



PCT

WELTORGANISATION FÜR GEISTIGES EIGENTUM
Internationales Büro

INTERNATIONALE ANMELDUNG VERÖFFENTLICHT NACH DEM VERTRAG ÜBER DIE
INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES PATENTWESENS (PCT)

<p>(51) Internationale Patentklassifikation ⁵ : C12N 15/15, C07K 13/00 A61K 37/64</p>	<p>A1</p>	<p>(11) Internationale Veröffentlichungsnummer: WO 93/09232 (43) Internationales Veröffentlichungsdatum: 13. Mai 1993 (13.05.93)</p>
<p>(21) Internationales Aktenzeichen: PCT/EP92/02450 (22) Internationales Anmeldedatum: 27. Oktober 1992 (27.10.92) (30) Prioritätsdaten: P 41 36 513.5 6. November 1991 (06.11.91) DE (71) Anmelder (für alle Bestimmungsstaaten ausser US): BASF AKTIENGESELLSCHAFT [DE/DE]; Carl-Bosch-Strasse 38, D-6700 Ludwigshafen (DE). (72) Erfinder; und (75) Erfinder/Anmelder (nur für US) : FRIEDRICH, Thomas [DE/DE]; Saalbaustrasse 22-24, D-6100 Darmstadt (DE). BIALOJAN, Siegfried [DE/DE]; Gartenstrasse 34, D-6836 Oftersheim (DE). KROEGER, Burkhard [DE/DE]; Donnersbergstrasse 13, D-6703 Limburgerhof (DE). KUENAST, Christoph [DE/DE]; Salierstrasse 2, D-6701 Otterstadt (DE).</p>		<p>(74) Gemeinsamer Vertreter: BASF AKTIENGESELLSCHAFT; Carl-Bosch-Strasse 38, D-6700 Ludwigshafen (DE). (81) Bestimmungsstaaten: CA, JP, US, europäisches Patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, SE). Veröffentlicht <i>Mit internationalem Recherchenbericht.</i></p>
<p>(54) Title: NOVEL THROMBIN INHIBITOR PROTEIN FROM <i>REDUVII</i> (54) Bezeichnung: NEUES THROMBININHIBITORISCHES PROTEIN AUS RAUBWANZEN (57) Abstract The description relates to a novel thrombin inhibiting protein from <i>reduvii</i> with a molecular weight of some 12000 Dalton and the N-terminus Glu-Gly-Gly-Glu-Pro-Cys-Ala-Cys-Pro-His-Ala-Leu-His-Arg-Val-Cys-Gly-Ser-Asp. The protein is useful in combatting diseases. (57) Zusammenfassung Es wird ein neues thrombininhibitorisches Protein aus Raubwanzen mit einem Molekulargewicht von etwa 12000 Dalton und dem N-Terminus Glu-Gly-Gly-Glu-Pro-Cys-Ala-Cys-Pro-His-Ala-Leu-His-Arg-Val-Cys-Gly-Ser-Asp beschrieben. Das Protein eignet sich zur Bekämpfung von Krankheiten.</p>		

LEDIGLICH ZUR INFORMATION

Code, die zur Identifizierung von PCT-Vertragsstaaten auf den Kopfhögen der Schriften, die internationale Anmeldungen gemäss dem PCT veröffentlichen.

AT	Österreich	FR	Frankreich	MR	Mauritanien
AU	Australien	GA	Gabon	MW	Malawi
BB	Barbados	GB	Vereinigtes Königreich	NL	Niederlande
BE	Belgien	GN	Guinea	NO	Norwegen
BF	Burkina Faso	GR	Griechenland	NZ	Neuseeland
BG	Bulgarien	HU	Ungarn	PL	Polen
BJ	Benin	IE	Irland	PT	Portugal
BR	Brasilien	IT	Italien	RO	Rumänien
CA	Kanada	JP	Japan	RU	Russische Föderation
CF	Zentrale Afrikanische Republik	KP	Demokratische Volksrepublik Korea	SD	Sudan
CG	Kongo	KR	Republik Korea	SE	Schweden
CH	Schweiz	KZ	Kasachstan	SK	Slowakische Republik
CI	Côte d'Ivoire	LI	Liechtenstein	SN	Senegal
CM	Kamerun	LK	Sri Lanka	SU	Sowjet Union
CS	Tschechoslowakei	LU	Luxemburg	TD	Tschad
CZ	Tschechische Republik	MC	Monaco	TG	Togo
DE	Deutschland	MG	Madagaskar	UA	Ukraine
DK	Dänemark	ML	Mali	US	Vereinigte Staaten von Amerika
ES	Spanien	MN	Mongolei	VN	Vietnam
FI	Finnland				

Neues thrombininhibitorisches Protein aus Raubwanzen

Beschreibung

5

Die vorliegende Erfindung betrifft ein neues thrombininhibitorisches Protein aus Raubwanzen sowie Verfahren zu dessen Herstellung.

- 10 Thrombin-Inhibitoren sind wichtige therapeutische Substanzen, die beispielsweise zur Prophylaxe oder Behandlung von Thrombosen oder arteriellen Reokklusionen verwendet werden.

- 15 In der deutschen Offenlegungsschrift DE 39 31 839 wird ein Thrombininhibitor, der aus der Lederzecke *Ornithodoros moubata* isoliert worden ist, beschrieben. Dieses Protein besitzt ein Molekulargewicht von etwa 15000 Dalton, einen isoelektrischen Punkt von pH 4-5 und die N-terminale Aminosäuresequenz SDYEFPPPKXRPG.

20

In der europäischen Offenlegungsschrift EP 345 614 wird der thrombininhibitorische Wirkstoff Amblyommin beschrieben, der aus Schildzecken isoliert wird. Es handelt sich dabei um ein Protein mit einem Molekulargewicht von 20 000 - 30 000 Dal-

25

ton und einem isoelektrischen Punkt zwischen 5,05 und 5,65.

30

Bisher wurde jedoch noch kein Protein mit thrombininhibitorischer Wirkung gefunden, das als Arzneimittel hinsichtlich hoher Wirksamkeit, fehlender Antigenizität, langer biologischer Halbwertszeit, geringer Nebenwirkung wie z.B. Blutungsneigung, vorteilhaft geeignet ist.

35

Der Erfindung lag daher die Aufgabe zugrunde, neue Thrombin-Inhibitoren zur Verfügung zu stellen, die als Arzneimittel hinsichtlich der oben genannten Eigenschaften geeignet sind.

Demgemäß wurde ein neues thrombininhibitorisches Protein aus Raubwanzen isoliert.

40

Das neue Protein besitzt folgende physikochemischen Eigenschaften. Durch Molekularsiebchromatographie wird ihm ein Molekulargewicht von 20000 - 24000 Dalton zugeordnet. In einem SDS-Polyacrylamidgel wird ein Molekulargewicht von

12000 ± 2000 Dalton bestimmt. Die Bestimmung des isoelektrischen Punktes ergibt einen pH-Wert zwischen 3,7 und 4,7.

5 Das Protein bindet spezifisch an eine Thrombin-Affinitäts-säule. Es hemmt die biologische Aktivität von Thrombin in einem in-vitro Enzymtest.

Folgende N-terminale Aminosäuresequenz wurde von dem Protein bestimmt (SEQ ID NO: 1):

10 Glu-Gly-Gly-Glu-Pro-Cys-Ala-Cys-Pro-His-Ala-Leu-His-Arg-Val-Cys-Gly-Ser-Asp

15 Das erfindungsgemäße Protein enthält eine Aminosäuresequenz von 103 Aminosäuren, die im Sequenzprotokoll SEQ ID NO: 3 angegeben ist.

Die für diese 103 Aminosäuren lange Proteinsequenz codierende DNA ist im Sequenzprotokoll SEQ ID NO: 2 aufgeführt.

20 Das erfindungsgemäße Protein kann darüber hinaus am C-Terminus noch weitere Aminosäuren enthalten. Ebenso kann das Protein am N-Terminus noch weitere Aminosäuresequenzen wie natürliche oder heterologe Leadersequenzen oder zusätzliche
25 Aminosäuren wie Methionin enthalten.

Aus der im Sequenzprotokoll SEQ ID NO: 3 angegebenen Aminosäuresequenz lassen sich zwei Domänen in diesem Molekül vorhersagen. Die erste Domäne ist im wesentlichen durch die Cysteinreste an den Positionen 6 und 48, die zweite Domäne durch die Cysteinreste an den Positionen 57 und 101 begrenzt.
30

Weiterhin sind Gegenstand der Erfindung DNA-Sequenzen, die für Proteine mit thrombininhibitorischer Wirkung codieren, und die aus der Gruppe, die von
35

a) DNA-Sequenzen mit der in SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12 und
40 SEQ ID NO: 14 beschriebenen Struktur, und

b) DNA-Sequenzen, die unter Standardbedingungen mit DNA Sequenzen a) hybridisieren, gebildet wird,

ausgewählt sind. Hierbei handelt es sich um DNA-Sequenzen
5 für Thrombininhibitoren mit bis zu sechs Domänen.

Auch kleinere Proteine, die nur eine dieser beiden Domänen
enthalten sind Gegenstand der Erfindung. Weiterhin sind als
erfindungsgemäße Proteine auch solche zu verstehen, die meh-
10 rere, bevorzugt bis zu 10, solcher Domänen enthalten.

Dabei sind die Mindestanforderungen für eine aktive Einzel-
domäne die Aminosäuren (aus SEQ ID NO: 7) ab Position 6 bis
48, 57 bis 101 oder 119 bis 160 oder die Aminosäuren aus SEQ
15 ID NO: 17 ab Position 6 bis 48, 57 bis 101, 120 bis 161, 190
bis 232, 241 bis 285 oder 303 bis 344.

N-terminale und C-terminale Extensionen stören die thrombi-
ninhibierende Aktivität nicht. Solche Verlängerungen können
20 z.B. Sequenzen beinhalten wie die Aminosäuren aus
SEQ ID NO: 7: 1 bis 5, 49 bis 56, 102 bis 118, 161 bis 170
oder aus SEQ ID NO: 17: 1 bis 5, 49 bis 56, 102 bis 119, 162
bis 189, 233 bis 240, 286 bis 302 oder 345 bis 354.

25 Diese Aminosäuresequenzen, die sich durch hohe Flexibilität
und Hydrophilie auszeichnen, können als Abstandhalter zwi-
schen den einzelnen Domänen eingesetzt werden. Solche Multi-
domänen-Proteine zeigen eine verlängerte Halbwertszeit, ohne
daß die Aktivität der Einzeldomäne wesentlich sinkt.

30 Durch die Insensitivität des neuen Proteins und seiner Ein-
zeldomänen gegenüber Verlängerungen am N-Terminus und C-Ter-
minus ist es auch möglich, den neuen Thrombininhibitor hete-
rolog mit anderen Proteinen zu verbinden. Solche Proteine
35 sind zum Beispiel Gewebe-Plasminogen-Aktivatoren, Streptoki-
nase, Urokinase, Anti-Thrombin III und aktiviertes Pro-
tein C. Ein wichtiges Protein für die Lysetherapie ist tPA,
seine Muteine und Derivate. Der Vorteil solcher Fusionspro-
teine ist eine erhöhte Thrombusspezifität bei gleichzeitiger
40 Verringerung der Wiederverschlußrate.

- Beispielsweise wird die Aminosäuresequenz des Gewebeplasminogenaktivators (EP 93619) zwischen Position 1 und 527 mit den Aminosäuren 1 bis 104 des doppelköpfigen Thrombininhibitors verlängert, indem hinter die für tPA 1 bis 157 kodierende Nukleotid-Sequenz die entsprechende Nukleotidsequenz aus SEQ ID NO: 2 über geeignete Restriktionsstellen verknüpft wird.
- 5
- Umgekehrt ist es aber auch möglich, zum Beispiel die Nukleotidsequenz ab Aminosäure 104 des Doppelkopfinhibitors mit der Sequenz für den Gewebeplasminogenaktivator zu verlängern, indem die Aminosäure 1 des tPA zur Aminosäure 105 des heterotrimeren Proteins wird.
- 10
- Durch solche Kombinationsproteine wird eine höhere Gerinnselspezifität erreicht, da der Thrombininhibitor selektiv das in den Thrombus eingeschlossene Thrombin bindet und seinen Fusionspartner so an das Gerinnsel heranführt.
- 15
- Eine alternative Strategie ist das oben beschriebene Fusionsprotein mit einer durch Proteasen spaltbaren Sequenz (z.B. einer Faktor Xa-Schnittstelle (ILE-GLU-GLY-ARG-X)) zu versehen, wodurch nach einer Aktivierung des Gerinnungssystems die Fusionspartner freigesetzt werden, um z.B. ihre fibrinolytische und antikoagulatorische Wirkung auszuüben. Es können mehrere, bevorzugt 2 bis 10 Fusionspartner so zusammengestellt werden.
- 20
- 25
- Solche Proteine, die mehrere dieser Domänen enthalten, werden zweckmäßigerweise durch gentechnische Verfahren hergestellt. Man verknüpft beispielsweise die DNA-Sequenzen die für die Domänen codieren, nach bekannten Verfahren zu einem synthetischen 'Multidomänen-Gen' und exprimiert dieses Gen in an sich bekannter Weise.
- 30
- 35
- Das neue Protein läßt sich aus Raubwanzen der Gattung Rhodnius isolieren. Hierzu werden die Raubwanzen zweckmäßigerweise in einem Puffer bei pH 6 bis 9, vorzugsweise pH 7 bis 8 aufgenommen und mit einem Homogenisator, vorzugsweise einem Mixer, homogenisiert. Anschließend werden die unlöslichen Bestandteile abgetrennt, bevorzugt abzentrifugiert.
- 40

Die weitere Reinigung des Proteins kann über chromatographische Methoden, bevorzugt Ionenaustauschchromatographie und/oder Affinitätschromatographie erfolgen. Besonders bevorzugt ist ein Reinigungsschritt über Thrombin-Affinitätschromatographie.

Die Reinigung des Proteins kann über einen Thrombin-Aktivitätstest verfolgt werden. Hierzu benutzt man zweckmäßigerweise einen optischen Test, bei dem ein chromogenes Substrat, beispielsweise Chromozym T, durch Thrombin umgesetzt wird. Die das neue Protein enthaltenden Fraktionen können bei Zugabe in diesen optischen Test an ihrer Thrombin inhibierenden Wirkung erkannt werden.

Besonders geeignet zur Herstellung des erfindungsgemäßen Proteins sind gentechnische Verfahren.

Hierzu wird in an sich bekannter Weise eine cDNA-Genbank aus der Raubwanze angelegt. Aus dieser Genbank kann das für das erfindungsgemäße Protein kodierende Gen isoliert werden, indem man beispielsweise eine DNA-Probe herstellt, deren Sequenz von der oben beschriebenen N-terminalen Aminosäuresequenz durch Rückübersetzung gemäß dem genetischen Code erhalten wird. Durch Hybridisierung mit dieser DNA-Probe läßt sich das entsprechende Gen auffinden und isolieren.

Für die Herstellung des entsprechenden Gens kann aber auch die Polymerase-Chain-Reaction (PCR)-Technik eingesetzt werden. Beispielsweise läßt sich mit Hilfe eines Primers, dessen Sequenz durch Rückübersetzung aus der oben beschriebenen N-terminalen Aminosäuresequenz erhalten wurde, und eines zweiten Primers, dessen Sequenz komplementär zum 3'-Ende des cDNA-Genfragments ist, bevorzugt mit der Sequenz Poly(dT), das cDNA-Genfragment für das erfindungsgemäße Protein durch PCR-Technik herstellen. Das entsprechende Gen läßt sich auch isolieren, indem man eine Expressionsgenbank von Raubwanzen anlegt und diese mit einem Antikörper der gegen das erfindungsgemäße Protein gerichtet ist, absucht.

Nachdem das entsprechende Gen isoliert worden ist, kann es durch gentechnische Verfahren in Organismen, z.B. in Bakterien, Hefen und anderen eukaryontischen Zellen, mit Hilfe

eines Expressionsvektors in an sich bekannter Weise exprimiert werden.

- 5 Bevorzugt wird in Prokaryonten, wie z.B. E. coli und mit Hilfe stark exprimierender Vektoren, z.B. unter der Kontrolle des induzierbaren tac-Promotors, wie z.B. in Plasmid pMal-p2 (Protein Fusion and Purification System, "GeneExpress", New England Biolabs) vorhanden, gearbeitet. Dabei
- 10 kommt es zur periplasmatischen Expression eines Fusionsproteins aus dem Maltose-Bindungsprotein und dem beschriebenen Thrombininhibitor. Der Fusionspartner kann nach Reinigung enzymatisch entfernt werden.

- 15 Die generelle Vorgehensweise zur gentechnischen Herstellung eines neuen Proteins bei bekannter Aminosäurepartialsequenz ist in Lehrbüchern der Gentechnologie, beispielsweise E.L. Winnacker, Gene und Klone, Verlag Chemie, Weinheim, 1984, beschrieben. Die experimentellen Bedingungen für die einzelnen Verfahren wie beispielsweise Anlegen einer Gen-
- 20 bank, Hybridisierung, Expression eines Gens sind bei T. Maniatis, "Molecular Cloning", Cold Spring Harbor Laboratory, 1990, beschrieben.

- 25 Unter Standardbedingungen sind beispielsweise Temperaturen zwischen 42 und 58°C in einer wäßrigen Pufferlösung mit einer Konzentration zwischen 0,1 und 1 x SSC (1 x SSC: 0,15M NaCl, 15mM Natriumcitrat pH 7,2) zu verstehen.

- 30 Das erfindungsgemäße Protein wird bevorzugt in Form seiner pharmazeutisch annehmbaren Salze verwendet.

- 35 Das neue Protein besitzt blutgerinnungshemmende Eigenschaften. Es kann beispielsweise zur Prophylaxe von Thrombosen oder arteriellen Reokklusionen, zur Behandlung von Thrombosen, zur Konservierung von Blut oder bei der extrakorporalen Zirkulation verwendet werden.

- 40 Die neuen Proteine sind wirksame Thrombininhibitoren. Sie können allein oder auch zusammen mit bekannter gerinnungshemmenden Faktoren als Arzneimittel verwendet werden. Als gerinnungshemmende Faktoren werden bevorzugt Thrombininhibitoren, beispielsweise Hirudin, Faktor Xa-Inhibitoren, beispielsweise TAP (Waxman et al., Science 248, 1990,

Seite 593-596) oder Plättchen-Aggregationshemmer, beispielsweise Kistrin (Dennis et al., Proc. Natl. Acad. Sci. USA, 87, 1989, Seite 2471-2475) eingesetzt.

- 5 Die Erfindung wird durch die folgenden Beispiele weiter veranschaulicht.

Beispiel 1

- 10 Reinigung des thrombininhibitorischen Proteins aus Raubwanzen

Eine Laborkultur der Raubwanzen (*Rhodnius prolixus*) wurde bei 28°C und 80 % relativer Luftfeuchtigkeit gehalten. In

- 15 30-tägigen Abständen wurden die Raubwanzen dadurch gefüttert, daß sie an Kaninchen saugen konnten. Die Raubwanzen wurden nach Erreichen des letzten Entwicklungsstadiums bei -20°C eingefroren.

- 20 25 g Raubwanzen wurden mit 75 ml 20 mM Natriumphosphatpuffer, 150 mM NaCl (pH 7,5) homogenisiert. Das Homogenat wurde 60 Minuten bei 20000 U/min (Sorvall RC-3B, Rotor SS-34) zentrifugiert. Der Niederschlag wurde verworfen.

- 25 Die Proteinlösung (Überstand) wurde auf eine Q-Sepharose Säule® (Pharmacia) aufgetragen (60 ml/h), die in 20 mM Natriumphosphatpuffer pH 8,0 äquilibriert worden war (Durchmesser 2,5 cm, Volumen 50 ml).

- 30 Es wurde mit 10 Säulenvolumen Äquilibrierungspuffer gewaschen.

Danach wurde ein linearer Gradient von 50 ml 20 mM Natriumphosphat (pH 8,0), nach 50 ml 20 mM Natriumphosphat

- 35 (pH 8,0), 1 M NaCl angelegt.

Aktive Fraktionen (gemessen durch Thrombininhibition) wurden gesammelt.

Die vereinigten aktiven Fraktionen wurden auf eine Affinitätssäule mit immobilisiertem Thrombin aufgetragen (Durchmesser 1,5 cm, Höhe 6,5 cm, 11,5 ml Säulenvolumen, 60 ml/h). Die Säule wurde gemäß Beispiel 3 hergestellt.

5

Die Säule wurde mit 20 mM Natriumphosphat pH 7,5 äquilibriert. Nach Auftragen der Proteinlösung wurde die Säule mit 10 Säulenvolumen Äquilibrierungspuffer gewaschen, bis die Absorption bei 280 nm auf Null zurückging.

10

Danach wurde mit 0,5 M NaCl, 20 mM Natriumphosphatpuffer pH 7,5 gewaschen. Unspezifisch adsorbiertes Material wurde so entfernt.

15

Spezifisch an Thrombin gebundenes Protein wurde mit 0,1 M Glycin, 0,5 M NaCl pH 2,8 eluiert. Die Säule wurde anschließend sofort wieder mittels Phosphatpuffer auf pH 7,5 eingestellt.

20

Die einzelnen Fraktionen wurden mit 0,1 M NaOH neutralisiert und auf ihre inhibitorische Wirkung gegen Thrombin untersucht.

Die durch Glycin/NaCl-Puffer pH 2,8 eluierten Fraktionen besaßen thrombininhibierende Wirkung.

25

Die gesammelten aktiven Fraktionen wurden, nachdem sie neutralisiert worden waren, mit Wasser verdünnt (1:10) und auf eine Mono-Q-Säule® gegeben (Pharmacia, 1 ml Säulenvolumen).

30

Die Säule wurde äquilibriert mit 20 mM Natriumphosphatpuffer pH 7,5, 150 mM NaCl (Puffer A). Es wurde mit Puffer A gewaschen bis die Absorption auf Null zurückging (10 Minuten). Dann wurde innerhalb von 50 Minuten auf 20 mM Natrium-

35

phosphat, pH 7,5, 800 mM NaCl (Puffer B) gewechselt (Fluß 0,5 ml/min).

Thrombin hemmende Fraktionen wurden gesammelt.

40

Die gesammelten Fraktionen wurden an einer RP 318® (Biorad) HPLC-Chromatographiesäule weiter gereinigt. Die Säule wurde äquilibriert mit 0,1 Gew.% Trifluoressigsäure (TFA) in dest. Wasser. Die vereinigten aktiven Fraktionen wurden auf die

- Säule aufgetragen. Mit einem Gradienten nach 0,1 Gew.-% TFA, 100 % Acetonitril während einer Stunde und einer Flußrate von 1 ml/min wurde die Säule eluiert. Die Absorption wurde bei 280 nm bestimmt. Es wurden 0,5 ml Fraktionen gesammelt.
- 5** Die gesammelten Fraktionen wurden zur Trockene einkonzentriert und in Phosphat gepufferter Salzlösung (PBS) (0,8 g/l NaCl; 0,2 g/l HCl; 0,144 g/l Natriumphosphat; 0,2 g/l Kaliumphosphat, pH 7,5) aufgenommen sowie die inhibitorische Aktivität bestimmt.
- 10** Die Proteinbestimmung erfolgte nach der Vorschrift von Bradford (Anal. Biochem., 72, 248-254 (1976)). Als Standardprotein diente Rinderserumalbumin (Boehringer Mannheim).
- 15** Beispiel 2
- Bestimmung der Hemmung von Thrombin durch den Inhibitor
- Thrombin (Boehringer Mannheim) wurde zu einer Endkonzentration von 25 mU/ml in Phosphat gepufferter Salzlösung (PBS) (0,8 g/l NaCl; 0,2 g/l HCl; 0,144 g/l Natriumphosphat; 0,2 g/l Kaliumphosphat, pH 7,5) gelöst.
- 20**
- Chromozym TH (Boehringer Mannheim) wurde in 20 ml H₂O/Flasche gelöst.
- 25**
- 50 µl Thrombinlösung und 100 µl Chromozym sowie 25 µl Probe oder Puffer wurden in die Näpfe einer Mikrotiterplatte gegeben. Sofort danach wurde zur Zeit 0 und nach 30 Minuten bei 37°C die Absorption bei 405 nm gemessen.
- 30**
- Bei starker Eigenfarbe der Probe wurde eine weitere Kontrolle ohne Thrombin wie oben behandelt.
- 35** Durch die Aktivität des Thrombins wird aus dem chromogenen Farbsubstrat ein bei 405 nm absorbierender Farbstoff freigesetzt. Die Hemmung des Thrombins durch einen Thrombininhibitor ist an einer geringeren Absorptionzunahme bei 405 nm erkennbar und wurde mit Hilfe einer Eichkurve quantifiziert.
- 40**

Beispiel 3

Herstellung einer Affinitätssäule mit Thrombin als Ligand

5 a) Kopplung:

2 g CNBr aktivierte Sepharose (Pharmacia) wurden mit 200 ml 1 mM HCl auf einer Nutsche gewaschen. Das Gel wurde in 100 mM NaHCO₃, 500 mM NaCl pH 8,3 aufgenommen und sofort mit 10000 Units Thrombin (Sigma) in 100 mM NaHCO₃, 500 mM NaCl, pH 8,3 gemischt.

Die Lösung wurde 24 Stunden bei 4°C vorsichtig geschüttelt.

15

b) Blockierung:

Das Gelmaterial wurde nach Absetzen mit 100 mM NaHCO₃, 500 mM NaCl, pH 8,3 gewaschen. Die Sepharose wurde dann mit 100 mM NaHCO₃, 500 mM NaCl, 1 M Ethanolamin pH 8,3 für 2 Stunden inkubiert.

20

c) Vorbereitung:

Zur Entfernung des ungebundenen Thrombin wird das Gelmaterial vor Gebrauch noch einmal in der Säule mit 20 Säulenvolumen PBS pH 7,4 gewaschen.

25

Beispiel 4

30

Bestimmung des Molekulargewichts durch Molekularsiebchromatographie.

Bei einer Flußgeschwindigkeit von 1 ml/min wurde durch Mono-Q®-Chromatographie gereinigtes Material in 20 mM Natriumphosphat, 150 mM NaCl, pH 7,5 auf einer Molekularsiebsäule des Typs TSK® (Pharmacia, Spherogel TSK 3000® SW, 7,5 mm Durchmesser, 60 cm Höhe) getrennt.

35

Ebenso wurde mit den Eichproteinen verfahren (Serumalbumin MW 67000 Da, Ovalbumin MW 45000 Da, Chymotrypsinogen A MW 25000 Da).

40

Der Logarithmus des Molekulargewichts der Eichproteine wurde gegen deren Elutionszeit in einem Diagramm aufgetragen.

- 5 In den fraktionierten Eluaten der Probe wurde die Thrombininhibition bestimmt.

Die Elutionszeit des Inhibitors ergab am Schnittpunkt mit der Eichgeraden den Logarithmus des gesuchten Molekulargewichts.

10

Durch diese Bestimmung wurde ein Molekulargewicht zwischen 20000 und 24000 Dalton bestimmt.

Beispiel 5

15

Bestimmung des Molekulargewichts durch Tricine-SDS-Polyacrylamid Gel Elektrophorese.

20

(Literatur: Analytical Biochemistry, 166, 368 - 379 (1987)
Tricine-Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis for the Separation of Proteins in the range from 1 - 1000 kDa, Schägger, H. and von Jagow, G.)

25

Die Gelelektrophorese wurde entsprechend der Literaturvorschrift bei 20 mA und 1400 V, 30 Watt, durchgeführt.

Das nach dieser Methode bestimmte Molekulargewicht betrug 12000 ± 2000 Dalton.

30

Als Eichproteine dienten (Intaktes Myoglobin 17,2 kDa, Bromcyanpeptid Myoglobin I+II 14,6 kDa, Bromcyanpeptid Myoglobin I 8,2 kDa, Bromcyanpeptid Myoglobin II 6,4 kDa, Bromcyanpeptid Myoglobin III 2,6 kDa, Myoglobin 1-14)

35

Beispiel 6

Sequenzbestimmung des Inhibitors
Reduktion und Carboxymethylierung.

40

2,8 ml Proteinlösung (0,029 mg/ml) wurden mit 0,28 ml Puffer (1 M Tris/HCl, 0,5 M Guanidinhydrochlorid, pH 8,6) gemischt. Danach wurden 0,116 ml Dithiothreitol (DTT, 10 mg/ml) zugegeben und 10 Minuten bei 37°C inkubiert. Danach wurde nach

Zugabe von 0,185 ml Jodacetamid (10 mg/ml) für 90 Minuten bei 37°C inkubiert. Die Reaktion wurde mit 0,073 ml DTT wie oben beendet.

- 5 Das Protein wurde durch erneute reversed phase HPLC an RP 318[®] gereinigt. Das Gemisch wurde auf eine Endkonzentration von 0,1 Gew.% Trifluoressigsäure (TFA) eingestellt und in einem HPLC System von Hewlett Packard (HP 1090 Liquid Chromatograph) getrennt. Mit Lösungsmittel A (0,1 Gew.-%
- 10 TFA, 100 % H₂O) wurde 5 Minuten gewaschen. Dann wurde der Anteil von Lösungsmittel B (90 % Acetonitril, 10 % H₂O, 0,1 Gew.% TFA) im Verlauf von 120 Minuten auf 50 % angehoben. Die Absorption des Eluats bei 214 und 280 nm wurde gemessen. Absorbierende Fraktionen wurden gesammelt. Das Protein wurde durch SDS-Gelelektrophorese identifiziert und
- 15 einer Sequenzanalyse auf einem Applied Biosystems 477 A Protein Sequencer nach Vorschrift des Geräteherstellers unterworfen.

- 20 Folgende aminoterminal Sequenz wurde erhalten (SEQ ID NO: 1):

Glu-Gly-Gly-Glu-Pro-Cys-Ala-Cys-Pro-His-Ala-Leu-His-Arg-Val-Cys-Gly-Ser-Asp

- 25 Beispiel 7

Bestimmung des isoelektrischen Punktes durch isoelektrische Fokussierung

- 30 Die Bestimmung wurde mit einem LKB Multiphor 2117 (Horizontalsystem) und einem LKB powersupply 2103 durchgeführt. Es wurden Fertiggele eingesetzt (Pharmacia Ampholine PAGplate pH 3,5 - 9,5). Als Standardproteine wurden Amyloglucosidase
- 35 pH 3,5; Soyabean Trypsin Inhibitor, pH 4,55; β -Lactoglobulin A, pH 5,2; Bovine carbonic anhydratase, pH 5,85; Human carbonic anhydratase, pH 6,55; Horse myoglobin, pH 6,85 und 7,35; Lentil Lectin, pH 8,15, 8,45, 8,65 und Trypsinogen, pH 9,3 eingesetzt.

- 40

Die Bedingungen der Fokussierung: 1500 Volt, 30 Watt.

Puffer: Anode 1M Phosphorsäure

Kathode 1M Natronlauge

- 5** Die Platten wurden 30 Minuten zur Ausbildung eines pH-Gradienten vorfokussiert. Die Proben wurden auf Filterplättchen aufgetragen, die auf dem Gel lagen. Für 30 Minuten wurde . weiterfokussiert, die Filterplättchen abgenommen, und nach weiteren 30 Minuten wurde die Fokussierung beendet. Die Gele
- 10** wurden sofort in 2 mm Scheiben geschnitten und in dest. Wasser überführt. Über Nacht eluierte das Protein aus den Gelscheiben. Durch einen Thrombininhibitionstest wurde die Lage des Thrombininhibitors bestimmt. Der pH-Wert kann auch direkt durch eine pH-Elektrode bestimmt werden. Hirudin, als
- 15** Vergleichssubstanz hatte einen isoelektrischen Punkt von pH 3,5 und niedriger. Der neue Inhibitor hatte einen isoelektrischen Punkt von pH $4,2 \pm 0,5$.

Beispiel 8

- 20** Herstellung einer DNA-Sequenz, die für ein thrombininhibitorisches Protein codiert.

a) Isolierung von RNA und Herstellung einer cDNA-Bank

- 25** Gesamt-RNA aus ganzen Tieren der Spezies *Rhodnius prolixus* wurde durch Aufschluß in Guanidiniumthiocyanat gewonnen. Dabei wurde mit Materialien und nach Anleitung des "RNA Isolation Kit" der Firma Stratagene, La Jolla, CA, USA (Catalog
- 30** Nr. 200345) gearbeitet. Die polyadenylierte messenger RNA wurde aus der o.a. Gesamt-RNA durch oligo(dT)-Affinitätsseparation selektiert. Dieses Verfahren wurde mit Materialien und nach Anleitung des "PolyATtract mRNA Isolation System" der Firma Promega, Madison, WI, USA (Catalog Nr. Z5200)
- 35** durchgeführt.

- Aus polyadenylierter messenger RNA wurde mit Materialien und nach Anleitung des "ZAP-cDNA Synthesis Kit" der Firma Stratagene, La Jolla, CA, USA (Catalog Nr. 200400) cDNA synthetisiert, die dann mit Materialien und nach Anleitung des
- 40** "Uni-ZAP XR GigapackII Cloning Kit" der Firma Stratagene,

La Jolla, CA, USA (Catalog Nr. 237611) in Lambda Phagen verpackt wurde.

b) Herstellung von Oligonukleotidproben für die PCR

- 5 Zur Klonierung von cDNA-Fragmenten mit Hilfe der Polymerase-Kettenreaktion (PCR, s. "Molecular Cloning", 2nd edition (1989), Sambrook, J. et al., CSH-Press, Seite 14.1 ff.) wurde von Peptiden der in Beispiel 6 beschriebenen aminoter-
- 10 minalen Sequenz (SEQ ID NO: 1) ausgegangen.

Unter Zugrundelegung des genetischen Codes läßt sich aus der Peptidsequenz SEQ ID NO: 18:

- 15 NH₂-Glu Gly Gly Glu Pro Cys Ala Cys Pro His Ala (Pos. 1-11)

die Nukleinsäuresequenz SEQ ID NO: 19:

5'- GAA GGT GGT GAA CCN TGY GCN TGY CCN CAY GC -3'

- 20 des kodierenden DNA-Stranges herleiten. Wegen der bekannten Degeneration des genetischen Codes sind an manchen Positionen mehrere Nukleotide (N: A, C, G, T; Y: C, T;) einsetzbar. Damit ergibt sich während der Oligonukleotid-Synthese eine
- 25 Komplexität von 512 verschiedenen Oligonukleotiden.

- Die Synthesen wurden mit einem Applied Biosystems Typ 360A DNA-Synthesizer durchgeführt. Die Oligonukleotide wurden nach Entfernung der Schutzgruppen gelelektrophoretisch über
- 30 ein Acrylamid/Harnstoff-Gel gereinigt.

c) Herstellung von DNA-templates für die PCR

- 5 µg Gesamt-RNA oder 1 µg Poly(A)⁺-RNA aus der unter a) auf-
- 35 geführten RNA-Präparation wurden mit dem Oligonukleotid A-B-T₁₈, (SEQ ID NO: 20):

5'-CGAGGGGGATGGTCGACGGAAGCGACCTTTTTTTTTTTTTTTTTT-3'

- 40 und mit Hilfe des Enzyms Reverse Transkriptase in einzelsträngige cDNA (1^ocDNA) übersetzt. Dabei wurde mit Materialien und nach Anleitung des "SuperScript Preamplification System" der Firma Gibco BRL, Eggenstein, Deutschland (Cata-

log Nr. 8089SA) gearbeitet. Nach Beendigung der Reaktion wurden die niedermolekularen Bestandteile über Biospin-30-Säulen der Firma BioRad, Richmond, CA, USA (Catalog Nr. 732-6006) abgetrennt.

5

d) PCR und Klonierung

Die Polymerase-Kettenreaktion wurde nach bekannten Protokollen durchgeführt (s. "Molekular Cloning", 2nd edition
10 (1989), Sambrook, J. et al., CSH-Press, Seite 14.1 ff.). Dazu wurde ein "DNA Thermal Cycler" der Firma Perkin Elmer benutzt. Dabei wurde das Prinzip der "verschachtelten Primer" nach Frohmann, M.A. et al. (Proc. Natl. Acad. Sci., USA (1988) 85, 8998-9002) modifiziert angewendet.

15

Im einzelnen wurde die 1^ocDNA aus c) mit den Oligonukleotiden SEQ ID NO: 19, s.o. und SEQ ID NO: 21 (aus A-B-T₁₈):

5'-CGAGGGGGATGGTCGACGG-3'

20

amplifiziert. Die PCR-Produkte wurden gelelektrophoretisch aufgetrennt und größenfraktioniert. Separierte Agarose-scheibchen mit DNA-Fragmenten steigender Molmasse wurden dann in eine zweite PCR mit Oligonukleotid SEQ ID NO: 19,
25 s.o. und SEQ ID NO: 22 (aus A-B-T₁₈):

5'-GATGGTCGACGGAAGCGACC-3'

30

eingesetzt. Die solchermaßen selektierten PCR-Produkte wurden ebenfalls gelelektrophoretisch aufgetrennt und nach Standardmethoden eluiert. Nach Subklonierung in die EcoRV-Schnittstelle des Vektors pBluescriptKS und Vermehrung des Plasmides in E. coli DH5alpha ergab die Sequenzanalyse eines
35 Klonen mit der SEQ ID NO: 2 ein offenes Leseraster von 103 Aminosäuren mit der SEQ ID NO: 3, das das in SEQ ID NO: 1 beschriebene Peptid beinhaltet. Der Klon wurde daraufhin pRPTI genannt.

40

e) Screening der cDNA-Bank

1x10⁶ rekombinante Phagen der Rhodnius prolixus cDNA-Bank wurden einem Screening mit einer SEQ ID NO: 2 entsprechenden Probe unterzogen. Dabei wurde nach bekannten Protokollen ge-

arbeitet (s. "Molekular Cloning", 2nd edition (1989), Sambrook, J. et al., CSH-Press). Die position Klone wurden sequenzanalysiert und sind als SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16 aufgeführt.

f) Heterologe Expression des Thrombininhibitors

Zur Herstellung des rekombinanten Thrombininhibitors wurde zunächst die cDNA Sequenz SEQ ID NO: 2 kodierend für die Aminosäuresequenz SEQ ID NO: 3 von Position 1 bis 103 mit Hilfe beschriebener Methoden PCR-amplifiziert und mit passenden Enden in die XmnI und BamHI Schnittstelle des bakteriellen Expressionsvektors pMAL-p2 (Protein Fusion and Purification System, GeneExpress, New England Biolabs Nr. 800) kloniert. Diese Klonierung führt zu einer Fusion des bakteriellen Maltose-Bindungsproteins mit dem beschriebenen Thrombininhibitor im gleichen Leseraster. E. coli DH5alpha Zellen wurden mit dem entstandenen Plasmid transformiert und das rekombinante Fusionsprotein wurde nach Angaben des Herstellers exprimiert und gereinigt. Unter Ausnutzung einer Faktor Xa-Schnittstelle an der Fusionsstelle kann der Thrombininhibitor vom Fusionspartner enzymatisch getrennt werden. Die Ausbeute des in E. coli periplasmatisch exprimierten Thrombininhibitors liegt bei 6000 Units/L Kultur (Standard: National Institute of Health).

30

35

40

17

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:

- (A) NAME: BASF Aktiengesellschaft
- (B) STREET: Carl-Bosch-Strasse 38
- (C) CITY: Ludwigshafen
- (E) COUNTRY: Bundesrepublik Deutschland
- (F) POSTAL CODE (ZIP): D-6700
- (G) TELEPHONE: 0621/6048526
- (H) TELEFAX: 0621/6043123
- (I) TELEX: 1762175170

(ii) TITLE OF INVENTION: Neue thrombininhibitorische Proteine
aus Raubwanzen

(iii) NUMBER OF SEQUENCES: 22

(iv) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) OPERATING SYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

(v) CURRENT APPLICATION DATA:

APPLICATION NUMBER:

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: N-terminal

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Rhodnius prolixus
- (D) DEVELOPMENTAL STAGE: adult
- (F) TISSUE TYPE: whole animals

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..19

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

18

Glu Gly Gly Glu Pro Cys Ala Cys Pro His Ala Leu His Arg Val Cys
 1 5 10 15
 Gly Ser Asp

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 350 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Rhodnius prolixus*
- (D) DEVELOPMENTAL STAGE: adult
- (F) TISSUE TYPE: whole animals

(vii) IMMEDIATE SOURCE:

- (B) CLONE: pBSKS-/RPTI#5.6

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..309

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

GAA GGT GGT GAA CCG TGC GCT TGC CCC CAC GCT CTG CAT AGA GTT TGC	48
Glu Gly Gly Glu Pro Cys Ala Cys Pro His Ala Leu His Arg Val Cys	
1 5 10 15	
GGC TCT GAT GGT GAA ACT TAT AGC AAC CCT TGT ACG CTG AAC TGT GCT	96
Gly Ser Asp Gly Glu Thr Tyr Ser Asn Pro Cys Thr Leu Asn Cys Ala	
20 25 30	
AAA TTC AAT GGA AAG CCA GAA CTT GTA AAA GTC CAT GAT GGT CCT TGC	144
Lys Phe Asn Gly Lys Pro Glu Leu Val Lys Val His Asp Gly Pro Cys	
35 40 45	
GAA CCG GAT GAG GAT GAA GAT GTT TGC CAA GAA TGT GAT GGT GAT GAA	192
Glu Pro Asp Glu Asp Glu Asp Val Cys Gln Glu Cys Asp Gly Asp Glu	
50 55 60	
TAC AAA CCA GTT TGC GGA TCT GAC GAC ATA ACT TAC GAT AAC AAC TGT	240
Tyr Lys Pro Val Cys Gly Ser Asp Asp Ile Thr Tyr Asp Asn Asn Cys	
65 70 75 80	

- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Rhodnius prolixus
 - (D) DEVELOPMENTAL STAGE: adult
 - (F) TISSUE TYPE: whole animals
- (vii) IMMEDIATE SOURCE:
 - (A) LIBRARY: R.prolixus unizAP/XR
 - (B) CLONE: unizAP#3
- (ix) FEATURE:
 - (A) NAME/KEY: CDS
 - (B) LOCATION: 3..275
- (ix) FEATURE:
 - (A) NAME/KEY: 3'UTR
 - (B) LOCATION: 276..447

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

AA ACT TAT AGC AAC CCT TGT ACG CTG AAC TGT GCT AAA CAC AAT GGA	47
Thr Tyr Ser Asn Pro Cys Thr Leu Asn Cys Ala Lys His Asn Gly	
1 5 10 15	
AAG CCA GGT CTT GTA AAA GTC CAT GAT GGT CCT TGC GAA CCG GAT GAG	95
Lys Pro Gly Leu Val Lys Val His Asp Gly Pro Cys Glu Pro Asp Glu	
20 25 30	
GAT GAA GAT GTT TGC CAA GAA TGT GAT GAT GTC GAT TAC GAA CCA GTT	143
Asp Glu Asp Val Cys Gln Glu Cys Asp Asp Val Asp Tyr Glu Pro Val	
35 40 45	
TGT GGA ACT GAC GAC AAA ACT TAC GAT AAC AAC TGT CGA CTA GAG TGT	191
Cys Gly Thr Asp Asp Lys Thr Tyr Asp Asn Asn Cys Arg Leu Glu Cys	
50 55 60	
GCC TCT ATC TCT TCC AGC CCA GGA CTT GAA CTG AAG CAC ACA GGA AAA	239
Ala Ser Ile Ser Ser Ser Pro Gly Leu Glu Leu Lys His Thr Gly Lys	
65 70 75	
TGT CTA CCC CAT TTG GAT TTT CCC GAC CCA GTT TAAAGCTTGC ACATAACGGA	292
Cys Leu Pro His Leu Asp Phe Pro Asp Pro Val	
80 85 90	
AAATGCACTA TAGCAGAGTT ATATCACGGT TTATTGTAAA AAAAGATTAT ATGAATTTAT	352
CATAATATCA ATAAAATAGC TTATTTTAAA AATATTGAAC CAATTTAAAT TTTCAACATA	412
TGTATATGTA AATAAATTTA AAAAAAAAAA AAAAA	447

(2) INFORMATION FOR SEQ ID NO:5:

21

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 90 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

```

Thr Tyr Ser Asn Pro Cys Thr Leu Asn Cys Ala Lys His Asn Gly Lys
 1           5           10           15
Pro Gly Leu Val Lys Val His Asp Gly Pro Cys Glu Pro Asp Glu Asp
           20           25           30
Glu Asp Val Cys Gln Glu Cys Asp Asp Val Asp Tyr Glu Pro Val Cys
           35           40           45
Gly Thr Asp Asp Lys Thr Tyr Asp Asn Asn Cys Arg Leu Glu Cys Ala
 50           55           60
Ser Ile Ser Ser Ser Pro Gly Leu Glu Leu Lys His Thr Gly Lys Cys
 65           70           75           80
Leu Pro His Leu Asp Phe Pro Asp Pro Val
           85           90

```

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 732 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Rhodnius prolixus
- (D) DEVELOPMENTAL STAGE: adult
- (F) TISSUE TYPE: whole animals

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: R.prolixus uniZAP/XR
- (B) CLONE: uniZAP#3.2

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 3..557

23

ATG AGA GAA GTT CCA GGA CTT GAA CTG AAG CAC ACA GGA AAA TGT CTA	527
Met Arg Glu Val Pro Gly Leu Glu Leu Lys His Thr Gly Lys Cys Leu	
150 155 160	
CCC CAT TTG GAT TTT CCC GAC CCA GTT TAAAGCTTGC ACATAACGGA	574
Pro His Leu Asp Phe Pro Asp Pro Val	
165 170	
AAATGCACTA TAGCAGAGTT ATATCACGGT TTATTGTAAA AAAAGATTAT ATGAATTTAT	634
CATAATATCA ATAAAATAGC TTATTTTAAA AATATTGAAC CAATTTAAAT TTTCAACATA	694
TGTATATGTA AATAAATTTA AAAAAAAAAA AAAAAAAAAA	732

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 184 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

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

Arg Leu Leu Leu Leu Leu Gly Leu Ala Ala Leu Val Ala Ala Glu Gly	
-14 -10 -5 1	
Gly Glu Pro Cys Ala Cys Pro His Ala Leu His Arg Val Cys Gly Ser	
5 10 15	
Asp Gly Glu Thr Tyr Ser Asn Pro Cys Thr Leu Asn Cys Ala Lys Phe	
20 25 30	
Asn Gly Lys Pro Glu Leu Val Lys Val His Asp Gly Pro Cys Glu Pro	
35 40 45 50	
Asp Glu Asp Glu Asp Val Cys Gln Glu Cys Asp Gly Asp Glu Tyr Lys	
55 60 65	
Pro Val Cys Gly Ser Asp Gly Ile Thr Tyr Asp Asn Asn Cys Arg Leu	
70 75 80	
Glu Cys Ala Ser Ile Ser Ser Ser Pro Gly Val Glu Leu Lys His Glu	
85 90 95	
Gly Ile Cys Arg Lys Glu Glu Lys Lys Leu Pro Lys Arg Ser Val Gly	
100 105 110	
Leu Glu His Thr Cys Val Cys Pro Tyr Asn Tyr Phe Pro Val Cys Gly	
115 120 125 130	
Thr Asp Gly Glu Thr Tyr Pro Asn Leu Cys Ala Leu Gln Cys Arg Met	
135 140 145	

24

Arg Glu Val Pro Gly Leu Glu Leu Lys His Thr Gly Lys Cys Leu Pro
 .50 155 160

His Leu Asn Phe Pro Asp Pro Val
 165 170

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 677 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Rhodnius prolixus*
- (D) DEVELOPMENTAL STAGE: adult
- (F) TISSUE TYPE: whole animals

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: R.prolixus unizAP/XR
- (B) CLONE: unizAP#5

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..561

(ix) FEATURE:

- (A) NAME/KEY: 3'UTR
- (B) LOCATION: 562..677

(ix) FEATURE:

- (A) NAME/KEY: sig_peptide
- (B) LOCATION: 1..48

(ix) FEATURE:

- (A) NAME/KEY: mat_peptide
- (B) LOCATION: 49..561

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

ATG AAG CGT CTA CTG TTG TTA CTC GGA TTG GCT GCA CTC GTT GCA GCT
 Met Lys Arg Leu Leu Leu Leu Gly Leu Ala Ala Leu Val Ala Ala
 -16 -15 -10 -5

25

GAA GGG GGG GAA CCA TGC GCA TGT CCA CAT GCT CTG CAT AGA GTT TGC	96
Glu Gly Gly Glu Pro Cys Ala Cys Pro His Ala Leu His Arg Val Cys	
1 5 10 15	
GGC TCT GAT GGT GAA ACT TAT AGC AAC CCT TGT ACG CTG AAC TGT GCT	144
Gly Ser Asp Gly Glu Thr Tyr Ser Asn Pro Cys Thr Leu Asn Cys Ala	
20 25 30	
AAA TTT AAT GGA AAG CCA GAA CTT GTA AAA GTC CAT GAT GGT CCT TGC	192
Lys Phe Asn Gly Lys Pro Glu Leu Val Lys Val His Asp Gly Pro Cys	
35 40 45	
GAA CCG GAT GAG GAT GAA GAT GTT TGC CAA GAA TGT GAT GGT GAT GAA	240
Glu Pro Asp Glu Asp Glu Asp Val Cys Gln Glu Cys Asp Gly Asp Glu	
50 55 60	
TAC AAA CCA GTT TGC GGA TCT GAC GGC ATA ACT TAC GAT AAC AAC TGT	288
Tyr Lys Pro Val Cys Gly Ser Asp Gly Ile Thr Tyr Asp Asn Asn Cys	
65 70 75 80	
CGA CTA GAG TGT GCC TCT ATC TCT TCC AGC CCA GGA GTT GAA CTG AAA	336
Arg Leu Glu Cys Ala Ser Ile Ser Ser Ser Pro Gly Val Glu Leu Lys	
85 90 95	
CAT GAA GGG ATA TGT AGA AAG GAG GAA AAG AAA CTT CCT AAA AGA TCT	384
His Glu Gly Ile Cys Arg Lys Glu Glu Lys Lys Leu Pro Lys Arg Ser	
100 105 110	
GTG GGA TTG GAA CAT ACA TGC GTC TGT CCT TAT AAT TAT TTC CCG GTT	432
Val Gly Leu Glu His Thr Cys Val Cys Pro Tyr Asn Tyr Phe Pro Val	
115 120 125	
TGC GGA ACA GAT GGG GAA ACC TAT CCC AAC TTG TGC GCC CTC CAA TGT	480
Cys Gly Thr Asp Gly Glu Thr Tyr Pro Asn Leu Cys Ala Leu Gln Cys	
130 135 140	
CGT ATG AGA GAA GTT CCA GGA CTT GAA CTG AAG CAC ACA GGA AAA TGT	528
Arg Met Arg Glu Val Pro Gly Leu Glu Leu Lys His Thr Gly Lys Cys	
145 150 155 160	
CTA CCC CAT TTG GAT TTT CCC GAC CCA GTT TAAAGCTTGC ACATAACGGA	578
Leu Pro His Leu Asp Phe Pro Asp Pro Val	
165 170	
AAATGCACTA TAGCAGAGTT ATATCACGGT TTATTGTAAA AAAAGGATTA TATGAATTTA	638
TCATAATATC AATAAAATAG CTTATTTTAA AAATATTGA	677

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 186 amino acids

26

(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

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

Met Lys Arg Leu Leu Leu Leu Leu Gly Leu Ala Ala Leu Val Ala Ala
-16 -15 -10 -5

Glu Gly Gly Glu Pro Cys Ala Cys Pro His Ala Leu His Arg Val Cys.
1 5 10 15

Gly Ser Asp Gly Glu Thr Tyr Ser Asn Pro Cys Thr Leu Asn Cys Ala
20 25 30

Lys Phe Asn Gly Lys Pro Glu Leu Val Lys Val His Asp Gly Pro Cys
35 40 45

Glu Pro Asp Glu Asp Glu Asp Val Cys Gln Glu Cys Asp Gly Asp Glu
50 55 60

Tyr Lys Pro Val Cys Gly Ser Asp Gly Ile Thr Tyr Asp Asn Asn Cys
65 70 75 80

Arg Leu Glu Cys Ala Ser Ile Ser Ser Ser Pro Gly Val Glu Leu Lys
85 90 95

His Glu Gly Ile Cys Arg Lys Glu Glu Lys Lys Leu Pro Lys Arg Ser
100 105 110

Val Gly Leu Glu His Thr Cys Val Cys Pro Tyr Asn Tyr Phe Pro Val
115 120 125

Cys Gly Thr Asp Gly Glu Thr Tyr Pro Asn Leu Cys Ala Leu Gln Cys
130 135 140

Arg Met Arg Glu Val Pro Gly Leu Glu Leu Lys His Thr Gly Lys Cys
145 150 155 160

Leu Pro His Leu Asp Phe Pro Asp Pro Val
165 170

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 566 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

28

CCA GTT TAAAGCTTGC ACATAACGGA AAATGCACTA TAGCAGAGGT ATATCACGGT 438
 Pro Val
 130
 TTATTGTAAA AAAAGGTTAT ATGAATTTAT CATAATATCA ATTAAAATAG CTTATTTTAA 498
 AAATATTGCC CATTTAAATT TTCAACATAT GTATATGTAA ATAAATTTAA AAAAAAAAAA 558
 AAAAAAAAAA 566

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 129 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

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

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

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 327 base pairs
 - (B) TYPE: nucleic acid

29

(C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: cdna to mRNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Rhodnius prolixus
 - (D) DEVELOPMENTAL STAGE: adult
 - (F) TISSUE TYPE: whole animals
- (vii) IMMEDIATE SOURCE:
 - (A) LIBRARY: R.prolixus unizAP/XR
 - (B) CLONE: unizAP#11
- (ix) FEATURE:
 - (A) NAME/KEY: CDS
 - (B) LOCATION: 1..327
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GTC GAT TAC GAA CCA GTT TGT GGA ACT GAC GAC AAA ACT TAC GAT AAC	48
Val Asp Tyr Glu Pro Val Cys Gly Thr Asp Asp Lys Thr Tyr Asp Asn	
1 5 10 15	
AAC TGT CGA CTA GAG TGT GCC TCT ATC TCT TCC AGC CCA GGA GTT GAA	96
Asn Cys Arg Leu Glu Cys Ala Ser Ile Ser Ser Ser Pro Gly Val Glu	
20 25 30	
CTG AAA CAT GAA GGG ATA TGT AGA AAG GAG GAA AAG AAA CTT CCT AAA	144
Leu Lys His Glu Gly Ile Cys Arg Lys Glu Glu Lys Lys Leu Pro Lys	
35 40 45	
AGA TCT GTG GGA TTG GAA CAT ACA TGC GTC TGT CCT TAT AAT TAT TTC	192
Arg Ser Val Gly Leu Glu His Thr Cys Val Cys Pro Tyr Asn Tyr Phe	
50 55 60	
CCG GTT TGC GGA ACA GAT GGG GAA ACC TAT CCC AAC TTG TGC GGC CTC	240
Pro Val Cys Gly Thr Asp Gly Glu Thr Tyr Pro Asn Leu Cys Gly Leu	
65 70 75 80	
CAA TGT CGT ATG AGA GAA GTT CCA GGA CTT GAA CTG AAG CAC ACA GGA	288
Gln Cys Arg Met Arg Glu Val Pro Gly Leu Glu Leu Lys His Thr Gly	
85 90 95	
AAA TGT CTA CCC CAT TTG GAT TTT CCC GAC CCA GTT TA	327
Lys Cys Leu Pro His Leu Asp Phe Pro Asp Pro Val	
100 105	

(2) INFORMATION FOR SEQ ID NO:13:

30

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 108 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

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

Val	Asp	Tyr	Glu	Pro	Val	Cys	Gly	Thr	Asp	Asp	Lys	Thr	Tyr	Asp	Asn
1				5					10					15	
Asn	Cys	Arg	Leu	Glu	Cys	Ala	Ser	Ile	Ser	Ser	Ser	Pro	Gly	Val	Glu
			20					25					30		
Leu	Lys	His	Glu	Gly	Ile	Cys	Arg	Lys	Glu	Glu	Lys	Lys	Leu	Pro	Lys
		35					40					45			
Arg	Ser	Val	Gly	Leu	Glu	His	Thr	Cys	Val	Cys	Pro	Tyr	Asn	Tyr	Phe
	50					55					60				
Pro	Val	Cys	Gly	Thr	Asp	Gly	Glu	Thr	Tyr	Pro	Asn	Leu	Cys	Gly	Leu
	65				70					75					80
Gln	Cys	Arg	Met	Arg	Glu	Val	Pro	Gly	Leu	Glu	Leu	Lys	His	Thr	Gly
				85					90					95	
Lys	Cys	Leu	Pro	His	Leu	Asp	Phe	Pro	Asp	Pro	Val				
			100						105						

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 616 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Rhodnius prolixus
- (D) DEVELOPMENTAL STAGE: adult
- (F) TISSUE TYPE: whole animals

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: R.prolixus unizAP/XR
- (B) CLONE: unizAP#12

32

GGA ACA GAT GGG GAA ACC TAT CCC AAC TTG TGC GCC CTC CAA TGT CGT 478
 Gly Thr Asp Gly Glu Thr Tyr Pro Asn Leu Cys Ala Leu Gln Cys Arg
 130 135 140 145

ATG AGA GAA GTT CCA GGA CTT GAA CTG AAG CAC ACA GGA AAA TGT CTA 526
 Met Arg Glu Val Pro Gly Leu Glu Leu Lys His Thr Gly Lys Cys Leu
 150 155 160

CCC CAT TTG GAT TTT CCC GAC CCA GTT TAAAGCTTGC ACATAACGGA 573
 Pro His Leu Asp Phe Pro Asp Pro Val
 165 170

AAATGCACTA TAGCAGAGTT ATATCACGGT TTATTGTAAA AAA 616

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 184 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Arg Leu Leu Leu Leu Leu Gly Leu Ala Ala Leu Val Ala Ala Glu Gly
 -14 -10 -5 1

Gly Glu Pro Cys Ala Cys Pro His Ala Leu His Arg Val Cys Gly Ser
 5 10 15

Asp Gly Glu Thr Tyr Ser Asn Pro Cys Thr Leu Asn Cys Ala Lys Phe
 20 25 30

Asn Gly Lys Pro Glu Leu Val Lys Val His Asp Gly Pro Cys Glu Pro
 35 40 45 50

Asp Glu Asp Glu Asp Val Cys Gln Glu Cys Asp Asp Val Asp Tyr Glu
 55 60 65

Pro Val Cys Gly Thr Asp Asp Lys Thr Tyr Asp Asn Asn Cys Arg Leu
 70 75 80

Glu Cys Ala Ser Ile Ser Ser Ser Pro Gly Val Glu Leu Lys His Glu
 85 90 95

Gly Ile Cys Arg Lys Glu Glu Lys Lys Leu Pro Lys Arg Ser Val Gly
 100 105 110

Leu Glu His Thr Cys Val Cys Pro Tyr Asn Tyr Phe Pro Val Cys Gly
 115 120 125 130

Thr Asp Gly Glu Thr Tyr Pro Asn Leu Cys Ala Leu Gln Cys Arg Met
 135 140 145

33

Arg Glu Val Pro Gly Leu Glu Leu Lys His Thr Gly Lys Cys Leu Pro
 150 155 160

His Leu Asp Phe Pro Asp Pro Val
 165 170

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1282 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cdna to mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Rhodnius prolixus
- (D) DEVELOPMENTAL STAGE: adult
- (F) TISSUE TYPE: whole animals

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: R.prolixus unizAP/XR
- (B) CLONE: unizAP#9

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 3..1109

(ix) FEATURE:

- (A) NAME/KEY: sig_peptide
- (B) LOCATION: 3..44

(ix) FEATURE:

- (A) NAME/KEY: mat_peptide
- (B) LOCATION: 45..1109

(ix) FEATURE:

- (A) NAME/KEY: 3'UTR
- (B) LOCATION: 1110..1282

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

AG CGT CTA CTG TTG TTA CTC GGA TTG GCT GCA CTC GTT GCA GCT GAA
 Arg Leu Leu Leu Leu Leu Gly Leu Ala Ala Leu Val Ala Ala Glu
 -14 -10 -5 1

34

GGG GGG GAA CCA TGC GCA TGT CCA CAT GCT CTG CAT AGA GTT TGC GGC	95
Gly Gly Glu Pro Cys Ala Cys Pro His Ala Leu His Arg Val Cys Gly	
5 10 15	
TCT GAT GGT GAA ACT TAT AGC AAC CCT TGT ACG CTG AAC TGT GCT AAA	143
Ser Asp Gly Glu Thr Tyr Ser Asn Pro Cys Thr Leu Asn Cys Ala Lys	
20 25 30	
TTC AAT GGA AAG CCA GAA CTT GTA AAA GTC CAT GAT GGT CCT TGC GAA	191
Phe Asn Gly Lys Pro Glu Leu Val Lys Val His Asp Gly Pro Cys Glu	
35 40 45	
CCG GAT GAG GAT GAA GAT GTT TGC CAA GAA TGT GAT GGT GAT GAA TAC	239
Pro Asp Glu Asp Glu Asp Val Cys Gln Glu Cys Asp Gly Asp Glu Tyr	
50 55 60 65	
AAA CCA GTT TGC GGA TCT GAC GAC ATA ACT TAC GAT AAC AAC TGT CGA	287
Lys Pro Val Cys Gly Ser Asp Asp Ile Thr Tyr Asp Asn Asn Cys Arg	
70 75 80	
CTA GAG TGT GCC TCT ATC TCT TCC AGC CCA GGA GTT GAA CTG AAA CAT	335
Leu Glu Cys Ala Ser Ile Ser Ser Ser Pro Gly Val Glu Leu Lys His	
85 90 95	
GAA GGA CCT TGT AGA ACC GAG GAA AAG AAA ATT CTT AAA AGA TCT GAT	383
Glu Gly Pro Cys Arg Thr Glu Glu Lys Lys Ile Leu Lys Arg Ser Asp	
100 105 110	
GAA TTC GAA ATG TAT AGA TGC GCA TGT CCG AAA ATA TAT TAT CCG GTT	431
Glu Phe Glu Met Tyr Arg Cys Ala Cys Pro Lys Ile Tyr Tyr Pro Val	
115 120 125	
TGC GGA ACA GAT GGT GAA ACC TAT CCC AAC TTG TGC GTC CTC GAA TGT	479
Cys Gly Thr Asp Gly Glu Thr Tyr Pro Asn Leu Cys Val Leu Glu Cys	
130 135 140 145	
CAT ATG AGA ATG AAT CCA GGA CTT CAA TTG CAC CAT TAT GGA CAT TGT	527
His Met Arg Met Asn Pro Gly Leu Gln Leu His His Tyr Gly His Cys	
150 155 160	
CAA CAT CAT CAT CAC CAT CAT CCT CCT CCT CAT CAC CAT CAT CAT CAT	575
Gln His His His His His His Pro Pro Pro His His His His His His	
165 170 175	
CAT CCT CAT CAC ACC ACT GAG AAA CCA GTA GAA CCA TGC GCA TGT CCA	623
His Pro His His Thr Thr Glu Lys Pro Val Glu Pro Cys Ala Cys Pro	
180 185 190	
CAT GCT CTG CAT AGA GTT TGC GGC TCT GAT GGT GAA ACT TAT AGC AAC	671
His Ala Leu His Arg Val Cys Gly Ser Asp Gly Glu Thr Tyr Ser Asn	
195 200 205	

35

CCT TGT ACG CTG AAC TGT GCT AAA CAC AAT GGA AAG CCA GGT CTT GTA	719
Pro Cys Thr Leu Asn Cys Ala Lys His Asn Gly Lys Pro Gly Leu Val	
210 215 220 225	
AAA GTC CAT GAT GGT CCT TGC GAA CCG GAT GAG GAT GAA GAT GTT TGC	767
Lys Val His Asp Gly Pro Cys Glu Pro Asp Glu Asp Glu Asp Val Cys	
230 235 240	
CAA GAA TGT GAT GAT GTC GAT TAC GAA CCA GTT TGT GGA ACT GAC GAC	815
Gln Glu Cys Asp Asp Val Asp Tyr Glu Pro Val Cys Gly Thr Asp Asp	
245 250 255	
AAA ACT TAC GAT AAC AAC TGT CGA CTA GAG TGT GCC TCT ATC TCT TCC	863
Lys Thr Tyr Asp Asn Asn Cys Arg Leu Glu Cys Ala Ser Ile Ser Ser	
260 265 270	
AGC CCA GGA GTT GAA CTG AAA CAT GAA GGG ATA TGT AGA AAG GAG GAA	911
Ser Pro Gly Val Glu Leu Lys His Glu Gly Ile Cys Arg Lys Glu Glu	
275 280 285	
AAG AAA CTT CCT AAA AGA TCT GTG GGA TTG GAA CAT ACA TGC GTC TGT	959
Lys Lys Leu Pro Lys Arg Ser Val Gly Leu Glu His Thr Cys Val Cys	
290 295 300 305	
CCT TAT AAT TAT TTC CCG GTT TGC GGA ACA GAT GGG GAA ACC TAT CCC	1007
Pro Tyr Asn Tyr Phe Pro Val Cys Gly Thr Asp Gly Glu Thr Tyr Pro	
310 315 320	
AAC TTG TGC GCC CTC CAA TGC CGT ATG AGA GAA GTT CCA GGA CTT GAA	1055
Asn Leu Cys Ala Leu Gln Cys Arg Met Arg Glu Val Pro Gly Leu Glu	
325 330 335	
CTG AAG CAC ACA GGA AAA TGT CTA CCC CAT TTG GAT TTT CCC GAC CCA	1103
Leu Lys His Thr Gly Lys Cys Leu Pro His Leu Asp Phe Pro Asp Pro	
340 345 350	
GTT TAAAGCTTGC ACATAACGGA AAATGCACTA TAGCAGAGTT ATATCACGGT	1156
Val	
355	
TTATTGTAAA AAAAGATTAT ATGAATTTAT CATAATATCA ATAAAATAGC TTATTTTAAA	1216
AATATTGAAC CAATTTAAAT TTTCAACATA TGTATATGTA AATAAATTTA AAAAAAAAAA	1276
AAAAAA	1282

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 368 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

36

(ii) MOLECULE TYPE: protein

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

Arg Leu Leu Leu Leu Leu Gly Leu Ala Ala Leu Val Ala Ala Glu Gly
 -14 -10 -5 1

Gly Glu Pro Cys Ala Cys Pro His Ala Leu His Arg Val Cys Gly Ser
 5 10 15

Asp Gly Glu Thr Tyr Ser Asn Pro Cys Thr Leu Asn Cys Ala Lys Phe.
 20 25 30

Asn Gly Lys Pro Glu Leu Val Lys Val His Asp Gly Pro Cys Glu Pro
 35 40 45 50

Asp Glu Asp Glu Asp Val Cys Gln Glu Cys Asp Gly Asp Glu Tyr Lys
 55 60 65

Pro Val Cys Gly Ser Asp Asp Ile Thr Tyr Asp Asn Asn Cys Arg Leu
 70 75 80

Glu Cys Ala Ser Ile Ser Ser Ser Pro Gly Val Glu Leu Lys His Glu
 85 90 95

Gly Pro Cys Arg Thr Glu Glu Lys Lys Ile Leu Lys Arg Ser Asp Glu
 100 105 110

Phe Glu Met Tyr Arg Cys Ala Cys Pro Lys Ile Tyr Tyr Pro Val Cys
 115 120 125 130

Gly Thr Asp Gly Glu Thr Tyr Pro Asn Leu Cys Val Leu Glu Cys His
 135 140 145

Met Arg Met Asn Pro Gly Leu Gln Leu His His Tyr Gly His Cys Gln
 150 155 160

His His His His His His Pro Pro Pro His His His His His His His
 165 170 175

Pro His His Thr Thr Glu Lys Pro Val Glu Pro Cys Ala Cys Pro His
 180 185 190

Ala Leu His Arg Val Cys Gly Ser Asp Gly Glu Thr Tyr Ser Asn Pro
 195 200 205 210

Cys Thr Leu Asn Cys Ala Lys His Asn Gly Lys Pro Gly Leu Val Lys
 215 220 225

Val His Asp Gly Pro Cys Glu Pro Asp Glu Asp Glu Asp Val Cys Gln
 230 235 240

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

37

Thr Tyr Asp Asn Asn Cys Arg Leu Glu Cys Ala Ser Ile Ser Ser Ser
 260 265 270
 Pro Gly Val Glu Leu Lys His Glu Gly Ile Cys Arg Lys Glu Glu Lys
 275 280 285 290
 Lys Leu Pro Lys Arg Ser Val Gly Leu Glu His Thr Cys Val Cys Pro
 295 300 305
 Tyr Asn Tyr Phe Pro Val Cys Gly Thr Asp Gly Glu Thr Tyr Pro Asn
 310 315 320
 Leu Cys Ala Leu Gln Cys Arg Met Arg Glu Val Pro Gly Leu Glu Leu
 325 330 335
 Lys His Thr Gly Lys Cys Leu Pro His Leu Asp Phe Pro Asp Pro Val
 340 345 350

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 11 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: N-terminal

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Rhodnius prolixus*
- (D) DEVELOPMENTAL STAGE: adult
- (F) TISSUE TYPE: whole animals

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION: 1..11

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

Glu Gly Gly Glu Pro Cys Ala Cys Pro His Ala
 1 5 10

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

38

- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: YES
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: *Rhodnius prolixus*
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

GAAGGTGGTG AACCNTGYGC NTGYCCNCAY GC

32

(2) INFORMATION FOR SEQ ID NO:20:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 45 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: YES
- (iv) ANTI-SENSE: NO
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

CGAGGGGGAT GGTTCGACGGA AGCGACCTTT TTTTTTTTTT TTTT

45

(2) INFORMATION FOR SEQ ID NO:21:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 19 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: YES
- (iv) ANTI-SENSE: NO
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

CGAGGGGGAT GGTTCGACGG

19

(2) INFORMATION FOR SEQ ID NO:22:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 20 base pairs
 - (B) TYPE: nucleic acid

39

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

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

GATGGTCGAC GGAAGCGACC

Patentansprüche

- 5 1. Protein mit thrombininhibitorischer Wirkung aus Raubwanzen der Gattung *Rhodnius*, mit einer Aminosäuresequenz, die aus der Gruppe, die von SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13 und SEQ ID NO: 15 gebildet wird, ausgewählt ist.
- 10 2. Protein mit thrombininhibitorischer Wirkung, enthaltend eine oder mehrere Domänen mit der in SEQ ID NO: 3 angegebenen Aminosäuresequenz.
- 15 3. DNA-Sequenzen, die für ein Protein gemäß Anspruch 1 codieren.
- 20 4. DNA-Sequenzen, die für Proteine mit thrombininhibitorischer Wirkung codieren, und die aus der Gruppe, die von
 - a) DNA-Sequenzen mit der in SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12 und SEQ ID NO: 14 beschriebenen Struktur, und
 - 25 b) DNA-Sequenzen, die unter Standardbedingungen mit DNA Sequenzen a) hybridisieren, gebildet wird,ausgewählt sind.
- 30 5. Expressionsvektor, der eine DNA-Sequenz gemäß Anspruch 3 oder 4 enthält.
- 35 6. Verwendung eines Proteins gemäß Anspruch 1 oder 2 zur Bekämpfung von Krankheiten.
- 40 7. Arzneimittel, enthaltend ein oder mehrere Proteine, die von DNA-Sequenzen gemäß Anspruch 4 codiert werden, und einen weiteren gerinnungshemmenden Faktor.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/EP 92/02450

A. CLASSIFICATION OF SUBJECT MATTER

IPC⁵ C12N15/15; C07K13/00; A61K37/64
 According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC⁵ C07K; C12N; A61K

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

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

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	GB, A, 1 092 421 (NATIONAL RESEARCH DEVELOPMENT CORPORATION) 22 November 1967 see the whole document --- -----	1-7

Further documents are listed in the continuation of Box C.

See patent family annex.

* Special categories of cited documents:

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

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

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

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

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

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

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

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

"&" document member of the same patent family

Date of the actual completion of the international search

06 January 1993 (06.01.93)

Date of mailing of the international search report

22 January 1993 (22.01.93)

Name and mailing address of the ISA/

EUROPEAN PATENT OFFICE

Authorized officer

Facsimile No.

Telephone No.

**ANNEX TO THE INTERNATIONAL SEARCH REPORT
ON INTERNATIONAL PATENT APPLICATION NO. EP 9202450
SA 65889**

This annex lists the patent family members relating to the patent documents cited in the above-mentioned international search report.
The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information. 06/01/93

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
GB-A-1092421		None	

EPO FORM P0079

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82

INTERNATIONALER RECHERCHENBERICHT

Internationales Aktenzeichen

PCT/EP 92/02450

I. KLASSIKATION DES ANMELDUNGSGEGENSTANDS (bei mehreren Klassifikationssymbolen sind alle anzugeben) ⁶		
Nach der Internationalen Patentklassifikation (IPC) oder nach der nationalen Klassifikation und der IPC Int.Kl. 5 C12N15/15; C07K13/00; A61K37/64		
II. RECHERCHIERTE SACHGEBIETE		
Recherchierter Mindestprüfstoff ⁷		
Klassifikationssystem	Klassifikationssymbole	
Int.Kl. 5	C07K ; C12N ; A61K	
Recherchierte nicht zum Mindestprüfstoff gehörende Veröffentlichungen, soweit diese unter die recherchierten Sachgebiete fallen ⁸		
III. EINSCHLAGIGE VERÖFFENTLICHUNGEN ⁹		
Art. ⁹	Kennzeichnung der Veröffentlichung ¹¹ , soweit erforderlich unter Angabe der maßgeblichen Teile ¹²	Betr. Anspruch Nr. ¹³
A	GB,A,1 092 421 (NATIONAL RESEARCH DEVELOPMENT CORPORATION) 22. November 1967 siehe das ganze Dokument -----	1-7
<p>¹⁰ Besondere Kategorien von angegebenen Veröffentlichungen :</p> <p>"A" Veröffentlichung, die den allgemeinen Stand der Technik definiert, aber nicht als besonders bedeutsam anzusehen ist</p> <p>"E" älteres Dokument, das jedoch erst am oder nach dem internationalen Anmeldedatum veröffentlicht worden ist</p> <p>"L" Veröffentlichung, die geeignet ist, einen Prioritätsanspruch zweifelhaft erscheinen zu lassen, oder durch die das Veröffentlichungsdatum einer anderen im Recherchenbericht genannten Veröffentlichung belegt werden soll oder die aus einem anderen besonderen Grund angegeben ist (wie ausgeführt)</p> <p>"O" Veröffentlichung, die sich auf eine mündliche Offenbarung, eine Benutzung, eine Ausstellung oder andere Maßnahmen bezieht</p> <p>"P" Veröffentlichung, die vor dem internationalen Anmeldedatum, aber nach dem beanspruchten Prioritätsdatum veröffentlicht worden ist</p> <p>"T" Spätere Veröffentlichung, die nach dem internationalen Anmeldedatum oder dem Prioritätsdatum veröffentlicht worden ist und mit der Anmeldung nicht kollidiert, sondern nur zum Verständnis des der Erfindung zugrundeliegenden Prinzips oder der ihr zugrundeliegenden Theorie angegeben ist</p> <p>"X" Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann nicht als neu oder auf erfinderischer Tätigkeit beruhend betrachtet werden</p> <p>"Y" Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann nicht als auf erfinderischer Tätigkeit beruhend betrachtet werden, wenn die Veröffentlichung mit einer oder mehreren anderen Veröffentlichungen dieser Kategorie in Verbindung gebracht wird und diese Verbindung für einen Fachmann nahelegend ist</p> <p>"&" Veröffentlichung, die Mitglied derselben Patentfamilie ist</p>		
IV. BESCHEINIGUNG		
Datum des Abschlusses der internationalen Recherche	Absenddatum des internationalen Recherchenberichts	
06. JANUAR 1993	22. 01. 93	
Internationale Recherchenbehörde	Unterschrift des bevollmächtigten Bediensteten	
EUROPAISCHES PATENTAMT	THIELE U.H.	

**ANHANG ZUM INTERNATIONALEN RECHERCHENBERICHT
 ÜBER DIE INTERNATIONALE PATENTANMELDUNG NR.**

EP 9202450
 SA 65889

In diesem Anhang sind die Mitglieder der Patentfamilien der im obengenannten internationalen Recherchenbericht angeführten Patentdokumente angegeben.
 Die Angaben über die Familienmitglieder entsprechen dem Stand der Datei des Europäischen Patentamts am
 Diese Angaben dienen nur zur Unterrichtung und erfolgen ohne Gewähr.

06/01/93

Im Recherchenbericht angeführtes Patentdokument	Datum der Veröffentlichung	Mitglied(er) der Patentfamilie	Datum der Veröffentlichung
GB-A-1092421		Keine	

EPO FORM P0473

Für nähere Einzelheiten zu diesem Anhang : siehe Amtsblatt des Europäischen Patentamts, Nr.12/82