



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<p>(51) International Patent Classification <sup>6</sup> : C12N 15/12, C07K 14/705, C12N 9/74, C07K 16/28, C12N 1/21, 1/19, 5/10, C12Q 1/00, G01N 33/68, A61K 38/48, 38/17</p>	A2	<p>(11) International Publication Number: <b>WO 99/43809</b></p> <p>(43) International Publication Date: 2 September 1999 (02.09.99)</p>
<p>(21) International Application Number: PCT/US99/02983</p> <p>(22) International Filing Date: 11 February 1999 (11.02.99)</p> <p>(30) Priority Data: 09/032,397 27 February 1998 (27.02.98) US</p> <p>(71) Applicant: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA [US/US]; 12th floor, 1111 Franklin Street, Oakland, CA 94607-5200 (US).</p> <p>(72) Inventors: COUGHLIN, Shaun, R.; 2 Turtle Rock Court, Tiburon, CA 94920 (US). KAHN, Mark; 1337 Carlos Avenue, Burlingame, CA 94010 (US).</p> <p>(74) Agent: DEVORE, Dianna, L.; Bozicevic, Field &amp; Francis LLP, Suite 200, 285 Hamilton Avenue, Palo Alto, CA 94301 (US).</p>	<p>(81) Designated States: AU, CA, JP, KR, NO, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).</p> <p><b>Published</b> <i>Without international search report and to be republished upon receipt of that report.</i></p>	
<p>(54) Title: PROTEASE-ACTIVATED RECEPTOR 4 AND USES THEREOF</p> <p>(57) Abstract</p> <p>Disclosed are cDNAs and genomic DNAs encoding protease-activated receptor 4 (PAR4) from mouse and human, and the recombinant polypeptides expressed from such cDNAs. The recombinant receptor polypeptides, receptor fragments and analogs expressed on the surface of cells are used in methods of screening candidate compounds for their ability to act as agonists or antagonists to the effects of interaction between thrombin and PAR4. Agonists are used as therapeutics to treat wounds, promote clotting, and as reagents to activate platelets in diagnostic tests. Antagonists are used as therapeutics to control blood coagulation, treat heart attack and stroke, and block inflammatory and proliferative responses to injury as occur in normal wound healing and variety of diseases including atherosclerosis, restenosis, pulmonary inflammation (ARDS) and glomerulosclerosis. Antibodies specific for a protease-activated receptor 4 (or receptor fragment or analog) and their use as a therapeutic are also disclosed.</p>		

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

## PROTEASE-ACTIVATED RECEPTOR 4 AND USES THEREOF

### FIELD OF THE INVENTION

This invention relates to nucleic acids, their encoded protease-activated receptor 4  
5 proteins, and screening assays for agonists and antagonists of the protease activated receptor  
4 proteins.

### BACKGROUND OF THE INVENTION

Thrombin, a coagulation protease generated at sites of vascular injury, activates  
10 platelets, leukocytes, and mesenchymal cells (Vu, T.-K.H. et al. (1991) Cell 64:1057-1068).  
Activation of platelets by thrombin is thought to be critical for hemostasis and thrombosis. In  
animal models, thrombin inhibitors block platelet-dependent thrombosis, which is the cause of  
most heart attacks and strokes in humans. Available data in humans suggests that thrombosis  
in arteries can be blocked by inhibitors of platelet function and by thrombin inhibitors. Thus it  
15 is likely that thrombin's actions on platelets contribute to the formation of clots that cause  
heart attack and stroke. Thrombin's other actions on vascular endothelial cells and smooth  
muscle cells, leukocytes, and fibroblasts may mediate inflammatory and proliferative  
responses to injury, as occur in normal wound healing and a variety of diseases  
(atherosclerosis, restenosis, pulmonary inflammation (ARDS), glomerulosclerosis, etc.). A  
20 thorough understanding of how thrombin activates cells is an important goal.

A receptor that mediates thrombin signaling has been previously identified (Vu, T.-  
K.H. et al. (1991) Cell 64:1057-1068; USP 5,256,766). This receptor revealed a novel  
proteolytic mechanism of activation and is referred to as PAR1 (protease-activated receptor  
1). PAR1 is activated by the binding of thrombin to and cleavage of PAR1's amino terminal  
25 exodomain at a specific site. Receptor cleavage unmask a new amino terminus, which then  
functions as a tethered peptide ligand by binding intramolecularly to the body of the receptor  
to effect transmembrane signaling (Vu, T.-K.H. et al. (1991) Cell 64:1057-1068). Synthetic  
peptides that mimic this tethered ligand domain function as PAR1 agonists and activate it  
independent of thrombin and receptor cleavage (Vu, T.-K.H. et al. (1991) Cell 64:1057-  
30 1068).

To identify which of thrombin's known cellular actions are mediated by PAR1, a  
PAR1 knockout mouse was recently generated (Connolly, A. et al. (1996) Nature 381:516-

519). Analysis of mice in which both alleles of the PAR1 gene were disrupted provided definitive evidence for a second platelet thrombin receptor and for tissue specific roles of distinct thrombin receptors. Specifically, in mice, PAR1 is not important for platelet responses but is critical for fibroblast responses.

5 Since the identification of PAR1, two other protease-activated receptors have been cloned. A second protease-activated receptor (PAR2) was cloned during a search for relatives of the Substance K receptor (Nystedt, S., et al. (1994) PNAS USA, 91:9208-9212). The physiological activator of PAR2 remains unknown; it is not activated by thrombin. A third protease-activated receptor (PAR3) has also been cloned, but available data suggests  
10 that this receptor is involved in thrombin-mediated platelet responses in mice but not in humans.

There is a need for a better understanding of thrombin-mediated platelet activation. There is also a need for the identification and characterization of factors involved in platelet-mediated pathologies, such as platelet-dependent arterial thrombosis. The understanding of  
15 the mechanisms of such offer new mechanisms for treating associated pathologies.

#### SUMMARY OF THE INVENTION

Protease-activated receptor 4 (PAR4) is disclosed. PAR4 is useful in assaying libraries of compounds for their activity as thrombin agonists and antagonists. DNA encoding  
20 PAR4 is also disclosed as is its insertion into a functional expression vector, DNA expressed in a cell line, and the use of the DNA expression product in an assay to identify compounds as agonists or antagonists of thrombin's effect on PAR4.

The invention comprises substantially pure DNA (cDNA or genomic DNA) encoding a protease-activated receptor 4 (PAR4) from vertebrate tissues (SEQ ID NO:1, SEQ ID  
25 NO:3, AND SEQ ID NO:4) and degenerate sequences thereof; substantially pure protease-activated receptor 4 polypeptides encoded thereby; as well as amino acid sequences substantially identical to the amino acid sequences SEQ ID NO:2 and SEQ ID NO:5 from mouse and human, respectively. The invention also features DNA sequences that hybridize under stringent conditions to DNA encoding PAR4, or to DNA complementary to DNA  
30 encoding PAR4. Such DNA sequences are preferably at least 25 nucleotides in length, more preferably 50 nucleotides in length. The invention further comprises fragments of the PAR4 receptor which are activated by thrombin. Such fragments may have the same amino acid

sequence as SEQ ID NO:2 and SEQ ID NO:5 or be at least 80% identical to the amino acid sequences SEQ ID NO:2 and SEQ ID NO:5. Such fragments are preferably at least 10 amino acids in length, more preferably at least 30 amino acids in length.

5 In various embodiments, the DNA, receptor or receptor fragment is derived from a vertebrate animal, preferably, human or mouse. However, the gene can be chemically synthesized.

An object of the invention is to provide a nucleotide sequence encoding a novel receptor, preferably PAR4 and its functional equivalents.

10 Another object is to provide a cell line genetically engineered to express the nucleotide sequence encoding PAR4.

Another object is to provide an antibody which selectively binds the PAR 4 receptor.

Another object is to provide a method whereby a compound or library of compounds can be assayed for their ability to activate or block the receptor expressed by the nucleotide sequence.

15 An advantage of the present invention is that a novel thrombin receptor PAR4 is disclosed making it possible to identify novel thrombin agonists and antagonists which may not be identifiable via PAR1, PAR2, or PAR3 receptors.

A feature of the invention is that it makes it possible to obtain additional information regarding thrombin activation and the sequence of biochemical events initiated by such.

20 These and other objects, advantages and features of the present invention will become apparent to those skilled in the art upon reading the disclosure.

#### BRIEF DESCRIPTION OF THE DRAWINGS

25 Figs. 1A and 1B are the complete nucleotide sequences (SEQ ID NO:1) of the mouse protease-activated receptor 4 gene coding region cDNA.

Fig. 2 is the deduced amino acid sequence (SEQ ID NO:2) of the receptor encoded by the nucleotide sequence of Fig. 1. The amino acid sequence encoding mouse PAR4 contains 397 amino acids. The deduced amino acid sequence begins at nucleotides 1-3 (ATG = Met) and ends at nucleotides 1192-1194 (TGA = stop).

30 Fig. 3 is the nucleotide sequence of genomic mouse PAR4 (SEQ ID NO:3) demonstrating a small intron (250 bp) between the signal peptide and cleavage site of the receptor coding sequence. The intronic sequence is underlined.

Figs. 4A and 4B are the complete nucleotide sequences (SEQ ID NO:4) of the human protease-activated receptor 4 gene coding region cDNA.

Fig. 5 is the deduced amino acid sequence (SEQ ID NO:5) of the receptor encoded by the nucleotide sequence of Fig. 4. The amino acid sequence encoding human PAR4 contains  
5 385 amino acids. The deduced amino acid sequence begins at nucleotides 3-5 (ATG = Met) and ends at nucleotides 1157-1159 (TGA = stop).

Figs. 6A and 6B show the alignment of the deduced amino acid sequences (SEQ ID NO:2, 6-8) of the mouse PAR4, mouse PAR3, mouse PAR2, and mouse PAR1 genes. To indicate homology, gaps (represented by blank spaces) have been introduced into the five  
10 sequences. Exact amino acid matches between at least three of the PARs have been enclosed in a shaded box.

Fig. 7 shows the structure of the mouse PAR genes, revealing a small (250bp) intron separating exon 1, which encodes the signal peptide, from exon 2, which encodes the mature receptor protein (SEQ ID NO:3, 9-11).

15 Fig. 8 pictorially describes the thrombin cleavage site and activating peptide of the mouse PAR4 receptor.

Fig. 9 is a bar graph showing the calcium response of oocytes expressing mouse PAR4 upon exposure to different agonists. Agonists include thrombin, the predicted activating peptide for PAR4 (GYPGKF), and the predicted activating peptide for PAR1  
20 (SFLLRN).

Fig. 10 is a bar graph showing the calcium response of oocytes expressing human PAR4 upon exposure to different agonists. Agonists include thrombin, the predicted activating peptide for human PAR4 (GYPGQV), the predicted activating peptide for mPAR4 (GYPGKF), and the predicted activating peptide for PAR1 (SFLLRN).

25 Fig. 11 is a bar graph showing the calcium response of oocytes expressing mouse PAR4 upon exposure to various serine proteases.

Fig. 12 is a bar graph representing the activation of PAR4 in *Xenopus laevis* oocytes cell types upon exposure to the tethered ligand peptides of the PARs.

Fig. 13 is a graph showing the aggregation of PAR3 knockout mouse platelets in  
30 response to mPAR4 activating peptide GYPGKF.

Fig. 14 is a graph showing ATP secretion (top) and aggregation (bottom) of wild-type mouse platelets in response to mPAR4 activating peptide GYPGKF.

Fig. 15 is a graph depicting the activation of human platelets desensitized to the PAR1 activating peptide by the PAR4 predicted activating peptide.

#### DESCRIPTION OF THE PREFERRED EMBODIMENTS

5 Before the present protease-activated receptor assays and methods of using such are described, it is to be understood that this invention is not limited to the particular DNA sequences, materials, methods, or processes described as such may, of course, vary. It is also to be understood that the terminology used herein is for the purpose of describing particular  
10 will be limited only by the appended claims.

It must be noted that as used in this specification and the appended claims, the singular forms "a", "and," and "the" include plural referents unless the contexts clearly dictates otherwise. Thus, for example, reference to "a DNA sequence" includes mixtures and large numbers of such sequences, reference to "an assay" includes assays of the same general  
15 type, and reference to "the method" includes one or more methods or steps of the type described herein.

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

20 Unless defined otherwise, all technical and scientific terms herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any methods and materials, similar or equivalent to those described herein, can be used in the practice or testing of the present invention, the preferred methods and materials are described herein. All publications cited herein are incorporated herein by  
25 reference for the purpose of disclosing and describing specific aspects of the invention for which the publication is cited in connection with.

#### DEFINITIONS

30 By "protease-activated receptor 4", "PAR4", "PAR4 receptor" and the like, is meant all or part of a vertebrate cell surface protein which is specifically activated by thrombin or a thrombin agonist thereby activating PAR4-mediated signaling events (e.g., phosphoinositide hydrolysis, Ca<sup>2+</sup> efflux, platelet aggregation). The polypeptide is characterized as having the

ligand activating properties (including the agonist activating and antagonist inhibiting properties) and tissue distribution described herein. Specifically, PAR4 receptors are expressed by the DNA sequences of SEQ ID NOs: 2, 4, and 5.

5 By a "polypeptide" is meant any chain of amino acids, regardless of length or post-translational modification (e.g., glycosylation).

By "substantially pure" is meant that the protease-activated receptor 4 polypeptide provided by the invention is at least 60%, by weight, free from the proteins and naturally-occurring organic molecules with which it is naturally associated. Preferably, the preparation is at least 75%, more preferably at least 90%, and most preferably at least 99%, by weight,  
10 PAR4 polypeptide. A substantially pure PAR4 polypeptide may be obtained, for example, by extraction from a natural source, by expression of a recombinant nucleic acid encoding a PAR4 polypeptide, or by chemically synthesizing the protein. Purity can be measured by any appropriate method, e.g., column chromatography, polyacrylamide gel electrophoresis, or HPLC analysis. The protein is substantially pure if it can be isolated to a band in a gel.

15 By a "substantially identical" amino acid sequence is meant an amino acid sequence which differs only by conservative amino acid substitutions, for example, substitution of one amino acid for another of the same class (e.g., valine for leucine, arginine for lysine, etc.) or by one or more non-conservative amino acid substitutions, deletions, or insertions located at positions of the amino acid sequence which do not destroy the biological activity of the  
20 receptor. Such equivalent receptors can be isolated by extraction from the tissues or cells of any animal which naturally produces such a receptor or which can be induced to do so, using the methods described below, or their equivalent; or can be isolated by chemical synthesis; or can be isolated by standard techniques of recombinant DNA technology, e.g., by isolation of cDNA or genomic DNA encoding such a receptor. Substantially identical receptors have the  
25 same biological function, e.g. are activated by the same compound.

By "derived from" is meant encoded by the genome of that organism and present on the surface of a subset of that organism's cells.

By "isolated DNA" is meant DNA that is not in its native environment in terms of not being immediately contiguous with (i.e., covalently linked to) the complete coding sequences  
30 with which it is immediately contiguous (i.e., one at the 5' end and one at the 3' end) in the naturally-occurring genome of the organism from which the DNA of the invention is derived. The term therefore includes, for example, recombinant DNA which is incorporated into a

vector; into an autonomously replicating plasmid or virus; or into the genomic DNA of a prokaryote or eukaryote; or which exists as a separate molecule (e.g., a cDNA or a genomic or cDNA fragment produced by PCR or restriction endonuclease digestion) independent of other sequences. It also includes any recombinant DNA which is part of a hybrid gene  
5 encoding additional polypeptide sequence.

"Isolated DNA" can mean the DNA is in vectors which are preferably capable of directing expression of the protein encoded by the DNA in a vector-containing cell and further includes cells containing such vectors (preferably eukaryotic cells, e.g., CHO cells (ATCC; Cat. No. CCL 61 or COS-7 cells (ATCC; Cat. No. CRL 1651; and the *Xenopus*  
10 oocytes of the type described in the above cited reference Vu, T.-K.H. et al. (1991) Cell 64:1057-1068). Preferably, such cells are stably transfected with such isolated DNA.

By "transformed cell" and "transfected cell", "genetically engineered cell", and the like, is meant a cell into which (or into an ancestor of which) has been introduced, by means of genetic engineering, a DNA molecule encoding a PAR4 (or DNA encoding a biologically  
15 active fragment or analog, thereof). Such a DNA molecule is "positioned for expression" meaning that the DNA molecule is positioned adjacent to a DNA sequence which directs transcription and translation of the sequence (i.e., facilitates the production of the PAR4 protein, or fragment or analog, thereof).

By "antibody" is meant an immunoglobulin protein which is capable of binding an  
20 antigen. Antibody as used herein is meant to include the entire antibody as well as any antibody fragments (e.g. F(ab')<sub>2</sub>, Fab', Fab, Fv) capable of binding the epitope, antigen or antigenic fragment of interest.

Antibodies of the invention are immunoreactive or immunospecific for and therefore specifically and selectively bind to a PAR4 protein. Antibodies for PAR4 are preferably  
25 immunospecific -- i.e., not substantially cross-reactive with related materials. Although the term "antibody" encompasses all types of antibodies (e.g., monoclonal) the antibodies of the invention are preferably produced using the phage display methodology described herein. The preferred antibody of the invention is a purified antibody. By purified antibody is meant one which is sufficiently free of other proteins, carbohydrates, and lipids with which it is naturally  
30 associated. Such an antibody "preferentially binds" to a PAR4 protein (or an antigenic fragment thereof), i.e., does not substantially recognize and bind to other antigenically-unrelated molecules.

By "specifically activates", as used herein, is meant an agent, such as thrombin, a thrombin analog, a PAR4 agonist or other chemical agent including polypeptides such as an antibody, which activates protease-activated receptor 4, receptor polypeptide or a fragment or analog thereof to initiate PAR4-mediated biological events as described herein, but which  
5 does not substantially bind other molecules in a sample, e.g., a biological sample, which naturally includes a protease-activated receptor 4 polypeptide.

By "specifically inhibits", as used herein, is meant an agent, such as a thrombin analog, a PAR4 antagonist or other chemical agent including polypeptides such as an antibody, which inhibits activation of protease-activated receptor 4, receptor polypeptide or a fragment or  
10 analog thereof, such as by inhibiting thrombin or by blocking activation of PAR4 by thrombin or other PAR4 activator. Preferably, the agent activates or inhibits the biological activity *in vivo* or *in vitro* of the protein to which it binds.

By "biological activity" is meant the ability of the protease-activated receptor 4 to bind thrombin or a PAR4 agonist and signal the appropriate cascade of biological events (e.g.,  
15 phosphoinositide hydrolysis, Ca<sup>2+</sup> efflux, and platelet aggregation, and the like, as described herein.

By "substantial increase" is meant an increase in activity or other measurable phenotypic characteristic that is at least approximately a 2-fold increase over control level (where control assays are performed in the absence of activator), preferably at least  
20 approximately a 5-fold increase, more preferably at least approximately a 10-fold increase in activity over a control assay.

By "substantial decrease" or "substantial reduction" is meant a decrease or reduction in activity or other measurable phenotypic characteristic that is approximately 80% or the control level, preferably reduced to approximately 50% of the control level, or more  
25 preferably reduced to approximately 10% or less of the control level.

The terms "screening method" and "assay method" are used to describe a method of screening a candidate compound for its ability to act as an agonist or antagonist of a PAR4 ligand. The method involves: a) contacting a candidate agonist compound with a  
30 recombinant protease-activated receptor 4 (or PAR4 agonist-binding fragment or analog); b) measuring activation of the receptor, the receptor polypeptide or the receptor fragment or analog; and c) identifying agonist compounds as those which interact with the recombinant receptor and trigger or block PAR4 activation. Interaction may be cleavage of the receptor

to unmask an intramolecular receptor activating peptide or by mimicking the intramolecular receptor-activating peptide. A tethered ligand may be more difficult to block than a free agonist. Thus, blocking thrombin is the acid test for an antagonist which will block responses by other thrombin substrates. These terms include assays that examine effects on unoccupied  
5 receptors as well as assays that utilize displacement of a ligand from an occupied receptor.

By an "agonist" is meant a molecule which mimics a particular activity, in this case, interacting with a PAR4 ligand in a manner which activates thereby triggering the biological events which normally result from the interaction (e.g., phosphoinositide hydrolysis, Ca<sup>2+</sup> efflux, and platelet aggregation). Preferably, an agonist initiates a substantial increase in  
10 receptor activity relative to control assays in the absence of activator or candidate agonist. An agonist may possess the same, less, or greater activity than a naturally-occurring PAR4 ligand.

The terms "antagonist assay", "antagonist screening" and the like, refer to a method of screening a candidate compound for its ability to antagonize interaction between a naturally-  
15 occurring activating ligand or an agonist and the PAR4. The method involves: a) contacting a candidate antagonist compound with a first compound which includes a recombinant PAR4 (or agonist-binding fragment or analog) on the one hand and with a second compound which includes thrombin or a PAR4 agonist on the other hand; b) determining whether the first and second compounds interact or are prevented from interaction by the candidate compound;  
20 and c) identifying antagonistic compounds as those which interfere with the interaction of the first compound (PAR4 receptor) to the second compound (PAR4 agonist) and which thereby substantially reduce thrombin or PAR4 agonist-activated biological events (e.g., phosphoinositide hydrolysis, Ca<sup>2+</sup> efflux, and platelet aggregation).

By an "antagonist" is meant a molecule which blocks activation of a PAR4 receptor.  
25 This can be done by inhibiting a particular activity such as the ability of thrombin, for example, to interact with a protease-activated receptor 4 thereby triggering the biological events resulting from such an interaction (e.g., phosphoinositide hydrolysis, Ca<sup>2+</sup> efflux, and platelet secretion, or platelet aggregation). An antagonist may bind to and thereby block activation of a PAR4 receptor.

30 The terms "treatment", "treating", "treat" and the like are used herein to generally mean obtaining a desired pharmacologic and/or physiologic effect. The effect may be prophylactic in terms of completely or partially preventing a disease or symptom thereof

and/or may be therapeutic in terms of a partial or complete cure for a disease and/or adverse effect attributable to the disease. "Treatment" as used herein covers any treatment of a disease in a mammal, particular a human, and includes:

- 5 (a) preventing the disease or symptom from occurring in a subject which may be predisposed to the disease or symptom but has not yet been diagnosed as having it;
- (b) inhibiting the disease symptom, i.e., arresting its development; or
- (c) relieving the disease symptom, i.e., causing regression of the disease.

#### PREFERRED EMBODIMENTS

10 In preferred embodiments of both screening methods, the recombinant PAR4 is stably expressed by a vertebrate cell which normally presents substantially no PAR4 on its surface (i.e., a cell which does not exhibit any significant thrombin-mediated phosphoinositide hydrolysis or  $\text{Ca}^{2+}$  efflux in the presence of a PAR activator); the vertebrate cell is a mammalian cell, is a Rat 1 cell, or a COS 7 cell; and the candidate antagonist or candidate  
15 agonist is a thrombin analog, PAR4 peptide fragment or analog or other chemical agent including a polypeptide such as an antibody.

The receptor proteins of the invention are likely involved in the activation of vertebrate platelet, leukocyte, and mesenchymal cells in response to wounding, as well as mediating signaling in embryonic development. Such proteins and in particular PAR4  
20 antagonists are useful therapeutics for the treatment of such conditions as thrombosis, atherosclerosis, restenosis, and inflammation associated with normal wound healing and a variety of diseases including atherosclerosis, restenosis, pulmonary inflammation (ARDS) and glomerulosclerosis. Preferred therapeutics include 1) agonists, e.g., thrombin analogs, PAR4 peptide fragments or analogs thereof, or other compounds which mimic the action of  
25 thrombin upon interaction with the protease-activated receptor 4 or mimic the action of an intramolecular receptor activating peptide; and 2) antagonists, e.g., thrombin analogs, antibodies, or other compounds, which block thrombin or protease-activated receptor 4 function by interfering with the thrombin:receptor interaction or by interfering with the receptor intramolecular activating peptide. The dosage would be expected to be comparable  
30 with current antiinflammatory drugs and should be adjusted based on the age, sex, weight and condition of the patient beginning with small doses and increasing gradually based on responsiveness and toxicity.

Because the receptor component may now be produced by recombinant techniques and because candidate agonists and antagonists may be screened using transformed, cultured cells, the instant invention provides a simple and rapid approach to the identification of useful therapeutics. Isolation of the PAR4 gene (as cDNA or genomic DNA) allows its expression  
5 in a cell type which does not normally bear PAR4 on its surface, providing a system for assaying a thrombin:receptor interaction and receptor activation.

### EXAMPLES

The following examples are put forth so as to provide those of ordinary skill in the art  
10 with a complete disclosure and description of how to make receptor proteins and sequences encoding such proteins and carry out the methodology for finding such DNA sequences and proteins, and are not intended to limit the scope of what is regarded as the invention. Efforts have been made to insure accuracy with respect to numbers used (e.g., amounts, temperature, etc.) but some experimental errors and deviations should be accounted for. Unless indicated  
15 otherwise, parts or parts by weight, molecular weight is weight average molecular weight; temperature is in degrees centigrade; and pressure is at or near atmospheric.

There now follows a description of the cloning and characterization of the cDNA, genomic DNA and the receptor protein of the protease-activated receptor 4 from mouse and human. Expression vectors containing and capable of expressing the PAR4 DNA, as well as  
20 transformed cells containing and expressing the DNA of the invention are also described. Also described are possible PAR4 agonists and antagonists as well as screening assays for receptor agonists and receptor antagonists.

#### 25 EXAMPLE 1: ISOLATION OF THE MOUSE AND HUMAN PROTEASE-ACTIVATED RECEPTOR 4

The public expressed sequence tag (EST) database was searched for potential protease-activated receptor sequences by identifying sequences with homology to PAR1, PAR2 and PAR3. One particular EST, clone 400689, was identified in a database search using PAR2 sequences, showing similarity over an eleven amino acid stretch. The EST was  
30 further characterized.

The EST sequence was used to obtain mouse and human cDNA and genomic clones by a combination of PCR and hybridization techniques (see, for example, Sambrook, J. et al.

(1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, New York). The nucleotide sequences for the mouse and human PAR4, and deduced amino acid sequences corresponding to these nucleotide sequences, are shown in Figs. 1-4. The mouse PAR4 cDNA contained an open reading frame encoding a 397 amino acid putative G protein-coupled receptor (Fig. 2). The 5' sequence of the PAR4 cDNAs encoded a predicted thrombin cleavage site, suggesting that this new receptor was a novel thrombin receptor. The gene was characterized further as a potential protease-activated receptor.

The predicted protein structure of the mouse PAR4 product showed significant homology in conserved regions when compared to other mouse protease-activated receptors. Homology between the deduced mouse PAR4 amino acid sequence and the amino acid sequences of mouse PAR1, PAR2, and PAR3 are shown by alignment in Figs. 6A and 6B. Amino acids shared by at least three of the receptor molecules are boxed and shaded.

The genomic region containing the mouse PAR4 gene was sequenced, revealing an exon organization characteristic of protease-activated receptors: a first exon encoding the signal peptide region separated from a second exon encoding the majority of the structure of the mature receptor molecule by a characteristic intron of approximately 250 bp (Fig. 7). The mPAR4 thrombin cleavage site is found at the amino terminus of the protein, and the amino acid sequence corresponding to the mPAR4 tethered ligand peptide, or "activating peptide", is directly adjacent to the cleavage site (Fig. 8).

#### *Cloning of mouse PAR4*

The mouse PAR4 cDNA used for the functional studies presented below was cloned from a mouse embryo day 14-15 library. The sequence of EST clone 400689 was used to generate two reverse primers, CCAGTCACAGAAGTGTAGAGGAGCAAATGG (R2) and CAAGCCAGACCCCCTTCCCAC (R3), for 5'RACE using cDNA from mouse embryo day 14-15 (Marathon cDNA, Clontech). The forward primers used were Clontech AP1 and AP2. Two "nested" PCR reactions were performed, the first using primers AP1 and R2, the second with AP2 and R3. The first PCR reaction conditions were as follows: 95°C for 5 seconds, 72°C for 4 minutes for the first five cycles, and 95°C for 5 seconds, 70°C for 4 minutes for the next five cycles, and 95°C for 2 seconds, 68°C for 4 minutes for the following 25 cycles. The second PCR reaction conditions were 95°C for 1 minute, followed by 95°C for 30 seconds, 68°C for 3 minutes for 25 cycles. A dominant band of 950 bp was

seen following the nested (AP2-R3) reaction. This product was sequenced, providing mPAR4 sequence 5' to the first transmembrane domain.

Sequence to the start codon was obtained with a second series of 5'RACE reactions again using the Clontech Marathon cDNA and AP1 and AP2 primers. The primary reaction used AP1 and reverse primer CCACAGCCACCACAAGCCCATAGAG (RACE 1) and the second reaction used AP2 and CCCAGCAAGCAGTGCTTGAGAGCTG (RACE 2). The reaction conditions for AP1-RACE1 were: 95°C for 30 seconds, 68°C for 3 minutes for 25 cycles. The conditions for the nested AP2-RACE 2 reaction were the same except the PCR went for 30 cycles. A dominant 300 bp band was obtained and sequenced.

The EST clone 400689 was sequenced from both ends to confirm the sequence in the EST database. This sequence, however, differed from sequence which was later obtained both from a cDNA obtained by hybridization screening of a bEND (mouse brain endothelial cell line) library and from a BAC genomic clone (Genome Systems). We believe the latter sequence to be correct.

The final functional clone was obtained by generating a PCR product from the start codon to the stop codon indicated by the EST clone using bEND cDNA as a template. The 3' end of this clone was subsequently replaced by subcloning with a genomic fragment at the Nco I site. Both the original product and the clone containing the 3' genomic fragment were tested in oocytes and found to be equally active.

#### *Human PAR4*

The human PAR4 sequence was determined using a degenerate PCR scheme. Human PAR4 sequence was obtained using degenerate PCR primers to amplify a 900 bp dominant product from total human genomic DNA. The primers used for the PCR reactions were: TA(A/G)TA(A/G)TA(A/G/T)AT(A/G)AAIGG(G/A)TCIAC(G/A)CA (designated DR1), GTIGGIA(C/T)TICCIGCIAA(C/T)GG(A/C/G/T)(C/T)T (designated DF1), where "I" designates Inosine. Reaction conditions were: 95°C for 4 minutes, followed by 95°C for 1 minute, 50°C for 2.5 minutes, 72°C for 1.5 minutes for 50 cycles. Sequencing of the 900 bp product revealed a novel amino acid sequence that was 88% identical to mouse PAR4.

Human megakaryocytic cell lines were screened for PAR4 expression by Northern using the above 900 bp product as a probe and the K562 erythroleukemia cell line found to be positive. 5'RACE was then performed using the GIBCO 5'RACE kit and K562 mRNA per

the manufacturer's instructions. Nested PCR reactions were performed using two reverse primers: CGAGGTTTCATCAGCAGCATGG (GSP2A) and TGCGTGTCCAGCAGGGACAG (GSP2B). Conditions for the first reaction using the GIBCO forward anchor primer and GSP2A were: 95°C for 4 minutes, followed by 95°C for 45 seconds, 56°C for 1 minute, and 72 for 1 minute for 35 cycles. The hemi-nested second reaction was performed with the GIBCO anchor primer and GSP2B as follows: 95°C for 4 minutes, followed by 95°C for 45 seconds, 50°C for 1 minute, and 72°C for 1 minute for 30 cycles. A dominant 350 bp band was observed, subcloned and sequenced, providing sequence to the hPAR4 start codon.

The same kit and template was used for 3'RACE. The forward primers used were: CCTTCTTCGTGCCAGCAAC (3'GSPA) and GCTGCTGCTGCATTACTCGG (3'GSPB). The GIBCO UAP primer was used as a reverse primer and hemi-nested PCR performed. The first reaction consisted of 95°C for 4 minutes followed by 95°C for 45 seconds, 50°C for 1 minute, and 72°C for 4 minutes for 35 cycles. The second reaction consisted of 95°C for 45 seconds, 50°C for 1 minute, and 72°C for 1 minute for 30 cycles. A dominant 1.6 kb band was observed, subcloned and sequenced, providing sequence to a stop codon.

A functional hPAR4 clone was created by PCR using Vent polymerase with primers from the start codon to 50 bp beyond the stop codon using 25 cycles of PCR. The template was K562 cDNA. The PCR product was sequenced and subcloned into an oocyte expression vector for generating cRNA (pFROGGY). Human PAR4 cRNA was microinjected into *Xenopus* oocytes to demonstrate function.

#### EXAMPLE 2: POLYPEPTIDE EXPRESSION

Polypeptides according to the invention may be produced by transformation of a suitable host cell with all or part of a PAR3 encoding cDNA fragment (e.g., the cDNAs described above) in a suitable expression vehicle, and expression of the receptor.

Those skilled in the field of molecular biology will understand that any of a wide variety of expression systems may be used to provide the recombinant receptor protein. The precise host cell used is not critical to the invention. The receptor may be produced in a prokaryotic host (e.g., *E. coli*) or in a eukaryotic host (e.g., *Saccharomyces cerevisiae* or mammalian cells, e.g., COS-6M, COS-7, NIH/3T3, or Chinese Hamster Ovary cells). Such cells are available from a wide range of sources (e.g., the American Type Culture Collection,

Rockville, MD). The method of transfection and the choice of expression vehicle will depend on the host system selected. Transformation and mammalian cell transfection methods are described, e.g., in Ausubel et al. (Current Protocols in Molecular Biology, John Wiley & Sons, New York, (1989)); expression vehicles may be chosen from those provided, e.g., in 5 Cloning Vectors: A Laboratory Manual (Pouwels, P.H. et al., (1985), Supp. 1987).

Particularly preferred expression systems are the *Xenopus* oocyte cells of Vu et al. (Vu et al., Cell (1991) *supra*) and insect cells (SF9-baculovirus) transfected with an expression vector containing and expressing a receptor protein or biologically active fragment thereof. DNA encoding the human or mouse PAR4 or an appropriate receptor fragment or 10 analog (as described above) is inserted into the expression vector in an orientation designed to allow expression. Alternatively, the PAR4 (or biologically active receptor fragment or analog) is expressed by a stably-transfected mammalian cell line. Other preferable host cells which may be used in conjunction with the expression vehicle include NIH/3T3 cells (ATCC Accession No. 1658). The expression may be used in a screening method of the invention 15 (described below) or, if desired, the recombinant receptor protein may be isolated as described below.

A number of vectors suitable for stable transfection of mammalian cells are available to the public, e.g., see Pouwels et al. (*supra*); methods for constructing such cell lines are also publicly available, e.g., in Ausubel et al. (*supra*). In one example, cDNA encoding the 20 receptor (or receptor fragment or analog) is cloned into an expression vector which includes the dihydrofolate reductase (DHFR) gene. Integration of the plasmid and, therefore, the PAR4-encoding gene into the host cell chromosome is selected for by inclusion of 0.01-300  $\mu$ M methotrexate in the cell culture medium (as described in Ausubel et al., *supra*). This dominant selection can be accomplished in most cell types. Recombinant protein expression 25 can be increased by DHFR-mediated amplification of the transfected gene. Methods for selecting cell lines bearing gene amplifications are described in Ausubel et al. (*supra*); such methods generally involve extended culture in medium containing gradually increasing levels of methotrexate. DHFR-containing expression vectors commonly used for this purpose include pCVSEII-DHFR and pAdD26SV(A) (described in Ausubel et al., *supra*). Any of the 30 host cells described above or, preferably, a DHFR-deficient CHO cell line (e.g., CHO DHFR cells, ATCC Accession No. CRL 9096) are among the host cells preferred for DHFR selection of a stably-transfected cell line or DHFR-mediated gene amplification.

One particularly preferred stable expression system is a Rat 1 cell (ATCC) stably transfected with a pcDNA1/NEO (InVitrogen, San Diego, CA) expression vector.

Expression of the recombinant receptor (e.g., produced by any of the expression systems described herein) may be assayed by immunological procedures, such as Western blot or immunoprecipitation analysis of recombinant cell extracts, or by immunofluorescence of intact recombinant cells (using, e.g. the methods described in Ausubel et al., supra).  
5  
Recombinant receptor protein is detected using an antibody directed to the receptor. Described below are methods for producing anti-protease-activated receptor 4 antibodies using, as an immunogen, the intact receptor or a peptide which includes a suitable protease-activate receptor 4 epitope. To detect expression of a PAR3 fragment or analog, the  
10 antibody is preferably produced using, as an immunogen, an epitope included in the fragment or analog.

Once the recombinant PAR4 protein (or fragment or analog, thereof) is expressed, it is isolated, e.g., using immunoaffinity chromatography. In one example, an anti-PAR4  
15 antibody may be attached to a column and used to isolate intact receptor or receptor fragments or analogs. Lysis and fractionation of receptor-harboring cells prior to affinity chromatography may be performed by standard methods (see, e.g., Ausubel et al., supra). Once isolated, the recombinant protein can, if desired, be further purified, e.g., by high performance liquid chromatography (see, e.g., Fisher, Laboratory Techniques In Biochemistry  
20 And Molecular Biology, eds., Work and Burdon, Elsevier, (1980)).

Receptors of the invention, particularly short receptor fragments, can also be produced by chemical synthesis (e.g., by the methods described in Solid Phase Peptide Synthesis, (1984) 2nd ed., The Pierce Chemical Co., Rockford, IL).

25 **EXAMPLE 3:           ACTIVATION STUDIES OF THE RECOMBINANT PROTEASE-ACTIVATED RECEPTOR 4**

PAR4 was demonstrated to be activated by thrombin when expressed in the *Xenopus* oocyte cells of Vu et al. (Vu et al., Cell (1991) *supra*). The ability of PAR4 to mediate signaling by  $\alpha$ -thrombin was tested. *Xenopus* oocytes were microinjected with cRNA  
30 encoding epitope-tagged mouse PAR4 (mPAR4). Thrombin-triggered  $^{45}\text{Ca}$  release was measured as described in Vu et al. (Vu, T.-K. H. et al. (1991) *supra*). The oocytes were microinjected with either 2.5 or 25 ng of the mPAR4 cRNA. A control set of oocytes were

injected with 25 ng of the mouse PAR1 cDNA, WT5. Surface expression of receptors may be confirmed by M1 antibody binding by the method of Ishii, K. et al. (Ishii, K. et al. (1995) J. Biol. Chem. 270:16435-16440; and Ishii, K. et al. (1993) J. Biol. Chem. 268:9780-9786, which references are herein incorporated by reference in their entirety).

5 Microinjection of *Xenopus* oocytes with mouse PAR4 cRNA conferred thrombin-dependent <sup>45</sup>Ca mobilization (Fig. 9) which reflects agonist-triggered phosphoinositide hydrolysis in this system.  $\alpha$ -thrombin was able to induce calcium mobilization in both the PAR1 and PAR4 expressing oocytes, even at injection levels of 1nM. mPAR4 expressing cells treated showed approximately a 25-fold increase of Ca release at both concentrations,  
10 and this did not vary significantly between the cells expressing higher or lower concentration of mPAR4. In addition, exposure of mPAR4-injected oocytes to mPAR4's activation peptide, GYPGKF (SEQ ID NO:12 ), induced PAR4-mediated <sup>45</sup>Ca mobilization to approximately a 33-fold increase, which is a 2-fold increase over PAR1-mediated mobilization. Exposure to the PAR1 activating peptide, SLLRN (SEQ ID NO:13), failed to activate PAR4 even at a  
15 concentration of 100 $\mu$ M.

Human PAR4 cRNA was also transcribed in vitro from hPAR4 cDNA using standard techniques as described above. *Xenopus* oocytes were microinjected with 2ng of cRNA/oocyte, and calcium signaling in response to agonist peptides was measured (Fig. 10). 10nM thrombin caused a 30-fold increase in calcium mobilization in oocytes expressing  
20 PAR4, consistent with hPAR4 as a functional thrombin receptor. The human PAR4 activating peptide GYPGQV (SEQ ID NO:14) also showed calcium mobilization at 100 $\mu$ M. The mouse activating peptide, GYPGKF, was even more potent than the human activating peptide, whereas the human PAR1 activating peptide showed no significant activity. The negative controls, uninjected cells and cells expressing irrelevant receptors, also showed no  
25 activity.

#### EXAMPLE 4: SPECIFICITY OF THE RECOMBINANT PROTEASE-ACTIVATED RECEPTOR 4

The specificity of activation of PAR4 was examined by the introduction of a number  
30 of the arginine/lysine specific serine proteases and PAR4 activating peptides to *Xenopus* oocytes expressing PAR4 (Fig. 11). Various concentrations of the arginine/lysine specific serine proteases plasmin, trypsin, tissue plasminogen activator (APC), Factor VIIa, Factor

Xa, and thrombin were tested. A variant of the PAR4 activating peptide was also tested. The mPAR4 receptor was activated upon treatment with thrombin, its activating peptide GYPGKF, and the relatively less specific serine protease trypsin. No significant mPAR4 activation was seen in response to the other proteases tested, nor with the treatment of the  
5 variant activating peptide GAPGKF, which is expected to lack activity.

The specificity of mPAR4, mPAR1, and mPAR2 signaling was also examined. Protease-triggered <sup>45</sup>Ca release was measured in *Xenopus* oocytes expressing mouse PAR1, PAR2 or PAR4 stimulated with the activating peptides of each receptor and the proteases thrombin and trypsin (Fig. 12). Each PAR was activated specifically by its respective  
10 activating peptide, and not by the activating peptides of the other mPARs. mPAR2 and mPAR3 showed a small level of activation upon treatment with trypsin, but PAR4 displayed a much greater response with a 24-fold increase in <sup>45</sup>Ca release. Finally, PAR1 and PAR4 showed a significant response to thrombin, with PAR4 expressing oocytes exhibiting a greater response to thrombin than thrombin receptor mPAR1.

15

#### EXAMPLE 5: PAR4 TISSUE EXPRESSION IN MOUSE AND HUMAN

Northern analysis of mouse tissues revealed that PAR4 mRNA was strongest in the mouse spleen cells. The levels of splenic expression of PAR4 are similar to the expression of PAR3 in megakaryocytes, the predicted site of action for PAR3. The role of PAR4 in mouse  
20 tissues awaits elucidation, but the finding of PAR4 in spleen is consistent with a role for PAR4 in mediating activation of platelets and other hematopoietic cells by thrombin.

In situ hybridization of mouse tissue reveals the presence of PAR4 mRNA in splenic megakaryocytes, the platelet precursor cells. Control samples in which hybridization is performed with a sense strand probe control are a negative control for all cell types.

25 Northern analysis of mouse tissues for PAR4 mRNA show signal in spleen, with low levels seen in brain, heart, and other tissues. The spleen is a hematopoietic organ in mouse, and both Northern and *in situ* hybridization data suggest that PAR4 is most abundantly expressed in megakaryocytes in the mouse.

The *in situ* hybridization studies are performed as follows. Anesthetized adult  
30 C57BL/6 mice are perfusion-fixed with 4% paraformaldehyde. Organs to be tested are dissected, trimmed, and immersion-fixed for 4 hours in 4% paraformaldehyde. Processed tissues are embedded in paraffin, and 5 mm sections were cut. Sense or antisense <sup>35</sup>S-

riboprobe are transcribed *in vitro* from mouse PAR2 cDNA subcloned into the *EcoR*I site of pBluescript II SK<sup>-</sup> (Stratagene, San Diego, CA). Hybridization, wash, and development conditions are as reported for mouse PAR1 (Soifer, S.J. et al. (1993) *Am. J. Pathol.* 144:60-69). To carry out Northern analysis, a <sup>32</sup>P-labeled probe for the mouse message is generated by random priming (Prime-It II kit; Stratagene) of PCR-amplified DNA fragments corresponding to mouse cDNA codons representing transmembrane domains 2 to 3. High stringency hybridizations and washes were performed as per the Clonetech protocol for Northern analysis.

#### 10 EXAMPLE 6: ASSAYS FOR PAR4 FUNCTION

Useful receptor fragments or analogs of the invention are those which interact with thrombin and are activated to initiate the cascade of events associated with thrombin receptor interaction. Such an interaction may be detected by an *in vitro* functional assay method (e.g., the phosphoinositide hydrolysis assay, <sup>45</sup>Ca efflux assay, or platelet aggregation assay described herein). This method includes, as components, thrombin and a recombinant protease-activated receptor 4 (or a suitable fragment or analog) configured to permit thrombin binding (e.g., those polypeptides described herein). Thrombin may be obtained from Sigma Chemical Co. (St. Louis, MO) or similar supplier.

Preferably, the protease-activated receptor 4 component is produced by a cell that naturally presents substantially no receptor on its surface, e.g., by engineering such a cell to contain nucleic acid encoding the receptor component in an appropriate expression system. Suitable cells are, e.g., those discussed above with respect to the production of recombinant receptor, such as Rat 1 cells or COS-7 cells.

#### 25 EXAMPLE 7: SCREENING FOR PROTEASE-ACTIVATED RECEPTOR 4 ACTIVATOR ANTAGONISTS AND AGONISTS

##### *Antagonists*

As discussed above, one aspect of the invention features screening for compounds that inhibit the interaction between thrombin (or other PAR4 activating compound) and the protease-activated receptor 4, thereby preventing or reducing the cascade of events that are mediated by that interaction. The elements of the screen are a PAR4 activator (such as thrombin), a candidate antagonist, and recombinant PAR4 (or a suitable receptor fragment or

analog, as outlined above) configured to permit detection of PAR4 activator, antagonist, and PAR4 function. An additional element may be  $^{45}\text{Ca}$ , Fura-2,  $^3\text{H}$ -inositol, or another indicator used to detect downstream signaling (Ishii, K. et al. (1993) *supra*; and Nanevicz, T. et al. (1996) *supra*).

5 Inhibition of thrombin-induced platelet aggregation may also be used as a means of monitoring an antagonist of PAR4 receptor activation. Thrombin is incubated with the candidate inhibitory compound (such as a peptide) for 5 minutes, then the mixture is added to washed platelets and platelet activation is followed as platelet ATP secretion by lumiaggregometry (see, for example, Connolly, A.J. et al. Nature 381:516-519 (1996); and  
10 USPN 5,256,766). Alternately, platelets are incubated with a candidate PAR 4 antagonist for 5 minutes. Thereafter the response to thrombin is measured.

Inclusion of potential antagonists in the screening assay along with thrombin allows for the screening and identification of authentic receptor antagonists as those which decrease thrombin-mediated events, such as platelet aggregation.

15 Appropriate candidate thrombin antagonists include PAR4 fragments, particularly, fragments of the protein predicted to be extracellular and therefore likely to bind thrombin or the tethered ligand; such fragments would preferably include five or more amino acids. Candidate PAR 4 antagonists include thrombin analogs as well as other peptide and non-peptide compounds and anti-PAR4 antibodies.

20

#### *Agonists*

Another aspect of the invention features screening for compounds that act as PAR4 agonists. Activation of the PAR4 with thrombin or an agonist leads to a cascade of events (such as phosphoinositide hydrolysis,  $\text{Ca}^{2+}$  efflux, and platelet aggregation), providing a  
25 convenient means for measuring thrombin or other agonist activity.

The agonist screening assay of the invention utilizes recombinant cells expressing recombinant PAR4 (or a suitable receptor fragment or analog, as outlined herein) configured to permit detection of PAR4 function. Alternatively, a cell such as a leukocyte, a platelet, or a mesenchymal cell that naturally expresses PAR4 may be used. Other elements of the screen  
30 include a detectable downstream substrate of the PAR4 activation, such as radiolabelled phosphoinositide, the hydrolysis of which to a detectable product indicates PAR4 activation by the candidate agonist.

<sup>45</sup>Ca efflux from a cell expressing PAR4 may be used as a means of measuring receptor activation by candidate agonists (Williams, J.A. et al., (1988) PNAS USA 85:4939-4943; Vu, T.-K. H., et al. (1991) Cell 64:1057-1068; and USPN 5,256,766, which references are herein incorporated by reference in their entirety). <sup>45</sup>Ca release by oocytes expressing cRNA encoding PAR4 are assessed as follows. Briefly, intracellular calcium pools are labeled by incubating groups of 30 oocytes in 300 µl calcium-free MBSH containing 50 µCi <sup>45</sup>CaCl<sub>2</sub> (10-40 mCi/mg Ca; Amersham) for 4 hours at room temperature. The labeled oocytes are washed, then incubated in MBSH II without antibiotics for 90 minutes. Groups of 5 oocytes are selected and placed in individual wells in a 24-well tissue culture plate (Falcon 3047) containing 0.5 ml/well MBSH II without antibiotics. This medium is removed and replaced with fresh medium every 10 minutes, the harvested medium is analyzed by scintillation counting to determine <sup>45</sup>Ca released by the oocytes during each 10-minute incubation. The 10-minute incubations are continued until a stable baseline of <sup>45</sup>Ca release per unit time is achieved. Two additional 10-minute collections are obtained, then test medium including agonist is added and agonist-induced <sup>45</sup>Ca release determined.

A voltage clamp assay provides an alternative method of monitoring agonist activity. Agonist-induced inward chloride currents are measured in voltage-clamped oocytes expressing thrombin receptor encoding cRNA essentially as previously described (Julius, D. et al. Science (1988) 241:558-563, herein incorporated by reference in its entirety) except that either the single electrode voltage-clamp technique or a two electrode technique may be employed. Platelet aggregation may also be used as a means of monitoring PAR4 receptor activation (see, for example, Connolly, A.J. et al. Nature 381:516-519 (1996). Human platelets may use both PAR 1 and PAR 4.

An agonist useful in the invention is one which imitates the normal thrombin-mediated signal transduction pathway leading, e.g., to an increase in phosphoinositide hydrolysis. Appropriate candidate agonists include thrombin analogs or PAR4 tethered ligand domains or other agents which mimic the action of thrombin or the PAR 4 tethered ligand domain. Agonists would be useful for aiding discovery of antagonists.

#### 30 EXAMPLE 8: ANTI-PROTEASE-ACTIVATED RECEPTOR 4 ANTIBODIES

Protease-activated receptor 4 (or immunogenic receptor fragments or analogs) may be used to raise antibodies useful in the invention. Receptor fragments preferred for the

production of antibodies are those fragments deduced or shown experimentally to be extracellular.

Antibodies directed to PAR4 peptides are produced as follows. Peptides corresponding to all or part of the PAR4 protein are produced using a peptide synthesizer by standard techniques. The peptides are coupled to KLH with m-maleimide benzoic acid N-hydroxysuccinimide ester. The KLH-peptide is mixed with Freund's adjuvant and injected into animals, e.g. guinea pigs or goats, to produce polyclonal antibodies. Monoclonal antibodies may be prepared using the PAR4 polypeptides described above and standard hybridoma technology (see, e.g., Kohler et al., Nature (1975) 256:495, 1975; Kohler et al., Eur. J. Immunol. (1976) 6:292; Kohler et al., Eur. J. Immunol. (1976) 6:511; Hammerling et al., in Monoclonal Antibodies and T Cell Hybridomas, Elsevier, NY, (1981); and Ausubel et al., supra). Antibodies are purified by peptide antigen affinity chromatography.

Once produced, antibodies are tested for their ability to bind PAR4 by specific binding to the surface of PAR4-transfected cells or by Western blot or immunoprecipitation analysis (such as by the methods described in Ausubel et al., supra).

Antibodies which specifically recognize PAR4 are considered to be likely candidates for useful antagonists; such candidates are further tested for their ability to specifically interfere with the interaction between thrombin and PAR4 (using the functional antagonist assays described herein). Antibodies which antagonize thrombin:PAR4 binding or PAR4 function are considered to be useful antagonists in the invention.

#### EXAMPLE 9:           ACTIVATION OF MOUSE PLATELETS USING THE PAR4                           ACTIVATING PEPTIDE

Blood was collected from mice anesthetized with pentobarbitol by cannulating the inferior vena cava at the level of the renal veins. The blood was mixed with 3.15% citrate (15% of total final volume) and spun at 200g for 7 minutes to obtain platelet rich plasma (PRP). EDTA was added to a final concentration of 10mM, and PGE1 added to a final concentration of 1uM. The PRP was spun at 500g for 10 minutes. The platelet pellet was resuspended in platelet buffer (20mM Tris-HCl pH 7.4, 140 mM NaCl, 2.5 mM KCL, 1.0 mM MgCl<sub>2</sub>, 1 mg.ml glucose, 0.2% BSA) and 1 mM EDTA and 1uM PGE1 in a volume equal to the original PRP volume, and again spun at 500 g for 10 minutes. The second platelet pellet was resuspended in platelet buffer without EDTA or PGE1. The volume was

adjusted such that the OD<sub>500</sub> was 1.0. After a minimum of 30 minutes on ice, 300 ul of platelet suspensions were used to measure aggregation and secretion in a Chronolog aggregometer according to manufacturer's instructions.

5 The PAR4 peptide GYPGKF was added to a final concentration of 500 uM. Secretion and aggregation in response to this peptide were measured over 4-10 minutes (Fig. 13). Since mouse platelets also express PAR3, responses to PAR4 peptide were tested in platelets from PAR3 gene knockout mice to exclude the possibility that PAR3 might be mediating responses to the PAR4 peptide. Note persistent responses to PAR4 peptide were noted in these platelets (Fig. 14).

10

EXAMPLE 10:           ACTIVATION OF HUMAN PLATELETS DESENSITIZED TO  
                                  THE PAR1 ACTIVATING PEPTIDE

To test whether PAR4 might be functionally expressed in human platelets, platelets were prepared by standard techniques and analyzed by lumiaggregometry as above. Because  
15 the mouse peptide GYPGKF was a stronger agonist at human PAR4 than the cognate human peptide GYPGQV, GYPGKF was used. Human platelets were activated by GYPGKF. Human platelets also express PAR1 (like PAR3 in mouse platelets). To exclude the possibility that the PAR4 peptide might be cross-reacting with PAR1, we desensitized human platelets with the PAR1 agonist SFLLRN. Note that these platelets failed to respond to a  
20 second challenge with SFLLRN, but did respond to GYPGKF (Fig. 15); thus, the PAR4 peptide is not acting through PAR1, strengthening the conclusion that PAR4 is functionally expressed in human platelets.

EXAMPLE 11:           THERAPEUTIC USES OF PAR4

25           Particularly suitable therapeutics for the treatment of wound healing, thrombosis, atherosclerosis, restenosis, inflammation, and other thrombin-mediated signaling disorders are the agonists and antagonists described above formulated in an appropriate buffer such as physiological saline. Where it is particularly desirable to mimic a receptor fragment conformation at the membrane interface, the fragment may include a sufficient number of  
30 adjacent transmembrane residues. In this case, the fragment may be associated with an appropriate lipid fraction (e.g., in lipid vesicles or attached to fragments obtained by disrupting a cell membrane). Alternatively, anti-PAR4 antibodies produced as described

above may be used as a therapeutic. Again, the antibodies would be administered in a pharmaceutically-acceptable buffer (e.g., physiological saline). If appropriate, the antibody preparation may be combined with a suitable adjuvant.

Antibodies to PAR 4 are useful antagonists which can be formulated as indicated above. Other therapeutically useful antagonists are peptides derived from PAR4 that bind to and block thrombin and include formulation comprising a pharmaceutically acceptable carrier and one or more of the following:

- a) the isolated sequence  
PAPRGYPGQVCANDSDTLELPD (SEQ ID NO:15);
- 10 b) uncleavable thrombin inhibitor  
PAPRPYPGQVCANDSDTLELPD (SEQ ID NO:16), wherein the  
PAR 4 cleavage site is mutated to block cleavage;
- c) uncleavable thrombin inhibitor  
15 PAP(hR)GYPGQVCANDSDTLELPD (SEQ ID NO:17), wherein the  
PAR 4 cleavage site P1 is mutated to block cleavage, and hR is beta-  
homoarginine (the extra methylene group is in the main chain);
- d) uncleavable thrombin inhibitor  
(dF)PRPYPGQVCANDSDTLELPD (SEQ ID NO:18), wherein the  
good active site binding sequence dFPR is substituted for PNPR and  
20 dF is D-Phenylalanine;
- e) any of a)-d) above where all or part of the  
sequence corresponding to GYPGQVCAN is replaced with spacer  
sequences such as GGG;
- f) variations and combinations of a)-e) which  
25 act as antagonists.

The therapeutic preparation is administered in accordance with the condition to be treated. Ordinarily, it will be administered intravenously, at a dosage, of a duration, and with the appropriate timing to elicit the desired response. Appropriate timing refers to, for  
30 example, time relative to wounding, time intervals between therapeutic administrations, and the like, at which administration of therapeutic preparation elicits the desired response. Alternatively, it may be convenient to administer the therapeutic orally, nasally, or topically,

e.g., as a liquid or a spray. The dosages are determined to be an amount of the therapeutic agent delivered to an animal that substantially reduces or alleviates disease symptoms. Treatment may be repeated as necessary for substantial reduction or alleviation of disease symptoms.

5           PAR4 activator agonists can be used for the treatment of bleeding. Antagonists may be useful in controlling the formation of clots that cause heart attack and stroke, mediating inflammation and the proliferative responses to injury in normal wound healing and a variety of diseases including atherosclerosis, restenosis, pulmonary inflammations (ARDS), glomerulosclerosis, etc.

10           The methods of the invention may be used to screen therapeutic receptor activator agonists and antagonists for their effectiveness in altering thrombin-mediated biological events, such as phosphoinositide hydrolysis or other cell signaling events by the assays described above. Where a non-human mammal is treated or where a therapeutic for a non-human animal is screened, the PAR4 or receptor fragment or analog or the antibody  
15 employed is preferably specific for that species.

#### OTHER EMBODIMENTS

Polypeptides according to the invention include any protease-activated 4 receptors (as described herein). Such receptors may be derived from any source, but are preferably derived  
20 from a vertebrate animal, e.g., a human or mouse. These polypeptides are used, e.g., to screen for antagonists which disrupt, or agonists which mimic, a thrombin:receptor interaction.

Polypeptides of the invention also include any analog or fragment of a PAR4 capable of interacting with thrombin. Such analogs and fragments may also be used to screen for  
25 PAR4 ligand antagonists or agonists. In addition, that subset of receptor fragments or analogs which bind thrombin and are, preferably, soluble (or insoluble and formulated in a lipid vesicle) may be used as antagonists to reduce the *in vivo* concentration of endogenous thrombin, either circulating concentration or local concentration. The efficacy of a receptor analog or fragment is dependent upon its ability to interact with thrombin; such an interaction  
30 may be readily assayed using PAR4 functional assays (e.g., those described herein).

Specific receptor analogs of interest include full-length or partial receptor proteins including an amino acid sequence which differs only by conservative amino acid substitutions,

for example, substitution of one amino acid for another of the same class (e.g., valine for leucine, arginine for lysine, etc.) or by one or more non-conservative amino acid substitutions, deletions, or insertions located at positions of the amino acid sequence which do not destroy the receptors' ability to signal thrombin-mediated events (e.g., as assayed above).

5           Specific receptor fragments of interest include any portion of the PAR4 which is capable of interacting with thrombin, for example, all or part of the extracellular domains predicted from the deduced amino acid sequence. Such fragments may be useful as antagonists (as described above), and are also useful as immunogens for producing antibodies which neutralize the activity of PAR4 *in vivo* (e.g., by interfering with the interaction between  
10 the receptor and thrombin). The area as illustrated in Fig. 8 is most likely to bind thrombin. For the human PAR4 protein, modification of the (R47/G48) cleavage site, e.g. substitution of proline for G48, will render peptides mimicking this site uncleavable. Such peptides will bind thrombin with high affinity.

          Extracellular regions of novel protease-activated receptors may be identified by  
15 comparison with related proteins of similar structure (e.g., other members of the G-protein-coupled receptor family); useful regions are those exhibiting homology to the extracellular domains of well-characterized members of the family.

          Alternatively, from the primary amino acid sequence, the secondary protein structure and, therefore, the extracellular domain regions may be deduced semi-empirically using a  
20 hydrophobicity/hydrophilicity calculation such as the Chou-Fasman method (see, e.g., Chou and Fasman, Ann. Rev. Biochem. (1978) 47:251). Hydrophilic domains, particularly ones surrounded by hydrophobic stretches (e.g., transmembrane domains) present themselves as strong candidates for extracellular domains. Finally, extracellular domains may be identified experimentally using standard enzymatic digest analysis, e.g., tryptic digest analysis.

25           Candidate fragments (e.g., any extracellular fragment) are tested for interaction with thrombin by the assays described herein (e.g., the assay described above). Such fragments are also tested for their ability to antagonize the interaction between thrombin and its endogenous receptor, such as PAR4, using the assays described herein. Analogs of useful receptor fragments (as described above) may also be produced and tested for efficacy as screening  
30 components or antagonists (using the assays described herein); such analogs are also considered to be useful in the invention.

Identification of the receptor(s) that mediate thrombin signaling provides potential targets for the development of drugs that block thrombin's undesirable actions or mimic its desirable activities. Thrombin receptor antagonists may be used for inhibition of platelet-dependent thrombosis in the setting of unstable angina and myocardial infarction or for  
5 blocking thrombin's proinflammatory actions on endothelial cells in the setting of vascular injury. Thrombin receptor agonists may be used to promote hemostasis and fibroblast proliferation at wound sites.

Unmasked tethered ligand domain peptides may provide lead structures for the development of PAR4 agonists or antagonists.

10 The instant invention is shown and described herein in what is considered to be the most practical, and preferred embodiments. It is recognized, however, that departures may be made therefrom, which are within the scope of the invention, and that obvious modifications will occur to one skilled in the art upon reading this disclosure.

15

## CLAIMS

That which is claimed is:

1. Substantially pure DNA encoding a protease-activated receptor 4 (PAR4).  
5
2. The substantially pure DNA of claim 1, wherein the DNA encodes the amino acid sequence of SEQ ID NO:2 or SEQ ID NO:5.
3. The substantially pure DNA of claim 2, wherein the DNA is SEQ ID NO:1,  
10 SEQ ID NO:4, or degenerate variants thereof.
4. Substantially pure DNA having 50% or greater sequence identity to the DNA sequence of claim 1, wherein the DNA selectively hybridizes to sequences complementary to the DNA of claim 1 under stringent conditions.  
15
5. Substantially pure DNA complementary to the DNA of claim 1, wherein the DNA has 50% or greater sequence identity to the DNA of claim 1 and the DNA selectively hybridizes to the DNA of claim 1 under stringent conditions.
- 20 6. An isolated PAR4 polypeptide.
7. The isolated polypeptide of claim 9 having an amino acid sequence of SEQ ID NO:2 or SEQ ID NO:5.
- 25 8. A fragment or analog of a polynucleotide according to Claim 6 or 7.
9. A substantially pure polypeptide having an amino acid sequence wherein said polypeptide is activated by thrombin; and said polypeptide mediates phosphoinositide hydrolysis in a cell expressing said polypeptide on its surface.  
30
10. A substantially pure PAR4 activating peptide.

11. An antibody which selectively binds to the polypeptide of claim 9.
12. A vector comprising the DNA of claims 1-5.
- 5 13. A cell comprising the vector of claim 12.
14. An assay device, comprising:  
a support surface;  
and a cell of claim 13 or membranes derived therefrom.
- 10 15. A therapeutic composition, comprising:  
a PAR4 ligand agonist and a physiologically-acceptable carrier.
- 15 16. A therapeutic composition, comprising:  
a PAR4 ligand antagonist; and a physiologically-acceptable carrier.

1/18

10 20 30 40 50 60  
ATGTGCTGGCCGCTGCTGTATCCTTTGGTGCTGGGGCTCAGCATCAGCCTGGCAGAGGGC

70 80 90 100 110 120  
ATCCAGACCCCCAGCATCTACGATGATGTAGAGAGTACCAGGGGAAGCCATGAAGGCCCT

130 140 150 160 170 180  
CTGGGTCCCACAGTAGAACTCAAGGAGCCGAAGTCCTCAGACAAGCCTAATCCACGAGGC

190 200 210 220 230 240  
TACCCGGGCAAATTCTGTGCCAACGACAGTGACACGCTGGAGCTCCCGGCCAGCTCTCAA

250 260 270 280 290 300  
GCACTGCTGCTGGGGTGGGTATCCACAAGGCTGGTACCTGCCCTCTATGGGCTTGTGGTG

310 320 330 340 350 360  
GCTGTGGGGCTGCCTGCCAATGGGCTGGCGCTGTGGGTGCTGGCCACAAGGGTGCCACGC

370 380 390 400 410 420  
CTGCCATCCACCATTCTGCTCACGAACCTGGCAGTGGCTGATTGCTGTGGCCCTGGTG

430 440 450 460 470 480  
CCGCCACCACGACTGGCTTACCACTTGCCTGGCCAGCGCTGGCCATTTGGTGAGGCTGCC

490 500 510 520 530 540  
TGCCGGGTGGCCACAGCTGCCCTCTATGGCCACATGTATGGTTCAGTGTGCTGCTGGCT

550 560 570 580 590 600  
GCAGTCAGCTTGGACAGATACCTGGCCCTGGTGCATCCTTTGCGGGCCCGTGCCTGCGT

610 620 630 640 650 660  
GGTCAACGCCTCACTACTGGACTCTGTTGGTGGCCTGGCTCTCTGCAGCCACCCTGGCC

670 680 690 700 710 720  
TTGCCCTCTCACTCTGCATCGGCAGACCTTCCGATTAGCTGGCTCCGATCGCATGCTGTGT

730 740 750 760 770 780  
CATGATGCGCTGCCCCTGACTGAGCAGACCTCCCACTGGAGACCGGCCTTCATCTGCCTG

790 800 810 820 830 840  
GCTGTCTGGGCTGCTTTCGTGCCACTGCTGGCCATGGGCCTGTGCTATGGAGCCACCCTT

850 860 870 880 890 900  
CGTGCCTGGCGGCAATGGCCAGCGCTACAGCCATGCACTCAGACTGACAGCCCTGGTA

910 920 930 940 950 960  
CTGTTCTCGGCAGTGGCCTCTTTACACCTAGCAATGTGCTGCTGGTGCTGCACTATTCA

970 980 990 1000 1010 1020  
AACCCGAGCCCTGAGGCCTGGGGCAATCTCTATGGAGCCTATGTGCCAGCCTGGCACTC

1030 1040 1050 1060 1070 1080  
AGCACCCCTCAACAGCTGCGTAGACCCTTTCATCTACTACTATGTGTCCCATGAGTTCAGG

1090 1100 1110 1120 1130 1140  
GAGAAGGTACGCGCTATGTTGTGTGCCAGCCGGAGGCCAGCAGCTCCTCTCAGGCCCTCC

Fig. 1A

2/18

1150 1160 1170 1180 1190 1200  
AGGGAGGCTGGAAGCCGAGGGACTGCCATTTGCTCCTCTACACTTCTGTGACTGGTAGCT

1210 1220 1230 1240 1250 1260  
GAGGTGGGAAGGGGGCATTCTGGCTTGACTGGGTCTCCCCTTAAACTACATCCCTCTTGA

1270 1280 1290 1300 1310 1320  
ACCCTCAGGACATGACCTTATTTGGATATGCAGTTGGTGCGACCTTCATTAGTGGAGCTG

1330 1340 1350 1360  
AGGTCCACTGGAAATGCTTTTGTAAAAGGTCTGGGTACTAT

Fig. 1B

3/18

10 20 30 40 50 60  
MCWPLLYPLVLGLSISLAEGIQTPSIYDDVESTGRGSHEGPLGPTVELKEPKSSDKPNPRG

70 80 90 100 110 120  
YFGKFCANDSDTLELPASSQALLLGWVSTRVLPALYGLVVAVGLPANGALWVLA TRVPR

130 140 150 160 170 180  
LPSTILLTNLAVADSLALVPPRLAYHLRGQRWPFGEAACRVATAALYGHMYGSVLLLA

190 200 210 220 230 240  
AVSLDRYLALVHPLRARALRGQLTTGLCLVAWLSAATLALPLTLHRQTFRLAGSDRMLC

250 260 270 280 290 300  
HDALPLTEQTSWHPAFICLAVLGCFVPLLAMGLCYGATLRALAANGQRYSHALRLTALV

310 320 330 340 350 360  
LFSAVASFTPSNVLLVLHYSNPSPEAWGNLYGAYVPSLALSTLNSCVDPFIYYVSHFR

370 380 390  
EKVRAMLCRQPEASSSSQASREAGSRGTAICSSTLL\*

Fig. 2

4/18

10 20 30 40 50 60  
GATATGTGCTGGCCGCTGCTGTATCCTTTGGTGCTGGGGCTCANCATCANCCTGGCANAG

70 80 90 100 110 120  
GGCNGGNACGANNCCCAGCATCTACGATGATGTAGANAGTACCAGGGGAANCCATGGTG

130 140 150 160 170 180  
ANTGACTGTNTCCCTTAAAGGGGTGAATCAGAAATGGAGCTANTGNTGAGCAGGNGNACA

190 200 210 220 230 240  
GNNTTTANGTCCCTAAAANCCATGCCTTTTGGGANTGGGTTGTATCCTTCCNTTAANTGA

250 260 270 280 290 300  
NTNNTGGANTGGGGACANTGAGGCACCCACAATGCCTAAGACTTTCAAGGATATTCTCCT

310 320 330 340 350 360  
TCATCNTGTATCCCTAAAGGCAGGGNAGAGCAGTGGNTGACTGATGTCCCCCTCTCTCCCA

370 380 390 400 410 420  
CAGAAGGCCCTNTGGGTCCACAGTAGAACTCAAGGAGCCGAAGTCCTCAGACAAGCCTA

430 440 450 460 470 480  
ATCCCCGAGGGTACCCGGGCAAATTCTGTGCCAANGACAGTGACACGCTGGAGCTCCCGG

490 500 510 520  
CCAGCTCTCAAGCACTGCNTGNTGGGGTGGGTCCCCACGANTTT

Fig. 3

5/18

10            20            30            40            50            60  
 TC-ATGTGGGGGCGACTGCTCCTGTGGCCCCGTGGTGTGGGGTTCAGCCTGTCTGGCGGG

70            80            90            100           110           120  
 ACCCAGACCCCCAGCGTCTACGACGAGAGCGGGAGCACCGGAGGTGGTGATGACAGCAGC

130           140           150           160           170           180  
 CCCTCAATCCTGCCTGCCCCCGCGGTACCCAGGCCAAGTCTGTGCCAATGACAGTGAC

190           200           210           220           230           240  
 ACCCTGGAGCTCCCGGACAGCTCACGGGCACTGCTTCTGGGCTGGGTGCCCACCAGGCTG

250           260           270           280           290           300  
 GTGCCCCGCCCTCTATGGGCTGGTCTGGTGGTGGGGCTGCCGGCCAATGGGCTGGCGCTG

310           320           330           340           350           360  
 TGGGTGCTGGCCACGCAGGCACCTCGGCTGCCCTCCACCATGCTGCTGATGAACCTCGCG

370           380           390           400           410           420  
 ACTGCTGACCTCCTGCTGGCCCTGGCGCTGCCCCCGGGATCGCCTACCACCTGCGTGGC

430           440           450           460           470           480  
 CAGCGCTGGCCCTTCGGGGAGGCCGCTGCCGCTGGCCACGGCCGCACTCTATGGTCAC

490           500           510           520           530           540  
 ATGTATGGCTCAGTGCTGCTGCTGGCCGCCGTCAGCCTGGATCGCTACCTGGCCCTGGTG

550           560           570           580           590           600  
 CACCCGCTGCGGGCCCGCGCCCTGCGTGGCCGGCGCTGGCCCTTGGACTCTGCATGGCT

610           620           630           640           650           660  
 GCTTGGCTCATGGCGGCCGCCCTGGCACTGCCCTGACACTGCAGCGGCAGACCTTCCGG

670           680           690           700           710           720  
 CTGGCGCGCTCCGATCGCGTGCTCTGCCATGACGCGCTGCCCTGGACGCACAGGCCTCC

730           740           750           760           770           780  
 CACTGGCAACCGGCCCTTACCTGCCTGGCGCTGTTGGGCTGTTTCTGCCCTGCTGGCC

790           800           810           820           830           840  
 ATGCTGCTGTGCTACGGGGCCACCCTGCACACGCTGGCGGCCAGCGGCCGGCGCTACGGC

850           860           870           880           890           900  
 CACGCGCTGAGGCTGACCGCAGTGGTGTGGCCTCCGCCGTGGCCTTCTTCGTGCCCAGC

910           920           930           940           950           960  
 AACCTGCTGCTGCTGCTGCATTACTCGGACCCGAGCCCCAGCGCCTGGGGCAACCTCTAT

970           980           990           1000           1010           1020  
 GGTGCCTACGTGCCCAGCCTGGCGCTGAGCACCCCTCAACAGCTGCGTGGATCCCTTCATC

1030           1040           1050           1060           1070           1080  
 TACTACTACGTGTCGGCCGAGTTCAGGGACAAGGTGCGGGCAGGGCTCTTCCAACGGTGC

1090           1100           1110           1120           1130           1140  
 CCGGGGGACACCGTGGCCTCCAAGGCCTCTGCGGAAGGGGGCAGCCGGGGCATGGGCACC

Fig. 4A

6/18

1150 1160 1170 1180 1190 1200  
CACTCCTCTTTGCTCCAGTGACACAAAGTGGGGAAGGCTGTACTGGGTCGAACAGGGTCC

1210 1220 1230 1240 1250 1260  
CTTCCCCACTTCACGTCCCTTCCTGGGACCTCAGAAATGTGACCTTATTTGAAATAGGGT

1270 1280 1290 1300 1310 1320  
TGTTACAACCTGTCACTAGCAGAGGTCACCTTTGGAGAAGGGTGGGCCTTACATCCAGTGTG

1330 1340 1350 1360 1370 1380  
GGTGGTGTCCATATAAGATAAGGAGAGGCCAGGCCTGGTGGCTCACGCCTGTAATCCCAG

1390 1400 1410 1420 1430 1440  
CACTTTAAGAGGCCAAGGCGGATGGATCACTTGAGCCCAGGAGTTCAACACCAGCCTGAG

1450 1460 1470 1480 1490 1500  
CAACATGGTAAAACCCCATCTCTACCAAAAATACAAAATTAGCTGGGCTTGGTGGCTGG

1510 1520 1530  
CGCCTGTAATCCCAGCTACTCANGAGACTGAGGCA

Fig. 4B

7/18

10 20 30 40 50 60  
MWGRLLWPLVLGFSLSGGTQTPSVYDESGSTGGGDDSTPSILPAPRGYPGQVCANDSDT

70 80 90 100 110 120  
LELPDSSRALLGWVPTLVPALYGLVVLVGLPANGGLALWVLATQAPRLPSTMLLMNLAT

130 140 150 160 170 180  
ADLLLALALPPRIAYHLRGQRWPFGEAACRLATAALYGHMYGSVLLLAASLDRYLALVH

190 200 210 220 230 240  
PLRARALRGRRALGLCMAAWLMAAALALPLTLQRQTFRLARSDRVLCHDALPLDAQASH

250 260 270 280 290 300  
WQPAFTCLALLGCFLPLLAMLLCYGATLHTLAASGRRYGHALRLTAVVLASAVAFFVPSN

310 320 330 340 350 360  
LLLLLHYSDPSPSAWGNLYGAYVPSLALSTLNSCVDPFIYYVSAEFRDKVRAGLFQRSP

370 380  
GDTVASKASAEGGSRGMGTHSSLLQ\*

Fig. 5

8/18

		10	20	30																											
mPAR4	AA	M	C	W	P	L	L	Y	P	L	V	L	G	L	S	I	S	L	A	E	G	-	I	Q	T	P	S	I	Y	D	D
mPAR3	AA	M	K	I	L	I	L	V	-	-	A	A	G	L	L	F	L	P	V	T	V	-	C	Q	S	G	I	N	V	S	D
mPAR2	AA	M	R	S	L	S	L	A	W	L	L	G	G	I	T	L	L	A	A	S	V	S	C	S	R	T	E	N	L	A	P
mPAR1	AA	M	G	P	R	R	L	L	I	V	A	L	G	L	S	L	C	G	P	L	L	S	S	R	V	P	M	S	Q	P	E
		40	50	60																											
mPAR4	AA	V	E	S	T	R	G	S	H	E	G	P	L	G	P	T	V	E	L	K	E	P	K	S	S	D	K	P	N	P	R
mPAR3	AA	-	-	N	S	A	K	P	T	L	T	I	K	S	F	N	G	G	P	-	Q	N	T	F	E	E	F	P	L	S	D
mPAR2	AA	G	R	N	S	K	G	R	S	L	I	G	R	L	E	T	Q	P	-	P	I	T	G	K	G	V	P	V	E	P	
mPAR1	AA	S	E	R	T	D	A	T	V	N	P	R	S	F	F	L	R	N	P	S	E	N	T	F	E	L	V	P	L	G	D
		70	80	90																											
mPAR4	AA	-	-	-	-	-	-	-	-	-	G	Y	P	G	K	F	C	A	N	D	S	D	-	-	-	-	-	-	-	-	T
mPAR3	AA	I	E	-	-	-	-	-	-	-	G	W	T	G	A	T	T	T	I	K	A	E	C	P	E	D	S	I	S	T	
mPAR2	AA	-	-	-	-	-	-	-	-	-	G	F	S	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
mPAR1	AA	E	E	E	E	E	K	N	E	S	V	L	L	E	G	R	A	V	Y	L	N	I	S	L	P	P	H	T	P	P	P
		100	110	120																											
mPAR4	AA	L	E	L	P	A	S	S	Q	A	L	L	G	W	V	S	T	R	L	V	P	A	L	Y	G	L	V	V	A	V	
mPAR3	AA	L	H	V	N	N	A	T	I	G	Y	L	R	S	S	L	S	T	Q	V	I	P	A	I	Y	I	L	L	F	V	V
mPAR2	AA	-	-	I	D	E	F	S	A	S	I	L	T	G	K	L	T	T	V	F	L	P	V	V	Y	I	I	V	F	V	I
mPAR1	AA	P	F	I	S	E	D	A	S	G	Y	L	T	S	P	W	L	T	L	F	M	P	S	V	Y	T	I	V	F	I	V
		130	140	150																											
mPAR4	AA	G	L	P	A	N	G	L	A	L	W	V	L	A	T	R	V	P	R	L	P	-	S	T	T	L	L	T	N	L	A
mPAR3	AA	G	V	P	S	N	I	V	T	L	W	K	L	S	L	R	T	K	S	I	S	-	L	V	I	F	H	T	N	L	A
mPAR2	AA	G	L	P	S	N	G	M	A	L	W	I	F	L	F	R	T	K	K	K	H	P	A	V	I	Y	M	A	N	L	A
mPAR1	AA	S	L	P	L	N	V	L	A	I	A	V	F	V	L	R	M	K	V	K	K	P	A	V	V	Y	M	L	H	L	A
		160	170	180																											
mPAR4	AA	V	A	D	S	L	L	A	L	V	P	P	P	R	L	A	Y	H	L	R	G	Q	R	W	P	F	G	E	A	A	C
mPAR3	AA	I	A	D	L	L	F	C	V	T	L	P	F	K	I	A	Y	H	L	N	G	N	N	W	V	F	G	E	V	M	C
mPAR2	AA	L	A	D	L	L	S	V	I	W	F	P	L	K	I	S	Y	H	L	H	G	N	N	W	V	Y	G	E	A	L	C
mPAR1	AA	M	A	D	V	L	F	V	S	V	L	P	F	K	I	S	Y	Y	F	S	G	T	D	W	Q	F	G	S	G	M	C
		190	200	210																											
mPAR4	AA	R	V	A	T	A	A	L	Y	G	H	M	Y	G	S	V	L	L	L	A	A	V	S	L	D	R	Y	L	A	L	V
mPAR3	AA	R	I	T	T	V	V	F	Y	G	N	M	Y	C	A	I	L	I	L	T	C	M	G	I	N	R	Y	L	A	T	A
mPAR2	AA	K	V	L	I	G	F	F	Y	G	N	M	Y	C	S	I	L	F	M	T	C	L	S	V	Q	R	Y	W	V	I	V
mPAR1	AA	R	F	A	T	A	A	F	Y	G	N	M	Y	A	S	I	M	L	M	T	V	I	S	I	D	R	F	L	A	V	V

**FIG. 6A**

**SUBSTITUTE SHEET (Rule 26)**

220 9/18

230

240

mPAR4 AA H P L R A R A L R G Q R L T T G L C L V A W L S A A T L A L  
 mPAR3 AA H P F T Y Q K L P K R S F S L L M C G I V W V M V F L Y M L  
 mPAR2 AA N P M G H P R K - K A N I A V G V S L A I W L L I F L V T I  
 mPAR1 AA Y P I Q S L S W R T L G R A N F T C V V I W V M A I M G V V

250

260

270

mPAR4 AA P L T L H R Q T F R L A G S D R M L C H D A L P - L T E Q T  
 mPAR3 AA P F V I L K Q E Y H L V H S E I T T C H D V V D A C E S P S  
 mPAR2 AA P L Y V M K Q T I Y I P A L N I T T C H D V L P - E E V L V  
 mPAR1 AA P L L L K E Q T T R V P G L N I T T C H D V L S - E N L M Q

280

290

300

mPAR4 AA S H W R P A F I C L A V L G C F V P L L A M G L C Y G A T L  
 mPAR3 AA S F R F Y Y F V S L A F F G F L I P F V I I F C Y T T L I  
 mPAR2 AA G D M F N Y F L S L A I G V F L F P A L L T A S A Y V L M I  
 mPAR1 AA G F Y S Y Y F S A F S A I F F L V P L I V S T V C Y T S I I

310

320

330

mPAR4 AA R A L A A N - - - - - G Q R Y S H A L R L T A L V L F S A  
 mPAR3 AA H K L K S - - - - - K D R I W L G Y I K A V L L I L V I F  
 mPAR2 AA K T L R S S A M D E H S E K K R Q R A I R L I I T V L A M Y  
 mPAR1 AA R C L S S S A V A N - - R S K K S R A L F L S A A V F C I F

340

350

360

mPAR4 AA V A S F T P S N V L L V L H Y S N P S P E - A W G N L Y G A  
 mPAR3 AA T I C F A P T N I I L V I H H A N Y Y Y H - N T D S L Y F M  
 mPAR2 AA F I C F A P S N L L L V V H Y F L I K T Q - R Q S H V Y A L  
 mPAR1 AA I V C F G P T N V L L I V H Y L F L S D S P G T E A A Y F A

370

380

390

mPAR4 AA Y V P S L A L S T L N S C V D P F I Y Y Y V S H E F R E K V  
 mPAR3 AA Y L I A L C L G S L N S C L D P F L Y F V M S K - - - - -  
 mPAR2 AA Y L V A L C L S T L N S C I D P F V Y Y F V S K D F R D H A  
 mPAR1 AA Y L L C V C V S S V S C C I D P L I Y Y Y A S S E C Q R H L

400

410

420

mPAR4 AA R - A M L C R Q P E A S S S S Q A S R E A G S R G T A I C S  
 mPAR3 AA - - - - - V V D Q L N P  
 mPAR2 AA R N A L L C R S V R T V N R M Q I S L S S N K F S R K S G S  
 mPAR1 AA Y S I L C C K E S S D P N S C N S T G Q L M P S K M D T C S

430

mPAR4 AA S T L L  
 mPAR3 AA  
 mPAR2 AA Y S S S S T S V K T S Y  
 mPAR1 AA S H L N N S I Y K K L L A

FIG. 6B



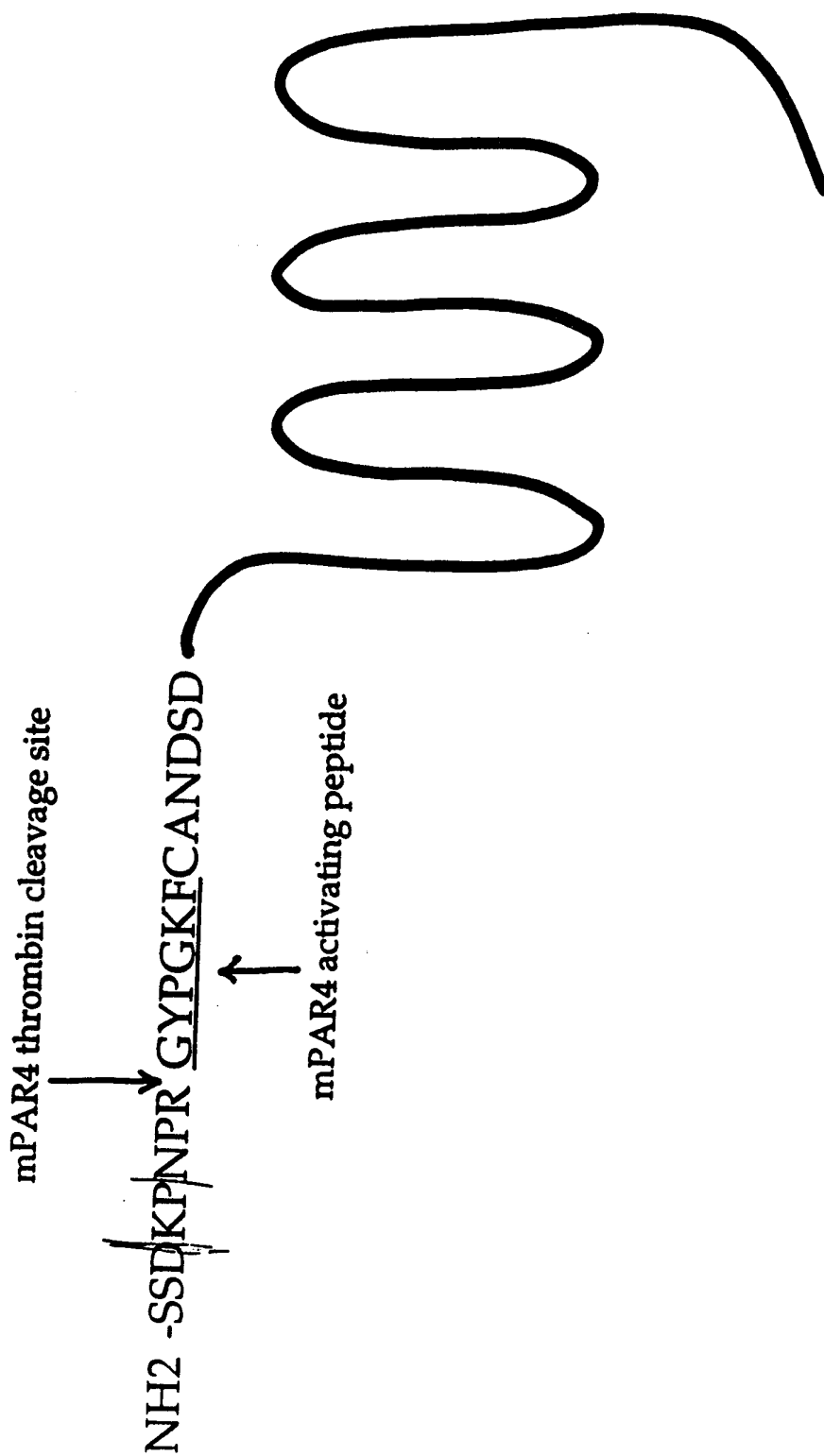


Fig. 8

12/18

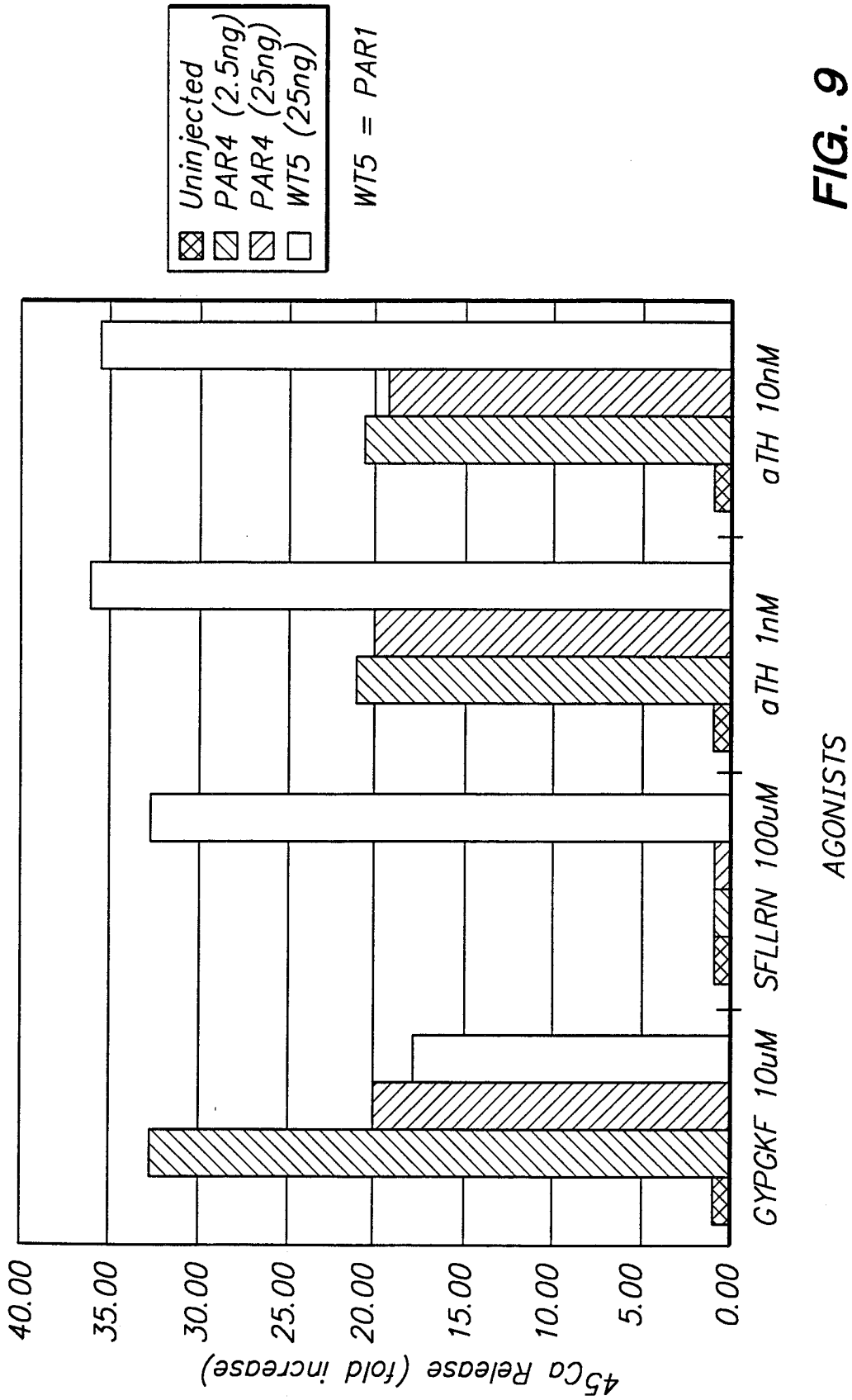


FIG. 9

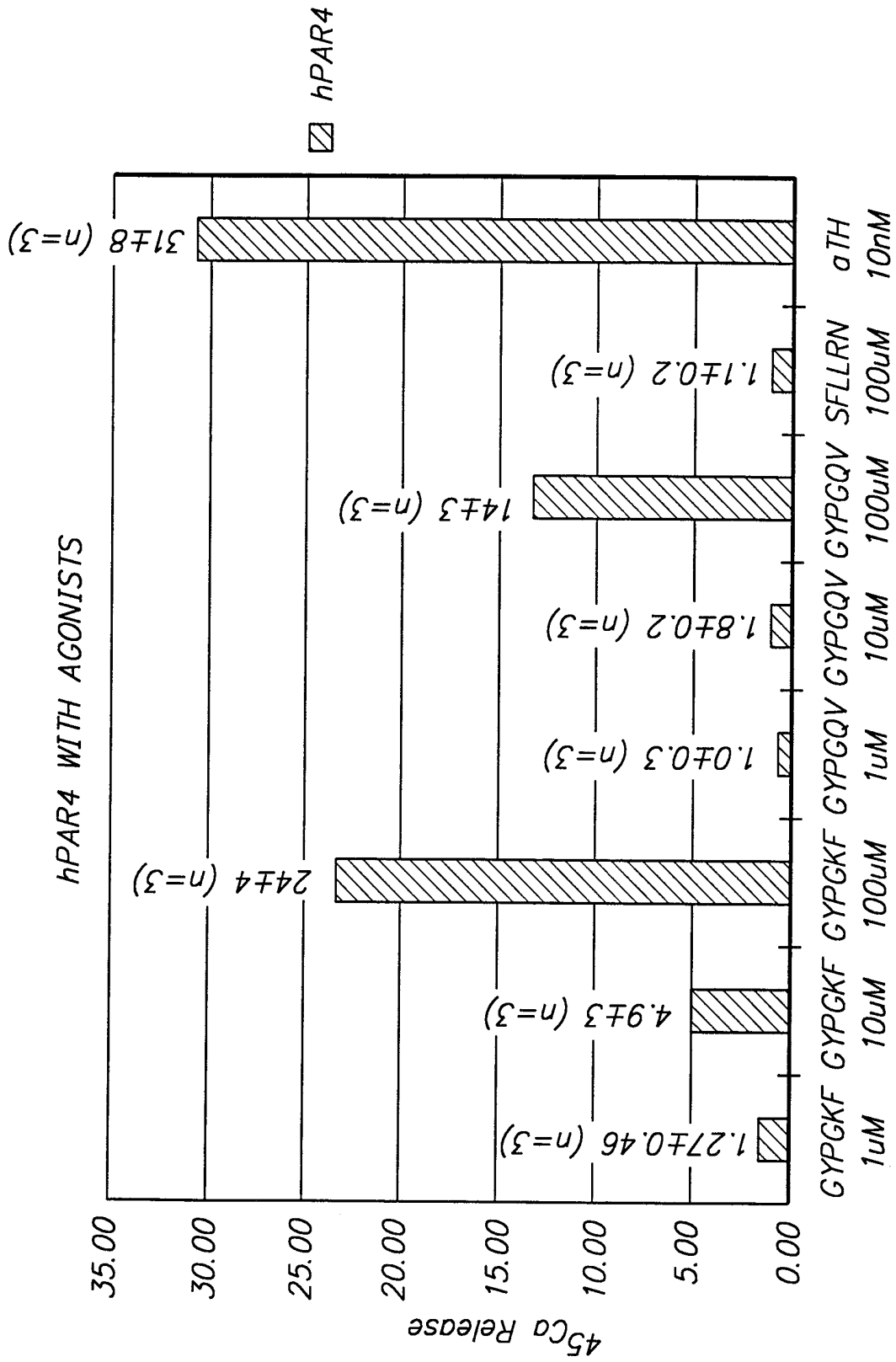


FIG. 10

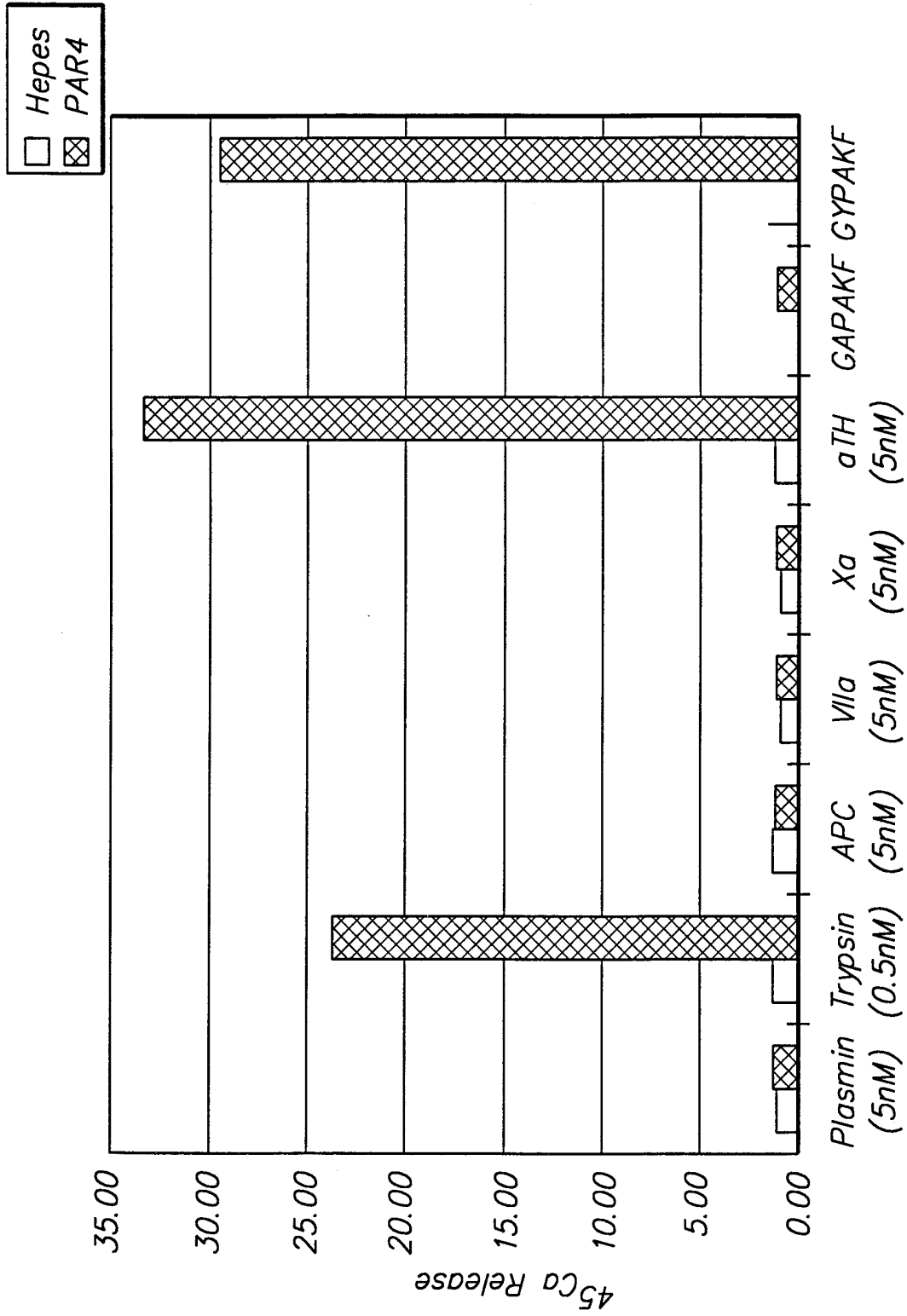
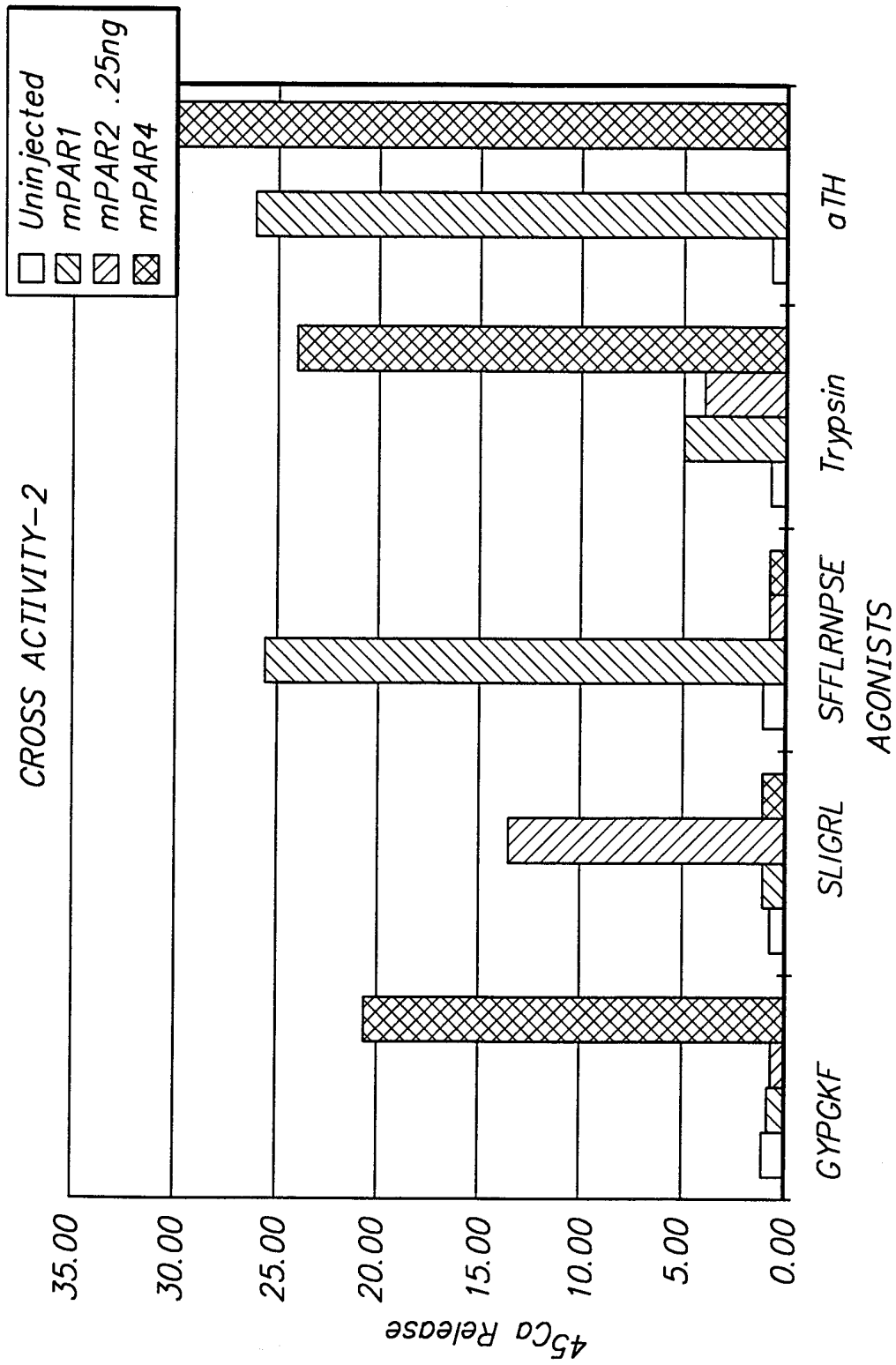


FIG. 11

AGONISTS



**FIG. 12**

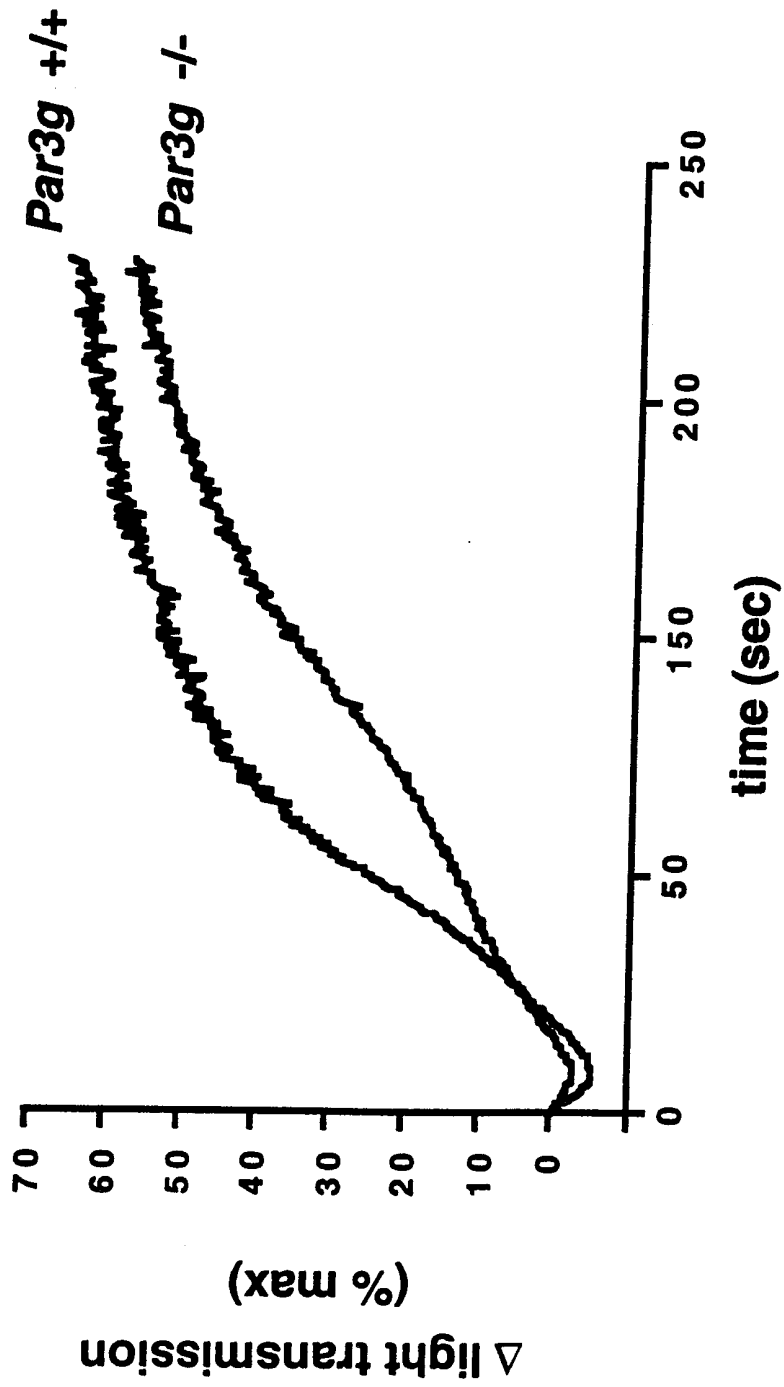


Fig. 13

17/18

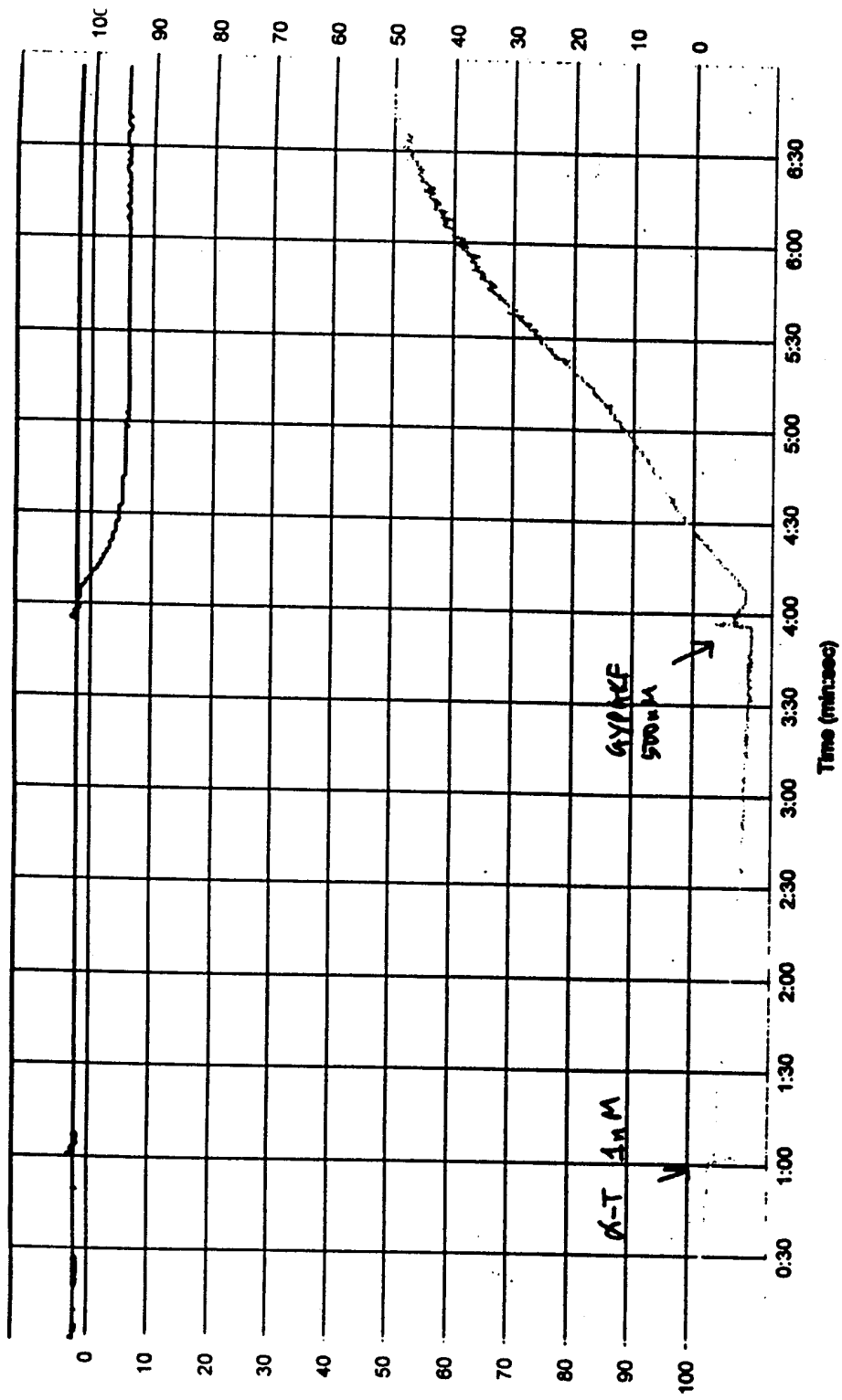


Fig. 14

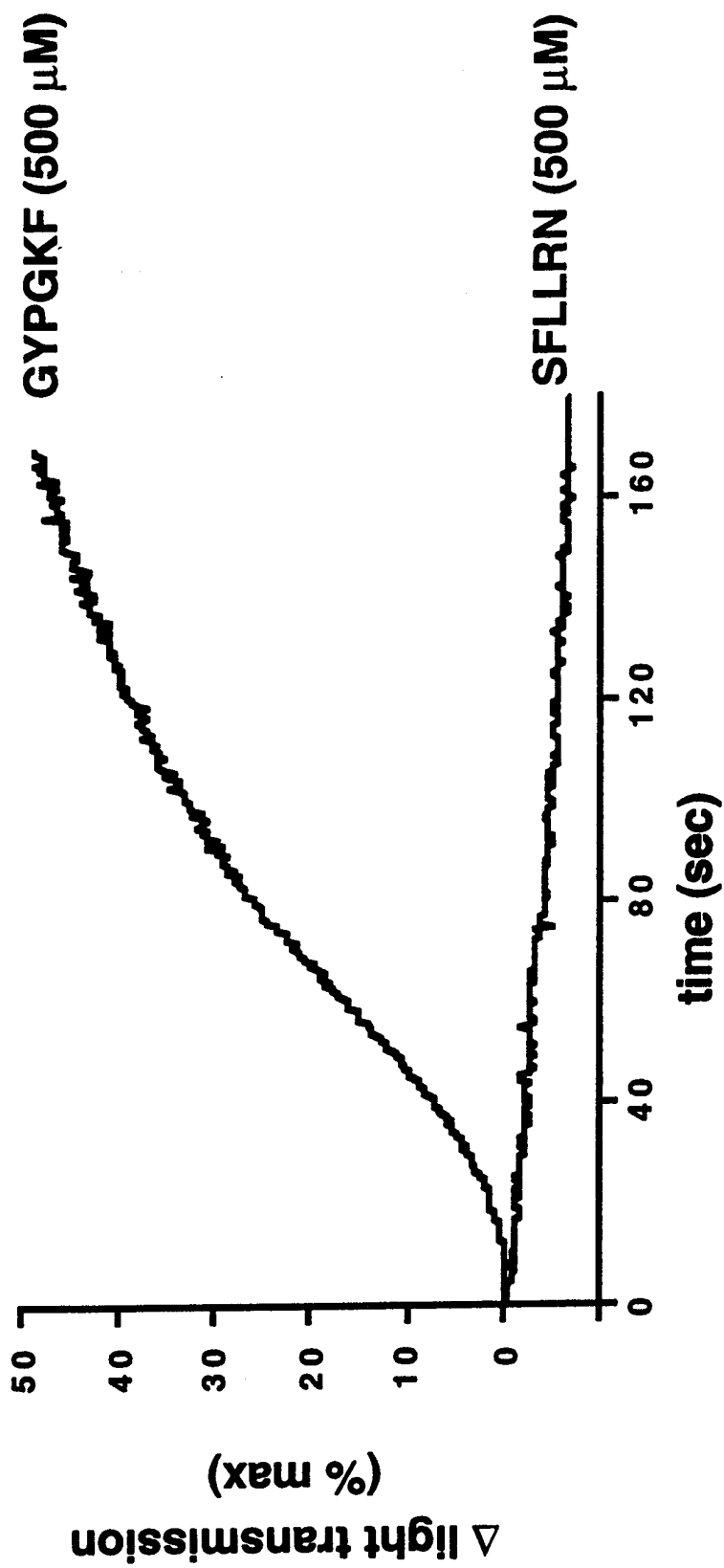


Fig. 15

## SEQUENCE LISTING

## (1) GENERAL INFORMATION

- (i) APPLICANT: The Regents of the University of California
- (ii) TITLE OF THE INVENTION: Protease Activated Receptor  
4 and Uses Thereof
- (iii) NUMBER OF SEQUENCES: 30
- (iv) CORRESPONDENCE ADDRESS:
  - (A) ADDRESSEE: Bozicevic, Field & Francis LLP
  - (B) STREET: 285 Hamilton Avenue, Suite 200
  - (C) CITY: Palo Alto
  - (D) STATE: CA
  - (E) COUNTRY: USA
  - (F) ZIP: 94301
- (v) COMPUTER READABLE FORM:
  - (A) MEDIUM TYPE: Diskette
  - (B) COMPUTER: IBM Compatible
  - (C) OPERATING SYSTEM: DOS
  - (D) SOFTWARE: FastSEQ for Windows Version 2.0
- (vi) CURRENT APPLICATION DATA:
  - (A) APPLICATION NUMBER: not yet assigned
  - (B) FILING DATE: herewith
  - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: USSN 09/032,397
  - (B) FILING DATE: 27 February 1998 (27/02/98)
- (viii) ATTORNEY/AGENT INFORMATION:
  - (A) NAME: DeVore, Dianna L
  - (B) REGISTRATION NUMBER: P-42,484
  - (C) REFERENCE/DOCKET NUMBER: 06510/093W01
- (ix) TELECOMMUNICATION INFORMATION:

- (A) TELEPHONE: 650-327-3400
- (B) TELEFAX: 650 327-3231
- (C) TELEX:

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1360 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATGTGCTGGC	CGCTGCTGTA	TCCTTTGGTG	CTGGGGCTCA	GCATCAGCCG	GCAGAGGGCA	60
TCCAGACCCC	CAGCATCTAC	GATGATGTAG	AGAGTACCAG	GGGAAGCCAT	GAAGGCCCTC	120
TGGGTCCCAC	AGTAGAACTC	AAGGAGCCGA	AGTCCTCAGA	CAAGCCTAAT	CCACGAGGCT	180
ACCCGGGCAA	ATTCTGTGCC	AACGACAGTG	ACACGCTGGA	GCTCCCGGCC	AGCTCTCAAG	240
CACTGCTGCT	GGGGTGGGTA	TCCACAAGGC	TGGTACCTGC	CCTCTATGGG	CTTGTGGTGG	300
CTGTGGGGCT	GCCTGCCAAT	GGGCTGGCGC	TGTGGGTGCT	GGCCACAAGG	GTGCCACGCC	360
TGCCATCCAC	CATTCTGCTC	ACGAACCTGG	CAGTGGCTGA	TTCGCTGTTG	GCCCTGGTGC	420
CGCCACCACG	ACTGGCTTAC	CACTTGCGTG	GCCAGCGCTG	GCCATTTGGT	GAGGCTGCCT	480
GCCGGGTGGC	CACAGCTGCC	CTCTATGGCC	ACATGTATGG	TTCAGTGTG	CTGCTGGCTG	540
CAGTCAGCTT	GGACAGATAC	CTGGCCCTGG	TGCATCCTTT	GCGGGCCCGT	GCACTGCGTG	600
GTCAACGCCT	CACTACTGGA	CTCTGTTTGG	TGGCCTGGCT	CTCTGCAGCC	ACCCTGGCCT	660
TGCCCTCTCAC	TCTGCATCGG	CAGACCTTCC	GATTAGCTGG	CTCCGATCGC	ATGCTGTGTC	720
ATGATGCGCT	GCCCCTGACT	GAGCAGACCT	CCCACTGGAG	ACCGGCCTTC	ATCTGCCTGG	780
CTGTCTGGG	CTGCTTCGTG	CCACTGCTGG	CCATGGGCCT	GTGCTATGGA	GCCACCCTTC	840
GTGCACTGGC	GGCCAATGGC	CAGCGCTACA	GCCATGCACT	CAGACTGACA	GCCCTGGTAC	900
TGTTCTCGGC	AGTGGCCTCT	TTCACACCTA	GCAATGTGCT	GCTGGTGCTG	CACTATTCAA	960
ACCCGAGCCC	TGAGGCCTGG	GGCAATCTCT	ATGGAGCCTA	TGTGCCCAGC	CTGGCACTCA	1020
GCACCCTCAA	CAGCTGCGTA	GACCCTTTCA	TCTACTACTA	TGTGTCCCAT	GAGTTCAGGG	1080
AGAAGGTACG	CGCTATGTTG	TGTCGCCAGC	CGGAGGCCAG	CAGCTCCTCT	CAGGCCTCCA	1140
GGGAGGCTGG	AAGCCGAGGG	ACTGCCATTT	GCTCCTCTAC	ACTTCTGTGA	CTGGTAGCTG	1200
AGGTGGGAAG	GGGGCATTCT	GGCTTGACTG	GGTCTCCCCT	TAAACTACAT	CCCTCTTGAA	1260
CCCTCAGGAC	ATGACCTTAT	TTGGATATGC	AGTTGGTGCG	ACCTTCATTA	GTGGAGCTGA	1320
GGTCCACTGG	AAATGCTTTT	GTAAGAGGTC	TGGGTACTAT			1360

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 397 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```

Met Cys Trp Pro Leu Leu Tyr Pro Leu Val Leu Gly Leu Ser Ile Ser
 1           5           10           15
Leu Ala Glu Gly Ile Gln Thr Pro Ser Ile Tyr Asp Asp Val Glu Ser
           20           25           30
Thr Arg Gly Ser His Glu Gly Pro Leu Gly Pro Thr Val Glu Leu Lys
           35           40           45
Glu Pro Lys Ser Ser Asp Lys Pro Asn Pro Arg Gly Tyr Pro Gly Lys
           50           55           60
Phe Cys Ala Asn Asp Ser Asp Thr Leu Glu Leu Pro Ala Ser Ser Gln
65           70           75           80
Ala Leu Leu Leu Gly Trp Val Ser Thr Arg Leu Val Pro Ala Leu Tyr
           85           90           95
Gly Leu Val Val Ala Val Gly Leu Pro Ala Asn Gly Leu Ala Leu Trp
           100          105          110
Val Leu Ala Thr Arg Val Pro Arg Leu Pro Ser Thr Ile Leu Leu Thr
           115          120          125
Asn Leu Ala Val Ala Asp Ser Leu Leu Ala Leu Val Pro Pro Pro Arg
           130          135          140
Leu Ala Tyr His Leu Arg Gly Gln Arg Trp Pro Phe Gly Glu Ala Ala
145          150          155          160
Cys Arg Val Ala Thr Ala Ala Leu Tyr Gly His Met Tyr Gly Ser Val
           165          170          175
Leu Leu Leu Ala Ala Val Ser Leu Asp Arg Tyr Leu Ala Leu Val His
           180          185          190
Pro Leu Arg Ala Arg Ala Leu Arg Gly Gln Arg Leu Thr Thr Gly Leu
           195          200          205
Cys Leu Val Ala Trp Leu Ser Ala Ala Thr Leu Ala Leu Pro Leu Thr
           210          215          220
Leu His Arg Gln Thr Phe Arg Leu Ala Gly Ser Asp Arg Met Leu Cys
225          230          235          240
His Asp Ala Leu Pro Leu Thr Glu Gln Thr Ser His Trp Arg Pro Ala
           245          250          255
Phe Ile Cys Leu Ala Val Leu Gly Cys Phe Val Pro Leu Leu Ala Met
           260          265          270

```

Gly Leu Cys Tyr Gly Ala Thr Leu Arg Ala Leu Ala Ala Asn Gly Gln  
 275 280 285  
 Arg Tyr Ser His Ala Leu Arg Leu Thr Ala Leu Val Leu Phe Ser Ala  
 290 295 300  
 Val Ala Ser Phe Thr Pro Ser Asn Val Leu Leu Val Leu His Tyr Ser  
 305 310 315 320  
 Asn Pro Ser Pro Glu Ala Trp Gly Asn Leu Tyr Gly Ala Tyr Val Pro  
 325 330 335  
 Ser Leu Ala Leu Ser Thr Leu Asn Ser Cys Val Asp Pro Phe Ile Tyr  
 340 345 350  
 Tyr Tyr Val Ser His Glu Phe Arg Glu Lys Val Arg Ala Met Leu Cys  
 355 360 365  
 Arg Gln Pro Glu Ala Ser Ser Ser Ser Gln Ala Ser Arg Glu Ala Gly  
 370 375 380  
 Ser Arg Gly Thr Ala Ile Cys Ser Ser Thr Leu Leu  
 385 390 395

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 516 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

ATGTGCTGGC CGCTGCTGTA TCCTTTGGTG CTGGGGCTCA GCATCAGCCG GCAGAGGGCA	60
TCCAGACCCC CAGCATCTAC GATGATGTAG AGAGTACCAG GGAAGCCAT GGTGANTGAC	120
TGTNTCCCTT AAAGGGGTGA ATCAGAAATG GAGCTANTGN TGAGCAGGNG NACAGNNTTT	180
ANGTCCCTAA AANCCATGCC TTTTGGGANT GGGTTGTATC CTTCNTTAA NTGANTNNTG	240
GANTGGGGAC ANTGAGGCAC CCACAATGCC TAAGACTTTC AAGGATATTC TCCTTCATCN	300
TGTATCCCTA AAGGCAGGNG AGAGCAGTGG NTGACTGATG TCCCCTCTCT CCCACAGAAG	360
GCCCTCTGGG TCCCACAGTA GAACTCAAGG AGCCGAAGTC CTCAGACAAG CCTAATCCAC	420
GAGGCTACCC GGGCAAATTC TGTGCCAACG ACAGTGACAC GCTGGAGCTC CCGGCCAGCT	480
CTCAAGCACT GCTGCTGGGG TGGGTATCCA CAAGGC	516

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1534 base pairs

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```

TCATGTGGGG GCGACTGCTC CTGTGGCCCC TGGTGCTGGG GTTCAGCCTG TCTGGCGGCA      60
CCCAGACCCC CAGCGTCTAC GACGAGAGCG GGAGCACCGG AGGTGGTGAT GACAGCACGC      120
CCTCAATCCT GCCTGCCCCC CGCGGCTACC CAGGCCAAGT CTGTGCCAAT GACAGTGACA      180
CCCTGGAGCT CCCGGACAGC TCACGGGCAC TGCTTCTGGG CTGGGTGCC ACCAGGCTGG      240
TGCCCGCCCT CTATGGGCTG GTCCTGGTGG TGGGGCTGCC GGCCAATGGG CTGGCGCTGT      300
GGGTGCTGGC CACGCAGGCA CCTCGGCTGC CCTCCACCAT GCTGCTGATG AACCTCGCGA      360
CTGCTGACCT CCTGCTGGCC CTGGCGCTGC CCCC GCGGAT CGCCTACCAC CTGCGTGGCC      420
AGCGCTGGCC CTTCGGGGAG GCCGCCTGCC GCCTGGCCAC GGCCGCACTC TATGGTCACA      480
TGTATGGCTC AGTGCTGCTG CTGGCCGCCG TCAGCCTGGA TCGCTACCTG GCCCTGGTGC      540
ACCCGCTGCG GGCCCGCGCC CTGCGTGGCC GGCGCCTGGC CCTTGGA CTGATGGCTG      600
CTTGGCTCAT GCGGGCCGCC CTGGCACTGC CCCTGACACT GCAGCGGCAG ACCTTCCGGC      660
TGGCGCGCTC CGATCGCGTG CTCTGCCATG ACGCGCTGCC CCTGGACGCA CAGGCCTCCC      720
ACTGGCAACC GGCCTTACC TGCCTGGCGC TGTTGGGCTG TTTCTGCC CTGCTGGCCA      780
TGCTGCTGTG CTACGGGGCC ACCCTGCACA CGCTGGCGGC CAGCGGCCGG CGCTACGGCC      840
ACGCGCTGAG GCTGACCGCA GTGGTGCTGG CCTCCGCCGT GGCCTTCTTC GTGCCAGCA      900
ACCTGCTGCT GCTGCTGCAT TACTCGGACC CGAGCCCAG CGCCTGGGGC AACCTCTATG      960
GTGCCTACGT GCCCAGCCTG GCGCTGAGCA CCCTCAACAG CTGCGTGGAT CCCTTCATCT     1020
ACTACTACGT GTCGGCCGAG TTCAGGGACA AGGTGCGGGC AGGGCTCTTC CAACGGTCGC     1080
CGGGGGACAC CGTGGCCTCC AAGGCCTCTG CGGAAGGGGG CAGCCGGGGC ATGGGCACCC     1140
ACTCCTCTTT GCTCCAGTGA CACAAAGTGG GGAAGGCTGT ACTGGGTCGA ACAGGGTCCC     1200
TTCCCCACT TCACGTCCTT CCTGGGACCT CAGAATGTGA CCTTATTTGG AAATAGGGTT     1260
GTTACAAC TGCTACTAGCAG AGGTCACTTT GGAGAAGGGT GGGCCTTACA TCCAGTGTGG     1320
GTGGTGTCTT CATAAGATAA GGAGAGGCCA GGCCTGGTGG CTCACGCCTG TAATCCCAGC     1380
ACTTTAAGAG GCCAAGGCGG ATGGATCACT TGAGCCCAGG AGTTCAACAC CAGCCTGAGC     1440
AACATGGTAA AACCCCATCT CTACCAAAA TACAAAAATT AGCTGGGCTT GGTGGCTGGC     1500
GCCTGTAATC CCAGCTACTC ANGAGACTGA GGCA                                     1534
    
```

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 386 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein



Ala Tyr Val Pro Ser Leu Ala Leu Ser Thr Leu Asn Ser Cys Val Asp  
 325 330 335  
 Pro Phe Ile Tyr Tyr Tyr Val Ser Ala Glu Phe Arg Asp Lys Val Arg  
 340 345 350  
 Ala Gly Leu Phe Gln Arg Ser Pro Gly Asp Thr Val Ala Ser Lys Ala  
 355 360 365  
 Ser Ala Glu Gly Gly Ser Arg Gly Met Gly Thr His Ser Ser Leu Leu  
 370 375 380  
 Gln  
 385

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 407 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

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

Leu Pro Phe Lys Ile Ala Tyr His Leu Asn Gly Asn Asn Trp Val Phe  
                                   165                                  170                                  175  
 Gly Glu Val Met Cys Arg Ile Thr Thr Val Val Phe Tyr Gly Asn Met  
                                   180                                  185                                  190  
 Tyr Cys Ala Ile Leu Ile Leu Thr Cys Met Gly Ile Asn Arg Tyr Leu  
                                   195                                  200                                  205  
 Ala Thr Ala His Pro Phe Thr Tyr Gln Lys Leu Pro Lys Arg Ser Phe  
                                   210                                  215                                  220  
 Ser Leu Leu Met Cys Gly Ile Val Trp Val Met Val Phe Leu Tyr Met  
 225                                  230                                  235                                  240  
 Leu Pro Phe Val Ile Leu Lys Gln Glu Tyr His Leu Val His Ser Glu  
                                   245                                  250                                  255  
 Ile Thr Thr Cys His Asp Val Val Asp Ala Cys Glu Ser Pro Ser Ser  
                                   260                                  265                                  270  
 Phe Arg Phe Tyr Tyr Phe Val Ser Leu Ala Phe Phe Gly Phe Leu Ile  
                                   275                                  280                                  285  
 Pro Phe Val Ile Ile Ile Phe Cys Tyr Thr Thr Leu Ile His Lys Leu  
                                   290                                  295                                  300  
 Lys Ser Lys Asp Arg Ile Trp Leu Gly Tyr Ile Lys Ala Val Leu Leu  
 305                                  310                                  315                                  320  
 Ile Leu Val Ile Phe Thr Ile Cys Phe Ala Pro Thr Asn Ile Ile Leu  
                                   325                                  330                                  335  
 Val Ile His His Ala Asn Tyr Tyr Tyr His Asn Thr Asp Ser Leu Tyr  
                                   340                                  345                                  350  
 Phe Met Tyr Leu Ile Ala Leu Cys Leu Gly Ser Leu Asn Ser Cys Leu  
                                   355                                  360                                  365  
 Asp Pro Phe Leu Tyr Phe Val Met Ser Lys Val Val Asp Gln Leu Asn  
                                   370                                  375                                  380  
 Pro Xaa Ser Ala Met Ala Arg Pro Leu Xaa Arg Pro Arg Arg Asp Ile  
 385                                  390                                  395                                  400  
 Trp Glu Asp Ile His Ala Trp  
                                   405

## (2) INFORMATION FOR SEQ ID NO:7:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 399 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Met Arg Ser Leu Ser Leu Ala Trp Leu Leu Gly Gly Ile Thr Leu Leu  
 1 5 10 15  
 Ala Ala Ser Val Ser Cys Ser Arg Thr Glu Asn Leu Ala Pro Gly Arg  
 20 25 30  
 Asn Asn Ser Lys Gly Arg Ser Leu Ile Gly Arg Leu Glu Thr Gln Pro  
 35 40 45  
 Pro Ile Thr Gly Lys Gly Val Pro Val Glu Pro Gly Phe Ser Ile Asp  
 50 55 60  
 Glu Phe Ser Ala Ser Ile Leu Thr Gly Lys Leu Thr Thr Val Phe Leu  
 65 70 75 80  
 Pro Val Val Tyr Ile Ile Val Phe Val Ile Gly Leu Pro Ser Asn Gly  
 85 90 95  
 Met Ala Leu Trp Ile Phe Leu Phe Arg Thr Lys Lys Lys His Pro Ala  
 100 105 110  
 Val Ile Tyr Met Ala Asn Leu Ala Leu Ala Asp Leu Leu Ser Val Ile  
 115 120 125  
 Trp Phe Pro Leu Lys Ile Ser Tyr His Leu His Gly Asn Asn Trp Val  
 130 135 140  
 Tyr Gly Glu Ala Leu Cys Lys Val Leu Ile Gly Phe Phe Tyr Gly Asn  
 145 150 155 160  
 Met Tyr Cys Ser Ile Leu Phe Met Thr Cys Leu Ser Val Gln Arg Tyr  
 165 170 175  
 Trp Val Ile Val Asn Pro Met Gly His Pro Arg Lys Lys Ala Asn Ile  
 180 185 190  
 Ala Val Gly Val Ser Leu Ala Ile Trp Leu Leu Ile Phe Leu Val Thr  
 195 200 205  
 Ile Pro Leu Tyr Val Met Lys Gln Thr Ile Tyr Ile Pro Ala Leu Asn  
 210 215 220  
 Ile Thr Thr Cys His Asp Val Leu Pro Glu Glu Val Leu Val Gly Asp  
 225 230 235 240  
 Met Phe Asn Tyr Phe Leu Ser Leu Ala Ile Gly Val Phe Leu Phe Pro  
 245 250 255  
 Ala Leu Leu Thr Ala Ser Ala Tyr Val Leu Met Ile Lys Thr Leu Arg  
 260 265 270  
 Ser Ser Ala Met Asp Glu His Ser Glu Lys Lys Arg Gln Arg Ala Ile  
 275 280 285  
 Arg Leu Ile Ile Thr Val Leu Ala Met Tyr Phe Ile Cys Phe Ala Pro  
 290 295 300  
 Ser Asn Leu Leu Leu Val Val His Tyr Phe Leu Ile Lys Thr Gln Arg  
 305 310 315 320

Gln Ser His Val Tyr Ala Leu Tyr Leu Val Ala Leu Cys Leu Ser Thr  
 325 330 335  
 Leu Asn Ser Cys Ile Asp Pro Phe Val Tyr Tyr Phe Val Ser Lys Asp  
 340 345 350  
 Phe Arg Asp His Ala Arg Asn Ala Leu Leu Cys Arg Ser Val Arg Thr  
 355 360 365  
 Val Asn Arg Met Gln Ile Ser Leu Ser Ser Asn Lys Phe Ser Arg Lys  
 370 375 380  
 Ser Gly Ser Tyr Ser Ser Ser Ser Thr Ser Val Lys Thr Ser Tyr  
 385 390 395

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 430 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

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

Pro Phe Lys Ile Ser Tyr Tyr Phe Ser Gly Thr Asp Trp Gln Phe Gly  
 165 170 175

Ser Gly Met Cys Arg Phe Ala Tyr Ala Ala Phe Tyr Gly Asn Met Tyr  
 180 185 190

Ala Ser Ile Met Leu Met Thr Val Ile Ser Ile Asp Arg Phe Leu Ala  
 195 200 205

Val Val Tyr Pro Ile Gln Ser Leu Ser Trp Arg Thr Leu Gly Arg Ala  
 210 215 220

Asn Phe Thr Cys Val Val Ile Trp Val Met Ala Ile Met Gly Val Val  
 225 230 235 240

Pro Leu Leu Leu Lys Glu Gln Thr Thr Arg Val Pro Gly Leu Asn Ile  
 245 250 255

Thr Thr Cys His Asp Val Leu Ser Glu Asn Leu Met Gln Gly Phe Tyr  
 260 265 270

Ser Tyr Tyr Phe Ser Ala Phe Ser Ala Ile Phe Phe Leu Val Pro Leu  
 275 280 285

Ile Val Ser Thr Val Cys Tyr Thr Ser Ile Ile Arg Cys Leu Ser Ser  
 290 295 300

Ser Ala Val Ala Asn Arg Ser Lys Lys Ser Arg Ala Leu Phe Leu Ser  
 305 310 315 320

Ala Ala Val Phe Cys Ile Phe Ile Val Cys Phe Gly Pro Thr Asn Val  
 325 330 335

Leu Leu Ile Val His Tyr Leu Phe Leu Ser Asp Ser Pro Gly Thr Glu  
 340 345 350

Ala Ala Tyr Phe Ala Tyr Leu Leu Cys Val Cys Val Ser Ser Val Ser  
 355 360 365

Cys Cys Ile Asp Pro Leu Ile Tyr Tyr Tyr Ala Ser Ser Glu Cys Gln  
 370 375 380

Arg His Leu Tyr Ser Ile Leu Cys Cys Lys Glu Ser Ser Asp Pro Asn  
 385 390 395 400

Ser Cys Asn Ser Thr Gly Gln Leu Met Pro Ser Lys Met Asp Thr Cys  
 405 410 415

Ser Ser His Leu Asn Asn Ser Ile Tyr Lys Lys Leu Leu Ala  
 420 425 430

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 81 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:

(A) NAME/KEY: intron

(B) LOCATION: 26...67

(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CTGTCTTCCC GCGTCCCTAT GAGCCAGCCA GGTAAGAGCT GCGGGNNNCT CAATTTTCTT	60
CTTTCAGAAT CAGAGAGGAC A	81

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 78 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:

(A) NAME/KEY: intron

(B) LOCATION: 29...57

(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

TGCAGCCGGA CCGAGAACCT TGCACCGGGT GAGCNNNGCT CCGCTTTCTT TGTACAGGAC	60
GCAACAACAG TAAAGGAA	78

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 38 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:

- (A) NAME/KEY: intron
- (B) LOCATION: 10...16
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

TTTCCTTTCA ATACAGGCAT AAATGTTTTTC AGACAACT

38

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Gly Tyr Pro Gly Lys Phe

1

5

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Ser Phe Leu Leu Arg Asn

1

5

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6 amino acids
- (B) TYPE: amino acid



## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (ix) FEATURE:

- (A) NAME/KEY: Other
- (B) LOCATION: 4...0
- (D) OTHER INFORMATION: Xaa is hR

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

Pro Ala Pro Xaa Gly Tyr Pro Gly Gln Val Cys Ala Asn Asp Ser Asp  
 1                                    5                                    10                                    15  
 Thr Leu Glu Leu Pro Asp  
                                   20

## (2) INFORMATION FOR SEQ ID NO:18:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (ix) FEATURE:

- (A) NAME/KEY: Other
- (B) LOCATION: 1...0
- (D) OTHER INFORMATION: Xaa is dF

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Xaa Pro Arg Pro Tyr Pro Gly Gln Val Cys Ala Asn Asp Ser Asp Thr  
 1                                    5                                    10                                    15  
 Leu Glu Leu Pro Asp  
                                   20

## (2) INFORMATION FOR SEQ ID NO:19:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 6 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

## (ix) FEATURE:

(A) NAME/KEY: Other

(B) LOCATION: 1...0

(D) OTHER INFORMATION: Xaa is Ala, Ser or Thr

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

Xaa Tyr Pro Gly Lys Phe

1

5

## (2) INFORMATION FOR SEQ ID NO:20:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 6 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

Gly Phe Pro Gly Lys Phe

1

5

## (2) INFORMATION FOR SEQ ID NO:21:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 6 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

## (ix) FEATURE:

- (A) NAME/KEY: Other
- (B) LOCATION: 2...0
- (D) OTHER INFORMATION: Xaa is parafluorophe

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Gly Xaa Pro Gly Lys Phe

1 5

## (2) INFORMATION FOR SEQ ID NO:22:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

Gly Leu Pro Gly Lys Phe

1 5

## (2) INFORMATION FOR SEQ ID NO:23:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

Gly Ile Pro Gly Lys Phe

1 5

## (2) INFORMATION FOR SEQ ID NO:24:



Gly Phe Pro Gly Gln Val

1 5

## (2) INFORMATION FOR SEQ ID NO:27:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 6 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: Other

(B) LOCATION: 2...0

(D) OTHER INFORMATION: Xaa is parafluoroPhe

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

Gly Xaa Pro Gly Gln Val

1 5

## (2) INFORMATION FOR SEQ ID NO:28:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 6 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

Gly Leu Pro Gly Gln Val

1 5

## (2) INFORMATION FOR SEQ ID NO:29:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 6 amino acids

(B) TYPE: amino acid

