhibitor

(51)⁷ International Patent Classification(s)
C07K 014/52 C12N 015/12
A61K 038/19 C12P 021/02
C07K 016/24 C12P 021/08
C12N 005/10

(21) Application No: **199734598**

(22) Application Date: **1997.07.11**

(87) WIPO No: **WO98/02458**

(30) Priority Data

(31) Number	(32) Date	(33) Country
8-186220	1996.07.16	JP
9-124063	1997.05.14	JP

(43) Publication Date : **1998.02.09**

(43) Publication Journal Date : **1998.04.02**

(44) Accepted Journal Date : **2000.06.15**

(71) Applicant(s)
Asahi Kasei Kogyo Kabushiki Kaisha

(72) Inventor(s)
Seiji Sakano; Akira Itoh

(74) Agent/Attorney
**WATERMARK PATENT and TRADEMARK ATTORNEYS, Locked Bag 5, HAWTHORN
VIC 3122**

PC

OPI DATE 09/02/98 APPLN. ID 34598/97
AOJP DATE 02/04/98 PCT NUMBER PCT/JP97/02414

AU9734598



(51) 国際特許分類6 C07K 14/52, C12P 21/02, A61K 38/19, C12N 5/10, 15/12, C07K 16/24, C12P 21/08 // (C12P 21/02, C12R 1:91) (C12N 5/10, C12R 1:91) (C12P 21/08, C12R 1:91)	A1	(11) 国際公開番号 WO98/02458 (43) 国際公開日 1998年1月22日(22.01.98)
(21) 国際出願番号 PCT/JP97/02414 (22) 国際出願日 1997年7月11日(11.07.97) (30) 優先権データ 特願平8/186220 1996年7月16日(16.07.96) JP 特願平9/124063 1997年5月14日(14.05.97) JP (71) 出願人 (米国を除くすべての指定国について) 旭化成工業株式会社 (ASAHI KASEI KOGYO KABUSHIKI KAISHA)[JP/JP] 〒530 大阪府大阪市北区堂島浜1丁目2番6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) 坂野誠治(SAKANO, Seiji)[JP/JP] 〒416 静岡県富士市鮫島351-1 Shizuoka, (JP) 伊藤 章(ITO, Akira)[JP/JP] 〒416 静岡県富士市中丸206-9 Shizuoka, (JP) (74) 代理人 弁理士 小林和憲(KOBAYASHI, Kazunori) 〒170 東京都豊島区北大塚2丁目25番1号 太陽生命大塚ビル3階 Tokyo, (JP)		(81) 指定国 AU, CA, JP, US, 欧州特許 (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE). 添付公開書類 国際調査報告書 請求の範囲の補正の期限前であり、補正書受領の際には再公開される。
(54)Title: DIFFERENTIATION INHIBITOR (54)発明の名称 分化抑制剤 (57) Abstract A novel human serrate-2 polypeptide consisting of a polypeptide containing the amino acid sequence described in SEQ ID NO:1 in the Sequence Listing and having the effect of regulating the differentiation of undifferentiated cells involving stem cells; its gene; a process for producing the same; and an antibody specifically recognizing the polypeptide.		

SPECIFICATION

Differentiation-suppressive polypeptide

Field of the invention

This invention relates to a novel bioactive substance which suppresses differentiation of undifferentiated cells.

Prior art

Human blood and lymph contain various types of cells and each cell plays important roles. For example, the erythrocyte carries oxygen; platelets have hemostatic action; and lymphocytes prevent infection. These various cells originate from hematopoietic stem cells in the bone marrow. Recently, it has been clarified that the hematopoietic stem cells are differentiated to various blood cells, osteoclasts and mast cells by stimulation of various cytokines in vivo, and environmental factors. In the cytokines, there have been found, for example, erythropoietin (EPO) for differentiation to erythrocytes; granulocyte colony stimulating factor (G-CSF) for differentiation to leukocytes; and platelet growth factor (mpl ligand) for differentiation to megakaryocytes which is a platelet producing cell, and the former two have already been clinically applied.

The undifferentiated blood cells are generally classified into two groups consisting of blood precursor cells which are destined to differentiate to specific blood series, and hematopoietic stem cells which have differentiation ability to all series, and self-replication activity. The blood precursor cells can be identified by various colony assays, however identification methods for the hematopoietic stem cells have not been established. In these cells, stem cell factor (SCF), interleukin-3 (IL-3), granulocyte-macrophage colony stimulating factor (GM-CSF), interleukin-6(IL-6), interleukin-1(IL-1), granulocyte colony stimulating factor (G-CSF) and oncostatin M have been reported to stimulate cell differentiation and proliferation.

Trials for expansion of hematopoietic stem cells in vitro have been examined in order to replace bone marrow transplantation with applying hematopoietic stem cell transplantation therapy or gene therapy. However, when the hematopoietic stem cells are cultured in the presence of the above



mentioned cytokines, multi-differentiation activities and self-replication activities, which are originally present in hematopoietic stem cells, gradually disappeared and the stem cells changed to blood cell precursors which only differentiate to specific series after 5 weeks of cultivation, and multi-differentiation activity, which is one of the specific features of the hematopoietic stem cells, is lost (Wagner et al. Blood 86, 512-523, 1995).

For proliferation of blood precursor cells, a single cytokine is not sufficient to cause the effect, rather synergistic action of several cytokines are important. Consequently, in order to proliferate the hematopoietic stem cells and maintain the specific features of hematopoietic stem cells, it is necessary to add cytokines which suppress differentiation, together with the cytokines which proliferate and differentiate the undifferentiated blood cells. In general, many cytokines which stimulate proliferation or differentiation of cells are known, but only small numbers of cytokines which suppress cell differentiation are known. For example, leukemia inhibitory factor (LIF) has an action of proliferation of mouse embryonic stem cells without differentiation, but it has no action against the hematopoietic stem cells or blood precursor cells. Transforming growth factor (TGF - β) has suppressive action for proliferation against various cells, but has no fixed actions against the hematopoietic stem cells or blood precursor cells.

Not only blood cells, but also undifferentiated cells, especially stem cells, are thought to be involved in tissue regeneration. The regeneration of tissues and proliferation of undifferentiated cells in each tissue can be achieved in various ways by referring to the known reference (Katsutoshi Yoshizato, Regeneration- a mechanism of regeneration, 1996. Yodosha Publ. Co.).

Notch is a receptor type membrane protein which is involved in the regulation of nerve cell differentiation, found in Drosophila. Homologues of Notch are found in various animal kinds in the invertebrate and vertebrate including nematoda (Lin-12). Xenopus laevis (Xotch), mouse (Motch) or human (TAN-1).

Ligands of Notch in Drosophila are known. These are Drosophila Delta (Delta) and Drosophila Serrate (Serrate). Notch ligand homologues are found in



various animal kinds similarly to the Notch receptor (Artavanis-Tsakonas et al., Science 268, 225-232, 1995).

Human Notch homologue, TAN-1 is found widely in tissues in vivo (Ellisen et al., Cell 66, 649-661, 1991). Two Notch analogous molecules other than TAN-1 are reported (Artavanis-Tsakonas et al., Science 268, 225-232, 1995). Expression of TAN-1 was also observed in CD34 positive cells in blood cells by PCR (Polymerase Chain Reaction) (Milner et al., Blood 83, 2057-2062, 1994). However, in relation to humans, gene cloning of human Delta and human Serrate, which are thought to be the Notch ligand, have not been reported.

In *Drosophila* Notch, binding with the ligand was studied and investigated in detail, and it was found that Notch can be bound to the ligand with Ca^{++} at the binding region, which is a repeated amino acid sequence No. 11 and No. 12 in the amino acid sequence repeat of Epidermal Growth Factor (EGF) like repeating sequence (Fehon et al., Cell 61, 523-534, 1990, Rebay et al., *ibid.* 67, 687-699, 1991 and Japan. Patent PCT Unexam. Pub1. 7-503123). EGF-like repeated sequences are conserved in Notch homologues in other species. Consequently, the same mechanism of binding with ligand is predicted.

An amino acid sequence which is called as DSL (Delta-Serrate-Lag-2) near the amino acid terminus, and EGF-like repeated sequence, like in the receptor, are conserved in the ligand (Artavanis-Tsakonas et al., Science 268, 225-232, 1995). EGF-like sequence has been found in thrombomodulin (Jackman et al., Proc. Natl. Acad. Sci. USA 83, 8834-8838, 1986), low density lipoprotein (LDL) receptor (Russell et al., Cell 37, 577-585, 1984), and blood coagulating factor (Furie et al., Cell 53, 505-518, 1988), and is thought to play important roles in extracellular coagulation and adhesion.

The vertebrate homologues of the cloned *Drosophila* Delta were found in chicken (C-Delta-1) and *Xenopus laevis* (X-Delta-1), and it has reported that X-Delta-1 had acted through Xotch in the generation of the protoneuron (Henrique et al., Nature 375, 787-790, 1995 and Chitnis et al., *ibid.* 375, 761-766, 1995).

Vertebrate homologue of *Drosophila* Serrate was found in rat as rat Jagged (Jagged)(Lindsell et al., Cell 80, 909-917, 1995). According to Lindsell et



al., mRNA of the rat Jagged is detected in the spinal cord of fetal rats. As a result of cocultivation of a myoblast cell line, that is, forced excess of expressed rat Notch with a rat Jagged expression cell line, suppression of differentiation of the myoblast cell line is found. However, the rat Jagged has no action against the myoblast cell line without forced expression of the rat Notch.

Considering in the above reports, the Notch and the ligand thereto may be involved in the differentiation regulation of nerve cells, however, except in some myoblast cells, their actions against cells including blood cells, especially primary cells, are unknown.

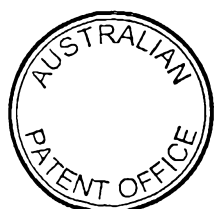
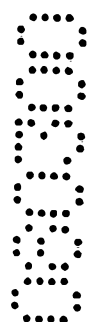
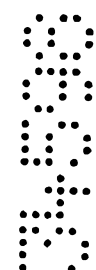
Problem to be solved by the invention

As above mentioned, concerning undifferentiated cells, proliferation for maintaining their specificities are not established. Major reasons are that factors suppressing differentiation of the undifferentiated cells have not been identified. The problem of the present invention is to provide a compound originating from novel factors, which can suppress differentiation of the undifferentiated cells.

Means for solving the problem

We have set up a hypothesis that Notch and its ligand have an action of differential regulation, not only for neurogenic cells, but also for various undifferentiated cells. However, for clinical application to humans, use of prior known different species such as chicken or *Xenopus laevis* type Notch ligand, will result in problems such as species specificities and antigenicities. Consequently, to obtain a prior unknown human Notch ligand is required. We postulated that ligands of human Notch (TAN-1 etc.), which are a human Delta homologue (hereinafter designated as human Delta), and human Serrate homologue (hereinafter designated as human Serrate), may be found. Also we postulated that these findings may result in a candidate drug useful for differential regulation of the undifferentiated cells. We have tried to identify the same.

In order to identify human Notch ligands, we have analyzed amino acid sequences which are conserved in animals other than humans, and sought to identify genes using PCR using mixed primers of the corresponding DNA sequence. As a result of extensive studies, we have succeeded in isolation of



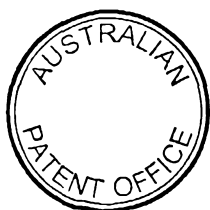
cDNAs encoding amino acid sequences of two new molecules, novel human Delta-1 and novel human Serrate-1, and have prepared expression systems using proteins having various forms using these cDNAs. Also, we have established purification methods for the proteins which were purified and isolated, and included these in patent application (International Publication WO 97/19172).

Further, we have sought to identify Drosophila Delta and Serrate analogous molecules other than human Delta and human Serrate (hereinafter designated as human Delta-1 and human Serrate-1, respectively) of the above patent application in the vertebrate.

We searched the data base of genetic sequences. Namely, based on the human Serrate-1 genetic sequence (amino acids sequence in SEQ ID NO: 5) which was first identified by the authors, we have found several of gene fragments (of length 200 - 350 bp) with high homology to EST (Expressed Sequence Tag), using the data base of gene fragments of random human cDNA sequences in the gene sequence data base GenBank, by using a gene sequence search software Genetyx/CD (Software Development Co.).

These short length gene fragments were cloned by PCR, and these gene fragments were used as probes to attempt to clone longer length gene fragments from human fetal cDNA libraries. The thus isolated longer gene fragments, of which the genetic sequences were determined, were again compared with the genetic sequence of human Serrate-1. As a result, a gene, which has relatively high homology with human Serrate-1, was identified and was designated human Serrate-2. The full length of Serrate-2 gene was isolated successfully.

Further, expression vectors of the said cloned Serrate-2 have been constructed. Purification methods of these proteins were established and the said protein was purified and isolated. Antibodies against human Serrate-2 were prepared using antigens of the said human Serrate-2, and a purification method of the said antibodies was established. The actions against undifferentiated blood cells were then confirmed. The present invention has completed accordingly.



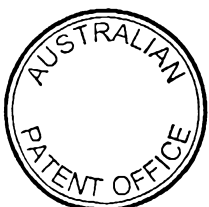
The present invention relates to a polypeptide comprising amino acid sequence of SEQ ID NO: 1, 2 or 3 of the sequence listing and the polypeptide having differentiation suppressive action against undifferentiated cells. Particularly, the polypeptide in which the undifferentiated cells affected are the undifferentiated cells, except for those of the brain and nervous system or muscular system cells, and the polypeptide in which the undifferentiated cells affected are the undifferentiated blood cells. The present invention also relates to the polypeptide having growth inhibitory action against the vascular endothelial cells, a pharmaceutical composition containing the said polypeptides, a cell culture medium containing the said polypeptides, and the cell culture medium in which the cell is the undifferentiated blood cell.

The present invention more particularly relates to a DNA encoding a polypeptide comprising amino acid sequence of SEQ ID NO: 1, 2 or 3 of the sequence listing, the DNA having DNA sequence 90 - 731 DNA sequence 90 - 3254, or DNA sequence 90 - 3725 of SEQ ID NO: 4 of the sequence listing. The present invention still further relates to a recombinant DNA comprising ligating the DNA encoding a polypeptide comprising amino acid sequence of SEQ ID NO: 1, 2 or 3 and a vector DNA which can express the protein in the host cell, and a cell comprising the recombinant DNA.

The present invention also relates to a process for production of the polypeptide by culturing cells and isolating the thus produced compounds, and an antibody specifically recognizing the polypeptide having the amino acid sequence of SEQ ID NO: 1, 2 or 3 of the sequence listing.

The present invention is explained in detail in the following.

Preparation of cDNA necessary for gene manipulation, expression analysis by Northern blotting, screening by hybridization, preparation of recombinant DNA, determination of DNA base sequence and preparation of cDNA library, are all a series of molecular biological experiments, which can be performed according to descriptions in the conventional textbook for the experiments. The above conventional textbook for the experiments is, for example, Maniatis et al. ed. Molecular Cloning. A laboratory manual, 1989, Eds.,



Sambrook, J., Fritsch, E. F. and Maniatis. T., Cold Spring Harbor Laboratory Press.

A novel compound of the present invention has at least polypeptides in the sequence listing SEQ ID NO: 1 - 3. Mutants and alleles which occur naturally are included in the polypeptides of the present invention, unless the polypeptides of the sequence listing, SEQ ID NO: 1, 2 or 3 lose their properties. Modification and substitution of amino acids are described in detail in the patent application by the name of Benntt et al. (National Unexam. Pub1. WO 96/2645) and can be prepared according to the description thereof. A modified polypeptide in the present invention means the modified polypeptide prepared by amino acid replacement and is defined as an amino acid sequence having an identity of more than 90 % in its amino acid sequence.

A DNA sequence encoding polypeptides of the sequence listing, SEQ ID NO:1 - 3 is shown in the sequence listing, SEQ ID NO:4 as well as its amino acid sequence. In these DNA sequences, even if amino acid level mutation is not generated, naturally isolated chromosomal DNA or cDNA thereof may mutate in the DNA base sequence as a result of degeneracy of genetic code without changing the amino acid sequence encoded by the DNA. A 5'-untranslated region and 3'-untranslated region are not involved in amino acid sequence determination of the polypeptide, so DNA sequences of these regions are easily mutated. The base sequence obtained by these degeneracies of genetic codes are included in the DNA of the present invention.

Undifferentiated cells in the present invention are defined as cells which can grow by specific stimulation, and cells which can be differentiated to the cells having specific functions as a result of the specific stimulations. These include undifferentiated cells of the skin tissues, undifferentiated cells of the brain and nervous systems, undifferentiated cells of the muscular systems and undifferentiated cells of the blood cells. These cells include the cells having self-replication activity which are known as stem cells, and the cells having an ability to generate the cells of these lines. The differentiation-suppressive action means suppressive action for autonomous or heteronomous differentiation of the



undifferentiated cells, and is an action for maintaining the undifferentiated condition. The brain and nervous undifferentiated cells can be defined as cells having the ability to differentiate to the cells of the brain or nerve, having specific functions, by specific stimulation. The undifferentiated cells of the muscular systems can be defined as cells having the ability to differentiate to the muscular cells, having specific functions, by specific stimulation. The blood undifferentiated cells in the present invention can be defined as the cell group consisting of the blood precursor cells, which are differentiated to the specific blood series identified by blood colony assay, and hematopoietic stem cells having differentiation to every series, and self-replication activities.

In the sequence listing, amino acid sequence in SEQ ID NO: 1 is a sequence of the active center of the present invention of human Serrate-2 with the signal peptide deleted, and corresponds to an amino acid No. 1 to 217 in SEQ ID NO: 3 of the matured full length amino acid sequence of human Serrate-2 of the present invention. The amino acid sequence in SEQ ID NO: 2 is the amino acid sequence of the extracellular domain of the present invention of human Serrate-2 with the signal peptide deleted, and corresponds to amino acids No. 1 to 1058 in SEQ ID NO: 3 of the matured full length amino acid sequence of human Serrate-2 of the present invention. The amino acid sequence of SEQ ID NO: 3 is the matured full length amino acid sequence of the human Serrate-2 of the present invention. The sequence of SEQ ID NO: 4 is the total amino acid sequence of human Serrate-2 of the present invention and cDNA encoding the same. The sequence of SEQ ID NO: 5 is the total amino acid sequence of human Serrate-1 used in the present invention and cDNA encoding the same.

The left and right ends of the amino acid sequences in the sequence listings indicate amino terminus (hereinafter designated the N-terminus) and carboxyl terminus (hereinafter designated the C-terminus), respectively, and the left and right ends of the nucleotide sequences are 5'-terminus and 3'-terminus, respectively.

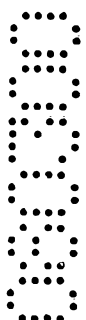
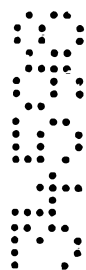
Cloning of the human Notch ligand gene can be performed using the following method. Part of amino acids sequences of the human Notch ligand is



evolutionarily conserved. DNA sequence corresponding to the conserved amino acid sequence is used as a primer for RT-PCR (Reverse Transcription Polymerase Chain Reaction), then a PCR template of the human origin is amplified using a PCR reaction, thereby fragments of human Notch ligand can be obtained. Furthermore, RT-PCR primer is carried out by using the known DNA sequence information of the Notch ligand homologue from organisms other than humans, and the known gene fragments may be obtained from a PCR template of the said organisms.

In order to perform PCR to obtain fragments of a human Notch ligand, PCR for the DSL sequence was considered, however a large number of combinations of DNA sequences corresponding to the amino acid sequence conserved in this region can be expected, and PCR design is therefore difficult. As a result, PCR of the EGF-like sequence has to be selected. As explained hereinbefore, since EGF-like sequence are conserved in a large number of molecules, obtaining the correct fragments and identification of the fragments is extremely difficult. We have designed and prepared about 50 PCR primer sets, for example the primer set of the sequence shown in Preferential example 1, PCR was performed with these primer sets using a PCR template of cDNA prepared from poly A⁺ RNA of various tissues of human origin, and more than 10 PCR products from each tissue were subcloned, as well as performing sequencing for more than 500 types. A clone having a desired sequence of human Serrate-1 was identified.

Namely, as shown in Preferential example 1, the obtained PCR product was cloned into the cloning vector, the host cells were transformed using a recombinant plasmid which contains the PCR product, the host cells containing the recombinant plasmid were cultured in large scale, the recombinant plasmid was purified and isolated, the DNA sequence of the PCR product inserted into the cloning vector was confirmed, and the gene fragment which may have the sequence of human Serrate-1 was compared with the sequence of the known Serrate homologue of other species in an attempt to identify such a sequence. We have succeeded in identifying the gene fragment which contains a part of the



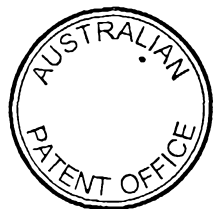
cDNA of human Serrate-1, the same sequence as DNA sequence from 1272 to 1737 described in the sequence listing, SEQ ID NO: 5. As shown in Preferential example 2, using the thus obtained human Serrate-1 gene fragment, full length cDNA is obtained from human cDNA library. We have already made patent applications in respect of these inventions (WO 97/19172).

In the present invention, we sought to identify ligands other than this human Serrate-1 molecule. Gene fragments showing high homology to the ligand to the gene sequence encoding amino acid sequence of the human Serrate-1 molecule, i.e. DNA sequence from 409 to 4062 in the sequence listing SEQ ID NO: 5, are screened in the data base of DNA sequence. Screening was performed by using gene sequence search software Genetyx/CD (Software Development Co.) on the DNA/fragments of random human cDNA sequence data base EST (Expressed Sequence Tag) of Genbank (release 91, 1995).

Recently, DNA sequencing technology has progressed, and analysis of total geneomic DNA and full cDNA sequence of humans, nematoda, Arabidopsis thaliana, Heynh. etc. were tried by random sequencing of genomic DNA and cDNA (Genome Directory, Nature, 377, 3S, 1995). For the human cDNA, EST project of TIGR (The Institute for Genomic Research), EST project of Washington Univ. - Merck, STS project of Colorado Univ. joined in these projects. Partial base sequences of cDNA provided by these organizations are registered in DNA data base of Genbank and EMBL and are disclosed. According to Genbank release 91 of October, 1995, cumulative registered numbers of EST clones are about 330, 000 clones, with average length 346 bp.

Based on data in these data bases, gene sequences or amino acid sequences of known or newly cloned novel molecules were searched using a homology search. The existence of analogues or similar family molecules to these molecules can be known and the sequence information of the partial DNA sequence can be obtained.

For analysis, commercially available analysis software can be used, or analysis software attached to the data base, for example BLAST of Genbank can be used. By using these, analysis can be performed by accessing to National



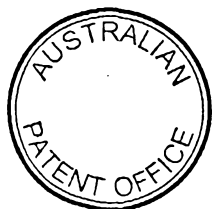
Center of Biotechnology Information, U.S.A., Institute of Chemistry, Kyoto Univ., Japan, through WWW (world wide web) or E-mail.

Gene sequence information of gene fragments (about 200 - 350 bp) with highly similarity to the objective gene can be obtained through these operations. The information of the obtained gene fragment includes general gene sequence information together with the clone name of the gene and organs or tissues from which the gene was extracted. The information on some DNA sequences, is raw data obtained by DNA sequencing, and the DNA sequence is unknown (marked N) and may contain incorrect DNA sequence information. Consequently, the DNA sequence information is not always exact.

From this gene information, DNA sequence not marked N is thought to be highly probable DNA sequence information. Further, the most probable DNA sequences within the DNA sequences identified are compared, and DNA fragments having significant homology were identified in this region (for genes with 200 bases or more, a similarity of the DNA sequence of above 40% is preferable). The thus identified DNA fragment can be obtained from Genom System Inc., U.S.A. etc., if the name of the clone is known. However, because the organ of origin is known, the DNA can be isolated using PCR from cDNA of commercially available expression organ libraries.

The thus identified gene information is partial information, and unless the total information is obtained, a full length amino acid sequence, which may be encoded by the said partial sequence of gene, is not always analogous to the similar molecule used in the original homology search. Exact information about the molecule cannot be determined only using that information. As shown hereinbelow, we have prepared a number of probes having homology to the required sequence and performed cloning by plaque hybridization, however almost all of the DNA fragments identified did not code the desired molecules. Consequently, this technique may be technologically possible, but does not work well.

We have prepared probes of about 50 DNA fragments which showed similarity with human Serrate-1 cDNA, by PCR. Finally, as shown in the



following, cloning was performed using a library screening technique. As a result of determining the DNA sequence, a gene was isolated using 3 types of DNA probes having the sequences in the SEQ ID NO: 8, 9 and 10, which were prepared based on the DNA sequence clones registered in Genbank (Reg. No. T08853, R50026 and R46751) which as described in Example 1, is a DNA fragment coding human Serrate-2 molecule.

Further, using the thus isolated cDNA fragment as a probe the cDNA library of the expression organs is screened, and then the longer gene having DNA sequence or full length cDNA gene can be screened. The full length clone can be made by isotope labeling and non-isotope labeling with the partial cloning gene, and screening the library by hybridization or other method. Isotope labeling can be performed by, for example, terminal labeling by using [³²P] γ -ATP and T4 polynucleotide kinase, or other labeling methods such as nick translation or primer extension method can be applied.

Further, cloning of the full length gene or longer gene fragments can be obtained by methods for extension of gene sequence with 5'-RACE or 3'-RACE method without using a library. In another method, a human origin cDNA library is ligated into the expression vector, expressed by COS-7 or other cells, and the ligated molecule is screened using the receptor Notch protein and expression cloned to isolate the cDNA of the ligand. For expression cloning, a cell sorter fractionation method is applied by binding the proteins with a polypeptide containing the amino acid sequence of prior known 4 Notches such as TAN-1, and a detection method by film emulsion using radioisotope can be used.

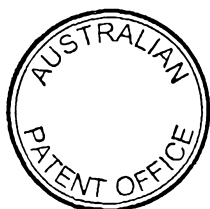
In this specification, a method for obtaining genes of human Serrate-2 is explained, and various methods clearly shown in this invention including PCR, by which clonings of human Delta-1 and human Serrate-1 were achieved, can be applied to obtain new Notch ligand family molecules which have never been cloned. For example, the conserved domains are found by comparison with amino acid sequence or DNA sequence of human Serrate-1 or human Serrate-2, and cloning thereof is performed after applying PCR, and also cloning can be performed by searching EST's based on human Delta-1 or human Serrate-2.



These cloned new Notch ligand family molecules can be used in the same way as the human Serrate-2 shown in the present invention, for example full length cloning, preparing expression vector, preparation of transformed cells, protein production, antibody production or screening the bioactive substances, and differentiation suppressive action for cells can be expected.

As shown in Example 2, these three gene fragments are labeled with radioisotope to prepare a hybridization probe, screened using cDNA of human fetal brain origin as the screening library, determined DNA sequence of the thus obtained clones, and the fragments were found to have high similarity with human Serrate-1 full length DNA nucleotide sequence. In these screenings, it was found that a full length cDNA sequence encoding the full length amino acid sequence could not be cloned. A further DNA probe was prepared based on the cloned DNA sequence, and again screening was performed, but the full length gene was still not identified. Finally, gene cloning containing the translation initiation Met codon was performed using the 5'-RACE method, the DNA sequence determined and finally the cloning of the cDNA encoding full length of gene sequence of human Serrate-2 was successful. The thus cloned cDNA was ligated as shown in Example 2 and cDNA encoding the full length of said human Serrate-2 was obtained.

Examples of plasmids integrated with cDNA are, for example, *E. coli* originated pBR322, pUC18, pUC19, pUC118 and pUC119 (Takara Shuzo Co., Japan), but other plasmids can be used if they can replicate and proliferate in the host cells. Examples of phage vectors integrated with cDNA are, λ gt10 and λ gt11, but the other vectors can be used if they can grow in the host cells. The thus obtained plasmids are transduced into suitable host cells such as genus Escherichia and genus Bacillus using the calcium chloride method. Examples of the above genus Escherichia are Escherichia coli K12HB101, MC1061, LE392 and JM109. Example of the above genus Bacillus is Bacillus subtilis MI114. Phage vector can be introduced into the proliferated *E. coli* by the in vitro packaging method (Enquist and Sternberg, Meth, Enzymol., 68.281-,1979).



The cloned full length DNA sequence was compared with the data base (Genbank release 93, 1996), and found that the total sequence is a novel sequence, although there are partial sequences as previously mentioned, three EST clones and several EST clone data entries, which are identical partial sequences other than the said three EST clones.

Further, the said amino acid sequence of human Serrate-2, i.e. amino acid sequences in SEQ ID NO: 1, 2 and 3, were compared with the data base of prior known amino acid sequences (SWISS-PROT, release 32, 1995 and Genbank CDS, release 93, 1996), and it was found that there are no identical amino acid sequences and therefore the sequences are novel sequences. According to a comparison with the amino acid sequence of human Serrate-1 and the Serrate homologue of the other organisms, the homologies with human Serrate-1, *Drosophila* Serrate, and rat jagged are 53.1 %, 34.3 % and 52.3 %, respectively. The substance of the present invention is different from these substances and is novel having a new amino acid sequence, and was first identified by the present inventors.

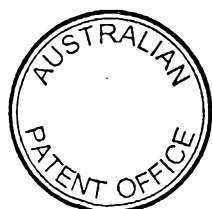
The said amino acid sequence was analyzed for hydrophilic and hydrophobic regions according to a method by Kyte-Doolittle (J. Mol. Biol. 157: 105, 1982). Results indicate that in the amino acid sequence listed in SEQ ID NO: 4, the amino acid sequence of a precursor of the full length gene consists of a 1238 amino acids residue, from -26 to 1212 in amino acid sequence, and the signal peptide domain is predicted to correspond to an amino acid sequence of 26 amino acid residues from No. -26 methionine to No. -1 proline; extracellular domain: 1055 amino acid residues from No. 1 methionine to No. 1055 glycine; transmembrane domain: 24 amino acid residues from No. 1056 leucine to No. 1079 tryptophan; and intracellular domain: region from No. 1080 threonine to No. 1212 glutamate. These domains are the predicted domain construction from the amino acid sequence, and the actual form may differ from the above structure, and the amino acids of each domain hereinabove defined may differ by 5 to 10 amino acids sequence.



In Example 6, identification of N-terminal (amino terminal) amino acid sequences of the purified ligand polypeptides EXS2Fc and EXS2FLAG of the present invention was performed and it was found that it was methionine of No. 1 in SEQ ID NO: 1 - 3. Consequently, signal peptide is at least from -26 methionine to -1 proline in SEQ ID NO: 4.

The family molecules of the Notch ligand have an evolutionally conserved common sequence in the extracellular domain, i.e. DSL sequence and repeated EGF-like sequence. As a result of comparison with the amino acid sequences of human Serrate-2 and human Serrate-1, these conserved sequences are predicted from the amino acid sequence Serrate-2. Namely, DSL sequence corresponds to a 43 amino acid residue from No. 172 cysteine to No. 214 cysteine of the amino acid sequence in the sequence listing, SEQ ID NO: 4.

The EGF-like sequence exists with 16 repeats wherein, in the amino acid sequence in the sequence listing, SEQ ID NO:4, the first EGF-like sequence is from No. 217 cysteine to No. 247 cysteine; the second EGF-like sequence is from No. 250 cysteine to No. 278 cysteine; the third EGF-like sequence is from No. 285 cysteine to No. 318 cysteine; the fourth EGF-like sequence is from No. 325 cysteine to No. 356 cysteine; the fifth EGF-like sequence is from No. 363 cysteine to No. 394 cysteine; the sixth EGF-like sequence is from No. 401 cysteine to No. 432 cysteine; the seventh EGF-like sequence is from No. 439 cysteine to No. 469 cysteine; the eighth EGF-like sequence is from No. 476 cysteine to No. 507 cysteine; the ninth EGF-like sequence is from No. 514 cysteine to No. 545 cysteine; the 10th EGF-like sequence is from No. 563 cysteine to No. 607 cysteine; the 11th EGF-like sequence is from No. 614 cysteine to No. 645 cysteine; the 12th EGF-like sequence is from No. 652 cysteine to No. 683 cysteine; the 13th EGF-like sequence is from No. 690 cysteine to No. 721 cysteine; the 14th EGF-like Sequence is from No. 729 cysteine to No. 760 cysteine; the 15th EGF-like sequence is from No. 767 cysteine to No. 798 cysteine; and the 16th EGF-like sequence is from No. 805 cysteine to No. 836 cysteine.

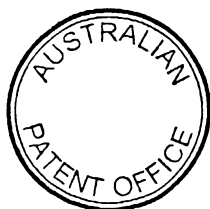


There are 2 cystein residues between the 9th EGF-like sequence and the 10th EGF-like sequence. Also, there are 6 cystein residues N-terminal of the DSL sequence and 16 cystein residues C-terminal of the 16th EGF-like sequence. These cystein residues including EGF-like sequence are conserved in almost identical position as in human Serrate-1.

A part of the sugar chain attached is predicted from the amino acid Sequence of human Serrate-2 to be No. 127, 544, 593, 726 and 1032 asparagine residues in the sequence listing, SEQ ID NO: 3 as possible binding sites of N-glycoside bonding for N-acetyl-D-glucosamine. O-glycoside bond of N-acetyl-D-galactosamine is predicted to be a serine or threonine residue rich part. Proteins bound with a sugar chain are generally thought to be stable in vivo and to have strong physiological activity. Consequently, in the amino acid sequence of the polypeptide having the sequence of the sequence listing, SEQ ID NO: 2 or 3, polypeptides having an N-glucoside or O-glucoside bond with the sugar chain of N-acetyl-D-glucosamine or N-acetyl-D-galactosamine are included in the present invention.

As a result of studies on the binding of Drosophila Notch and its ligand, the amino acid region necessary for binding with ligand of Drosophila Notch with Notch is from the N-terminus to the DSL sequence of the matured protein, in which signal peptide is removed (Japan. Pat. PCT Unexam. Publ. No. 7-503121). Further, similar studies using nematoda by Fitzgerald and Greenwald (Development, 121, 4275-4282, 1995) clearly indicated that the full length of the Notch ligand-like molecule, APX-1 from amino terminal to DSL domain, was sufficient for activation of the Notch-like receptor. These facts indicate that a domain necessary for expression of ligand action of the human Serrate-2 molecule is the novel amino acid sequence of the sequence listing, SEQ ID NO: 1.

Nothern blotting can be performed by using DNA encoding a part or all of gene sequence in the sequence listing, SEQ ID NO: 4. Consequently, a method for detection of expression of these genes can be achieved using hybridization or PCR, using complementary nucleic acids of above 12mer, preferably above



16mer, more preferably above 18mer, having a nucleic acid sequence of a part of the sequence in the sequence listing, SEQ ID NO: 4, i.e. antisense DNA or antisense RNA, its methylated, methylphosphated, deaminated, or thiophosphated derivatives. By the same method, detection of homologues of the gene from other organisms such as mice or gene cloning can be achieved.

Further cloning of genes from genomes including humans can be achieved. Using the genes cloned by such methods, further detailed functions of human Serrate-2 of the present invention can be identified. For example, using modern gene manipulation techniques, methods including transgenic mouse, gene targeting mouse or double knockout mouse, in which genes relating to the gene of the present invention are inactivated, can be applied. If abnormalities in the genome of the present gene are found, application to gene diagnosis and gene therapy can be made.

As described in Example 3, expression in the human normal tissues is observed in many tissues, and the length of the expressed mRNA is, for one type of the mRNA, about 5 kb length. By this means, detection of the expression of mRNA of the said molecule can be used for diagnosis of detection of malignant tumors in the part of normal organs in which expression of the mRNA is not observed. Further, by referring to patterns of the expressed organs, a use of human Serrate-2, can be identified.

A transformant in which vector pUCSR-2, which contains the cDNA encoding the total amino acid sequence of human Serrate-2 of the present invention, is transformed into E. coli JM109, has been deposited in the National Institute of Bioscience and Human-Technology, Agency of Industrial Science and Technology, MITI, of 1-1-3, Higasi, Tsukuba-shi, Ibaragi-ken, Japan, as E. coli: JM109-pUCSR-2. Date of deposit was October 28, 1996, and deposition No. is FBRM BP-5727.

Methods for the expression and purification of various forms of human Serrate-2 using the cDNA encoding the amino acid sequence of human Serrate-2, isolated by the above methods, are known in the references (Kriegler, Gene Transfer and Expression-A Laboratory Manual Stockton Press, 1990 and Yokota



et al. Biomanual Series 4, Gene transfer and expression and analysis, Yodosha Co., 1994). A cDNA encoding the amino acid sequence of the isolated said human Serrate-2 is ligated to a preferable expression vector and is produced in the host cells of eukaryotic cells such as animal cells and insect cells or prokaryotic cells such as bacteria.

In the expression of the molecule of the present invention, DNA encoding the polypeptide of the present invention may have the translation initiation codon at the 5'-terminus and the translation termination codon at the 3'-terminus. These translation initiation codon and translation termination codon can be added using preferable synthetic DNA adapters. Further, for expression of the said DNA, the promoter is linked upstream of the DNA sequence. Examples of vectors are plasmids originating from *Bacillus*, plasmids originating from yeast or bacteriophage such as λ -phage and animal viruses such as retrovirus and vaccinia virus.

Examples of promoters used in the present invention are any promoters preferable for corresponding to the host cells used in gene expression.

In the case that the host cell in the transformation is genus *Escherichia*, tac-promoter, trp-promoter and lac-promoter are preferable, and in case of host of genus *Bacillus*, SP01 promoter and SP02 promoter are preferable, and in case of host of yeast, PGK promoter, GAP promoter and ADH promoter are preferable.

In the case that the host cell is animal cells, a promoter originated from the SV40, promoter of retrovirus, metallothionein promoter or heatshock promoter can be applied.

Expression of the polypeptide of the present invention can be made using only DNA encoding the amino acid sequence of the sequence listing, SEQ ID NO: 1, 2 or 3. However, a protein including specific functions can be produced using DNA, to which cDNA encoding the known antigen epitope for easier detection of the produced polypeptide is added, or cDNA encoding the immunoglobulin Fc for forming a multimer of the said human Serrate-2 is added.



As shown in Example 4, we have prepared expression vectors, which express extracellular proteins, as follows.

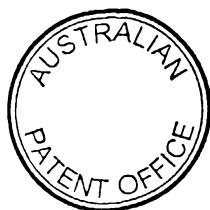
- 1) DNA encoding the amino acids from No. 1 to 1055 in amino acid sequence in the sequence listing, SEQ ID NO: 2,
- 2) DNA encoding chimeric protein to which was added a polypeptide having 8 amino acids, i.e. an amino acid sequence consisting of Asp Tyr Lys Asp Asp Asp Asp Lys (hereinafter designated the FLAG sequence, the sequence listing, SEQ ID NO: 22), at the C-terminus of the amino acids from No. 1 to 1055 in amino acid sequence in the sequence listing, SEQ ID NO: 2, and
- 3) DNA encoding a chimeric protein to which added Fc sequence below the hinge region of human IgG1 (refer to International Patent Unexam. Publ. WO 96/11212) at the C-terminus of the amino acids from No. 1 to 1055 in the amino acid sequence in the sequence listing, SEQ ID NO:2, and having a dimer structure by a disulfide bond in the hinge region,

are ligated individually with the expression vector pMKITNeo (Maruyama et al. Japan Molecular Biology Soc. Meeting Preliminary lecture record, obtainable from Dr. Maruyama in Tokyo Medical and Dental College) to prepare extracellular expression vectors of human Serrate-1.

The expression vectors, which expresses the full-length protein, can be prepared as follows.

- 4) DNA encoding the amino acids from No. 1 to 1212 in the sequence listing, SEQ ID NO: 3 and
- 5) DNA encoding a chimeric protein to which was added a polypeptide having the FLAG sequence to the C-terminus of amino acids from No. 1 to 1212 in the sequence listing, SEQ ID NO: 3

are ligated individually with the expression vector pMKITNeo to prepare the full-length expression vector of human Serrate-2. The transformant is prepared by using the expression plasmid containing DNA encoding the thus constructed said human Serrate-2.



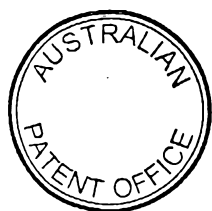
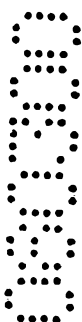
Examples of the host are genus *Escherichia*, genus *Bacillus*, yeast and animal cells. Examples of animal cells are simian cell COS-7 and Vero, Chinese hamster cell CHO and silk worm cell SF9.

As shown in Example 5, the above 5 type expression vectors of the above are transduced individually; the human Serrate-2 is expressed in COS-7 cells (obtainable from the Institute of Physical and Chemical Research, Cell Development Bank, RCB0539); and the transformants which were transformed by these expression plasmids, can be obtained. Further, the human Serrate-2 polypeptide can be produced by culturing the transformants under preferable culture conditions in medium using known culture methods.

The human Serrate-2 polypeptide can be isolated and purified from the above cultured mass, in general, by the following methods.

For extraction of the substance from cultured microbial cells or cells, microbial cells or cells are collected by a known method such as centrifugation after the cultivation, suspended in preferable buffer solution, the microbial cells or cells are disrupted by means of ultrasonication, lysozyme and/or freeze-thawing, and the crude extract of human Serrate-2 protein collected by centrifugation or filtration. The buffer solution may contain protein-denaturing agents such as urea and guanidine hydrochloride or surface active agents such as Triton-X. In the case of secretion in the cultured solution, the cultured mass is separated by a known method such as centrifugation, to separate from microbial cells or cells, and the supernatant solution is collected.

The thus obtained human Serrate-2, which are contained in the cell extracts or cell supernatants, can be purified by known protein purification methods. During the purification process, for confirmation of existence of the protein, in the case of the fused protein of the above FLAG and human IgGFc, the protein can be detected by immunoassay using antibodies against the known antigen epitope, and can be purified. In the case of proteins not expressed as fusion proteins, the antibody against human Serrate-2 in Example 7 can be used for detection.



A more useful purification method is affinity chromatography using an antibody. Antibodies used in this case are antibodies described in Example 7. For the fused protein, antibodies against other than human Serrate-2, for example, an antibody against FLAG in the case of a FLAG fusion protein, and protein G or protein A in the case of a human IgGFc fusion protein as shown in Example 6.

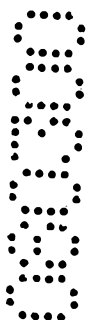
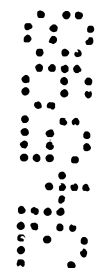
Physiological functions of the thus purified human Serrate-2 protein or human Serrate-2 can be identified by various assay methods, for example, physiological activity assaying methods using cell lines and animals such as mice and rats, assay methods of intracellular signal transduction based on molecular biological means, and binding with Notch receptor etc.

For that action, mainly an action suppressing cell differentiation will be expected, and actions such as stimulating tissue regeneration, etc. can be expected.

We have found that, as shown in Example 8, in the umbilical cord blood derived blood undifferentiated cells, in which the CD34 positive cell fraction is concentrated, polypeptides of the present invention have suppressive action on colony forming action against blood undifferentiated cells, which show colony formation in the presence of cytokines.

Further, as shown in Example 9, we have found that as a result of adding an IgG1 chimeric protein of human Serrate-2 to the liquid culture, in the presence of cytokines, the human Serrate-2 had the activity of significantly decreasing LTC-IC (Long-Term Culture-Initiating Cells) counts, which was positioned most undifferentiated blood stem cells in the human blood undifferentiated cells.

These results indicate that human Serrate-2 suppresses differentiation of blood undifferentiated cells, and that these actions spread from blood stem cells to colony forming cells. Further pharmaceuticals containing the polypeptide of the present invention have an action in protection and release of the bone marrow suppressive action, which is observed in adverse effects of antitumor agents.



Further, as shown in example 10, we have performed studies on action against vascular cells, for which an action of the molecule of the present invention has never been known, except for blood cells, and found that the molecule of the present invention has an action to suppress growth of the human vascular endothelia cells. Consequently, the present invention includes a growth suppressive agent for vascular cells and a therapeutic agent for diseases (refer to Folkman and Klagsbrun, SCIENCE 235, 442 - 447, 1987), in which the effect is expected to involve suppressing vascularization, containing polypeptide having amino acid sequence of SEQ ID NO: 1 - 3. The molecule of the present invention can be used for treatment of such diseases.

For pharmaceutical use, polypeptides of the present invention, having the above form, are lyophilized with the addition of preferable stabilizing agents such as human serum albumin and is used for injection, in the dissolved or suspended condition with distilled water. For example, preparation for injection or infusion at the concentration of 0.1 - 1000 μ g/ml may be provided. A mixture of the compound of the present invention 1 mg/ml and human serum albumin 5 mg/ml divided in a vial may maintain activity of the said compound for long term. For culturing and activating cells in vitro, the lyophilized preparation or liquid preparation of the polypeptide of the present invention are prepared and are added to the medium or immobilized in the vessel for culture. Toxicity of the polypeptide of the present invention was tested. Any polypeptide, 10 mg/kg was administered intraperitoneally in mice, but no death of mice was observed.

In vitro physiological activity of the polypeptide of the present invention can be evaluated by administering to disease model mice or its resembled disease rats or monkeys, and examining the recovery of physical and physiological functions, and abnormal findings. For example, in case of searching abnormality in relation to hemopoietic cells, bone marrow suppressive model mice are prepared by administering 5-FU series of antitumor agents, and bone marrow cell counts, peripheral blood cell counts and physiological functions are examined in the administered group or the non administered group of mice. Further, in case of searching in vitro cultivation and growth of hemopoietic



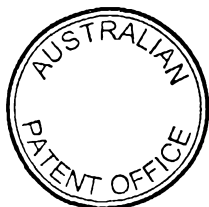
undifferentiated cells including hemopoietic stem cells, the bone marrow cells of mice are cultured in the groups with or without addition of the compound of the present invention, and the cultured cells are transferred into the lethal dose irradiated mice. The result of recovery is observed with the indications of survival rate and variation of blood counts. These results can be extrapolated to humans, and accordingly useful effective data for evaluation of the pharmacological activities of the compound of the present invention can be obtained.

Applications of the compound of the present invention for pharmaceuticals include diseases with abnormal differentiation of cells, for example leukemia and malignant tumors. These are a cell therapy, which is performed by culturing human derived cells in vitro maintaining their original functions or adding new functions, and therapy, which is performed by regenerating, without damage the functions which originally existed in the tissues by administering the compound of the present invention during regeneration after tissue injury. Amount of administration may differ according to the type of preparation and ranges from 10 μ g/kg to 10 mg/kg.

Further strong physiological activity can be achieved by expression, forming a multimer, of the polypeptide of the present invention.

Human Serrate-2 having a multimer structure can be produced by expressing a chimeric protein with the human IgG Fc region as described in the example, and expressing the multimer having a disulfide bond with the hinge region of the antibody, or a method of expressing a chimeric protein, in which the antibody recognition region is expressed at the C-terminus or N-terminus, and reacting with the polypeptide containing the extracellular part of the thus expressed said human Serrate and the antibody which recognize specifically the antibody recognition region at the C-terminus or N-terminus.

As an example of other methods, a method in which the fused protein bound with only the hinge region of the antibody is expressed, and the dimer is formed by constructing with the disulfide bond, can be mentioned. The multimer of human Serrate-2 having a higher specific activity than the dimer can be obtained. The said multimer is constructed using a fusion protein which is

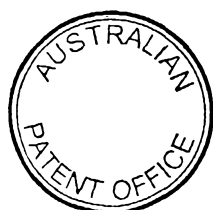
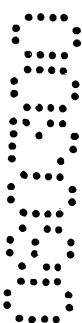
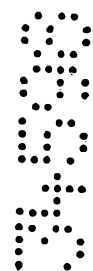


prepared for expressing the peptide of the C-terminus, N-terminus or other region. The protein forms a disulfide bond without effecting any of the activities of the other human Serrate-2. The multimer structure can also be expressed by arranging one or more peptide containing amino acids sequence of the sequence listing, SEQ ID NO: 1 or 2, using genetic engineering methods in series or in parallel. Other known methods for providing multimer structures of a dimer or more can be applied. Accordingly, the present invention includes any polypeptides containing amino acid sequences described in the sequence listing, SEQ ID NO: 1 or 2 in the form of dimer or more structure prepared by genetic engineering techniques.

Another example method is the multimerization method using a chemical cross-linker. For example, dimethylsuberimidate dihydrochloride for cross-linking lysine residues, N-(γ -maleimidebutyryloxy) succinimide for cross-linking thiol groups of cysteine residue and glutaraldehyde for cross-linking between amino groups, can be mentioned. The multimer with dimer or more can be synthesized by applying these cross-linking reactions. Accordingly, the present invention includes any polypeptides containing amino acid sequences described in the sequence listing, SEQ ID NO: 1 or 2 in the form of dimer or more structure prepared using chemical cross-linking agents.

In the application of medical care in which cells are proliferated and activated in vitro and are returned to the body, human Serrate-2 of the form hereinabove can be added directly in the medium, alternately an immobilization of the protein can be carried out. Immobilization method includes applying an amino group or carboxyl group in the human Serrate-2, using suitable spacers or the above mentioned cross-linkers, and the ligand can be covalently bound to the culture vessels. Accordingly, the present invention includes any polypeptides containing amino acid sequences described in the sequence listing, SEQ ID NO: 1 or 2 in the form of existing on the solid surface.

The human Serrate-2 molecule binds specifically with a receptor, a Notch receptor molecule. For example, expression of Notch receptor can be detected using a fusion protein with the above extracellular region of human Serrate-2 and



human IgGFc. Notch is known to be involved in some types of leukemia (Ellisen ry al., Cell 66, 649 -661, 1991). Accordingly, the polypeptide having the amino acid sequence in the sequence listing, SEQ ID NO: 1 or 2 can be used as a diagnostic reagent for in vitro or in vivo.

An antibody which specifically recognizes the said human Serrate-2 can be prepared as shown in Example 7. Also, it can be prepared by various methods as described in the reference (Antibodies a laboratory manual, E. Harlow et al., Cold Spring Harbor Laboratory), and using a recombinant antibody expressed in cells using the immunoglobulin gene isolated by gene cloning. These antibodies can be used for the purification of human Serrate-2. Namely, the detection and measurement of human Serrate-2 of the present invention can be performed by using an antibody, which specifically recognizes the human Serrate-2 shown in Example 7, and can be applied as a diagnostic agent for diseases such as malignant tumors accompanied with abnormal cell differentiation.

Embodiments of the invention

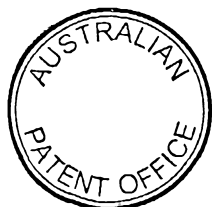
Following examples illustrate the embodiments of the present invention but are not construed as limiting the invention to these examples.

Preferential example 1

Preparation of human Serrate-1 gene probe

A mixed primer corresponding to the amino acid sequence conserved in Drosophila Serrate and rat Jagged, i.e. sense primer (sequence listing, SEQ ID NO: 6) and antisense primer (sequence listing, SEQ ID NO: 7), were used. The following designations are used in these sequences; i.e. S: C and G, Y: T and C, W: T and A, K: G and T, R: A and G and N: C, G, T and A.

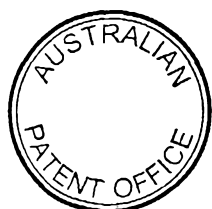
A synthetic oligonucleotide was prepared using an automatic DNA synthesizer which used a method based on immobilization. The automatic DNA synthesizer used was the 391PCR-MATE from Applied Biosystems Inc., U.S.A. Nucleotide, carrier immobilized with 3'-nucleotide, solution and reagents were used according to the instructions by the same corporation. The oligonucleotide was isolated from the carrier after finishing the designated coupling reaction and



treating the oligonucleotide carrier, from which the protective group of 5'-terminus was removed, with concentrated liquid ammonia at room temperature for one hour. To remove the protective groups of nucleic acid and phosphoric acid, the reactant solution containing the nucleic acid was allowed to stand in a concentrated ammonium solution in a sealed vial at 55°C for over 14 hours. Each oligonucleotide, from which the carrier and protective groups were removed, was purified using an OPC cartridge from Applied Biosystems Inc., and detritylated by using 2% trifluoroacetic acid. The primer was dissolved in deionized water to a final concentration of 100 pmol/ μ l after purification, and was used for PCR. Synthesis of other oligonucleotide were performed in the same manner.

Amplification by PCR was performed as follows.

Human fetal brain originated cDNA mixed solution (QUICK-Clone cDNA, CLONTEC Inc., U.S.A.) 1 μ l was used. 10xbuffer solution [500 mM KCl, 100 mM Tris-HCl (pH 8.3), 15 mM MgCl₂, 0.01% gelatin] 5 μ l, dNTP mixture (Takara Shuzo Co., Japan) 4 μ l, sense primer DLTS1 (100 pmol/ μ l) 5 μ l and antisense primer DLTA2 (100 pmol/ μ l) 5 μ l, which were specific to the above Serrate homologue, and TaqDNA polymerase (AmpliTaq, Takara Shuzo Co., Japan, 5U/ μ l) 0.2 μ l were added thereto. Deionized water was added to a total volume of 50 μ l. PCR was performed using 5 cycles of a cycle consisting of treatment at 95°C for 45 seconds, at 42°C for 45 seconds and 72°C for 2 minutes, further 35 cycles of a cycle consisting of treatment at 95°C for 45 seconds, at 50°C for 45 seconds and 72°C for 2 minutes, and finally allowed to stand at 72°C for 7 minutes. An aliquot of the PCR products was subjected to 2% agarose gel electrophoresis, stained with ethidium bromide (Nippon Gene Co., Japan), and observed under ultraviolet light to confirm amplification of an about 500 bp fragment of cDNA. The remainder of the thus obtained PCR product was subjected to electrophoresis with 2% agarose gel prepared by low melting point agarose (GIBCO BRL Inc., U.S.A.), stained using ethidium bromide, excising the about 500 bp bands under UV light, adding distilled water of the same volume as



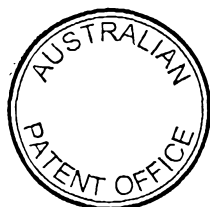
the gel, heating at 65°C for 10 minutes, and completely dissolving the gel. An equal volume of TE saturated phenol (Nippon Gene Co., Japan) was added and the dissolved gel was centrifuged at 15000 rpm for 5 minutes to separate supernatant solution. The same separation operation was performed after adding sequentially TE saturated phenol : chloroform (1 : 1) solution and chloroform. DNA was recovered from the final solution by ethanol precipitation.

A vector, pCRII vector (Invitrogen Inc., U.S.A., hereinafter designated pCRII) was used. The vector and the above DNA were mixed in a molar ratio of 1 : 3 and the DNA was ligated into the vector by using T4 DNA ligase (Invitrogen Inc., U.S.A.). The pCRII, to which DNA was integrated, was subjected to gene transduction into E. coli one shot competent cells (Invitrogen Inc., U.S.A.) and was spread on a semi-solid medium plate of L-Broth (Takara Shuzo Co., Japan) containing ampicillin (Sigma Corp., U.S.A.) 50 μ g/ml and allowed to stand at 37°C for about 12 hours. The appeared colonies were randomly elected, inoculated in L-Broth liquid medium (2 ml), containing same concentration of ampicillin, and shake cultured at 37°C for about 18 hours. The cultured bacterial cells were recovered and the plasmid was separated using Wizard Miniprep (Promega Inc., U.S.A.) according to the manufacturing instructions. The plasmid was digested using the restriction enzyme EcoRI. Integration of the said PCR product was confirmed by incision of a DNA fragment of about 400 bp. Base sequence of the incorporated DNA in the confirmed clone was determined by fluorescent DNA sequencer (Model 373S, Applied System Inc., U.S.A.). The cloned gene fragment was compared with the amino acid sequence of the Notch ligand molecule, i.e. Drosophila Serrate and rat Jagged, and significant analogous sequence was found. The sequence was confirmed to be a cDNA fragment encoding human Serrate-1.

Preferential example 2

Cloning of full length human Serrate-1 gene

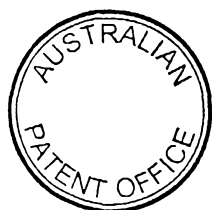
Screening of clones having the full length cDNA was performed by hybridization from a human placenta origin cDNA library (inserted cDNA in λ gt-11, CLONTEC INC, U.S.A.) in plaques corresponding to 1×10^6 plaques.



Generated plaques were transferred to a nylon filter (Hybond N+: Amersham Inc., U.S.A.). The nylon filter was subjected to alkaline treatment [allow to stand for 7 minutes on a filter paper permeated with a mixture of 1.5 M NaCl and 0.5M NaOH], followed by two neutralizing treatments [allow to stand for 3 minutes on a filter paper permeated with a mixture of 1.5 M NaCl, 0.5 M Tris-HCl (pH 7.2) and 1 mM EDTA]. Subsequently, the nylon filter was shaken for 5 minutes in 2-fold concentrated SSPE solution [0.36 M NaCl, 0.02 M sodium phosphate (pH 7.7) and 2 mM EDTA], washed and air-dried. Then the nylon filter was allowed to stand for 20 minutes on a filter paper, which was permeated with 0.4 M NaOH, and was shaken for 5 minutes with 5-fold concentrated SSPE solution and washed, and then again air-dried. Screening was conducted using the human Serrate-1 probe labeled with radioisotope ^{32}P using the filter.

DNA probe previously prepared was labeled with ^{32}P as follows. A DNA fragment was excised using by EcoRI from pCRII, to which purified PCR product by human Serrate primers (about 500 bp) were inserted and the DNA fragments were isolated from a low melting point agarose gel. The thus obtained DNA fragment was labeled using a DNA labeling kit (Megaprime DNA labeling system: Amersham, U.S.A.). The primer solution, $5\mu\text{l}$, and deionized water were added to DNA (25 ng) to a total volume of $33\mu\text{l}$, which was treated for 5 minutes in boiling water bath. Reaction buffer solution ($10\mu\text{l}$) containing dNTP, $\alpha\text{-}^{32}\text{P}$ -dCTP ($5\mu\text{l}$) and T4 DNA polynucleotide kinase solution ($2\mu\text{l}$) were added thereto, and treated at 37°C for 10 minutes in water bath. Subsequently, the mixture was purified using a Sephadex column (Quick Spin Column Sephadex G-50: Boehringer Mannheim Inc., Germany), then treated for 5 minutes in a boiling water bath and ice-cooled for 2 minutes immediately prior to use.

Hybridization was performed as follows. The prepared filter hereinabove was immersed into the prehybridization solution consisting of SSPE solution, in which the final concentration of each component was 5-fold concentration, 5-fold concentration of Denhardt's solution (Wako Pure Chemicals), 0.5 % SDS (sodium dodecyl sulfate) and salmon sperm (Sigma Co.) $10\mu\text{g/ml}$ denatured by

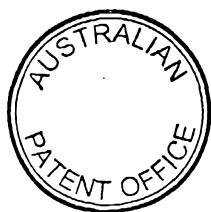


boiling water. The filter was shaken at 65°C for 2 hours, then the filter was immersed into the hybridization solution, which was the same composition with the above prehybridization solution, containing the probe labeled with ^{32}P by the above mentioned method, and the filter was shaken at 55°C for 16 hours to perform hybridization.

The filter was washed by immersing into SSPE solution containing 0.1% SDS, shaken at 55°C twice, and further immersed into a 10-fold dilution of SSPE solution containing 0.1% SDS, four times at 55°C. The washed filter was treated with autoradiography using an intensifying screen. Clones from the strongly exposed part were collected and the plaques obtained were again spread and screened by the same method hereinbefore to isolate a single clone.

22 phage clones were thus isolated. Phage from all of these clones was prepared to about 1×10^9 pfu, the phage DNA was purified, digested using restriction enzyme EcoRI and inserted into pBluescript (Stratagene Inc., U.S.A.) which was digested using EcoRI in the same way. The DNA sequences of the both ends of these clones were analyzed using a DNA sequencer. Two clones being S16 and S20 were the clones containing DNA sequence from No. 1 to 1873 in the sequence listing, SEQ ID NO: 5. Clones S5 and S16 were the clone containing DNA sequence from No. 990 to 4005 in the sequence listing, SEQ ID NO: 5. The deletion mutants of these clones were prepared by using a kilosequence deletion kit (Takara Shuzo Co.) according to the manufacturers instructions. Full-length cDNA base sequence encoding the polypeptide of the present invention was determined using the DNA sequencer (Applied Biosystem Inc.) from both the 5' and 3'-directions.

As a result, about 100 bp encoding the C-terminal amino acid sequence was absent from the clones, accordingly, cloning of the full-length gene was performed by using GIBCO-BRL, 3' RACE system kit according to the attached manual. Namely, cDNA cloning was performed using human origin poly A⁺ RNA (CLONTECH Corp.) in the 3'-direction and the resulting gene sequence was determined.



The thus cloned three gene fragments were inserted into a plasmid containing the full length DNA sequence of SEQ ID NO: 5 using the restriction enzyme Bgl 2 site at DNA sequence No. 1293 in SEQ ID NO: 5 and Accl site at No. 3943, between the EcoRI and XbaI of multicloning site in pUC18, to prepare pUCSR-1. The sequence of this gene together with the amino acid sequence, is shown in SEQ ID NO: 5.

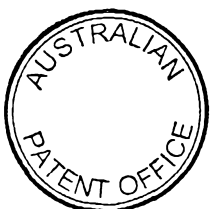
Example 1

Preparation of probe by PCR

Gene probes used for screening, i.e. the gene described in SEQ ID NO: 8, 9 and 10, were obtained as follows. These sequences correspond to Genbank registered numbers T08853, R50026 and R45751. Hereinafter, a probe having gene sequence SEQ ID NO: 8 is designated as ¥1, a probe having gene sequence SEQ ID NO: 9 is designated as ¥2, and a probe having gene sequence SEQ ID NO: 10 is designated as ¥4.

A gene having SEQ ID NO: 8 was isolated by PCR using oligonucleotide primers having SEQ ID NO: 11 and 12; a gene SEQ ID NO: 9 was isolated by PCR using oligonucleotide primers having SEQ ID NO: 13 and 14; and a gene SEQ ID NO: 10 was isolated by PCR using oligonucleotide primers having SEQ ID NO: 15 and 16.

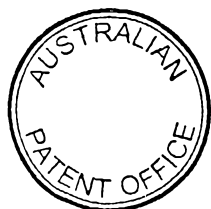
Amplification by PCR was performed as follows. Human fetal brain originated cDNA mixed solution (QUICK-Clone cDNA, CLONTEC Inc., U.S.A.) 1 μ l was used. 10 x buffer solution [500 mM KCl, 100 mM Tris-HCl (pH 8.3), 15 mM MgCl₂, 0.01 % gelatin] 5 μ l, dNTP mixture (Takara Shuzo Co., Japan) 4 μ l, the primers hereinbefore (20 pmol/ μ l) each 1 μ l and TaqDNA polymerase (AmpliTaq, Takara Shuzo Co., Japan, 5U/ μ l) 0.2 μ l were added thereto, and finally deionized water was added to a final total volume of 50 μ l. PCR was performed by 40 cycles of a cycle consisting of treatment at 95°C for 1 minute, at 55°C for 5 minutes and 72°C for 3 minutes, and finally allowed to stand at 72°C for 7 minutes. An aliquot of the PCR products was subjected to 2% agarose gel electrophoresis, stained with ethidium bromide (Nippon Gene Co., Japan), and



observed under ultraviolet light to confirm amplification of a gene of the predicted size.

Total amount of the thus obtained PCR product was subjected to electrophoresis using a 2% agarose gel prepared using low melting point agarose (GIBCO BRL Inc., U.S.A.), stained by ethidium bromide, excising each band under UV light, adding distilled water of equal volume to the gel, heating at 65°C for 10 minutes, to completely dissolve the gel. The dissolved gel was centrifuged at 15000 rpm for 5 minutes to separate supernatant solution, after adding an equal volume of TE saturated phenol (Nippon Gene Co., Japan). The same separation operation was performed after adding TE saturated phenol: chloroform (1 : 1) solution and chloroform sequentially. DNA was recovered from the final solution by ethanol precipitation.

A vector, pCRII vector (Invitrogen Inc., U.S.A., hereinafter designated pCRII) was used. The vector and the above DNA in molar ratio of 1 : 3 were mixed and DNA was ligated into the vector by using T4 DNA ligase (Invitrogen Inc., U.S.A.). The pCRII, to which DNA was integrated, was subjected to gene transduction into E. coli one shot competent cells (Invitrogen Inc., U.S.A.) and was spread on the semi-solid medium plate of L-Broth (Takara Shuzo Co., Japan) containing ampicillin (Sigma Corp., U.S.A.) 50 μ g/ml and allowed to stand at 37°C for about 12 hours. The appeared colonies were randomly selected, inoculated in the L-Broth liquid medium 2 ml containing same concentration of ampicillin and shake cultured at 37°C for about 18 hours. The cultured bacterial cells were recovered and the plasmid was separated using Wizard Miniprep (Promega Inc., U.S.A.) according to the attached explanation sheet. The plasmid was digested using the restriction enzyme EcoRI. Integration of the said PCR product was confirmed by detection of DNA of the predicted size. Base sequence of the incorporated DNA in the confirmed clone was determined by fluorescent DNA sequencer (Model 373S, Applied System Inc., U.S.A.). The sequence was compared with DNA sequence registered in Genbank hereinbefore. Isolation of a gene having DNA sequences SEQ ID NO:8, 9 and 10 as confirmed.



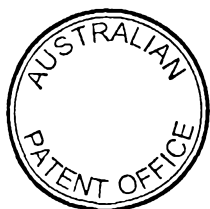
Example 2

Cloning of full length human Serrate-2 gene

A screening of clones having the full length cDNA was performed using hybridization from human fetal brain origin cDNA library (inserted cDNA in λ gt-11, CLONTECH Inc.) in plaques corresponding to 1×10^6 plaques. Appeared plaques were fixed with alkali by the same method as described in Preferential example 2. Using these filters, and using three types of probes, which were isolated in Example 1 and labeled with ^{32}P by the method described in Preferential example 2, screening was performed individually on each clone.

The isolated phage clones were: 2 clones when using ¥1 as a probe; 6 clones when using ¥2 as a probe; and 4 clones when using ¥4 as a probe. The clones isolated by ¥4 were all included in the clones isolated from ¥2. All of phages of these clones were prepared to about 1×10^9 pfu, and phage DNA was purified by using Wizard Lambda preps (Promega Inc.) according to the attached explanation sheet, and digested by EcoRI, then incorporated into pBluescript (Stratagene Inc.) or pUC18 (Pharmacia Inc.) digested by EcoRI.

Full DNA sequence of these clones was determined by DNA sequencing as described in Preferential example 2, and a part of the identical sequence was compared. The results indicate that: #5 clone was a clone containing DNA sequence from No. 484 to 2025 in SEQ ID NO: 4; #21 clone was a clone containing DNA sequence from No. 1882 to 3537 in SEQ ID NO: 4; and #86 clone was a clone containing DNA sequence from No. 2455 to 3955 in SEQ ID NO: 4. The remaining clones were those only having short insert consisting of a part identical to the sequences of these clones. This result indicated that, as a result of comparison with the amino acid sequence of human Serrate-1, the area encoding the N-terminal amino acid sequence was not cloned. Consequently, a probe having the DNA sequence of SEQ ID NO: 17 was prepared and screening was performed for the second time to cloning the cDNA in the 5'-region. A probe was prepared as described in Example 1, namely PCR primers having DNA sequence of SEQ ID NO: 18 and 19 were subjected to PCR using #5 clone as a



template. The library used was prepared as for the first screening and conditions were performed by the same method as before.

Six clones were isolated in the second screening. Phages of these all clones were prepared to about 1×10^9 pfu, purified by using Wizard Lambda Preps (Promega Inc.) according to the attached explanation sheet, digested using EcoRI and inserted into pUC18 which was also digested with EcoRI. The full DNA sequence of these clones was determined using DNA sequencing as in Preferential example 2. As a result of a comparison with a part of identical sequence, the S43-1 clone was identified and was considered to contain most of the 5' sequence. This clone was a clone containing DNA sequence from No. 38 to 1538 in SEQ ID NO: 4. The remaining clones were clones, with only a short insert consisting of a part identical with sequence determined already in the other clones, or isolated for the first time.

Although an ATG sequence, which encodes the translation initiation codon methionine, was not found in the second screening clones, further cloning of cDNA sequence in the 5'-direction was performed using the 5' RACE method. 5' RACE was performed by using 5' RACE system kit (GIBCO-BRL Inc.) according to the attached manual. Cloning of cDNA in the 5' direction in the gene using human heart origin poly A⁺ RNA (CLONTECH Inc.) was performed, and the gene sequence from DNA sequence No. 1 to No. 37 in SEQ ID NO: 4 was identified.

As a result, the DNA sequence in SEQ ID NO: 4, i.e. cDNA sequence encoding the full length of human Serrate-2, was determined.

In order to prepare cDNA encoding the full-length gene, and to obtain cDNA of the 5'-terminus, which could ligate with the other clone, the following PCR was performed for cloning. Namely, using the oligonucleotide having DNA sequence in SEQ ID NO: 20 and the oligonucleotide having DNA sequence in SEQ ID NO: 21, PCR was performed using S43-1 clone as a template, according to the method described in example 1. Similarly, it was subcloned into pCRII and the gene sequence was determined to identify the clone having DNA sequence from No. 1 to No. 503 in SEQ ID NO: 4. This clone is designated S2-5.



Among the above clones, i.e. gene S2-5, S43-1, #5, #21 and #86, S2-5 and S-43-1 were ligated to the restriction enzyme Sph I site at DNA sequence No. 217 in SEQ ID NO: 4; S43-1 and #5 were the same as in Kpn I site at No. 1453; #5 and #21 were the same as in Sac I site at No. 2016; and #21 and #86 were the same as in BamHI site at No. 2991, and finally DNA having DNA sequence SEQ ID NO: 4 was inserted between EcoRI site and Hind III site of the multicloning site of pUC18 to prepare pUCSR-2.

Example 3

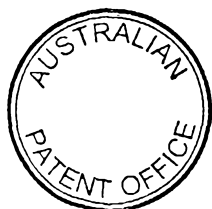
Expression of human Serrate-2 in organs

In order to examine the expression of mRNA of human Serrate-2, using filters, which were previously transferred with mRNA, i.e. Human Multiple Tissue Northern Blot, Human Multiple Tissue Northern Blot II, Human Multiple Tissue Northern Blot III and Human Fetal Multiple Tissue Northern Blot II (CLONTECH Inc.), ³²P labeling was performed by the previously described method using DNA labeling kit (Mega Prime DNA labeling system: Amersham Inc.) hereinbefore mentioned, with a probe DNA having sequence in SEQ ID NO: 17 as described in Example 2, and expression was examined by performing hybridization according to description of instruction for use attached to the above filters.

As a result, the length of the expressed mRNA was about 5kb. Strong expressions in human adult tissues were observed in heart, skeletal muscle, thyroid gland, spinal cord and trachea; clear expression observed in pancreas, prostate, testis, small intestine and adrenal gland, very weak expression observed in brain, placenta, kidney, thymus, ovary, stomach and lymph node, and no expression observed was lung, liver, spleen, colon, peripheral lymphocytes and bone marrow. In the human fetal tissues, strong expression was observed in the fetal lung, clear expression: fetal brain and fetal kidney and no expression: fetal liver.

Example 4

Preparation of an expression vector of human Serrate-2



Using the gene consisting of the DNA sequence described in the sequence listing, SEQ ID NO: 4, expression vectors of human Serrate-2 and its chimeric protein mentioned in the following 1)- 5) were prepared.

1) Expression vector of secretory extracellular human Serrate-2

The cDNA encoding polypeptide of amino acid sequence from No. 1 to 1055 in the sequence listing, SEQ ID NO: 2 was ligated with the expression vector pMKITNeo (Maruyama et al., Preliminary Paper, Japan Molecular Biology Soc. 1991, obtainable from Prof. Maruyama, Tokyo Medical and Dental Univ.), which has a $SR\alpha$ promoter and Neomycin resistance gene, to prepare an expression vector.

Namely, vector pUCSR-2 containing the DNA sequence in SEQ ID NO: 4 was used as template and an oligonucleotide having sequence in SEQ ID NO: 23 and an oligonucleotide having sequence in SEQ ID NO: 24 were used as primer, and PCR was performed according to a method described hereinbefore. The PCR product was ligated into the cloning vector pCRII, determined the gene sequence of the PCR product was determined to prepare DNA having gene sequence from No. 2986 to 3254 of the DNA sequence in SEQ ID NO: 4, in which the termination codon and restriction enzyme Sal I site were attached to the 3' end,

An about 3 kbp gene fragment obtained by digesting the pUCSR-2 with restriction enzyme EcoRI and BamHI, an about 250 bp gene fragment obtained by digesting the pCRII vector with restriction enzymes BamHI and Sal I, (the vector of which contained the above PCR product as an insert), and an about 4.3 kb gene fragment obtained by digesting the pMKITneo with restriction enzymes EcoRI and Xho I, were simultaneously ligated to obtain the expression vector containing a gene fragment of DNA sequence from No. 1 to 3254 in SEQ ID NO: 4. Secretory extracellular human Serrate-2 protein (hereinafter designated as EXS2) expression vector pMEXS2 was obtained.

2) Expression vector of FLAG chimera protein of secretory extracellular human Serrate-2



The cDNA encoding the chimeric protein, to which the cDNA encoding the FLAG sequence (SEQ ID NO: 22) was added to the C-terminus of the polypeptide from No. 1 to 1055 of the amino acid sequence in the sequence listing, SEQ ID NO: 2, was ligated to the expression vector pMKITNeo to prepare the expression vector.

Namely, vector pUCSR-2 containing the DNA sequence in SEQ ID NO: 4 was used as a template and an oligonucleotide having the sequence in SEQ ID NO: 23 and an oligonucleotide having sequence in SEQ ID NO: 25 were used as primers. PCR was performed according to a method described hereinbefore. The PCR product was ligated into the cloning vector pCRII, the gene sequence of the PCR product was determined to prepare DNA having gene sequence from No. 2986 to 3254 of DNA sequence in SEQ ID NO: 4, in which the DNA sequence encoding the FLAG sequence in the 3' end (DNA sequence in SEQ ID NO: 22), termination codon and restriction enzyme Sal I site were attached to the 3' terminus.

An about 3 kbp gene fragment obtained by digesting the pUCSR-2 with restriction enzyme EcoRI and BamHI, an about 300 bp gene fragment obtained by digesting the pCRII vector with restriction enzymes BamHI and Sal I, (the vector of which contained the above PCR product as an insert), and an about 4.3 kb gene fragment obtained by digesting the pMKITneo with restriction enzymes EcoRI and Xho I, were simultaneously ligated. Although the recognition sequence of the restriction enzymes Xho I and Sal I is different, they can be ligated due to complementary terminal gene sequences. The expression vector containing the gene fragment of DNA sequence from No. 1 to 3254 in SEQ ID NO: 4 and a gene fragment encoding the FLAG sequence was obtained. The secretory extracellular human Serrate-2 FLAG chimeric protein (hereinafter designated as EXS2FLAG) expression vector pMEXS2FLAG was obtained.

3) Expression vector of IgG1Fc chimera protein of secretory extracellular human Serrate-2

A cDNA encoding the chimeric protein, to which the cDNA encoding the amino acid sequence of the Fc region below the hinge part of human IgG1, was



added to the C-terminus of a polypeptide having the amino acid sequence from No. 1 to 1055 in the sequence listing, SEQ ID NO: 2, was ligated to the expression vector pMKITNeo to prepare the expression vector.

Preparation of a fusion protein with the immunoglobulin Fc protein was performed according to the method of Zettlmeissl et al. (Zettlmeissl et al., DNA cell Biol., 9, 347-354, 1990). A gene was prepared using genomic DNA (including introns) using PCR.

Human genomic DNA was used as a template. Oligonucleotides of the sequence in the sequence listing, SEQ ID NO: 28 with a restriction enzyme BamHI site, and oligonucleotides of the sequence in the sequence listing, SEQ ID NO: 29 with a restriction enzyme XbaI site were used as primers. PCR of the gene sequence encoding human IgG1F was performed using the primers and human genomic DNA as template. An about 1.4 kbp band was purified, treated using restriction enzyme BamHI and XbaI (Takara Shuzo Co., Japan), the genes were ligated into pBluescript using T4 DNA ligase, which was similarly treated with restriction enzymes. Later, the plasmid DNA was purified and sequenced to confirm the gene sequence, and the said gene sequence was confirmed to be genomic DNA in the hinge region of the heavy chain of human IgG1. (The sequence is referred to in Kabat et al., Sequence of Immunological Interest, NIH Publication No. 91-3242, 1991). Hereinafter this plasmid is designated as pBSIgFc.

A vector pUCSR-2 having DNA sequence in SEQ ID NO: 4 was used as template, an oligonucleotide having the sequence in SEQ ID NO: 23 and an oligonucleotide having the sequence in SEQ ID NO: 26, were used as primers, and PCR was performed according to the method described hereinbefore. The PCR product was ligated to cloning vector pCRII, the gene sequence of the PCR product was determined, to prepare DNA having the gene sequence from No. 2986 to 3254 of DNA sequence in SEQ ID NO: 4, in which a restriction enzyme Bgl II site was attached in the 3' end.

An about 250 bp gene fragment obtained by digesting the pCRII vector, which contained the above PCR product as an insert, with restriction enzymes



EcoRI and BamHI, and a fragment containing the gene encoding the above human IgG1FC as an insert, were ligated. In this case, BamHI and Bgl 2 sites can be ligated due to the complementary digested terminal gene sequence. The ligated fragment can not be digested using these restriction enzymes.

An about 1.5 kbp gene fragment obtained by digesting this vector with the restriction enzymes BamHI and Not 1, an about 3 kbp gene fragment obtained by digesting pUCSR-2 with the restriction enzymes EcoRI and BamHI, and an about 4.3 kb gene fragment obtained by digesting pMKITneo with the restriction enzymes EcoRI and Not 1, were simultaneously ligated. Restriction enzymes Xho I and Sal I have different recognition sequence, however they can be ligated due to complementary digested terminal gene sequence. An expression vector containing the gene fragment from DNA sequence No. 1 to 3254 in SEQ ID NO: 4 and a gene fragment encoding the human IgG1Fc was obtained. An expression vector pMEXS2Fc of Ig chimera protein of secretory extracellular human Serrate-2 (hereinafter this protein is designated as EXS2Fc) was obtained.

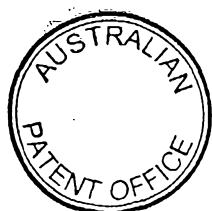
4) Expression vector of full length human Serrate-2 protein

The cDNA encoding the polypeptide from No. 1 to 1212 of the amino acid sequence in the sequence listing, SEQ ID NO: 3, was ligated to the expression vector pMKITNeo to prepare the expression vector.

Namely, an about 4 kbp gene fragment, which was excised by digesting pUCSR-2 with restriction enzymes EcoRI and Hind III, was ligated into pBluescript, which was digested with the same restriction enzymes. Subsequently, an about 4 kbp gene fragment cut from this vector by digesting with EcoRI and XhoI, was ligated with an about 4.3 kb gene fragment obtained by digesting the expression vector pMKITneo with restriction enzymes EcoRI and Not 1, to prepare an expression vector containing gene fragment of DNA sequence from No. 1 to 3955 in SEQ ID NO: 4. Full length human Serrate-2 protein (hereinafter designated as FS2) expression vector pMFS2 was obtained.

5) Expression vector of FLAG chimera protein of full length human

Serrate-2



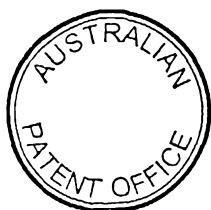
The cDNA encoding a chimeric protein, to which cDNA encoding the FLAG sequence (SEQ ID NO: 22) was added to the C-terminus of a polypeptide from No. 1 to 1212 of amino acid sequence in the sequence listing, SEQ ID NO: 3, was ligated to the expression vector pMKITNeo to prepare the expression vector.

Namely, vector pUCSR-2 having DNA sequence in SEQ ID NO: 4 was used as a template and an oligonucleotide having sequence in SEQ ID NO: 23 and an oligonucleotide having the sequence in SEQ ID NO: 27 were used as primers, and PCR was performed according to the method described hereinbefore. The PCR product was ligated into the cloning vector pCRII, the gene sequence of the PCR product was determined to prepare DNA having the gene sequence from No. 2986 to 3725 of DNA sequence in SEQ ID NO: 4, in which the DNA sequence encoding the FLAG sequence in the 3' end (DNA sequence in SEQ ID NO: 22), termination codon and restriction enzyme Sal I site were attached to the 3' end.

An about 3 kbp gene fragment obtained by digesting the pUCSR-2 with restriction enzymes EcoRI and BamHI, an about 700 bp gene fragment obtained by digesting the pCRII vector by restriction enzymes BamHI and Sal I, (the vector of which contained the above PCR product as an insert), and an about 4.3 kb gene fragment obtained by digesting the pMKITneo with restriction enzymes EcoRI and Xho I, were simultaneously ligated. Although the recognition sequence of restriction enzymes Xho I and Sal I is different, they can be ligated due to complementary terminal gene sequence. The expression vector containing the gene fragment of DNA sequence from No. 1 to 3725 in SEQ ID NO: 4 and gene fragment encoding the FLAG sequence was thus obtained. Full length human Serrate-2 FLAG chimeric protein (hereinafter designated as FS2FLAG) expression vector pMFS2FLAG was obtained.

Example 5

Expression and gene transfer of the human Serrate-2 expression vectors into cells



The expression vectors prepared in Example 4 were gene transduced into COS-7 cells (obtained from RIKEN Cell Bank, Physical and Chemical Research Institute, Japan, RCB0539).

Before the gene transduction was performed the cells were cultured in D-MEM (Dulbecco modified Eagle's medium, GIBCO-BRL Inc., U.S.A.) 10% FCS. On the day before gene transduction, the medium of cells was changed to achieve a cell count of 5×10^7 cells/ml and the cells cultured overnight. On the day of the gene transduction, the cells were sedimented by centrifugation, centrifugally washed twice with PBS (-) and made up to 1×10^7 cells/ml in PBS (-), 1 mM $MgCl_2$. The gene transfer was performed by electroporation using a gene transduction device Gene-pulsar (Bio-Rad Inc., U.S.A.). The above cell suspension (500 μ l) was collected in the cell for electroporation (0.4 mm), the expression vector (20 μ g) was added, and allowed to stand in ice for 5 minutes. Electroporation was performed under the condition 3 μ F, 450V twice, in between the electroporations cell mixture was allowed to stand at room temperature for 1 minute. Subsequent to the electroporations the mixture was incubated for 5 mins on ice, and then the cells were spread in 10ml of the culture medium described hereinbefore. The diameter of the dish was 10cm and the cells were cultured at 37°C in a 5% carbon dioxide incubator.

The next day, the culture supernatant solution was removed, the cells adhered to the dish were washed twice with PBS (-) 10 ml and 10 ml of serum-free D-MEM was added, and the cells cultured for 4 days. In the case of gene transduction into expression vectors pMEXS2, pMEXS2FLAG and pMEXS2Fc, the culture supernatant solution was recovered and a Centricon 30 (Amicon Inc., U.S.A.) was used to replace the medium with PBS(-) and simultaneously concentrate the solution.

In the case of gene transduction of pMSF2 and pMFS2FLAG, after 4 days culture, the cells were washed with PBS (-) 10 ml. The cells were scraped using cell scraper (Corster Corp.), and 10 ml of PBS (-) was added again. The cells were centrifuged at 1500 rpm for 5 minutes and washed. The cell precipitates



were suspended in the cell lysis buffer [50 mM Hepes (pH 7.5), 1% Triton X100, 10% glycerol, 4 mM EDTA, 50 μ g/ml Aprotinin, 100 μ M Leupeptin, 25 μ M Pepstatin A and 1 mM PMSF] 500 μ l, and allowed to stand in ice for 20 minutes, and centrifuged at 15000 rpm for 20 minutes to collect the supernatant solution to obtain cell lysates.

Using these samples, the expression of the FLAG chimeric and immunoglobulin chimeric proteins were detected using Western blotting. Namely, concentrated cultured supernatants or cell lysates were subjected to SDS-PAGE using an electrophoresis tank and polyacrylamide gel (gradient gel 5 - 15%)(ACI Japan Inc., Japan) according to the manufacturer's instructions. Samples were prepared by treatment in boiling water for 5 minutes with 2-mercaptoethanol (2-ME) for reduction, and for the non-reduced condition without boiling. Rainbow Markers (high molecular weight, Amersham Inc.) were used. Sample buffer solution and electrophoresis buffer were prepared with reference to the attached leaflet. When the SDS-PAGE was finished, the acrylamide gel was transferred to a PVDF membrane filter (BioRad Inc., U.S.A.) using the Mini Trans Blot Cell (BioRad Inc.).

The thus prepared filter was shaken overnight at 4°C in Blockace (Dainippon Pharm. Co., Japan), TBS-T [20 mM Tris, 137 mM NaCl (pH 7.6) and 0.1% Tween 20] to block. According to the manufacturers instructions of the ECL Western blotting detection system (Amersham Inc., U.S.A.); in case that protein was a FLAG chimera, anti-FLAG M2 mouse monoclonal antibody (Eastman Kodak, U.S.A.) was used as the primary antibody, and peroxidase labeled anti-mouse Ig sheep antibodies (Amersham Inc., U.S.A.) as a secondary antibody. In the case of a human IgG1Fc chimera, peroxidase labeled anti-human Ig sheep antibodies (Amersham Inc., U.S.A.) were reacted. Reaction time for the antibodies was 1 hour at room temperature, and in between each reaction, washing was performed by shaking in TBS-T at room temperature for 10 minutes, three times. After the final washing, the filter was immersed in the reaction solution of ECL-Western blotting detection system (Amersham Inc.,



U.S.A.) for 1 minute, and wrapped in polyvinylidene chloride wrap for exposure to X-ray film.

As a result, the bands showing the protein obtained by transduction of the expression vector pMEXS2FLAG was detected from the COS supernatant at about 135 kD by the anti-FLAGM2 antibody; and production of the desired protein EXS2FLAG was confirmed, from transduced cells containing the expression vector pMEXS2FLAG. Molecular weight changes as a result of the reduction treatment during SDS-PAGE were not observed. About 20 kD in sugar chains was added to the molecules, evidenced by comparing with the molecular weight estimated from the amino acid sequence.

Further, a band having a molecular weight of about 165 kD was detected in the supernatant solution of COS cells, to which the expression vector pMEXS2Fc was gene transduced, on SDS-PAGE using anti-human Ig sheep antibody under reducing conditions. A band having a molecular weight of about 330 kD was detected under non-reducing conditions. These results indicate that the desired protein EXS2Fc was produced, and consequently cells transformed by the expression vector pMEXS2Fc could be obtained. As the molecular weight of EXS2Fc under reducing conditions is about half of that under non-reducing conditions, the EXS2Fc is predicted to be a dimer connected by a disulfide bond. Further, the molecular weight of the band is about 40 kD larger than that calculated from the amino acid sequence. This indicates the addition of a sugar chain to the molecule.

Further, a band having a molecular weight of about 150 kD was detected from the extract of COS cells, to which the expression vector pMFS2FLAG was gene transduced, on SDS-PAGE using an anti-FLAG M2 antibody under reducing conditions. The results indicated that the desired protein FS2FLAG was produced, and consequently cells, which were transformed with the expression vector pMFS2FLAG, were obtained. As the molecular weight of the band of FS2FLAG is about 20 kD larger than that calculated based on the amino acid sequence, sugar chains may have been added to the extracellular region.



For proteins other than chimeric proteins, detection was conducted using anti-human Serrate-2 mouse monoclonal antibody and anti-human Serrate -2 rabbit polyclonal antibody, which were described in Example 7, as primary antibodies in Western blotting. Anti-mouse Ig sheep antibody (Amersham Inc.) or peroxidase labeled rabbit Ig sheep antibody (Amersham Inc.) were used as secondary antibodies.

As a result, a band having molecular weight of about 135 kD was detected in the supernatant solution of COS cells, to which the expression vector pMEXS2 was gene transduced. This indicated that the desired protein EXS2 was produced, and cells transformed by an expression vector pMEXS2 could be obtained. No changes in the molecular weight were observed due to reduction treatment during SDS-PAGE. A band having molecular weight about 150 kD was detected under reducing conditions from the COS cell extract, to which the expression vector pMFS2 was gene transduced. These results indicated that the desired protein FS2 was produced, and consequently cells transformed by the expression vector pMFS2 could be obtained. Further, in every case, the molecular weight of the band is about 20 kD larger than that calculated from amino acid sequence. This indicates the addition of sugar chains to the molecule in the extracellular region.

In the control experiments, cell lysate and cultured supernatant solution of COS-7 cells, to which pMKITNeo vector was transformed, were tested. No bands reacting with the anti-FLAG antibody, anti-human Ig antibody or anti-human Serrate-2 antibody could be detected.

Example 6

Purification of secretory extracellular human Serrate-2 chimera proteins of gene transduction cells

The cultured supernatant from COS-7 cells transformed by the expression vector pMEXS2FLAG or pMEXS2Fc (by the method described in Example 5), were prepared in large scale, and the chimeric protein, i.e. EXS2FLAG or EXS2Fc, was purified using an affinity column.



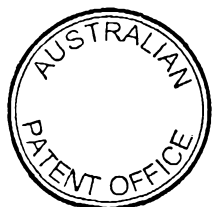
In case of EXS2FLAG, 2 liters of the cultured supernatant obtained by the method described in Example 5 was passed through a column packed with Anti-FLAG M2 Affinity Gel (Eastman Kodak, U.S.A.). The chimeric protein was adsorbed to the column due to the affinity of anti-FLAG antibody in the gel and the FLAG sequence of the chimeric protein. A disposable column, with an inner diameter 10 mm, (BioRad Inc., U.S.A.) was used packed with 5ml of the above gel. A circulation system consisting of medium bottle → column → peristaltic pump → medium bottle was set up. The circulation was run using a flow of 1 ml/min. for 72 hours. Thereafter the column was washed with PBS (-) 35 ml and eluted using 0.5 M Tris-glycine (pH 3.0) 50 ml. The eluate of 25 fractions, each 2 ml, was collected into the tube (Farcon Inc. 2063), and each fraction was neutralized by 200 transformed μ l of 0.5 M Tris-HCl (pH 9.5) previously added to each tube.

Each 10 μ l fraction of the EXS2FLAG which was purified by the above method was subjected to reduction treatment described in Example 5. SDS-PAGE electrophoresis using a 5 - 10 % gradient polyacrylamide gel was performed. After completion of the electrophoresis, silver staining was conducted by using Wako silver stain kit II according to the manufacturers instructions. Fractions from No. 4 to 8 showed detectable bands of EXS2FLAG. The size is identical with the result of Western blotting with anti-FLAG antibody obtained in Example 5. Therefore, purified EXS2FLAG was obtained.

For EXS2Fc, the cultured supernatant solution (2 liters) was adsorbed to the Protein A Sepharose column (Pharmacia Inc., Sweden), according to the same method as above, to collect the eluate fractions. Using the same quantity as for EXS2FLAG, a determination of the eluate fraction, identification of the size, and detection of the purity were performed using SDS-PAGE electrophoresis and silver staining under reducing conditions. Bands were detected in fractions from No. 4 to 15. The molecular weight thereof is identical with the result in Example 5 and therefore, purified EXS2Fc was obtained.

Example 7

Preparation of antibodies recognizing human Serrate-2



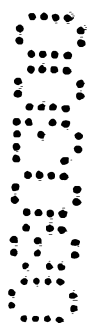
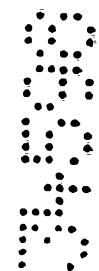
EXS2FLAG, purified by the method in Example 6, was used as an immunogen, and rabbits were immunized. After assaying the antibody titer, whole blood was collected and serum was obtained. Anti-human Serrate-2 rabbit polyclonal antibody was purified using Econopack serum IgG purification kit (BioRad Inc., U.S.A.) with reference to the manufacturers instructions.

EXS2FLAG purified by a method described in Example 6, was used as an immunogen, and mouse monoclonal antibodies were prepared according to methods known to skilled persons. The purified HSFLAG was administered to Balb/c mice (Nippon SLC Co., Japan), 10 μ g/mouse, immunized intracutaneously and subcutaneously. After the second immunization, an increased serum titer was confirmed by collecting blood ophthalmologically, the third immunization was performed. Subsequently, the spleen of the mice was collected and fused with mouse myeloma cells P3X63Ag8 (ATCC TIB9) using polyethylene glycol. Hybridoma's were selected using HAT medium (Immunological and Biological Research Institute, Japan), and the hybridoma strains, which produced antibodies specifically recognizing the extracellular region of human Serrate-2 in the medium, were isolated by enzyme immunoassay. The hybridoma strains producing a mouse monoclonal antibody, which specifically recognized human Serrate-2, were established.

Anti-human Serrate-2 monoclonal antibody was purified and prepared using Mab Trap GII (Pharmacia Inc., Sweden) and according to the manufacturers instructions, from the supernatant of the thus established hybridoma.

An affinity column was prepared using the monoclonal antibody. Preparation of the affinity column was performed according to the manufacturers instructions attached to the CNBr activated Sephadex 4B (Pharmacia Inc., Sweden). Coupling efficiency was 99.6%. A column, 2 cm² x 1 cm, containing gel 2 ml, was prepared.

A supernatant of the cultured cells, which contained EXS2, was passed through the column. The supernatant solution was passed at 20 ml/hr, subsequently PBS (-) 15 ml was passed at the same flow rate to wash the



column. Finally, the products were eluted using a mixture of 0.1 M sodium acetate and 0.5 M NaCl (pH 4.0). The eluate, each 1 ml fraction, was collected, and was neutralized by adding 1 M Tris-HCl (pH 9.1) 200 μ l to each fraction.

SDS-PAGE of the purified protein was conducted under reducing conditions according to the method described in Example 5, followed by silver staining and Western blotting to estimate the molecular weight. An about 140 kD band was detected. Consequently, Western blotting can be carried out using the said monoclonal antibodies and human Serrate-2 can be purified using the affinity columns.

Example 8

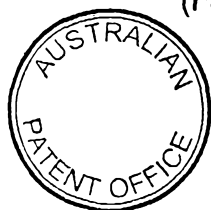
Effects of human Serrate-2 protein on colony formation of blood undifferentiated cells

In order to observe physiological action of human Serrate-2 on blood undifferentiated cells, CD34 positive cells were cultured in serum-free semi solid medium in the presence of EXS2Fc and known cytokines, and the number of colony forming cells were observed.

CD34 positive cells of human umbilical cord blood or human normal bone marrow blood were isolated from the mononuclear cells, which were treated using silica solution (Immunological and Biological Research Institute, Japan) according to the manufacturers instructions and fractionated from the low density cellular fraction (<1.077 g/ml) by densitometric centrifugation of Ficoll pack (Pharmacia Inc., Sweden).

Separation of CD34 positive cells was performed by using Dynabeads M-45 CD34 or DETACH a BEADS CD34 (Dynal Inc., Norway) and according to the manufacturers instructions. After separation, the purity was measured as follows. Cells were stained using FITC labeled CD34 antibody HPCA2 (Beckton-Deckinson Inc., U.S.A.) and examined using a flow-cytometer (FACS Calibur, Beckton-Deckinson, U.S.A.). Purity above 85% was confirmed for use.

The thus isolated CD34 positive cells were suspended homogeneously to form 400 cells/ml of the medium hereinbelow, and spread in a 35 mm dish (Falcon Inc., U.S.A.), then cultured for 2 weeks in a carbon dioxide incubator at



37°C under 5% carbon dioxide, 5% oxygen, 90% nitrogen and 100% humidity. The formed blood colonies were counted under the invert microscope.

The medium used was α -medium (GIBCO-BRL, Y.S.A.), containing 2% deionized bovine serum albumin (BSA, Sigma, U.S.A.), 10 μ g/ml human insulin (Sigma, U.S.A.), 200 μ g/ml transferrin (Sigma, U.S.A.), 10^{-5} M 2-mercaptoethanol (Nakarai Tesk Co., Japan), 160 μ g/ml soybean lecithin (Sigma, U.S.A.), 96 μ g/ml cholesterol (Sigma, U.S.A.) and 0.9 % methylcellulose (Wako Pure Chemicals, Japan).

To the above medium, with the concentration of cytokines described below, human Serrate-2 extracellular Ig chimera protein (EXS2Fc) was added to a final concentration of 1 μ g/ml. As a control, human IgG1 (Athens Research and Technology Inc., U.S.A.) was added to the same concentration in order to observe the effect of the IgGFc region.

Concentration of cytokines was as follows.

100 ng/ml, human SCF, 10 ng/ml human IL-3, 100 ng/ml human IL-6, 2U/ml Epo (Chugai Seiyaku Co., Japan) and 10 ng/ml human G-CSF (Chugai Seiyaku Co., Japan).

Results are shown in Table 1. Number of colonies/CD 34⁺ cells 400 cells are shown in mean of n = 3. Four different origin human umbilical cord blood CD34 positive cells were used.

Table 1

	EXS2Fc not added	EXS2Fc added
Experiment 1	30.1	12.8
Experiment 2	48.3	40.7
Experiment 3	38.2	28.1
Experiment 4	50.9	37.1

As shown in Table 1, human Serrate-2 has an action against four different origin human umbilical cord blood CD34 positive cells. Therefore, human



Serrate-2 of the present invention has suppressive action for differentiation of blood undifferentiated cells including blood cells.

Example 9

Effect of human Serrate-2 on blood undifferentiated cell LTC-IC in liquid culture

In order to observe the physiological action of human Serrate-2 on the blood undifferentiated cells, umbilical cord blood CD34 positive cells were cultured for two weeks in the serum-free liquid medium in the presence of EXS2Fc and known cytokines, and numbers of LTC-IC, which is presently thought to be most undifferentiated cells, were observed.

The umbilical cord blood mononuclear CD34 positive cells, 100000 to 20000 cells, separated by a method described in Example 8 were cultured in the following medium for 2 weeks. Numbers of LTC-IC in 3 experimental groups, which include a group before cultivation, a group of EXS2Fc and a control group, were determined.

The liquid culture medium was α medium with 2% BSA, 10 μ g/ml human insulin, 200 μ g/ml transferrin, 40 μ g/ml low density lipoprotein, and 10^{-5} M 2-mecaptoethanol, and 100 ng/ml human SCF, 10 ng/ml human IL-3, and 100 ng/ml human IL-6. EXS2Fc 1 μ g/ml was added to the above medium. In the control group, human IgG1 was added in an equal concentration.

Preparation of human bone marrow stromal cell layer used for LTC-IC, and quantitative assay of frequency of LTC-IC by a limiting dilution were performed according to the method of Sutherland et al. (Blood, 74, 1563-, 1989 and Proc. Natl. Acad. Sci. USA, 87, 3584-, 1990).

The bone marrow mononuclear cells, $1 - 2 \times 10^7$ cells, obtained in Example 8, before the separation and without the liquid silica treatment, were cultured in LTC medium (MyeloCult, Stem Cell Technologies Inc., Canada) 5 ml with hydrocortisone 1 μ M (Upjohn Japan Co., Japan), in T - 25 flask (Falcon Inc., U.S.A.) at 37°C under 5% carbon dioxide and 100% humidity in the carbon dioxide incubator. Culturing was carried out until the adhesive cell layers of the



stromal cell formation covered more than 80% of the bottom of the culture flask. Detachment of the cell layer was performed by treating it with EDTA solution (Cosmobio Co., Japan). The cells were plated into a 96 well plate (Beckton-Deckinson Inc., U.S.A.), about 2×10^4 cells/well and re-cultivation was continued in the same medium. X-ray, at 15 Gy, 250 KV peak was used to irradiate the reconstituted stromal cell layer. The growth of stromal cells was stopped and blood cells in the stromal cells were removed. The thus prepared stromal cells were used as the stromal cell layer for the experiments.

In the assay of LTC-IC, cell counts in each group were adjusted within the ranges of 25 - 400 cells/well for CD34 positive cells before cultivation, and 625 - 20000 cells/well for the cells after cultivation, and cells were diluted in a six step-dilution within these ranges. Each dilute step of cells was cocultured with the above stromal cell layer in the 96 well plate, 16 wells for each dilution step. Culture was performed in the same medium as used in stromal formation, at 37°C under 5% carbon dioxide and 100% humidity in the carbondioxide gas incubator for 5 weeks. Cells in suspension and in attachment after cultivation were recovered in each well. Collected cells were transferred to the semi-solid culture medium consisting of α -medium added with 0.9% methylcellulose, 30% fetal calf serum (FCS, ICN Biomedical Japan), 1% BSA, 10^{-5} M 2-mercaptoethanol, 100 ng/ml human SCF, 10 ng/ml human IL-3, 100 ng/ml human IL-6, 2 U/ml Epo and 10 ng/ml human G-CSF. After 2 weeks of cultivation, colony forming cells were detected as the same way as described in Example 8, and the numbers of wells, in which colony forming cells were found, were recorded. Incidence of LTC-IC was calculated according to the method of Taswell et al. (J. Immunol. 126, 1614-, 1981) based on the above data. Results are shown in Table 2.

Table 2

	Before cultivation	EXS2Fc not added	EXS2Fc added
Total number of cells	19475	1210000	930000
Number of LTC-IC	185	23	5



From the result shown in Table 2, human Serrate-2 has an action against LTC-IC and reduces its number.

Example 10

Effect of human Serrate-2 on growth of vascular endothelial cells

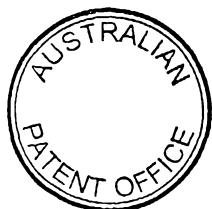
Vascular endothelial cells used were passage cultures of four generations of normal human aortic endothelial cells and normal human pulmonary arterial endothelial cells (Kurabo Inc., Japan).

Cells were plated at 500 cells/well in a 96 well plate for tissue culture (Falcon Inc., U.S.A.) in the tertiary passage culture, and cultured in a medium with low serum level for growth of vascular endothelial cells (HuMedia-EG2, Kurabo Inc., Japan) containing human recombinant EGF (Kurabo Inc., Japan) 100 ng/ml and human recombinant FGF-B 5 ng/ml. Human Serrate-2 extracellular chimera protein (EXS2Fc) was added to the final concentration of 1 μ g/ml. As a control, human IgG1 (Athens Research and Technology Inc., U.S.A.) was added with the same concentration in order to observe the effect of the IgGFc region. Control experiments were conducted without adding protein except for HuMedia-EG2. Culture was performed at 37°C, under 5% carbon dioxide and 100% humidity for 3 days and the number of cells was calculated.

Vascular endothelial cell counts were performed by using NR reagent set (Kurabo Inc., Japan). The method was developed by Borenfreund and Puerner (Journal of Tissue Culture Methods, 9(1), 7 - 9, 1984), i.e. the neutral red method in which vital stain pigment neutral red (3-amino-7-dimethylamino-2-methylphenazine hydrochloride) passed through the plasma membrane of living cells and was accumulated in the lysosome.

Absorption at 540 nm was measured using an immuno reader (NJ-2000, Japan Intermed Inc., Japan).

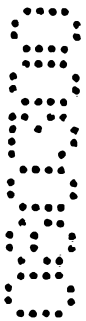
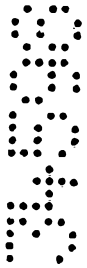
Results showed that in case of aortic endothelial cells, absorption in the control group was optical density (OD) 0.21 +/- 0.02, which is almost the same level as the human IgG1 group 0.20 +/- 0.01, and in the EXS2Fc group, it was 0.10 +/- 0.02 which was significantly lower than in the control.



In case of pulmonary arterial endothelial cells, control group showed 0.15 +/- 0.01 and the human IgG1 group showed almost same level 0.16 +/- 0.02, whereas the EXS2Fc group shows a significantly lower level of 0.07 +/- 0.02. This result indicated that EXS2Fc suppresses growth of vascular endothelial cells.

Effect of the invention

Human Serrate-2 of the present invention has an action for regulating differentiation of undifferentiated cells, and can be used as a novel regulating agent for differentiation of cells.



EDITORIAL NOTE

NO. 34598/97

There are no pages numbered 52 to 54.

SEQUENCE LISTING

INFORMATION FOR SEQ ID NO : 1

LENGTH : 214

TYPE : amino acid

STRANDEDNESS : single stranded

TOPOLOGY : linear

MOLECULE TYPE : amino acid

ORIGINAL SOURCE

ORGANISM : human

SEQUENCE DESCRIPTION

Met Gly Tyr Phe Glu Leu Gln Leu Ser Ala Leu Arg Asn Val Asn Gly

1 5 10 15

Glu Leu Leu Ser Gly Ala Cys Cys Asp Gly Asp Gly Arg Thr Thr Arg

20 25 30

Ala Gly Gly Cys Gly His Asp Glu Cys Asp Thr Tyr Val Arg Val Cys

35 40 45

Leu Lys Glu Tyr Gln Ala Lys Val Thr Pro Thr Gly Pro Cys Ser Tyr

50 55 60

Gly His Gly Ala Thr Pro Val Leu Gly Gly Asn Ser Phe Tyr Leu Pro

65 70 75 80

Pro Ala Gly Ala Ala Gly Asp Arg Ala Arg Ala Arg Ala Gly

85 90 95

Gly Asp Gln Asp Pro Gly Leu Val Val Ile Pro Phe Gln Phe Ala Trp

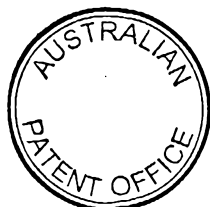
100 105 110

Pro Arg Ser Phe Thr Leu Ile Val Glu Ala Trp Asp Trp Asp Asn Asp

115 120 125

Thr Thr Pro Asn Glu Glu Leu Leu Ile Glu Arg Val Ser His Ala Gly

130 135 140



Met Ile Asn Pro Glu Asp Arg Trp Lys Ser Leu His Phe Ser Gly His
145 150 155 160
Val Ala His Leu Glu Leu Gln Ile Arg Val Arg Cys Asp Glu Asn Tyr
165 170 175
Tyr Ser Ala Thr Cys Asn Lys Phe Cys Arg Pro Arg Asn Asp Phe Phe
180 185 190
Gly His Tyr Thr Cys Asp Gln Tyr Gly Asn Lys Ala Cys Met Asp Gly
195 200 205
Trp Met Gly Lys Glu Cys
210 214

INFORMATION FOR SEQ ID NO : 2

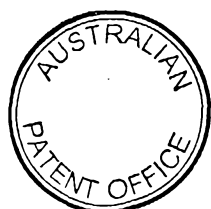
LENGTH : 1055
TYPE : amino acid
STRANDEDNESS : single stranded
TOPOLOGY : linear
MOLECULE TYPE : amino acid

ORIGINAL SOURCE

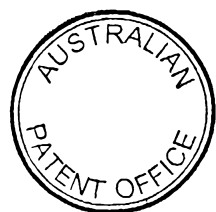
ORGANISM : human

SEQUENCE DESCRIPTION

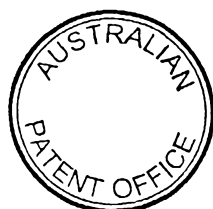
Met Gly Tyr Phe Glu Leu Gln Leu Ser Ala Leu Arg Asn Val Asn Gly
1 5 10 15
Glu Leu Leu Ser Gly Ala Cys Cys Asp Gly Asp Gly Arg Thr Thr Arg
20 25 30
Ala Gly Gly Cys Gly His Asp Glu Cys Asp Thr Tyr Val Arg Val Cys
35 40 45
Leu Lys Glu Tyr Gln Ala Lys Val Thr Pro Thr Gly Pro Cys Ser Tyr
50 55 60



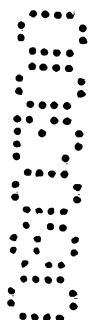
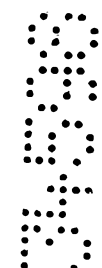
Gly	His	Gly	Ala	Thr	Pro	Val	Leu	Gly	Gly	Asn	Ser	Phe	Tyr	Leu	Pro
65					70					75					80
Pro	Ala	Gly	Ala	Ala	Gly	Asp	Arg	Ala	Arg	Ala	Arg	Ala	Arg	Ala	Gly
					85					90					95
Gly	Asp	Gln	Asp	Pro	Gly	Leu	Val	Val	Ile	Pro	Phe	Gln	Phe	Ala	Trp
				100					105					110	
Pro	Arg	Ser	Phe	Thr	Leu	Ile	Val	Glu	Ala	Trp	Asp	Trp	Asp	Asn	Asp
			115				120						125		
Thr	Thr	Pro	Asn	Glu	Glu	Leu	Leu	Ile	Glu	Arg	Val	Ser	His	Ala	Gly
	130					135						140			
Met	Ile	Asn	Pro	Glu	Asp	Arg	Trp	Lys	Ser	Leu	His	Phe	Ser	Gly	His
145					150					155				160	
Val	Ala	His	Leu	Glu	Leu	Gln	Ile	Arg	Val	Arg	Cys	Asp	Glu	Asn	Tyr
					165					170				175	
Tyr	Ser	Ala	Thr	Cys	Asn	Lys	Phe	Cys	Arg	Pro	Arg	Asn	Asp	Phe	Phe
			180					185					190		
Gly	His	Tyr	Thr	Cys	Asp	Gln	Tyr	Gly	Asn	Lys	Ala	Cys	Met	Asp	Gly
		195					200						205		
Trp	Met	Gly	Lys	Glu	Cys	Lys	Glu	Ala	Val	Cys	Lys	Gln	Gly	Cys	Asn
	210					215						220			
Leu	Leu	His	Gly	Gly	Cys	Thr	Val	Pro	Gly	Glu	Cys	Arg	Cys	Ser	Tyr
225					230					235				240	
Gly	Trp	Gln	Gly	Arg	Phe	Cys	Asp	Glu	Cys	Val	Pro	Tyr	Pro	Gly	Cys
				245						250				255	
Val	His	Gly	Ser	Cys	Val	Glu	Pro	Trp	Gln	Cys	Asn	Cys	Glu	Thr	Asn
			260					265					270		
Trp	Gly	Gly	Leu	Leu	Cys	Asp	Lys	Asp	Leu	Asn	Tyr	Cys	Gly	Ser	His
		275					280						285		



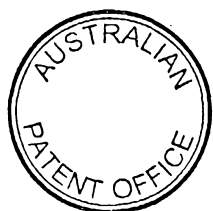
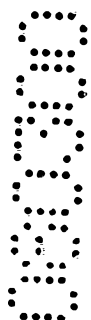
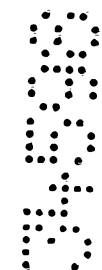
His Pro Cys Thr Asn Gly Gly Thr Cys Ile Asn Ala Glu Pro Asp Gln
 290 295 300
 Tyr Arg Cys Thr Cys Pro Asp Gly Tyr Ser Gly Arg Asn Cys Glu Lys
 305 310 315 320
 Ala Glu His Ala Cys Thr Ser Asn Pro Cys Ala Asn Gly Gly Ser Cys
 325 330 335
 His Glu Val Pro Ser Gly Phe Glu Cys His Cys Pro Ser Gly Trp Ser
 340 345 350
 Gly Pro Thr Cys Ala Leu Asp Ile Asp Glu Cys Ala Ser Asn Pro Cys
 355 360 365
 Ala Ala Gly Gly Thr Cys Val Asp Gln Val Asp Gly Phe Glu Cys Ile
 370 375 380
 Cys Pro Glu Gln Trp Val Gly Ala Thr Cys Gln Leu Asp Ala Asn Glu
 385 390 395 400
 Cys Glu Gly Lys Pro Cys Leu Asn Ala Phe Ser Cys Lys Asn Leu Ile
 405 410 415
 Gly Gly Tyr Tyr Cys Asp Cys Ile Pro Gly Trp Lys Gly Ile Asn Cys
 420 425 430
 His Ile Asn Val Asn Asp Cys Arg Gly Gln Cys Gln His Gly Gly Thr
 435 440 445
 Cys Lys Asp Leu Val Asn Gly Tyr Gln Cys Val Cys Pro Arg Gly Phe
 450 455 460
 Gly Gly Arg His Cys Glu Leu Glu Arg Asp Lys Cys Ala Ser Ser Pro
 465 470 475 480
 Cys His Ser Gly Gly Leu Cys Glu Asp Leu Ala Asp Gly Phe His Cys
 485 490 495
 His Cys Pro Gln Gly Phe Ser Gly Pro Leu Cys Glu Val Asp Val Asp
 500 505 510



Leu Cys Glu Pro Ser Pro Cys Arg Asn Gly Ala Arg Cys Tyr Asn Leu
 515 520 525
 Glu Gly Asp Tyr Tyr Cys Ala Cys Pro Asp Asp Phe Gly Gly Lys Asn
 530 535 540
 Cys Ser Val Pro Arg Glu Pro Cys Pro Gly Gly Ala Cys Arg Val Ile
 545 550 555 560
 Asp Gly Cys Gly Ser Asp Ala Gly Pro Gly Met Pro Gly Thr Ala Ala
 565 570 575
 Ser Gly Val Cys Gly Pro His Gly Arg Cys Val Ser Gln Pro Gly Gly
 580 585 590
 Asn Phe Ser Cys Ile Cys Asp Ser Gly Phe Thr Gly Thr Tyr Cys His
 595 600 605
 Glu Asn Ile Asp Asp Cys Leu Gly Gln Pro Cys Arg Asn Gly Gly Thr
 610 615 620
 Cys Ile Asp Glu Val Asp Ala Phe Arg Cys Phe Cys Pro Ser Gly Trp
 625 630 635 640
 Glu Gly Glu Leu Cys Asp Thr Asn Pro Asn Asp Cys Leu Pro Asp Pro
 645 650 655
 Cys His Ser Arg Gly Arg Cys Tyr Asp Leu Val Asn Asp Phe Tyr Cys
 660 665 670
 Ala Cys Asp Asp Gly Trp Lys Gly Lys Thr Cys His Ser Arg Glu Phe
 675 680 685
 Gln Cys Asp Ala Tyr Thr Cys Ser Asn Gly Gly Thr Cys Tyr Asp Ser
 690 695 700
 Gly Asp Thr Phe Arg Cys Ala Cys Pro Pro Gly Trp Lys Gly Ser Thr
 705 710 715 720
 Cys Ala Val Ala Lys Asn Ser Ser Cys Leu Pro Asn Pro Cys Val Asn
 725 730 735



Gly Gly Thr Cys Val Gly Ser Gly Ala Ser Phe Ser Cys Ile Cys Arg			
740	745	750	
Asp Gly Trp Glu Gly Arg Thr Cys Thr His Asn Thr Asn Asp Cys Asn			
755	760	765	
Pro Leu Pro Cys Tyr Asn Gly Gly Ile Cys Val Asp Gly Val Asn Trp			
770	775	780	
Phe Arg Cys Glu Cys Ala Pro Gly Phe Ala Gly Pro Asp Cys Arg Ile			
785	790	795	800
Asn Ile Asp Glu Cys Gln Ser Ser Pro Cys Ala Tyr Gly Ala Thr Cys			
805	810	815	
Val Asp Glu Ile Asn Gly Tyr Arg Cys Ser Cys Pro Pro Gly Arg Ala			
820	825	830	
Gly Pro Arg Cys Gln Glu Val Ile Gly Phe Gly Arg Ser Cys Trp Ser			
835	840	845	
Arg Gly Thr Pro Phe Pro His Gly Ser Ser Trp Val Glu Asp Cys Asn			
850	855	860	
Ser Cys Arg Cys Leu Asp Gly Arg Arg Asp Cys Ser Lys Val Trp Cys			
865	870	875	880
Gly Trp Lys Pro Cys Leu Leu Ala Gly Gln Pro Glu Ala Leu Ser Ala			
885	890	895	
Gln Cys Pro Leu Gly Gln Arg Cys Leu Glu Lys Ala Pro Gly Gln Cys			
900	905	910	
Leu Arg Pro Pro Cys Glu Ala Trp Gly Glu Cys Gly Ala Glu Glu Pro			
915	920	925	
Pro Ser Thr Pro Cys Leu Pro Arg Ser Gly His Leu Asp Asn Asn Cys			
930	935	940	
Ala Arg Leu Thr Leu His Phe Asn Arg Asp His Val Pro Gln Gly Thr			
945	950	955	960



Thr Val Gly Ala Ile Cys Ser Gly Ile Arg Ser Leu Pro Ala Thr Arg
 965 970 975
 Ala Val Ala Arg Asp Arg Leu Leu Val Leu Leu Cys Asp Arg Ala Ser
 980 985 990
 Ser Gly Ala Ser Ala Val Glu Val Ala Val Ser Phe Ser Pro Ala Arg
 995 1000 1005
 Asp Leu Pro Asp Ser Ser Leu Ile Gln Gly Ala Ala His Ala Ile Val
 1010 1015 1020
 Ala Ala Ile Thr Gln Arg Gly Asn Ser Ser Leu Leu Leu Ala Val Thr
 1025 1030 1035 1040
 Glu Val Lys Val Glu Thr Val Val Thr Gly Gly Ser Ser Thr Gly
 1045 1050 1055

INFORMATION FOR SEQ ID NO : 3

LENGTH : 1215

TYPE : amino acid

STRANDEDNESS : single stranded

TOPOLOGY : linear

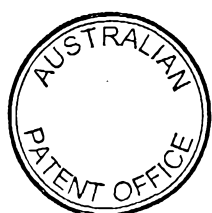
MOLECULE TYPE : amino acid

ORIGINAL SOURCE

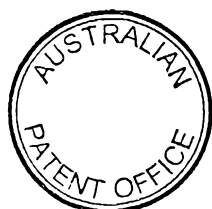
ORGANISM : human

SEQUENCE DESCRIPTION

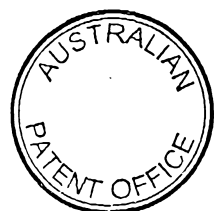
Met Gly Tyr Phe Glu Leu Gln Leu Ser Ala Leu Arg Asn Val Asn Gly
 1 5 10 15
 Glu Leu Leu Ser Gly Ala Cys Cys Asp Gly Asp Gly Arg Thr Thr Arg
 20 25 30
 Ala Gly Gly Cys Gly His Asp Glu Cys Asp Thr Tyr Val Arg Val Cys
 35 40 45



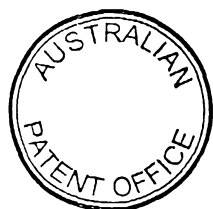
Leu Lys Glu Tyr Gln Ala Lys Val Thr Pro Thr Gly Pro Cys Ser Tyr
 50 55 60
 Gly His Gly Ala Thr Pro Val Leu Gly Gly Asn Ser Phe Tyr Leu Pro
 65 70 75 80
 Pro Ala Gly Ala Ala Gly Asp Arg Ala Arg Ala Arg Ala Arg Ala Gly
 85 90 95
 Gly Asp Gln Asp Pro Gly Leu Val Val Ile Pro Phe Gln Phe Ala Trp
 100 105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



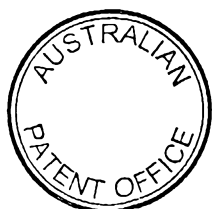
Trp Gly Gly Leu Leu Cys Asp Lys Asp Leu Asn Tyr Cys Gly Ser His	275	280	285
His Pro Cys Thr Asn Gly Gly Thr Cys Ile Asn Ala Glu Pro Asp Gln	290	295	300
Tyr Arg Cys Thr Cys Pro Asp Gly Tyr Ser Gly Arg Asn Cys Glu Lys	305	310	315
Ala Glu His Ala Cys Thr Ser Asn Pro Cys Ala Asn Gly Gly Ser Cys	325	330	335
His Glu Val Pro Ser Gly Phe Glu Cys His Cys Pro Ser Gly Trp Ser	340	345	350
Gly Pro Thr Cys Ala Leu Asp Ile Asp Glu Cys Ala Ser Asn Pro Cys	355	360	365
Ala Ala Gly Gly Thr Cys Val Asp Gln Val Asp Gly Phe Glu Cys Ile	370	375	380
Cys Pro Glu Gln Trp Val Gly Ala Thr Cys Gln Leu Asp Ala Asn Glu	385	390	395
Cys Glu Gly Lys Pro Cys Leu Asn Ala Phe Ser Cys Lys Asn Leu Ile	405	410	415
Gly Gly Tyr Tyr Cys Asp Cys Ile Pro Gly Trp Lys Gly Ile Asn Cys	420	425	430
His Ile Asn Val Asn Asp Cys Arg Gly Gln Cys Gln His Gly Gly Thr	435	440	445
Cys Lys Asp Leu Val Asn Gly Tyr Gln Cys Val Cys Pro Arg Gly Phe	450	455	460
Gly Gly Arg His Cys Glu Leu Glu Arg Asp Lys Cys Ala Ser Ser Pro	465	470	475
Cys His Ser Gly Gly Leu Cys Glu Asp Leu Ala Asp Gly Phe His Cys	485	490	495



His Cys Pro Gln Gly Phe Ser Gly Pro Leu Cys Glu Val Asp Val Asp			
500	505	510	
Leu Cys Glu Pro Ser Pro Cys Arg Asn Gly Ala Arg Cys Tyr Asn Leu			
515	520	525	
Glu Gly Asp Tyr Tyr Cys Ala Cys Pro Asp Asp Phe Gly Gly Lys Asn			
530	535	540	
Cys Ser Val Pro Arg Glu Pro Cys Pro Gly Gly Ala Cys Arg Val Ile			
545	550	555	560
Asp Gly Cys Gly Ser Asp Ala Gly Pro Gly Met Pro Gly Thr Ala Ala			
565	570	575	
Ser Gly Val Cys Gly Pro His Gly Arg Cys Val Ser Gln Pro Gly Gly			
580	585	590	
Asn Phe Ser Cys Ile Cys Asp Ser Gly Phe Thr Gly Thr Tyr Cys His			
595	600	605	
Glu Asn Ile Asp Asp Cys Leu Gly Gln Pro Cys Arg Asn Gly Gly Thr			
610	615	620	
Cys Ile Asp Glu Val Asp Ala Phe Arg Cys Phe Cys Pro Ser Gly Trp			
625	630	635	640
Glu Gly Glu Leu Cys Asp Thr Asn Pro Asn Asp Cys Leu Pro Asp Pro			
645	650	655	
Cys His Ser Arg Gly Arg Cys Tyr Asp Leu Val Asn Asp Phe Tyr Cys			
660	665	670	
Ala Cys Asp Asp Gly Trp Lys Gly Lys Thr Cys His Ser Arg Glu Phe			
675	680	685	
Gln Cys Asp Ala Tyr Thr Cys Ser Asn Gly Gly Thr Cys Tyr Asp Ser			
690	695	700	
Gly Asp Thr Phe Arg Cys Ala Cys Pro Pro Gly Trp Lys Gly Ser Thr			
705	710	715	720



Ala Arg Leu Thr Leu His Phe Asn Arg Asp His Val Pro Gln Gly Thr			
945	950	955	960
Thr Val Gly Ala Ile Cys Ser Gly Ile Arg Ser Leu Pro Ala Thr Arg			
965	970	975	
Ala Val Ala Arg Asp Arg Leu Leu Val Leu Leu Cys Asp Arg Ala Ser			
980	985	990	
Ser Gly Ala Ser Ala Val Glu Val Ala Val Ser Phe Ser Pro Ala Arg			
995	1000	1005	
Asp Leu Pro Asp Ser Ser Leu Ile Gln Gly Ala Ala His Ala Ile Val			
1010	1015	1020	
Ala Ala Ile Thr Gln Arg Gly Asn Ser Ser Leu Leu Leu Ala Val Thr			
1025	1030	1035	1040
Glu Val Lys Val Glu Thr Val Val Thr Gly Gly Ser Ser Thr Gly Leu			
1045	1050	1055	
Leu Val Pro Val Leu Cys Gly Ala Phe Ser Val Leu Trp Leu Ala Cys			
1060	1065	1070	
Val Val Leu Cys Val Trp Trp Thr Arg Lys Arg Arg Lys Glu Arg Glu			
1075	1080	1085	
Arg Ser Arg Leu Pro Arg Glu Glu Ser Ala Asn Asn Gln Trp Ala Pro			
1090	1095	1100	
Leu Asn Pro Ile Arg Asn Pro Ile Glu Arg Pro Gly Gly His Lys Asp			
1105	1110	1115	1120
Val Leu Tyr Gln Cys Lys Asn Phe Thr Pro Pro Pro Arg Arg Ala Asp			
1125	1130	1135	
Glu Ala Leu Pro Gly Pro Ala Gly His Ala Ala Val Arg Glu Asp Glu			
1140	1145	1150	
Glu Asp Glu Asp Leu Gly Arg Gly Glu Glu Asp Ser Leu Glu Ala Glu			
1155	1160	1165	



Lys Phe Leu Ser His Lys Phe Thr Lys Asp Pro Gly Arg Ser Pro Gly
 1170 1175 1180
 Arg Pro Ala His Trp Ala Ser Gly Pro Lys Val Asp Asn Arg Ala Val
 1185 1190 1195 1200
 Arg Ser Ile Asn Glu Ala Arg Tyr Ala Gly Lys Glu
 1205 1210 1212

INFORMATION FOR SEQ ID NO : 4

LENGTH : 3955 and 1238

TYPE : nucleic acid and amino acid

STRANDEDNESS : double stranded and single stranded

TOPOLOGY : linear

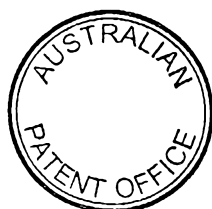
MOLECULE TYPE : cDNA to mRNA and amino acid

ORIGINAL SOURCE

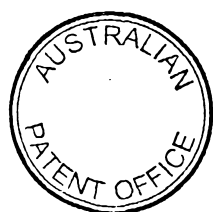
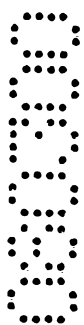
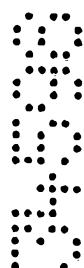
ORGANISM : human

SEQUENCE DESCRIPTION

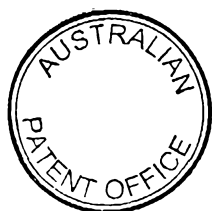
T CGCGGGGGCA ATG CGG GCG CAG GGC CGG GGG CGC CTT CCC 41
 Met Arg Ala Gln Gly Arg Gly Arg Leu Pro
 -26 -25 -20
 CGG CGG CTG CTG CTG CTG CTG GCG CTC TGG GTG CAG GCG GCG CGG CCC 89
 Arg Arg Leu Leu Leu Leu Leu Ala Leu Trp Val Gln Ala Ala Arg Pro
 -15 -10 -5 -1
 ATG GGC TAT TTC GAG CTG CAG CTG AGC GCG CTG CGG AAC GTG AAC GGG 137
 Met Gly Tyr Phe Glu Leu Gln Leu Ser Ala Leu Arg Asn Val Asn Gly
 1 5 10 15
 GAG CTG CTG AGC GCG GCC TGC TGT GAC GCG GAC GCG CGG ACA ACC CGC 185
 Glu Leu Leu Ser Gly Ala Cys Cys Asp Gly Asp Gly Arg Thr Thr Arg
 20 25 30



CCG GGG GGC TGC GGC CAC GAC GAG TGC GAC ACG TAC GTG CGC GTG TGC	233
Ala Gly Gly Cys Gly His Asp Glu Cys Asp Thr Tyr Val Arg Val Cys	
35 40 45	
CTT AAG GAG TAC CAG GCC AAG GTG ACG CCC ACG GGG CCC TGC AGC TAC	281
Leu Lys Glu Tyr Gln Ala Lys Val Thr Pro Thr Gly Pro Cys Ser Tyr	
50 55 60	
GGC CAC GGC GCC ACG CCC GTG CTG GGC GGC AAC TCC TTC TAC CTG CCG	329
Gly His Gly Ala Thr Pro Val Leu Gly Gly Asn Ser Phe Tyr Leu Pro	
65 70 75 80	
CCG GCG GGC GCT GCG GGG GAC CGA GCG CGG GCG CGG GCC CGG GCC GGC	377
Pro Ala Gly Ala Ala Gly Asp Arg Ala Arg Ala Arg Ala Arg Ala Gly	
85 90 95	
GGC GAC CAG GAC CCG GGC CTC GTC GTC ATC CCC TTC CAG TTC GCC TGG	425
Gly Asp Gln Asp Pro Gly Leu Val Val Ile Pro Phe Gln Phe Ala Trp	
100 105 110	
CCG CGC TCC TTT ACC CTC ATC GTG GAG GCC TGG GAC TGG GAC AAC GAT	473
Pro Arg Ser Phe Thr Leu Ile Val Glu Ala Trp Asp Trp Asp Asn Asp	
115 120 125	
ACC ACC CCG AAT GAG GAG CTG CTG ATC GAG CGA GTG TCG CAT GCC GGC	521
Thr Thr Pro Asn Glu Glu Leu Leu Ile Glu Arg Val Ser His Ala Gly	
130 135 140	
ATG ATC AAC CCG GAG GAC CGC TGG AAG AGC CTG CAC TTC AGC GGC CAC	569
Met Ile Asn Pro Glu Asp Arg Trp Lys Ser Leu His Phe Ser Gly His	
145 150 155 160	
GTG GCG CAC CTG GAG CTG CAG ATC CGC GTG CGC TGC GAC GAG AAC TAC	617
Val Ala His Leu Glu Leu Gln Ile Arg Val Arg Cys Asp Glu Asn Tyr	
165 170 175	
TAC AGC GCC ACT TGC AAC AAG TTC TGC CGG CCC CGC AAC GAC TTT TTC	665



Tyr Ser Ala Thr Cys Asn Lys Phe Cys Arg Pro Arg Asn Asp Phe Phe	
180 185 190	
GGC CAC TAC ACC TGC GAC CAG TAC GGC AAC AAG GCC TGC ATG GAC GGC	713
Gly His Tyr Thr Cys Asp Gln Tyr Gly Asn Lys Ala Cys Met Asp Gly	
195 200 205	
TGG ATG GGC AAG GAG TGC AAG GAA GCT GTG TGT AAA CAA GGG TGT AAT	761
Trp Met Gly Lys Glu Cys Lys Glu Ala Val Cys Lys Gln Gly Cys Asn	
210 215 220	
TTG CTC CAC GGG GGA TGC ACC GTG CCT GGG GAG TGC AGG TGC AGC TAC	809
Leu Leu His Gly Gly Cys Thr Val Pro Gly Glu Cys Arg Cys Ser Tyr	
225 230 235 240	
GGC TGG CAA GGG AGG TTC TGC GAT GAG TGT GTC CCC TAC CCC GGC TGC	857
Gly Trp Gln Gly Arg Phe Cys Asp Glu Cys Val Pro Tyr Pro Gly Cys	
245 250 255	
GTG CAT GGC AGT TGT GTG GAG CCC TGG CAG TGC AAC TGT GAG ACC AAC	905
Val His Gly Ser Cys Val Glu Pro Trp Gln Cys Asn Cys Glu Thr Asn	
260 265 270	
TGG GGC GGC CTG CTC TGT GAC AAA GAC CTG AAC TAC TGT GGC AGC CAC	953
Trp Gly Gly Leu Leu Cys Asp Lys Asp Leu Asn Tyr Cys Gly Ser His	
275 280 285	
CAC CCC TGC ACC AAC GGA GGC ACG TGC ATC AAC GCC GAG CCT GAC CAG	1001
His Pro Cys Thr Asn Gly Gly Thr Cys Ile Asn Ala Glu Pro Asp Gln	
290 295 300	
TAC CGC TGC ACC TGC CCT GAC GGC TAC TCG GGC AGG AAC TGT GAG AAG	1049
Tyr Arg Cys Thr Cys Pro Asp Gly Tyr Ser Gly Arg Asn Cys Glu Lys	
305 310 315 320	
GCT GAG CAC GCC TGC ACC TCC AAC CCG TGT GCC AAC GGG GGC TCT TGC	1097
Ala Glu His Ala Cys Thr Ser Asn Pro Cys Ala Asn Gly Gly Ser Cys	



325	330	335	
CAT GAG GTG CCG TCC GGC TTC GAA TGC CAC TGC CCA TCG GGC TGG AGC			1145
His Glu Val Pro Ser Gly Phe Glu Cys His Cys Pro Ser Gly Trp Ser			
340	345	350	
GGG CCC ACC TGT GCC CTT GAC ATC GAT GAG TGT GCT TCG AAC CCG TGT			1193
Gly Pro Thr Cys Ala Leu Asp Ile Asp Glu Cys Ala Ser Asn Pro Cys			
355	360	365	
GGC GCC GGT GGC ACC TGT GTG GAC CAG GTG GAC GGC TTT GAG TGC ATC			1241
Ala Ala Gly Gly Thr Cys Val Asp Gln Val Asp Gly Phe Glu Cys Ile			
370	375	380	
TGC CCC GAG CAG TGG GTG GGG GCC ACC TGC CAG CTG GAC GCC AAT GAG			1289
Cys Pro Glu Gln Trp Val Gly Ala Thr Cys Gln Leu Asp Ala Asn Glu			
385	390	395	400
TGT GAA GGG AAG CCA TGC CTT AAC GCT TTT TCT TGC AAA AAC CTG ATT			1337
Cys Glu Gly Lys Pro Cys Leu Asn Ala Phe Ser Cys Lys Asn Leu Ile			
405	410	415	
GGC GGC TAT TAC TGT GAT TGC ATC CCG GGC TGG AAG GGC ATC AAC TGC			1385
Gly Gly Tyr Tyr Cys Asp Cys Ile Pro Gly Trp Lys Gly Ile Asn Cys			
420	425	430	
CAT ATC AAC GTC AAC GAC TGT CGC GGG CAG TGT CAG CAT GGG GGC ACC			1433
His Ile Asn Val Asn Asp Cys Arg Gly Gln Cys Gln His Gly Gly Thr			
435	440	445	
TGC AAG GAC CTG GTG AAC GGG TAC CAG TGT GTG TGC CCA CGG GGC TTC			1481
Cys Lys Asp Leu Val Asn Gly Tyr Gln Cys Val Cys Pro Arg Gly Phe			
450	455	460	
GGA GGC CGG CAT TGC GAG CTG GAA CGA GAC AAG TGT GCC AGC AGC CCC			1529
Gly Gly Arg His Cys Glu Leu Glu Arg Asp Lys Cys Ala Ser Ser Pro			
465	470	475	480



TGC CAC AGC GGC GGC CTC TGC GAG GAC CTG GCC GAC GGC TTC CAC TGC 1577
Cys His Ser Gly Gly Leu Cys Glu Asp Leu Ala Asp Gly Phe His Cys

485 490 495

CAC TGC CCC CAG GGC TTC TCC GGC CCT CTC TGT GAG GTG GAT GTC GAC 1625
His Cys Pro Gln Gly Phe Ser Gly Pro Leu Cys Glu Val Asp Val Asp

500 505 510

CTT TGT GAG CCA AGC CCC TGC CGG AAC GGC GCT CGC TGC TAT AAC CTG 1673
Leu Cys Glu Pro Ser Pro Cys Arg Asn Gly Ala Arg Cys Tyr Asn Leu

515 520 525

GAG GGT GAC TAT TAC TGC GCC TGC CCT GAT GAC TTT GGT GGC AAG AAC 1721
Glu Gly Asp Tyr Tyr Cys Ala Cys Pro Asp Asp Phe Gly Gly Lys Asn

530 535 540

TGC TCC GTG CCC CGC GAG CCG TGC CCT GGC GGG GCC TGC AGA GTG ATC 1769
Cys Ser Val Pro Arg Glu Pro Cys Pro Gly Gly Ala Cys Arg Val Ile

545 550 555 560

GAT GGC TGC GGG TCA GAC GCG GGC CCT GGC ATG CCT GGC ACA GCA GCC 1817
Asp Gly Cys Gly Ser Asp Ala Gly Pro Gly Met Pro Gly Thr Ala Ala

565 570 575

TCC GGC GTG TGT GGC CCC CAT GGA CGC TGC GTC AGC CAG CCA GGC GGC 1865
Ser Gly Val Cys Gly Pro His Gly Arg Cys Val Ser Gln Pro Gly Gly

580 585 590

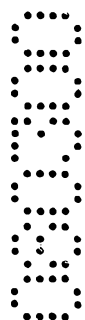
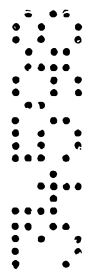
AAC TTT TCC TGC ATC TGT GAC AGT GGC TTT ACT GGC ACC TAC TGC CAT 1913
Asn Phe Ser Cys Ile Cys Asp Ser Gly Phe Thr Gly Thr Tyr Cys His

595 600 605

GAG AAC ATT GAC GAC TGC CTG GGC CAG CCC TGC CGC AAT GGC GGC ACA 1961
Glu Asn Ile Asp Asp Cys Leu Gly Gln Pro Cys Arg Asn Gly Gly Thr

610 615 620

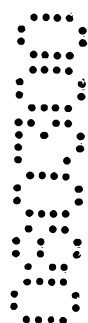
TGC ATC GAT GAG GTG GAC GCC TTC CGC TGC TTC TGC CCC AGC GGC TGC 2009



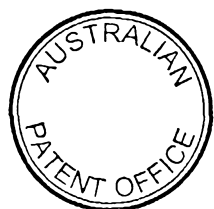
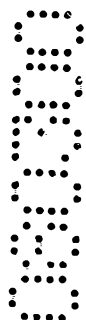
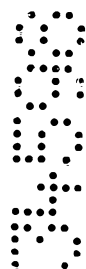
Cys Ile Asp Glu Val Asp Ala Phe Arg Cys Phe Cys Pro Ser Gly Trp	
625 630 635 640	
GAG GGC GAG CTC TGC GAC ACC AAT CCC AAC GAC TGC CTT CCC GAT CCC	2057
Glu Gly Glu Leu Cys Asp Thr Asn Pro Asn Asp Cys Leu Pro Asp Pro	
645 650 655	
TGC CAC AGC CGC GGC CGC TGC TAC GAC CTG GTC AAT GAC TTC TAC TGT	2105
Cys His Ser Arg Gly Arg Cys Tyr Asp Leu Val Asn Asp Phe Tyr Cys	
660 665 670	
GCG TGC GAC GAC GGC TGG AAG GGC AAG ACC TGC CAC TCA CGC GAG TTC	2153
Ala Cys Asp Asp Gly Trp Lys Gly Lys Thr Cys His Ser Arg Glu Phe	
675 680 685	
CAG TGC GAT GCC TAC ACC TGC AGC AAC GGT GGC ACC TGC TAC GAC AGC	2201
Gln Cys Asp Ala Tyr Thr Cys Ser Asn Gly Gly Thr Cys Tyr Asp Ser	
690 695 700	
GGC GAC ACC TTC CGC TGC GCC TGC CCC CCC GGC TGG AAG GGC AGC ACC	2249
Gly Asp Thr Phe Arg Cys Ala Cys Pro Pro Gly Trp Lys Gly Ser Thr	
705 710 715 720	
TGC GCC GTC GCC AAG AAC AGC AGC TGC CTG CCC AAC CCC TGT GTG AAT	2297
Cys Ala Val Ala Lys Asn Ser Ser Cys Leu Pro Asn Pro Cys Val Asn	
725 730 735	
GGT GGC ACC TGC GTG GGC AGC GGG GCC TCC TTC TCC TGC ATC TGC CGG	2345
Gly Gly Thr Cys Val Gly Ser Gly Ala Ser Phe Ser Cys Ile Cys Arg	
740 745 750	
GAC GGC TGG GAG GGT CGT ACT TGC ACT CAC AAT ACC AAC GAC TGC AAC	2393
Asp Gly Trp Glu Gly Arg Thr Cys Thr His Asn Thr Asn Asp Cys Asn	
755 760 765	
CCT CTG CCT TGC TAC AAT GGT GGC ATC TGT GTT GAC GGC GTC AAC TGG	2441
Pro Leu Pro Cys Tyr Asn Gly Gly Ile Cys Val Asp Gly Val Asn Trp	



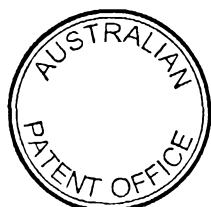
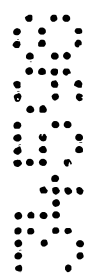
770	775	780	
TTC CGC TGC GAG TGT GCA CCT GGC TTC GCG GGG CCT GAC TGC CGC ATC			2489
Phe Arg Cys Glu Cys Ala Pro Gly Phe Ala Gly Pro Asp Cys Arg Ile			
785	790	795	800
AAC ATC GAC GAG TGC CAG TCC TCG CCC TGT GCC TAC GGG GCC ACG TGT			2537
Asn Ile Asp Glu Cys Gln Ser Ser Pro Cys Ala Tyr Gly Ala Thr Cys			
805	810	815	
GTG GAT GAG ATC AAC GGG TAT CGC TGT AGC TGC CCA CCC GGC CGA GCC			2585
Val Asp Glu Ile Asn Gly Tyr Arg Cys Ser Cys Pro Pro Gly Arg Ala			
820	825	830	
GGC CCC CGG TGC CAG GAA GTG ATC GGG TTC GGG AGA TCC TGC TGG TCC			2633
Gly Pro Arg Cys Gln Glu Val Ile Gly Phe Gly Arg Ser Cys Trp Ser			
835	840	845	
CGG GGC ACT CCG TTC CCA CAC GGA AGC TCC TGG GTG GAA GAC TGC AAC			2681
Arg Gly Thr Pro Phe Pro His Gly Ser Ser Trp Val Glu Asp Cys Asn			
850	855	860	
AGC TGC CGC TGC CTG GAT GGC CGC CGT GAC TGC AGC AAG GTG TGG TGC			2729
Ser Cys Arg Cys Leu Asp Gly Arg Arg Asp Cys Ser Lys Val Trp Cys			
865	870	875	880
GGA TGG AAG CCT TGT CTG CTG GCC GGC CAG CCC GAG GCC CTG AGC GCC			2777
Gly Trp Lys Pro Cys Leu Leu Ala Gly Gln Pro Glu Ala Leu Ser Ala			
885	890	895	
CAG TGC CCA CTG GGG CAA AGG TGC CTG GAG AAG GCC CCA GGC CAG TGT			2825
Gln Cys Pro Leu Gly Gln Arg Cys Leu Glu Lys Ala Pro Gly Gln Cys			
900	905	910	
CTG CGA CCA CCC TGT GAG GCC TGG GGG GAG TGC GGC GCA GAA GAG CCA			2873
Leu Arg Pro Pro Cys Glu Ala Trp Gly Glu Cys Gly Ala Glu Glu Pro			
915	920	925	



CCG AGC ACC CCC TGC CTG CCA CGC TCC GGC CAC CTG GAC AAT AAC TGT	2921
Pro Ser Thr Pro Cys Leu Pro Arg Ser Gly His Leu Asp Asn Asn Cys	
930 935 940	
GCC CGC CTC ACC TTG CAT TTC AAC CGT GAC CAC GTG CCC CAG GGC ACC	2969
Ala Arg Leu Thr Leu His Phe Asn Arg Asp His Val Pro Gln Gly Thr	
945 950 955 960	
ACG GTG GGC GCC ATT TGC TCC GGG ATC CGC TCC CTG CCA GCC ACA AGG	3017
Thr Val Gly Ala Ile Cys Ser Gly Ile Arg Ser Leu Pro Ala Thr Arg	
965 970 975	
GCT GTG GCA CGG GAC CGC CTG CTG GTG TTG CTT TGC GAC CGG GCG TCC	3065
Ala Val Ala Arg Asp Arg Leu Leu Val Leu Leu Cys Asp Arg Ala Ser	
980 985 990	
TCG GGG GCC AGT GCC GTG GAG GTG GCC GTG TCC TTC AGC CCT GCC AGG	3113
Ser Gly Ala Ser Ala Val Glu Val Ala Val Ser Phe Ser Pro Ala Arg	
995 1000 1005	
GAC CTG CCT GAC AGC AGC CTG ATC CAG GGC GCG GCC CAC GCC ATC GTG	3161
Asp Leu Pro Asp Ser Ser Leu Ile Gln Gly Ala Ala His Ala Ile Val	
1010 1015 1020	
GCC GCC ATC ACC CAG CGG GGG AAC AGC TCA CTG CTC CTG GCT GTC ACC	3209
Ala Ala Ile Thr Gln Arg Gly Asn Ser Ser Leu Leu Leu Ala Val Thr	
1025 1030 1035 1040	
GAG GTC AAG GTG GAG ACC GTT GTT ACG GGC GGC TCT TCC ACA GGT CTG	3257
Glu Val Lys Val Glu Thr Val Val Thr Gly Gly Ser Ser Thr Gly Leu	
1045 1050 1055	
CTG GTG CCT GTG CTG TGT GGT GCC TTC AGC GTG CTG TGG CTG GCG TGC	3305
Leu Val Pro Val Leu Cys Gly Ala Phe Ser Val Leu Trp Leu Ala Cys	
1060 1065 1070	
GTG GTC CTG TGC GTG TGG TGG ACA CGC AAG CGC AGG AAA GAG CGG GAG	3353



Val Val Leu Cys Val Trp Trp Thr Arg Lys Arg Arg Lys Glu Arg Glu	
1075 1080 1085	
AGG AGC CGG CTG CCG CGG GAG GAG AGC GCC AAC AAC CAG TGG GCC CCG	3401
Arg Ser Arg Leu Pro Arg Glu Glu Ser Ala Asn Asn Gln Trp Ala Pro	
1090 1095 1100	
CTC AAC CCC ATC CGC AAC CCC ATC GAG CGG CCG GGG GGC CAC AAG GAC	3449
Leu Asn Pro Ile Arg Asn Pro Ile Glu Arg Pro Gly Gly His Lys Asp	
1105 1110 1115 1120	
GTG CTC TAC CAG TGC AAG AAC TTC ACG CCG CCG CCG CCG AGG GCG GAC	3497
Val Leu Tyr Gln Cys Lys Asn Phe Thr Pro Pro Pro Arg Arg Ala Asp	
1125 1130 1135	
GAG GCG CTG CCC GGG CCG GCC GGC CAC GCG GCC GTC AGG GAG GAT GAG	3545
Glu Ala Leu Pro Gly Pro Ala Gly His Ala Ala Val Arg Glu Asp Glu	
1140 1145 1150	
GAG GAC GAG GAT CTG GGC CGC GGT GAG GAG GAC TCC CTG GAG CCG GAG	3593
Glu Asp Glu Asp Leu Gly Arg Gly Glu Glu Asp Ser Leu Glu Ala Glu	
1155 1160 1165	
AAG TTC CTC TCA CAC AAA TTC ACC AAA GAT CCT GGC CGC TCG CCG GGG	3641
Lys Phe Leu Ser His Lys Phe Thr Lys Asp Pro Gly Arg Ser Pro Gly	
1170 1175 1180	
AGG CCG GCC CAC TGG GCC TCA GGC CCC AAA GTG GAC AAC CGC GCG GTC	3689
Arg Pro Ala His Trp Ala Ser Gly Pro Lys Val Asp Asn Arg Ala Val	
1185 1190 1195 1200	
AGG AGC ATC AAT GAG GCC CGC TAC GCC GGC AAG GAG	3725
Arg Ser Ile Asn Glu Ala Arg Tyr Ala Gly Lys Glu	
1205 1210 1212	
TAGGGGCGGC TGGCAGCTGG GCCGGGACCC AGGGCCCTCG GTGGGAGCCA TGGCGTCTGC	3785
CGGACCCCGA GGCCGAGGCC ATGTGCATAG TTTCTTTATT TTGTGTAAAA AAACCACCAA	3845



AAACAAAAC CAAATGTTTA TTTTCTACGT TTCTTTAACC TTGTATAAAT TATTCAGTAA 3905
CTGTCAGGCT GAAAACAATG GAGTATTCTC GGAAAAA AAAA AAAAAA 3955

INFORMATION FOR SEQ ID NO : 5

LENGTH : 4005 and 1199
TYPE : nucleic acid and amino acid
STRANDEDNESS : double stranded and single stranded
TOPOLOGY : linear and unknown
MOLECULE TYPE : cDNA to mRNA and amino acid
ORIGINAL SOURCE

ORGANISM : human

SEQUENCE DESCRIPTION

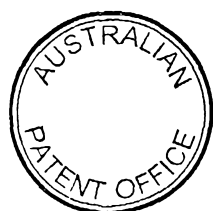
GGCCGGCC CGCGAGCTAG GCTGGTTTTT TTTTTTCTCC CCTCCCTCCC 48
CCCTTTTTCC ATGCAGCTGA TCTAAAAGGG AATAAAAGGC TGGCATAAT CATAATAATA 108
AAAGAAGGGG AGCGCGAGAG AAGGAAAGAA AGCCGGGAGG TGAAGAGGA GGGGAGCGT 168
CTCAAAGAAG CGATCAGAAT AATAAAAGGA GGCCGGGCTC TTTGCCTTCT GGAACGGGCC 228
GCTCTTGAAA GGGCTTTTGA AAAGTGGTGT TGTTTTCCAG TCGTGCATGC TCCAATCGGC 288
GGAGTATATT AGAGCCGGGA CGCGCGGGCC GCAGGGGCAG CGGCGACGGC AGCACC GGCG 348
GCAGCACCAG CGCGAACAGC AGCGGCGGGC TCCCAGTGC CCGCGGCGCG CGGCGCAGCG 408
ATG CGT TCC CCA CGG ACG CGC GGC CGG TCC GGG CGC CCC CTA AGC 453
Met Arg Ser Pro Arg Thr Arg Gly Arg Ser Gly Arg Pro Leu Ser
-31 -30 -25 -20
CTC CTG CTC GCC CTG CTC TGT GCC CTG CGA GCC AAG GTG TGT GGG GCC 501
Leu Leu Leu Ala Leu Leu Cys Ala Leu Arg Ala Lys Val Cys Gly Ala
-15 -10 -5 -1
TCG GGT CAG TTC GAG TTG GAG ATC CTG TCC ATG CAG AAC GTG AAC GGC 549
Ser Gly Gln Phe Glu Leu Glu Ile Leu Ser Met Gln Asn Val Asn Gly

1

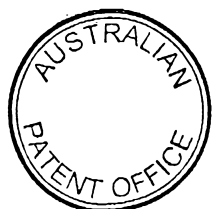
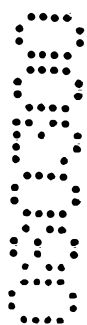
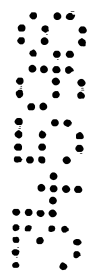
5

10

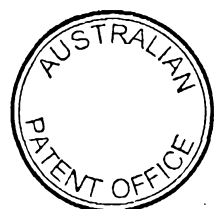
15



GAG CTG CAG AAC GGG AAC TGC TGC GGC GGC GCC CGG AAC CCG GGA GAC	597
Glu Leu Gln Asn Gly Asn Cys Cys Gly Gly Ala Arg Asn Pro Gly Asp	
20 25 30	
CGC AAG TGC ACC CGC GAC GAG TGT GAC ACA TAC TTC AAA GTG TGC CTC	645
Arg Lys Cys Thr Arg Asp Glu Cys Asp Thr Tyr Phe Lys Val Cys Leu	
35 40 45	
AAG GAG TAT CAG TCC CGC GTC ACG GCC GGG GGG CCC TGC AGC TTC GGC	693
Lys Glu Tyr Gln Ser Arg Val Thr Ala Gly Gly Pro Cys Ser Phe Gly	
50 55 60	
TCA GGG TCC ACG CCT GTC ATC GGG GGC AAC ACC TTC AAC CTC AAG GCC	741
Ser Gly Ser Thr Pro Val Ile Gly Gly Asn Thr Phe Asn Leu Lys Ala	
65 70 75 80	
AGC CGC GGC AAC GAC CGC AAC CGC ATC GTG CTG CCT TTC AGT TTC GCC	789
Ser Arg Gly Asn Asp Arg Asn Arg Ile Val Leu Pro Phe Ser Phe Ala	
85 90 95	
TGG CCG AGG TCC TAT ACG TTG CTT GTG GAG GCG TGG GAT TCC AGT AAT	837
Trp Pro Arg Ser Tyr Thr Leu Leu Val Glu Ala Trp Asp Ser Ser Asn	
100 105 110	
GAC ACC GTT CAA CCT GAC AGT ATT ATT GAA AAG GCT TCT CAC TCG GGC	885
Asp Thr Val Gln Pro Asp Ser Ile Ile Glu Lys Ala Ser His Ser Gly	
115 120 125	
ATG ATC AAC CCC AGC CGG CAG TGG CAG ACG CTG AAG CAG AAC ACG GGC	933
Met Ile Asn Pro Ser Arg Gln Trp Gln Thr Leu Lys Gln Asn Thr Gly	
130 135 140	
GTT GCC CAC TTT GAG TAT CAG ATC CGC GTG ACC TGT GAT GAC TAC TAC	981
Val Ala His Phe Glu Tyr Gln Ile Arg Val Thr Cys Asp Asp Tyr Tyr	
145 150 155 160	
TAT GGC TTT GGC TGC AAT AAG TTC TGC CGC CCC AGA GAT GAC TTC TTT	1029



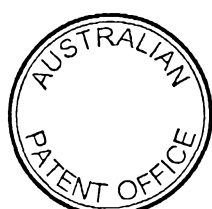
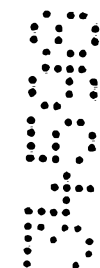
Tyr Gly Phe Gly Cys Asn Lys Phe Cys Arg Pro Arg Asp Asp Phe Phe	
165 170 175	
GGA CAC TAT GCC TGT GAC CAG AAT GGC AAC AAA ACT TGC ATG GAA GGC	1077
Gly His Tyr Ala Cys Asp Gln Asn Gly Asn Lys Thr Cys Met Glu Gly	
180 185 190	
TGG ATG GGC CCC GAA TGT AAC AGA GCT ATT TGC CGA CAA GGC TGC AGT	1125
Trp Met Gly Pro Glu Cys Asn Arg Ala Ile Cys Arg Gln Gly Cys Ser	
195 200 205	
CCT AAG CAT GGG TCT TGC AAA CTC CCA GGT GAC TGC AGG TGC CAG TAC	1173
Pro Lys His Gly Ser Cys Lys Leu Pro Gly Asp Cys Arg Cys Gln Tyr	
210 215 220	
GGC TGG CAA GGC CTG TAC TGT GAT AAG TGC ATC CCA CAC CCG GGA TGC	1221
Gly Trp Gln Gly Leu Tyr Cys Asp Lys Cys Ile Pro His Pro Gly Cys	
225 230 235 240	
GTC CAC GGC ATC TGT AAT GAG CCC TGG CAG TGC CTC TGT GAG ACC AAC	1269
Val His Gly Ile Cys Asn Glu Pro Trp Gln Cys Leu Cys Glu Thr Asn	
245 250 255	
TGG GGC GGC CAG CTC TGT GAC AAA GAT CTC AAT TAC TGT GGG ACT CAT	1317
Trp Gly Gly Gln Leu Cys Asp Lys Asp Leu Asn Tyr Cys Gly Thr His	
260 265 270	
CAG CCG TGT CTC AAC GGC GGA ACT TGT AGC AAC ACA GGC CCT GAC AAA	1365
Gln Pro Cys Leu Asn Gly Gly Thr Cys Ser Asn Thr Gly Pro Asp Lys	
275 280 285	
TAT CAG TGT TCC TGC CCT GAG GGC TAT TCA GGA CCC AAC TGT GAA ATT	1413
Tyr Gln Cys Ser Cys Pro Glu Gly Tyr Ser Gly Pro Asn Cys Glu Ile	
290 295 300	
GCT GAG CAC GCC TGC CTC TCT GAT CCC TGT CAC AAC AGA GGC AGC TGT	1461
Ala Glu His Ala Cys Leu Ser Asp Pro Cys His Asn Arg Gly Ser Cys	



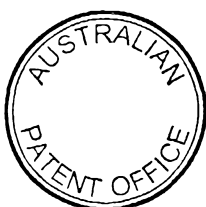
305	310	315	320	
AAG GAG ACC TCC CTG GGC TTT GAG TGT GAG TGT TCC CCA GGC TGG ACC	1509			
Lys Glu Thr Ser Leu Gly Phe Glu Cys Glu Cys Ser Pro Gly Trp Thr				
325	330	335		
GGC CCC ACA TGC TCT ACA AAC ATT GAT GAC TGT TCT CCT AAT AAC TGT	1557			
Gly Pro Thr Cys Ser Thr Asn Ile Asp Asp Cys Ser Pro Asn Asn Cys				
340	345	350		
TCC CAC GGG GGC ACC TGC CAG GAC CTG GTT AAC GGA TTT AAG TGT GTG	1605			
Ser His Gly Gly Thr Cys Gln Asp Leu Val Asn Gly Phe Lys Cys Val				
355	360	365		
TGC CCC CCA CAG TGG ACT GGC AAA ACG TGC CAG TTA GAT GCA AAT GAA	1653			
Cys Pro Pro Gln Trp Thr Gly Lys Thr Cys Gln Leu Asp Ala Asn Glu				
370	375	380		
TGT GAG GCC AAA CCT TGT GTA AAC GCC AAA TCC TGT AAG AAT CTC ATT	1701			
Cys Glu Ala Lys Pro Cys Val Asn Ala Lys Ser Cys Lys Asn Leu Ile				
385	390	395	400	
GCC AGC TAC TAC TGC GAC TGT CTT CCC GGC TGG ATG GGT CAG AAT TGT	1749			
Ala Ser Tyr Tyr Cys Asp Cys Leu Pro Gly Trp Met Gly Gln Asn Cys				
405	410	415		
GAC ATA AAT ATT AAT GAC TGC CTT GGC CAG TGT CAG AAT GAC GCC TCC	1797			
Asp Ile Asn Ile Asn Asp Cys Leu Gly Gln Cys Gln Asn Asp Ala Ser				
420	425	430		
TGT CGG GAT TTG GTT AAT GGT TAT CGC TGT ATC TGT CCA CCT GCC TAT	1845			
Cys Arg Asp Leu Val Asn Gly Tyr Arg Cys Ile Cys Pro Pro Gly Tyr				
435	440	445		
GCA GGC GAT CAC TGT GAG AGA GAC ATC GAT GAA TGT GCC AGC AAC CCC	1893			
Ala Gly Asp His Cys Glu Arg Asp Ile Asp Glu Cys Ala Ser Asn Pro				
450	455	460		



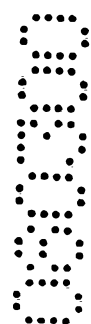
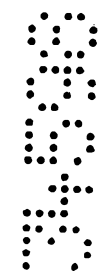
TGT TTG AAT GGG GGT CAC TGT CAG AAT GAA ATC AAC AGA TTC CAG TGT	1941
Cys Leu Asn Gly Gly His Cys Gln Asn Glu Ile Asn Arg Phe Gln Cys	
465 470 475 480	
CTG TGT CCC ACT GGT TTC TCT GGA AAC CTC TGT CAG CTG GAC ATC GAT	1989
Leu Cys Pro Thr Gly Phe Ser Gly Asn Leu Cys Gln Leu Asp Ile Asp	
485 490 495	
TAT TGT GAG CCT AAT CCC TGC CAG AAC GGT GCC CAG TGC TAC AAC CGT	2037
Tyr Cys Glu Pro Asn Pro Cys Gln Asn Gly Ala Gln Cys Tyr Asn Arg	
500 505 510	
GCC AGT GAC TAT TTC TGC AAG TGC CCC GAG GAC TAT GAG GGC AAG AAC	2085
Ala Ser Asp Tyr Phe Cys Lys Cys Pro Glu Asp Tyr Glu Gly Lys Asn	
515 520 525	
TGC TCA CAC CTG AAA GAC CAC TGC CGC ACG ACC CCC TGT GAA GTG ATT	2133
Cys Ser His Leu Lys Asp His Cys Arg Thr Thr Pro Cys Glu Val Ile	
530 535 540	
GAC AGC TGC ACA GTG GCC ATG GCT TCC AAC GAC ACA CCT GAA GGG GTG	2181
Asp Ser Cys Thr Val Ala Met Ala Ser Asn Asp Thr Pro Glu Gly Val	
545 550 555 560	
CGG TAT ATT TCC TCC AAC GTC TGT GGT CCT CAC GGG AAG TGC AAG AGT	2229
Arg Tyr Ile Ser Ser Asn Val Cys Gly Pro His Gly Lys Cys Lys Ser	
565 570 575	
CAG TCG GGA GGC AAA TTC ACC TGT GAC TGT AAC AAA GGC TTC ACG GGA	2277
Gln Ser Gly Gly Lys Phe Thr Cys Asp Cys Asn Lys Gly Phe Thr Gly	
580 585 590	
ACA TAC TGC CAT GAA AAT ATT AAT GAC TGT GAG AGC AAC CCT TGT AGA	2325
Thr Tyr Cys His Glu Asn Ile Asn Asp Cys Glu Ser Asn Pro Cys Arg	
595 600 605	
AAC GGT GGC ACT TGC ATC GAT GGT GTC AAC TCC TAC AAG TGC ATC TGT	2373



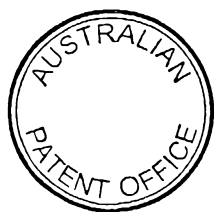
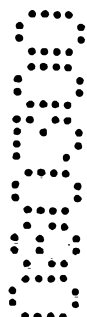
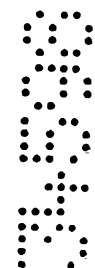
Asn Gly Gly Thr Cys Ile Asp Gly Val Asn Ser Tyr Lys Cys Ile Cys	
610 615 620	
AGT GAC GGC TGG GAG GGG GCC TAC TGT GAA ACC AAT ATT AAT GAC TGC	2421
Ser Asp Gly Trp Glu Gly Ala Tyr Cys Glu Thr Asn Ile Asn Asp Cys	
625 630 635 640	
AGC CAG AAC CCC TGC CAC AAT GGG GGC ACG TGT CGC GAC CTG GTC AAT	2469
Ser Gln Asn Pro Cys His Asn Gly Gly Thr Cys Arg Asp Leu Val Asn	
645 650 655	
GAC TTC TAC TGT GAC TGT AAA AAT GGG TGG AAA GGA AAG ACC TGC CAC	2517
Asp Phe Tyr Cys Asp Cys Lys Asn Gly Trp Lys Gly Lys Thr Cys His	
660 665 670	
TCA CGT GAC AGT CAG TGT GAT GAG GCC ACG TGC AAC AAC GGT GGC ACC	2565
Ser Arg Asp Ser Gln Cys Asp Glu Ala Thr Cys Asn Asn Gly Gly Thr	
675 680 685	
TGC TAT GAT GAG GGG GAT GCT TTT AAG TGC ATG TGT CCT GGC GGC TGG	2613
Cys Tyr Asp Glu Gly Asp Ala Phe Lys Cys Met Cys Pro Gly Gly Trp	
690 695 700	
GAA GGA ACA ACC TGT AAC ATA GCC CGA AAC AGT ACC TGC CTG CCC AAC	2661
Glu Gly Thr Thr Cys Asn Ile Ala Arg Asn Ser Ser Cys Leu Pro Asn	
705 710 715 720	
CCC TGC CAT AAT GGG GGC ACA TGT GTG GTC AAC GGC GAG TCC TTT ACC	2709
Pro Cys His Asn Gly Gly Thr Cys Val Val Asn Gly Glu Ser Phe Thr	
725 730 735	
TGC GTC TGC AAG GAA GGC TGG GAG GGG CCC ATC TGT GCT CAG AAT ACC	2757
Cys Val Cys Lys Glu Gly Trp Glu Gly Pro Ile Cys Ala Gln Asn Thr	
740 745 750	
AAT GAC TGC AGC CCT CAT CCC TGT TAC AAC ACC GGC ACC TGT GTG GAT	2805
Asn Asp Cys Ser Pro His Pro Cys Tyr Asn Ser Gly Thr Cys Val Asp	



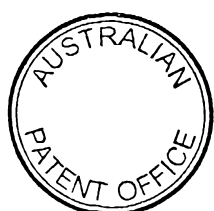
755	760	765	
GGA GAC AAC TGG TAC CGG TGC GAA TGT GCC CCG GGT TTT GCT GGG CCC			2853
Gly Asp Asn Trp Tyr Arg Cys Glu Cys Ala Pro Gly Phe Ala Gly Pro			
770	775	780	
GAC TGC AGA ATA AAC ATC AAT GAA TGC CAG TCT TCA CCT TGT GCC TTT			2901
Asp Cys Arg Ile Asn Ile Asn Glu Cys Gln Ser Ser Pro Cys Ala Phe			
785	790	795	800
GGA GCG ACC TGT GTG GAT GAG ATC AAT GGC TAC CCG TGT GTC TGC CCT			2949
Gly Ala Thr Cys Val Asp Glu Ile Asn Gly Tyr Arg Cys Val Cys Pro			
805	810	815	
CCA GGG CAC AGT GGT GCC AAG TGC CAG GAA GTT TCA GGG AGA CCT TGC			2997
Pro Gly His Ser Gly Ala Lys Cys Gln Glu Val Ser Gly Arg Pro Cys			
820	825	830	
ATC ACC ATG GGG AGT GTG ATA CCA GAT GGG GCC AAA TGG GAT GAT GAC			3045
Ile Thr Met Gly Ser Val Ile Pro Asp Gly Ala Lys Trp Asp Asp Asp			
835	840	845	
TGT AAT ACC TGC CAG TGC CTG AAT GGA CCG ATC GCC TGC TCA AAG GTC			3093
Cys Asn Thr Cys Gln Cys Leu Asn Gly Arg Ile Ala Cys Ser Lys Val			
850	855	860	
TGG TGT GGC CCT CGA CCT TGC CTG CTC CAC AAA GGG CAC AGC GAG TGC			3141
Trp Cys Gly Pro Arg Pro Cys Leu Leu His Lys Gly His Ser Glu Cys			
865	870	875	880
CCC AGC GGG CAG AGC TGC ATC CCC ATC CTG GAC GAC CAG TGC TTC GTC			3189
Pro Ser Gly Gln Ser Cys Ile Pro Ile Leu Asp Asp Gln Cys Phe Val			
885	890	895	
CAC CCC TGC ACT GGT GTG GGC GAG TGT CCG TCT TCC AGT CTC CAG CCG			3237
His Pro Cys Thr Gly Val Gly Glu Cys Arg Ser Ser Ser Leu Gln Pro			
900	905	910	



GTG AAG ACA AAG TGC ACC TCT GAC TCC TAT TAC CAG GAT AAC TGT GCG	3285
Val Lys Thr Lys Cys Thr Ser Asp Ser Tyr Tyr Gln Asp Asn Cys Ala	
915 920 925	
AAC ATC ACA TTT ACC TTT AAC AAG GAG ATG ATG TCA CCA GGT CTT ACT	3333
Asn Ile Thr Phe Thr Phe Asn Lys Glu Met Met Ser Pro Gly Leu Thr	
930 935 940	
ACG GAG CAC ATT TGC AGT GAA TTG AGG AAT TTG AAT ATT TTG AAG AAT	3381
Thr Glu His Ile Cys Ser Glu Leu Arg Asn Leu Asn Ile Leu Lys Asn	
945 950 955 960	
GTT TCC GCT GAA TAT TCA ATC TAC ATC GCT TGC GAG CCT TCC CCT TCA	3429
Val Ser Ala Glu Tyr Ser Ile Tyr Ile Ala Cys Glu Pro Ser Pro Ser	
965 970 975	
GCG AAC AAT GAA ATA CAT GTG GCC ATT TCT GCT GAA GAT ATA CGG GAT	3477
Ala Asn Asn Glu Ile His Val Ala Ile Ser Ala Glu Asp Ile Arg Asp	
980 985 990	
GAT GGG AAC CCG ATC AAG GAA ATC ACT GAC AAA ATA ATC GAT CTT GTT	3525
Asp Gly Asn Pro Ile Lys Glu Ile Thr Asp Lys Ile Ile Asp Leu Val	
995 1000 1005	
AGT AAA CGT GAT GGA AAC AGC TCG CTG ATT GCT GCC GTT GCA GAA GTA	3573
Ser Lys Arg Asp Gly Asn Ser Ser Leu Ile Ala Ala Val Ala Glu Val	
1010 1015 1020	
AGA GTT CAG AGG CGG CCT CTG AAG AAC AGA ACA GAT TTC CTT GTT CCC	3621
Arg Val Gln Arg Arg Pro Leu Lys Asn Arg Thr Asp Phe Leu Val Pro	
1025 1030 1035 1040	
TTG CTG AGC TCT GTC TTA ACT GTG GCT TGG ATC TGT TGC TTG GTG ACG	3669
Leu Leu Ser Ser Val Leu Thr Val Ala Trp Ile Cys Cys Leu Val Thr	
1045 1050 1055	
GCC TTC TAC TGG TGC CTG CGG AAG CGG CGG AAG CCG GCC AGC CAC ACA	3717



Ala Phe Tyr Trp Cys Leu Arg Lys Arg Arg Lys Pro Gly Ser His Thr	
1060	1065
1070	
CAC TCA GCC TCT GAG GAC AAC ACC ACC AAC AAC GTG CGG GAG CAG CTG	3765
His Ser Ala Ser Glu Asp Asn Thr Thr Asn Asn Val Arg Glu Gln Leu	
1075	1080
1085	
AAC CAG ATC AAA AAC CCC ATT GAG AAA CAT GGG GCC AAC ACG GTC CCC	3813
Asn Gln Ile Lys Asn Pro Ile Glu Lys His Gly Ala Asn Thr Val Pro	
1090	1095
1100	
ATC AAG GAT TAT GAG AAC AAG AAC TCC AAA ATG TCT AAA ATA ACG ACA	3861
Ile Lys Asp Tyr Glu Asn Lys Asn Ser Lys Met Ser Lys Ile Arg Thr	
1105	1110
1115	1120
CAC AAT TCT GAA GTA GAA GAG GAC GAC ATG GAC AAA CAC CAG CAG AAA	3909
His Asn Ser Glu Val Glu Glu Asp Asp Met Asp Lys His Gln Gln Lys	
1125	1130
1135	
GCC CGG TTT GCC AAG CAG CCG GCG TAC ACG CTG GTA GAC AGA GAA GAG	3957
Ala Arg Phe Ala Lys Gln Pro Ala Tyr Thr Leu Val Asp Arg Glu Glu	
1140	1145
1150	
AAG CCC CCC AAC GGC ACG CCG ACA AAA CAC CCA AAC TGG ACA AAC AAA	4005
Lys Pro Pro Asn Gly Thr Pro Thr Lys His Pro Asn Trp Thr Asn Lys	
1155	1160
1165	
CAG GAC AAC AGA GAC TTG GAA AGT GCC CAG AGC TTA AAC CGA ATG GAG	4053
Gln Asp Asn Arg Asp Leu Glu Ser Ala Gln Ser Leu Asn Arg Met Glu	
1170	1175
1180	
TAC ATC GTA	4062
Tyr Ile Val	
1185	1187
TAGCAGACCG CGGGCACTGC CGCCGCTAGG TAGAGTCTGA GGGCTTGTAG TTCTTTAAAC	4122
TGTCGTGTCA TACTCGAGTC TGAGGCCGTT GCTGACTTAG AATCCCTGTG TTAATTTAAG	4182



TTTTGACAAG CTGGCTTACA CTGGCA

4208

INFORMATION FOR SEQ ID NO : 6

LENGTH : 20

TYPE : nucleic acid

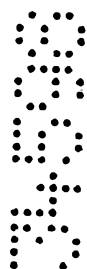
STRANDEDNESS : single stranded

TOPOLOGY : linear

MOLECULE TYPE : other nucleic acid synthetic DNA

SEQUENCE DESCRIPTION

TGCSTSTGYG ANACCAACTG



INFORMATION FOR SEQ ID NO : 7

LENGTH : 20

TYPE : nucleic acid

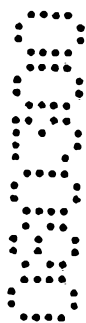
STRANDEDNESS : single stranded

TOPOLOGY : linear

MOLECULE TYPE : other nucleic acid synthetic DNA

SEQUENCE DESCRIPTION

TTTATKTCRC AWKTCKGWCC



INFORMATION FOR SEQ ID NO : 8

LENGTH : 180

TYPE : nucleic acid

STRANDEDNESS : double stranded

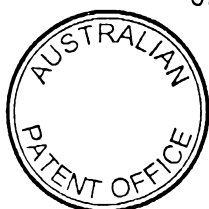
TOPOLOGY : linear

MOLECULE TYPE : cDNA to mRNA

ORIGINAL SOURCE

ORGANISM : human

SEQUENCE DESCRIPTION



GAGTGTGCAC CTGGCTTCGC GGGGCCTGAC TCCCGCATCA ACATCGACGA GTGCCAGTCC 60
 TCGCCCTGTG CCTACGGGGC CACGTGTGTG GATGAGATCA ACGGGTATCG CTGTAGCTGC 120
 CCACCCGGCC GAGCCGGCCC CCGGTGCCAG GAAGTGATCG GGTTCGGGAG ATCCTGCTGG 180

INFORMATION FOR SEQ ID NO : 9

LENGTH : 252
 TYPE : nucleic acid
 STRANDEDNESS : double stranded
 TOPOLOGY : linear
 MOLECULE TYPE : cDNA to mRNA

ORIGINAL SOURCE

ORGANISM : human

SEQUENCE DESCRIPTION

GCATCAACTG CCATATCAAC GTCAACGACT GTCGCGGGCA GTGTCAGCAT GGGGGCACCT 60
 GCAAGGACCT GGTGAACGGG TACCACTGTG TGTGCCCACG GGGCTTCGGA GGCCGGCATT 120
 GCGAGCTGGA ACGAGACAAG TGTGCCAGCA GCCCCTGCCA CAGCGGCGGC CTCTGCGAGG 180
 ACCTGGCCGA CGGCTTCCAC TGCCACTGCC CCCAGGGCTT CTCCGGGCCT CTCTGTGAGG 240
 TGGATGTCCA CC 252

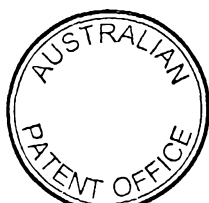
INFORMATION FOR SEQ ID NO : 10

LENGTH : 184
 TYPE : nucleic acid
 STRANDEDNESS : double stranded
 TOPOLOGY : linear
 MOLECULE TYPE : cDNA to mRNA

ORIGINAL SOURCE

ORGANISM : human

SEQUENCE DESCRIPTION



CAACTTTTCC TGCATCTGTG ACAGTGGCTT TACTGGCACC TACTGCCATG AGAACATTGA 60
 CGACTGCCTG GGCCAGCCCT GCCGCAATGG GGGCACATGC ATCGATGAGG TGGACGCCTT 120
 CCGCTGCTTC TGCCCCAGCG GCTGGGAGGG CGAGCTCTGC GACACCAATC CCAACGACTG 180
 CCTT 184

INFORMATION FOR SEQ ID NO : 11

LENGTH : 20
 TYPE : nucleic acid
 STRANDEDNESS : single stranded
 TOPOLOGY : linear
 MOLECULE TYPE : other nucleic acid synthetic DNA
 SEQUENCE DESCRIPTION
 GAGTGTGCAC CTGGCTTCGC

INFORMATION FOR SEQ ID NO : 12

LENGTH : 20
 TYPE : nucleic acid
 STRANDEDNESS : single stranded
 TOPOLOGY : linear
 MOLECULE TYPE : other nucleic acid synthetic DNA
 SEQUENCE DESCRIPTION
 CCAGCAGGAT CTCCCGAACC

INFORMATION FOR SEQ ID NO : 13

LENGTH : 20
 TYPE : nucleic acid
 STRANDEDNESS : single stranded



TOPOLOGY : linear

MOLECULE TYPE : other nucleic acid synthetic DNA

SEQUENCE DESCRIPTION

GCATCAACTG CCATATCAAC

INFORMATION FOR SEQ ID NO : 14

LENGTH : 20

TYPE : nucleic acid

STRANDEDNESS : single stranded

TOPOLOGY : linear

MOLECULE TYPE : other nucleic acid synthetic DNA

SEQUENCE DESCRIPTION

GGTCGACATC CACCTCACAG

INFORMATION FOR SEQ ID NO : 15

LENGTH : 20

TYPE : nucleic acid

STRANDEDNESS : single stranded

TOPOLOGY : linear

MOLECULE TYPE : other nucleic acid synthetic DNA

SEQUENCE DESCRIPTION

CAACTTTTCC TGCATCTGTG

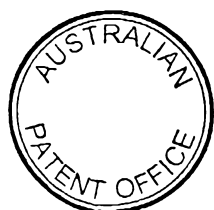
INFORMATION FOR SEQ ID NO : 16

LENGTH : 20

TYPE : nucleic acid

STRANDEDNESS : single stranded

TOPOLOGY : linear



MOLECULE TYPE : other nucleic acid synthetic DNA

SEQUENCE DESCRIPTION

AAGGCAGTCG TTGGGATTGG

INFORMATION FOR SEQ ID NO : 17

LENGTH : 397

TYPE : nucleic acid

STRANDEDNESS : double stranded

TOPOLOGY : linear

MOLECULE TYPE : cDNA to mRNA

ORIGINAL SOURCE

ORGANISM : human

SEQUENCE DESCRIPTION

ATGAGGAGCT GCTGATCGAG CGAGTGTCCG ATGCCGGCAT GATCAACCCG GAGGACCGCT 60
GGAAGAGCCT GCACTTCAGC GGCCACGTGG CGCACCTGGA GCTGCAGATC CGCGTGCGCT 120
GCGACGAGAA CTACTACAGC GCCACTTGCA ACAAGTTCTG CCGGCCCCCG AACGACTTTT 180
TCGGCCACTA CACCTGCGAC CAGTACGGCA ACAAGGCCTG CATGGACGGC TGGATGGGCA 240
AGGAGTGCAA GGAAGCTGTG TGTAACAAG GGTGTAATTT GCTCCACGGG GGATGCACCG 300
TGCCTGGGGA GTGCAGGTGC AGCTACGGCT GGCAAGGGAG GTTCTGCGAT GAGTGTGTCC 360
CCTACCCCGG CTGCGTGCAT GGCAGTTGTG TGGAGCC 397

INFORMATION FOR SEQ ID NO : 18

LENGTH : 20

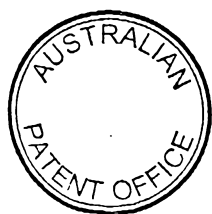
TYPE : nucleic acid

STRANDEDNESS : single stranded

TOPOLOGY : linear

MOLECULE TYPE : other nucleic acid synthetic DNA

SEQUENCE DESCRIPTION



ATGAGGAGCT GCTGATCGAG

INFORMATION FOR SEQ ID NO : 19

LENGTH : 20

TYPE : nucleic acid

STRANDEDNESS : single stranded

TOPOLOGY : linear

MOLECULE TYPE : other nucleic acid synthetic DNA

SEQUENCE DESCRIPTION

GGCTCCACAC AACTGCCCATG

INFORMATION FOR SEQ ID NO : 20

LENGTH : 60

TYPE : nucleic acid

STRANDEDNESS : single stranded

TOPOLOGY : linear

MOLECULE TYPE : other nucleic acid synthetic DNA

SEQUENCE DESCRIPTION

TCCGCGGGGGC AATGCGGGCG CAGGGCCGGG GGCGCCTTCC CCGGCGGCTG CTGCTGCTGC

INFORMATION FOR SEQ ID NO : 21

LENGTH : 22

TYPE : nucleic acid

STRANDEDNESS : single stranded

TOPOLOGY : linear

MOLECULE TYPE : other nucleic acid synthetic DNA

SEQUENCE DESCRIPTION

CACACGCGCA CGTACGTGTC GC



INFORMATION FOR SEQ ID NO : 22

LENGTH : 27 and 8

TYPE : nucleic acid and amino acid

STRANDEDNESS : single stranded

TOPOLOGY : linear

MOLECULE TYPE : other nucleic acid synthetic DNA and amino acid

SEQUENCE DESCRIPTION

GAT TAT AAA GAT GAT GAT GAT AAA TGA 27

Asp Tyr Lys Asp Asp Asp Asp Lys

1 5 8

INFORMATION FOR SEQ ID NO : 23

LENGTH : 20

TYPE : nucleic acid

STRANDEDNESS : single stranded

TOPOLOGY : linear

MOLECULE TYPE : other nucleic acid synthetic DNA

SEQUENCE DESCRIPTION

GCTCCGGGAT CCGCTCCCTG

INFORMATION FOR SEQ ID NO : 24

LENGTH : 33

TYPE : nucleic acid

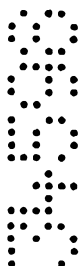
STRANDEDNESS : single stranded

TOPOLOGY : linear

MOLECULE TYPE : other nucleic acid synthetic DNA

SEQUENCE DESCRIPTION

CTCGAGCAAC CTGTGGAAGA GCCGCCCGTA ACA



INFORMATION FOR SEQ ID NO : 25

LENGTH : 58

TYPE : nucleic acid

STRANDEDNESS : single stranded

TOPOLOGY : linear

MOLECULE TYPE : other nucleic acid synthetic DNA

SEQUENCE DESCRIPTION

CTCGAGTCAT TTATCATCAT CATCTTTATA ATCACCTGTG GAAGAGCCGC CCGTAACA

INFORMATION FOR SEQ ID NO : 26

LENGTH : 31

TYPE : nucleic acid

STRANDEDNESS : single stranded

TOPOLOGY : linear

MOLECULE TYPE : other nucleic acid synthetic DNA

SEQUENCE DESCRIPTION

AGATCTCCTG TCGAAGAGCC GCCCGTAACAA

INFORMATION FOR SEQ ID NO : 27

LENGTH : 58

TYPE : nucleic acid

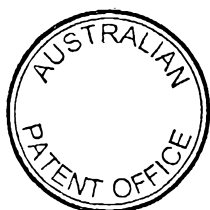
STRANDEDNESS : single stranded

TOPOLOGY : linear

MOLECULE TYPE : other nucleic acid synthetic DNA

SEQUENCE DESCRIPTION

CTCGAGTCAT TTATCATCAT CATCTTTATA ATCCTCCTTG CCGGCGTAGC GGCCTCA



INFORMATION FOR SEQ ID NO : 28

LENGTH : 36

TYPE : nucleic acid

STRANDEDNESS : single stranded

TOPOLOGY : linear

MOLECULE TYPE : other nucleic acid synthetic DNA

SEQUENCE DESCRIPTION

AAGGATCCCG AGGGTGTCTG CTGGAAGCCA GGCTCA

INFORMATION FOR SEQ ID NO : 29

LENGTH : 33

TYPE : nucleic acid

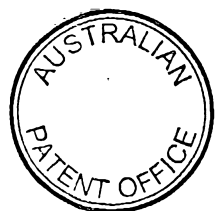
STRANDEDNESS : single stranded

TOPOLOGY : linear

MOLECULE TYPE : other nucleic acid synthetic DNA

SEQUENCE DESCRIPTION

CCTCTAGAGT CGCGGCCGTC GCACTCATTT ACC



EDITORIAL NOTE

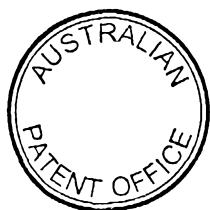
NO. 34598/97

The following claim pages are numbered

52 to 54.

THE CLAIMS DEFINING THE INVENTION ARE AS FOLLOWS:

1. An isolated polypeptide including the amino acid sequence of SEQ ID NO. 1 of the sequencing list.
2. The polypeptide according to claim 1 including the amino acid sequence of SEQ ID NO. 2 of the sequencing list.
3. The polypeptide according to claim 1 including the amino acid sequence of SEQ ID NO. 3 of the sequencing list.
4. The polypeptide according to any one of claims 1 to 3 having differentiation suppressive action against undifferentiated cells.
5. The polypeptide according to claim 4 wherein the undifferentiated cells are the undifferentiated cells except for those of the brain and nervous system or muscular system cells.
6. The polypeptide according to claim 4 wherein the undifferentiated cells are undifferentiated blood cells.
7. The polypeptide according to any one of claims 1 to 3 having suppressive action against proliferation of vascular endothelial cells.
8. A pharmaceutical composition comprising the polypeptide of any one of claims 1 and 2.
9. A cell culture medium including the polypeptide of any one of claims 1 and 2.

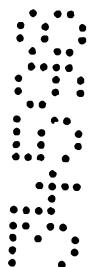


10. The cell culture medium according to claim 9 wherein the cell is the undifferentiated blood cell.

11. An isolated DNA molecule encoding a polypeptide including the amino acid sequence of SEQ ID NO. 1 of the sequencing list.

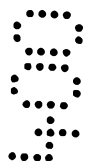
12. The DNA according to claim 11 including the DNA sequence 90 - 731 of SEQ ID NO. 4 of the sequencing list.

13. An isolated DNA molecule encoding a polypeptide including the amino acid sequence of SEQ ID NO. 2 of the sequencing list.



14. The DNA according to claim 13 including the DNA sequence 90 - 3254 of SEQ ID NO. 4 of the sequencing list.

15. An isolated DNA molecule encoding a polypeptide including the amino acid sequence of SEQ ID NO. 3 of the sequencing list.

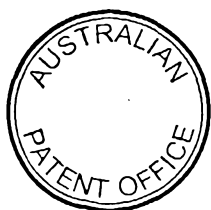


16. The DNA according to claim 15 including the DNA sequence 90 - 3725 of SEQ ID NO. 4 of the sequencing list.



17. A recombinant DNA including a DNA encoding a polypeptide including the amino acid sequence of SEQ ID NO. 1 of the sequencing list and a vector DNA which can express said DNA in the host cell.

18. A recombinant DNA including a DNA encoding a polypeptide including the amino acid sequence of SEQ ID NO. 2 of the sequencing list and a vector DNA which can express said DNA in the host cell.



19. A recombinant DNA including a DNA encoding a polypeptide including the amino acid sequence of SEQ ID NO. 3 of the sequencing list and a vector DNA which can express said DNA in the host cell.
20. A cell including the recombinant DNA of claim 17.
21. A cell including the recombinant DNA of claim 18.
22. A cell including the recombinant DNA of claim 19.
23. A process for production of the polypeptide of any one of claims 1 to 3 including culturing cells of any one of claims 20 to 22 in a medium and isolating the compound produced in the cultured mass.
24. An antibody specifically recognizing a polypeptide including the amino acid sequence of any one of SEQ ID NOs. 1 to 3 of the sequencing list.

DATED this 7th day of January 2000

ASAHI KASEI KOGYO KABUSHIKI KAISHA

WATERMARK PATENT AND TRADEMARK ATTORNEYS
290 BURWOOD ROAD
HAWTHORN VICTORIA 3122
AUSTRALIA

LCG:KMH:VRH P5715AU00.DOC

