

PCTWORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C12N 15/15, C07K 14/815, 19/00, C12N 15/62, A61K 38/58, C12N 15/70, 15/85, 1/21, 5/16 // (C12N 1/21, C12R 1:19)	A1	(11) International Publication Number: WO 95/03409 (43) International Publication Date: 2 February 1995 (02.02.95)
(21) International Application Number: PCT/EP94/01979 (22) International Filing Date: 17 June 1994 (17.06.94) (30) Priority Data: MI93A001627 22 July 1993 (22.07.93) IT (71) Applicant: FARMITALIA CARLO ERBA S.R.L. [IT/TT]; Via Carlo Imbonati, 24, I-20159 Milan (IT). (72) Inventors: SCACHERI, Emanuela; Via Flore, 99, I-20025 Legnano (IT). VISCO, Carlo; Via Console Marcello, 1, I-20156 Milan (IT). MOLINARI, Antonio; Via Belgirate, 20, I-20125 Milan (IT). GERNA, Marco; Via Palmaria, 1, I-20161 Milan (IT). BOLIS, Giorgio; Via Romagna, 12, I-20015 Parabiago (IT).		(81) Designated States: JP, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE). Published <i>With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>
(54) Title: ANALOGUES OF AN ANTI-THROMBIN POLYPEPTIDE AND PROCESS FOR THEIR PREPARATION (57) Abstract The present invention relates to analogues of anti-thrombin polypeptides and to a process for their preparation. The analogues have been obtained by a C-terminal modification of anti-thrombin polypeptides isolated from the leech <i>Hirudinaria manillensis</i> and are characterized by an improved biological activity in comparison to the polypeptides from which they are derived. The analogues of the invention may be prepared by: 1) isolating the anti-thrombin polypeptide from the tissue or secretions of the leech <i>Hirudinaria manillensis</i> and 2) subsequently modifying the C-terminal end of such polypeptide. Alternatively, the analogues may also be synthesized by recombinant DNA methods. According to this latter aspect, the invention provides DNA sequences, expression vectors and host cell lines for the recombinant preparation of the analogues. The analogues of the invention find a useful application in the treatment of venous thrombosis, vascular shunt occlusion and thrombin-induced disseminated intravascular coagulation.		

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.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgyzstan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LU	Luxembourg	TD	Chad
CS	Czechoslovakia	LV	Latvia	TG	Togo
CZ	Czech Republic	MC	Monaco	TJ	Tajikistan
DE	Germany	MD	Republic of Moldova	TT	Trinidad and Tobago
DK	Denmark	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	US	United States of America
FI	Finland	MN	Mongolia	UZ	Uzbekistan
FR	France			VN	Viet Nam
GA	Gabon				

Title: ANALOGUES OF AN ANTI-THROMBIN POLYPEPTIDE AND PROCESS
FOR THEIR PREPARATION

5 The present invention relates to analogues of anti-thrombin polypeptides and to a process for their preparation. The analogues have been obtained by C-terminal modification of anti-thrombin polypeptides from the leech Hirudinaria manillensis and have an improved biological activity in
10 comparison to the polypeptides from which they are derived.

 The most popular anticoagulant peptides are probably those belonging to the family of hirudins. Hirudin, originally isolated from the medicinal leech, Hirudo medicinalis, is a well known and well characterized
15 polypeptidic inhibitor of thrombin^{1,2}. More particularly, it binds thrombin by ionic interactions thus preventing the cleavage of fibrinogen to fibrin and the subsequent fibrin-clot formation. In animal studies hirudin has demonstrated efficacy in preventing venous thrombosis, vascular shunt
20 occlusion and thrombin-induced disseminated intravascular coagulation. In addition, hirudin exhibits low toxicity, little or no antigenicity and a very short clearance time from circulation.³

 Polypeptides with anticoagulant properties have been
25 isolated from a different leech species, Hirudinaria manillensis (EP-A-0347376 and WO 90/05143). This leech is evolutionarily more advanced than Hirudo medicinalis and could therefore synthesize anticoagulant peptides whose amino acid sequences may be different from those of hirudin and
30 other known hirudin variants.

 We have analysed a preparation obtained from Hirudinaria manillensis leeches. We have found three polypeptides having anti-thrombin activity, characterized by the following amino acid sequences (European Patent Application No. 92301721.4)

35 P1 (SEQ ID NO: 1)

Val Ser Tyr Thr Asp Cys Thr Glu Ser Gly Gln Asn Tyr Cys Leu

1

5

10

15

- 2 -

Cys Val Gly Gly Asn Leu Cys Gly Gly Gly Lys His Cys Glu Met
 20 25 30
 Asp Gly Ser Gly Asn Lys Cys Val Asp Gly Glu Gly Thr Pro Lys
 5 35 40 45
 Pro Lys Ser Gln Thr Glu Gly Asp Phe Glu Glu Ile Pro Asp Glu
 50 55 60
 Asp Ile Leu Asn;
 64
 10 P2 (SEQ ID NO: 2)
 Val Ser Tyr Thr Asp Cys Thr Glu Ser Gly Gln Asn Tyr Cys Leu
 1 5 10 15
 Cys Val Gly Ser Asn Val Cys Gly Glu Gly Lys Asn Cys Gln Leu
 20 25 30
 15 Ser Ser Ser Gly Asn Gln Cys Val His Gly Glu Gly Thr Pro Lys
 35 40 45
 Pro Lys Ser Gln Thr Glu Gly Asp Phe Glu Glu Ile Pro Asp Glu
 50 55 60
 Asp Ile Leu Asn; or
 20 64
 P3 (SEQ ID NO: 3)
 Val Ser Tyr Thr Asp Cys Thr Glu Ser Gly Gln Asn Tyr Cys Leu
 1 5 10 15
 Cys Val Gly Ser Asn Val Cys Gly Glu Gly Lys Asn Cys Gln Leu
 25 20 25 30
 Ser Ser Ser Gly Asn Gln Cys Val His Gly Glu Gly;
 35 40

Amino acid residues are presented according to the three letter code (Eur. J. Biochem. 138, 9-37, 1984).

30 The invention provides a polypeptide comprising a sequence represented by formula (I)



wherein P_{1-63} is the amino acid sequence from position 1 to position 63 of the polypeptide P1 (SEQ ID NO: 1) or P2 (SEQ

35 ID NO: 2), Xaa is any amino acid residue and Z is -OH, -NH₂, -Gly-OH or a dipeptide consisting of:

- 3 -

- a glycine or a polar amino acid residue, followed by an amino acid selected from the group consisting of: alanine, glycine, valine, leucine, isoleucine, proline, cysteine, methionine, serine, threonine, lysine, arginine, histidine, asparagine, glutamine, aspartic acid and glutamic acid, with the proviso that when Z is -OH, -NH₂ or -Gly-OH, Xaa can not be asparagine
and a pharmaceutically acceptable salt thereof.
- 10 The pharmaceutically acceptable salts of the polypeptides may be acid addition salts. They may be salts with an inorganic acid such as a hydrohalic acid, such as hydrochloric acid; sulphuric acid; phosphoric acid; or pyrophosphoric acid. The salts may be salts with an organic acid such as
- 15 benzenesulphonic, p-toluenesulphonic, methanesulphonic, acetic, lactic, palmitic, stearic, malic, tartaric, ascorbic or citric acid. The analogues also contain free carboxyl groups and may therefore be present as sodium, calcium, potassium, magnesium or ammonium salts or salts with a
- 20 physiologically tolerable organic nitrogen-containing base. The analogues can also be in the form of inner salts.

The polypeptides of the invention are able to inhibit the proteolytic effects of thrombin, for example fibrin formation from fibrinogen and the activation of factor V and

25 factor VIII. The polypeptides are therefore useful, for example, as anticoagulant and antithrombotic drugs, for the treatment of thromboembolic pathologies such as disseminate intravascular coagulation, cerebral embolisms, deep venous thrombosis and arterial thrombosis. They are also useful in

30 prophylactic treatment of coagulation disorders, for example for the therapy of acute myocardial infarction.

The polypeptides of the invention consist essentially of the sequence P₁₋₆₃-Xaa-Z. However, the polypeptides may be preceded by all or part, typically a C-terminal part, of a

35 leader sequence. The leader sequence may be a native leader sequence of a natural polypeptide from which the polypeptides

- 4 -

of the invention are derived, or may be a foreign leader sequence. The leader sequence may be capable of directing
5 secretion of the polypeptides of the invention. Two of the natural anti-thrombin polypeptides are expressed with a leader sequence which is cleaved subsequently. All or part of this leader sequence may therefore be present in the polypeptide of the invention, the sequence being:

10 Met Phe Ser Leu Lys Leu Phe Val Val Phe Leu Ala Val Cys Ile
Cys Val Ser Gln Ala (SEQ ID NO: 32).

A polypeptide according to the invention, or a salt thereof, may be prepared by: 1) isolating an anti-thrombin polypeptide from the tissue or secretions of a leech of the
15 species Hirudinaria manillensis and 2) subsequently modifying the C-terminal end of such a polypeptide. More specifically, the polypeptide can be obtained by obtaining a preparation of leech extract, subjecting the preparation to high pressure liquid chromatography, and modifying the C-terminal end of
20 the polypeptide.

A polypeptide according to the invention or a salt thereof can then be prepared by:

(a) providing a host, transformed with an expression vector comprising a DNA sequence encoding a said polypeptide,
25 under such conditions that the said polypeptide is expressed;

(b) isolating the said polypeptide thus obtaining a polypeptide of the invention wherein X is -OH or Gly-OH or a pharmaceutically acceptable salt thereof, and if desired,

(c) converting the polypeptide in which Z is -OH or Gly-
30 OH into the corresponding C-terminal amidated derivative by way of an amidating enzyme, so obtaining a polypeptide of the invention wherein Z is -NH₂ or a pharmaceutically acceptable salt thereof.

This approach is typically based on obtaining a
35 nucleotide sequence encoding the polypeptide it is wished to express and expressing the polypeptide in a recombinant

- 5 -

organism. The cultivation of the genetically modified organism leads to the production of the desired product displaying biological activity. The present invention therefore further provides

- an expression vector comprising a DNA sequence encoding a polypeptide of the invention;
- a host transformed with a compatible expression vector according to the invention; and
- 10 - a DNA sequence encoding a polypeptide according to the invention.

A host in which a polypeptide according to the invention is able to be expressed is prepared by transforming a host with a compatible expression vector of the invention. The expression vector can be prepared by:

(a) chemically synthesising a DNA sequence encoding a polypeptide of the invention; and

(b) inserting the said DNA into an expression vector.

Alternatively, an expression vector can be prepared by:

20 (a) producing and isolating a cDNA encoding the polypeptide P1 or P2 from mRNA of a leech of the species Hirudinaria manillensis;

(b) mutagenising the isolated cDNA so as to obtain a nucleotide sequence encoding a polypeptide of the invention; and

25 (c) inserting the said nucleotide sequence into an expression vector.

A polypeptide according to the invention is consequently prepared by providing a transformed host under such conditions that the polypeptide is expressed. When a eucaryotic host is employed, the polypeptide can be obtained glycosylated. The polypeptide can be isolated as such or in the form of a pharmaceutically acceptable salt. In this way, a polypeptide or salt according to the invention may be

35 obtained in essentially pure form.

The polypeptides of the invention may be modified by way

- 6 -

of amino acid extension at either or each end. A polypeptide composed of such an extended sequence must of course still exhibit anti-thrombin activity. For example, a short
5 sequence of up to 30 amino acid residues may be provided at either or each terminus.

The polypeptides of the invention may be subjected to one or more post-translational modification such as sulphation, COOH- amidation, acylation or chemical alteration
10 of the polypeptide chain. For example a polypeptide having a glycine residue at its carboxy terminus may be subjected to enzymatic amidation with peptidyl-glycine α -amidating monooxygenase (PAM enzyme). PAM enzyme catalyzes the C-terminal oxydative cleavage of glycine-extended peptides to
15 give the corresponding amidated product and glyoxylate⁴.

The polypeptide P_{1-63} -Val-NH₂ constitutes a particular embodiment of the present invention. P_{1-63} -Val-NH₂ is the polypeptide of formula (I) wherein Xaa is Val and Z is -NH₂. The polypeptide may be obtained by amidating the polypeptide
20 of formula P_{1-63} -Val-Gly-OH.

The C-terminal α -amide structure contributes to the biological stability of the amidated polypeptide conferring protection from the carboxypeptidase-mediated degradation, increases its in vivo half-life and confers a more selective
25 receptor binding⁵. For these reasons, α -amidation represents an important and characteristic post-translational processing of many secreted peptides in eukaryotic organisms.

The amidating activity of PAM enzyme has been shown to be dependent upon the presence of molecular oxygen and copper
30 ions and to be stimulated by cofactors such as ascorbate, catalase or potassium iodide⁶.

As an alternative to the use of PAM enzyme, other amidating enzymes may be used. For example, carboxypeptidase Y and, more generally the class of carboxypeptidases has been proved
35 to be an efficient amidating enzyme^{7,8}. Contrary to the PAM enzyme, carboxypeptidases are able to amidate polypeptides

- 7 -

having any amino acid residue at their C-terminal end.

In order to produce a polypeptide of the invention by recombinant DNA technology, a gene encoding a said
5 polypeptide is prepared. The invention includes DNA sequences consisting essentially of the following sequences:

GTT TCT TAC ACC GAC TGC ACC GAA TCT GGC CAG AAC TAC TGC CTG
TGC GTT GGT TCT AAC GTT TGC GGT GAA GGT AAA AAC TGC CAG CTG
10 TCT TCT TCT GGT AAC CAG TGC GTT CAC GGT GAA GGT ACC CCG AAA
CCG AAA TCT CAG ACT GAA GGT GAC TTC GAA GAA ATT CCG GAC GAA
GAC ATC CTG GTT GGT TAG (SEQ ID NO: 16);

GTT TCT TAC ACC GAC TGC ACC GAA TCT GGC CAG AAC TAC TGC CTG
15 TGC GTT GGT GGT AAC CTG TGC GGT GGT GGT AAA CAC TGC GAA ATG
GAT GGT TCT GGT AAC AAA TGC GTT GAT GGT GAA GGT ACC CCG AAA
GCG AAA TCT GAG ACT GAA GGT GAC TTC GAA GAA ATT CCG GAC GAA
GAC ATC CTG GTT GGT TAG (SEQ ID NO: 17);

GTT TCT TAC ACC GAC TGC ACC GAA TCT GGC CAG AAC TAC TGC CTG
20 TGC GTT GGT TCT AAC GTT TGC GGT GAA GGT AAA AAC TGC CAG CTG
TCT TCT TCT GGT AAC CAG TGC GTT CAC GGT GAA GGT ACC CCG AAA
CCG AAA TCT CAG ACT GAA GGT GAC TTC GAA GAA ATT CCG GAC GAA
GAC ATC CTG AAC GGT GCT TAG TAA (SEQ ID NO: 33).

25

The DNA coding sequence typically does not include introns. The DNA sequence is isolated and purified. The gene is inserted in an expression vector able to drive production of the recombinant product. The DNA sequence may be a
30 mutagenized cDNA sequence. The DNA sequence may be a synthetic DNA sequence. The synthetic gene is typically prepared by chemically synthesising oligonucleotides which, in total, correspond to the desired gene. The oligonucleotides are then assembled to obtain the gene.

35 A gene may therefore be constructed from six chemically synthesised oligonucleotides, each oligonucleotide

- 8 -

representing about one third of one strand of a double-stranded DNA gene. The oligonucleotides are ligated and annealed to obtain the desired gene. Typically, a gene is
 5 constructed with restriction sites at each end to facilitate its subsequent manipulation.

A DNA sequence may be provided which further encodes a leader peptide as mentioned above. The leader peptide is capable of directing secretion of the polypeptide from cells
 10 in which the polypeptide is to be expressed. The sequence encoding the leader peptide is typically fused to the 5'-end of the DNA sequence encoding the polypeptide.

The leader peptide may be the OmpA leader peptide when expression in a bacterial host, such as E. coli is required.
 15 The leader peptide may be the leader peptide of vesicular stomatitis virus G protein (VSV G protein) when expression is to be in insect cells. Appropriate DNA sequences encoding the OmpA and VSV G protein leader sequences are:

OmpA leader:

20 ATG AAA AAG ACA GCT ATC GCG ATT GCA GTG GCA CTG GCT GGT 42
 TTC GCT ACC GTA GCG CAG GCC (SEQ ID NO: 4) 63

VSV G protein leader:

ATG AAG TGC CTT TTG TAC TTA GCC TTT TTA TTC ATT GGG GTG 42
 AAT TGC (SEQ ID NO: 5) 48

25 A DNA sequence may be provided which encodes a fusion protein which is cleavable to release a polypeptide of the invention. A DNA sequence may be used which encodes a carrier polypeptide sequence fused via a cleavable linkage to the N-terminus of a polypeptide of the invention. The
 30 cleavable linkage may be one cleavable by cyanogen bromide.

For expression of the polypeptide, an expression vector is constructed which comprises a DNA sequence encoding the polypeptide and which is capable of expressing the polypeptide when provided in a suitable host. Appropriate
 35 transcriptional and translational control elements are provided, including a promoter for the DNA sequence, a

- 9 -

transcriptional termination site, and translational start and stop codons. The DNA sequence is provided in the correct frame such as to enable expression of the polypeptides to occur in a host compatible with the vector.

The expression vector typically comprises an origin of replication and, if desired, a selectable marker gene such as an antibiotic resistance gene. A promoter is operably linked to the DNA sequence encoding the polypeptide. The expression vector may be a plasmid. In that case, preferably a promoter selected from the P_{trp} and $P_{lpp/lac}$ promoters is operably linked to the DNA sequence. Alternatively, the expression vector may be a virus. The virus may be a recombinant baculovirus in which the polyhedrin promoter is operably linked to the DNA sequence encoding the polypeptide.

An expression vector capable of expressing the polypeptide may be prepared in any convenient fashion. A DNA fragment encoding the polypeptide may be inserted into an appropriate restriction site of an expression vector, for example a plasmid vector. A recombinant baculovirus may be prepared by:

- (i) cloning a gene encoding the polypeptide into a baculovirus transfer vector at a restriction site downstream of the polyhedrin promoter; and
- (ii) co-transfecting insect cells susceptible to baculovirus infection with the recombinant transfer vector from step (i) and intact wild-type baculovirus DNA.

Homologous recombination occurs, resulting in a recombinant baculovirus harbouring the polypeptide gene downstream of the polyhedrin promoter. The baculovirus transfer vector may be one having a unique cloning site downstream of the polyhedrin ATG start codon. The product that is then expressed by the resulting recombinant baculovirus will be a fusion protein in which a N-terminal portion of the polyhedrin protein is fused to the N-terminus of the polypeptide of the invention. As indicated above, a

- 10 -

cleavable linkage may be provided at the fusion junction.

The insect cells employed in step (ii) are typically Spodoptera frugiperda cells. The wild-type baculovirus is typically Autographa californica nuclearpolyhedrosis virus (AcNPV).

An expression vector encoding the polypeptide is provided in an appropriate host. Cells are transformed with the gene of the polypeptide. A transformed host is provided under such conditions that the polypeptide is expressed. Transformed cells, for example, are cultivated so as to enable expression to occur. Any compatible host-vector system may be employed.

The transformed host may be a prokaryotic or eukaryotic host. A bacterial or yeast host may be employed, for example E. coli or S. cerevisiae. Gram positive bacteria may be employed. A preferred bacterial host is a strain of E. coli type B. Insect cells can alternatively be used, in which case a baculovirus expression system is appropriate. The insect cells are typically Spodoptera frugiperda cells. As a further alternative, cells of a mammalian cell line may be transformed. A transgenic animal, for example a non-human mammal, may be provided in which the polypeptide is produced.

The polypeptide that is expressed may be isolated and purified. A polypeptide having the amino acid sequence depicted in formula (I) above preceded by a Met residue attributable to a translation start codon can be obtained. Alternatively, as mentioned above, a fusion protein may be obtained comprising the amino acid sequence of formula (I) fused to a carrier sequence. The carrier sequence is typically fused to the N-terminus of the polypeptide of formula (I).

Where a suitable linkage is provided in the fusion protein between the amino acid sequence of formula (I) and the carrier sequence, a polypeptide having an amino acid sequence according to formula (I) can be released by cleavage with a

- 11 -

suitable agent.

A polypeptide of the invention or a pharmaceutically acceptable salt thereof can also be prepared by:

- 5 (a) chemically synthesising the said polypeptide; and
- (b) isolating the said polypeptide thus obtained or a pharmaceutically acceptable salt thereof.

The polypeptides can therefore be built up by chemical synthesis from single amino acids and/or preformed peptides
10 of two or more amino acids in the order of the sequence of the desired polypeptide. Solid-phase or solution methods may be employed. The resultant polypeptide may be converted into a pharmaceutically acceptable salt if desired.

In solid-phase synthesis, the amino acid sequence of the
15 desired polypeptide is built up sequentially from the C-terminal amino acid which is bound to an insoluble resin. When the desired polypeptide has been produced, it is cleaved from the resin. When solution-phase synthesis is employed, the desired polypeptide may again be built up from the C-
20 terminal amino acid. The carboxy group of this acid remains blocked throughout by a suitable protecting group, which is removed at the end of the synthesis.

Whichever technique, solid-phase or solution-phase, is employed each amino acid added to the reaction system
25 typically has a protected amino group and an activated carboxy group. Functional side-chain groups are protected too. After each step in the synthesis, the amino-protecting group is removed. Side-chain functional groups are generally removed at the end of the synthesis.

30 A polypeptide may be converted into a pharmaceutically acceptable salt. It may be converted into an acid addition salt with an organic or inorganic acid. Suitable acids include acetic, succinic and hydrochloric acid. Alternatively, the polypeptide may be converted into
35 a carboxylic acid salt such as the ammonium salt or an alkali metal salt such as the sodium or potassium salt.

- 12 -

A polypeptide or pharmaceutically acceptable salt thereof may be used in a pharmaceutical composition, together with a pharmaceutically acceptable carrier or excipient
5 therefor. Such a formulation is typically for intravenous administration (in which case the carrier is generally sterile saline or water of acceptable purity). A polypeptide according to the invention is an anti-thrombin polypeptide and is suitable for treatment of thromboembolic events, such
10 as the coagulation of blood, typically in a human patient. A polypeptide according to the present invention displays improved biological characteristics when compared to the corresponding native polypeptide from which it is derived. A polypeptide can therefore be used for the therapy and
15 prophylaxis of thromboses and thromboembolisms, including the prophylaxis of post-operative thromboses, for acute shock therapy (for example for septic or polytraumatic shock), for the therapy of consumption coagulopathies, in haemodialyses, haemoseparations and in extracorporeal blood circulation. In
20 one embodiment of the invention, the polypeptide or salt thereof can be coadministered with a plasminogen activator, such as tissue plasminogen activator. The dosage depends especially on the specific form of administration and on the purpose of the therapy or prophylaxis. The size of the
25 individual doses and the administration regime can best be determined by way of an individual judgement of the particular case of illness; the methods of determining relevant blood factors required for this purpose are familiar to the person skilled in the art. Normally, in the case of an
30 injection the therapeutically effective amount of the compounds according to the invention is in a dosage range of from approximately 0.005 to approximately 0.1 mg/kg body weight. A range of from approximately 0.01 to approximately 0.05 mg/kg body weight is preferred. The administration is
35 effected by intravenous, intramuscular or subcutaneous injection. Accordingly, pharmaceutical compositions for

- 13 -

parenteral administration in single dose form contain per dose, depending on the mode of administration, from approximately 0.4 to approximately 7.5 mg of the compound according to the invention. In addition to the active ingredient these pharmaceutical compositions usually also contain a buffer, for example a phosphate buffer, which is intended to keep the pH value between approximately 3.5 and 7, and also sodium chloride, mannitol or sorbitol for adjusting the isotonicity. They may be in freeze-dried or dissolved form, it being possible for solutions advantageously to contain an antibacterially active preservative, for example from 0.2 to 0.3% 4-hydroxybenzoic acid methyl ester or ethyl ester.

A composition for topical application can be in the form of an aqueous solution, lotion or gel, an oily solution or suspension or a fat-containing or, especially, emulsified ointment. A composition in the form of an aqueous solution is obtained, for example, by dissolving the active ingredients according to the invention, or a therapeutically acceptable salt thereof, in an aqueous buffer solution of from e.g., pH 4 to pH 6.5 and, if desired, adding a further active ingredient, for example an anti-inflammatory agent, and/or a polymeric binder, for example polyvinylpyrrolidone, and/or a preservative. The concentration of active ingredient is from approximately 0.1 to approximately 1.5mg, preferably from 0.25 to 1.0 mg, in 10 ml of a solution or 10 g of a gel.

An oily form of administration for topical application is obtained, for example, by suspending the active ingredient according to the invention, or a therapeutically acceptable salt thereof, in an oil, optionally with the addition of swelling agents, such as aluminium stearate, and/or surfactants (tensides) having an HLB value ("hydrophilic-lipophilic balance") of below 10, such as fatty acid monoesters of polyhydric alcohols, for example glycerin

- 14 -

monostearate, sorbitan monolaurate, sorbitan monostearate or sorbitan monooleate. A fat-containing ointment is obtained, for example, by suspending the active ingredient according to
5 the invention, or a salt thereof, in a spreadable fatty base, optionally with the addition of a tenside having an HLB value of below 10.

An emulsified ointment is obtained by triturating an aqueous solution of the active ingredient according to the invention,
10 or a salt thereof, in a soft, spreadable fatty base with the addition of a tenside having an HLB value of below 10. All these forms for topical application can also contain preservatives. The concentration of active ingredient is from approximately 0.1 to approximately 1.5mg, preferably
15 from 0.25 to 1.0 mg, in approximately 10 g of base.

In addition to the compositions described above and pharmaceutical compositions analogous thereto that are intended for direct medicinal use in the body of a human or a mammal, the present invention relates also to
20 pharmaceutical compositions and preparations for medicinal use outside the living body of humans or mammals. Such compositions and preparations are used especially as anticoagulant additives to blood that is being subjected to circulation or treatment outside the body (for example extra
25 corporeal circulation or dialysis in artificial kidneys), preservation or modification (for example haemoseparation). Such preparations, such as stock solutions or alternatively preparations in single dose form, are similar in composition to the injection preparations described above; however, the
30 amount or concentration of active ingredient is advantageously based on the volume of blood to be treated or, more precisely, on its thrombin content. In this connection it must be borne in mind that the active ingredient according to the invention (in free form) completely deactivates
35 approximately 5 times the amount by weight of thrombin, are physiologically harmless even in relatively large amounts,

- 15 -

and are eliminated from the circulating blood rapidly even in high concentrations so that there is no risk of overdose, even, for example, during transfusions. Depending on the specific purpose, the suitable dose is from approximately 0.01 to approximately 1.0 mg of the active ingredient/litre of blood, although the upper limit may still be exceeded without risk. The following Examples illustrate the invention. In the accompanying drawings:

10 Figure 1 is a chromatogram showing the results of the HPLC analysis of Example 1. P1 to P3 denote the three different peaks obtained from the preparation according to Example 1, FT is for flow through and 4-AB is for 4-aminobenzamidine.

15 Figure 2 shows the elution profiles obtained in Example 2(b) for trypsin-digested PE-P1 (A) and PE-P2 (B).

Figure 3 shows the nucleotide sequence of the six oligos coding for the analogue P2₁₋₆₃-Val-Gly-OH linked to the OmpA leader peptide. P2₁₋₆₃-Val-Gly-OH is an analogue of Formula (I) wherein P is the anti-thrombin polypeptide P2, Xaa is Val and Z is Gly-OH. The sequence shown in bold face indicates the OmpA leader peptide; the Hind III, Bal I and BamHI restriction sites are underlined and the codons for Val₆₄ and Gly₆₅ are in italics.

25 Figure 4 shows the scheme of the construction of the plasmid, named OMP-P2VG, which is the source of a Bal I-BamHI DNA fragment for further genetic constructions.

Figure 5 shows schematically the construction of pFC-P2VG which is the plasmid used for the production of P2₁₋₆₃-Val-Gly-OH protein in E. coli.

Figure 6 shows the general structure of the plasmid pOMPA-P2VG used for the production of P2₁₋₆₃-Val-Gly-OH in E. coli. We employed traditional gene manipulation techniques to prepare this new plasmid where the P2₁₋₆₃-Val-Gly-OH gene is under transcriptional control of the hybrid promoter P_{lpp/lac}. Even in this case, the OmpA leader peptide drives secretion

- 16 -

of P2₁₋₆₃-Val-Gly-OH to the periplasm of E. coli.

Figure 7 shows the nucleotide sequence and assembling of the synthetic oligos used for the secretion of P2₁₋₆₃-Val-Gly-OH from insect cells. The sequence shown in bold face indicates the VSV G protein leader peptide.

Figure 8 is a schematic representation of the construction of a new recombinant M13, named VSV-P2VG, where the complete P2₁₋₆₃-Val-Gly-OH gene is linked to the VSV G protein leader peptide.

Figure 9 shows schematically the construction of pAc-P2VG which has been used as transfer vector to the baculovirus genome. pAcYM1 is the starting plasmid widely used as acceptor of heterologous sequences to be transferred to the virus.

Figure 10 shows the nucleotide sequence and assembling of the synthetic oligos coding for the beginning of the P2₁₋₆₃Val-Gly-OH chain. The ATG codon coding for the additional methionine residue is shown in bold face.

Figure 11 shows schematically the construction of pAcFT1, which has been used for intracellular expression. Figure 12 is a schematic representation of a new transfer plasmid, named pAcFT1-P2VG, which carries the complete P2₁₋₆₃-Val-Gly-OH sequence linked to the first 18 amino acids of polyhedrin. This plasmid has been used to transfer the heterologous sequence to the baculovirus genome.

Figure 13 shows the nucleotide sequence of the six oligos coding for the analogue P1₁₋₆₃-Val-Gly-OH linked to the OmpA leader peptide. P1₁₋₆₃-Val-Gly-OH in an analogue of Formula (I) wherein P is the anti-thrombin polypeptide P1, Xaa is Val and Z is Gly-OH. The sequence shown in bold face indicates the OmpA leader peptide; the Hind III, Bal I and BamHI restriction sites are underlined and the codons for Val₆₄ and Gly₆₅ are in italics.

Example 1

An antithrombin preparation was prepared from

- 17 -

Hirudinaria manillensis leeches according to the procedure illustrated under a) to d) below:

a) Acetone extraction

5 Ethanol dried leech heads (2920 g) were finely chopped into small pieces and treated with a mixture 40:60 acetone/water (7.5 l). After homogenisation with stirring at room temperature, the mixture was spun for 15 min at 2,700 rpm and the supernatant was
10 decanted; the pellet was again resuspended in 40:60 acetone:water mixture, stirred for 30 min and the mixture centrifuged for 15 min at 2,700rpm. The supernatant was pooled with the initial one and acidified to pH 4.5 with glacial acetic acid (vol. 8.5 l). The mixture was spun at
15 2,700 rpm for 15 minutes, then the supernatant was decanted and the pH of the solution adjusted to pH 6.0 by adding 30% ammonia. Following rotary evaporation at 35°C, the pH of the concentrated solution was lowered to 1.8; precipitated contaminants were removed by centrifuging and the raw anti-
20 thrombin material was precipitated from the mixture using a 9-fold acetone excess. The mixture was then spun down, the pellet resuspended in acetone and again centrifuged. The precipitated material was then lyophilized.

b) Ionic exchange purification

25 The raw anti-thrombin material was reconstituted in water, dialyzed against 10 mM ammonium acetate buffer at pH 4.0 and loaded onto a carboxymethyl Sepharose column (CM-Sepharose, Pharmacia, 2.6 x 30 cm) pre-equilibrated in the same buffer. Following a 100 ml washing with starting
30 buffer, anti-thrombin active fractions were eluted with 20mM ammonium acetate pH 4.5, collected and pooled (1.3 l). For further purification steps, pooled fractions were concentrated to 0.5 l in a Minitan apparatus (Millipore); the concentrated solution was neutralised with NaOH and then
35 applied on to a Q Sepharose column equilibrated in 20 mM Tris-HCl pH 7.0. The bound material was eluted with a linear

- 18 -

gradient of 0 - 1 M NaCl in the starting buffer. The fractions containing with anti-thrombin activity were pooled, concentrated and desalted on a Superdex S-200 column eluted
5 with 20 mM Tris-HCl pH 7.5 at a flow rate of 4ml/min. Active pool from gel filtration was concentrated by Minitan and further purified by weak anion exchange chromatography (DEAE FPLC). The active material was loaded onto a Protein Pak DEAE-5PW column (Waters) and eluted with a gradient of 0 - 1
10 M NaCl in 20 mM Tris-HCl pH 6.5, at a flow rate of 1.0 ml/min. Active fractions were pooled, characterized for protein content and activity (specific activity: 800 ATU/mg), and freeze-dried in a Speed Vacconcentrator (Savant).

The thus obtained partially purified material (specific
15 activity 800 ATU/mg) was then subjected to two additional chromatographic steps, in order to get homogenous polypeptides, as described below under c) and d):

c) Thrombin-Sepharose

20 Commercial bovine thrombin (Sigma) was further purified according to the procedure described by Lundblad⁹ and then was attached to activated Sepharose CL 6B (Pharmacia) following manufacturer's instructions. The column (1.7 ml) was equilibrated with 50 mM Tris-HCl pH 8.3 and the freeze-
25 dried material from DEAE-FPLC (reconstituted in buffer) was loaded. The column was subjected to three washings, in starting buffer, then in the same buffer containing 3.0 M NaCl and again with starting buffer (each washing was three times the column volume). Flow rate was 0.3 ml/min. The
30 bound material was eluted with 10 ml of 0.1M 4-aminobenzamidine in 25 mM HCl. The active fractions were pooled and buffer exchanged in 50 mM Tris-HCl pH 8.3 onto a PD-10 column (Pharmacia).

Unbound material eluted from the column by washing in
35 starting buffer and still containing anti-thrombin activity, was reloaded onto the column until all the activity was bound

- 19 -

and chromatographed.

d) RP-HPLC

5 Material obtained after affinity chromatography was finally purified by reverse phase high performance chromatography (RP-HPLC) on a C4 Vydac column (4.6 x 250 mm.5 μ) using 20 mM sodium phosphate pH 7.5 as first eluent and 50% acetonitrile in water as modified. Anti-thrombin
10 polypeptides were eluted with a linear gradient from 5% to 55% of eluent B in 45 minutes, at room temperature with a flow rate of 1.0 ml/min. The resulting chromatogram is shown in Figure 1.

15 Peaks of protein (detected at 220 nm) were manually collected, concentrated under vacuum and re-chromatographed under the same conditions.

20 Pure anti-thrombin polypeptides after C4 HPLC were characterized for protein content, amino acid composition, N-terminal sequence, C-terminal end and their activity determined by in vitro assays (ATU/NIH test and "thrombin time" test). Each of the three peaks of protein has been found to be endowed with anti-thrombin activity.

25 The complete amino acid sequences of the polypeptides labeled P1 and P2 in Figure 1 were determined by N-terminal sequencing of the peptides obtained from tryptic and V8 protease digests. The sequences are reported above together with the partial amino acid sequence of the other polypeptide (P3).

30 Example 2 - Tryptic digestion and peptide mapping of pyridylethylated (PE) P1 and P2

a) Reduction/Alkylation - Active fractions purified by affinity chromatography on Thrombin-Sepharose (Example 1c)
35 were pooled and buffer exchanged in 10 mM Tris-HCl pH 8.3 onto a PD-10 column. The active pool (about 50 μ g) was

- 20 -

concentrated in a Speed-Vac centrifuge (Savant) and treated with 100 μ l of 1% b-mercaptoethanol in 6M guanidine-HCl / 50 mM Tris-HCl pH 8.5, under nitrogen, in the dark, for 2 hours at room temperature. Then 4 μ l of 4-vinyl-pyridine (neat) were added and the mixture incubated again for 2 hours as above¹⁰.

Pyridylethylated polypeptides were first recovered from the reaction mixture by RP-HPLC on a C4 Vydac (4.6x250mm, 5 μ m) column eluted with a 90 min. linear gradient from 5-65% acetonitrile in 0.1% TFA, at a flow rate of 1.0ml/min. Under such conditions the mixture of anti-thrombin polypeptides is poorly resolved so that they have to be re-chromatographed on the same column using the elution system sodium phosphate/acetonitrile with the conditions already described in Example 1d.

b) Trypsin digestion and peptide mapping of PE-P1 and PE-P2 - Purified PE-P1 and PE-P2 (respectively 10 and 20 μ g) were digested with TPCK-treated trypsin (Sigma) in 200 μ l of 1% ammonium bicarbonate pH 8.0 in the presence of 0.2 M sodium phosphate. Trypsin was added at an enzyme-to-substrate ratio of 1:20 (w/w) and incubation was carried out for 4 hours at 37°C. Digestion was stopped by freeze-drying in Savant.

Peptides obtained by tryptic digestion were separated on a μ Bondapak C18 column (3.9x300 mm, 10 μ Waters) or on a C4-Vydac (4.6x250 mm, 5 μ m) column eluted using a 60 min. linear gradient from 5-65% acetonitrile in 0.1% TFA, at a flow rate of 1.0 ml/min. Eluted peaks were manually collected, concentrated in Savant and then subjected either to amino acid analysis and to N-terminal sequence analysis on a pulsed liquid-phase mod.477A Sequencer (Applied Biosystems).

The results of C4-HPLC peptide mapping of trypsin-digested PE-P1 (A) and PE-P2 (B) are shown below.

- 21 -

<u>Fragment</u>			<u>Amino acid sequence</u>	
5	A	1-13	VSYTDCTESGQNY	(SEQ ID NO: 6)
		14-26	CLCVGGNLCGGGK	(SEQ ID NO: 7)
		27-36	HCEMDGSGNK	(SEQ ID NO: 8)
		37-47	CVDGEGTPKPK	(SEQ ID NO: 9)
		37-47 (*)	CVDGEGX*PKPK	(SEQ ID NO: 10)
10		48-64	SQTEGDFEEIPDEDILN	(SEQ ID NO: 11)

15	B	1-13	VSYTDCTESGQNY	(SEQ ID NO: 12)
		14-26	CLCVGSNVCGEK	(SEQ ID NO: 13)
		27-47	NCQLSSSGNQCVHGEGX*PKPK	(SEQ ID NO: 14)
		48-64	SQTEGDFEEIPDEDILN	(SEQ ID NO: 15)

(*) X = residue not detected by amino acid sequencing; X = T
 20 by amino acid analysis.

The complete amino acid sequences of P1 and P2 are therefore:

P1 (SEQ ID NO: 1)

Val Ser Tyr Thr Asp Cys Thr Glu Ser Gly Gln Asn Tyr Cys Leu
 25 Cys Val Gly Gly Asn Leu Cys Gly Gly Gly Lys His Cys Glu Met
 Asp Gly Ser Gly Asn Lys Cys Val Asp Gly Glu Gly Thr Pro Lys
 Pro Lys Ser Gln Thr Glu Gly Asp Phe Glu Glu Ile Pro Asp Glu
 Asp Ile Leu Asn

P2 (SEQ ID NO: 2)

30 Val Ser Tyr Thr Asp Cys Thr Glu Ser Gly Gln Asn Tyr Cys Leu
 Cys Val Gly Ser Asn Val Cys Gly Glu Gly Lys Asn Cys Gln Leu
 Ser Ser Ser Gly Asn Gln Cys Val His Gly Glu Gly Thr Pro Lys
 Pro Lys Ser Gln Thr Glu Gly Asp Phe Glu Glu Ile Pro Asp Glu
 Asp Ile Leu Asn

- 22 -

Example 3: Chemical synthesis of the gene coding for the analogue P2₁₋₆₃-Val-Gly-OH

5 As already said, the analogue P2₁₋₆₃-Val-Gly-OH is one of formula (I) wherein P is the antithrombin polypeptide P2, Xaa is Val and Z is Gly-OH.

 The nucleotide sequence coding for the analogue P2₁₋₆₃-Val-Gly-OH was designed on the basis of the Escherichia coli preferred codons¹¹. Moreover, a BalI restriction site was engineered very close to the 5' end of the synthetic P2₁₋₆₃-Val-Gly-OH coding sequence to allow insertion of such sequence in different expression vectors. Indeed, the same synthetic gene was used for expression of recombinant P2₁₋₆₃-Val-Gly-OH protein in bacterial and insect cells. All plasmid DNA manipulations were carried out as described by Maniatis et al¹². In the case of insect cells methods were developed which yielded protein P2₁₋₆₃-Val-Gly-OH as a secreted or cytoplasmic product. In the case of bacterial cells, methods were developed to obtain secretion to the periplasm of the recombinant product. In both cases, when a secreted product is to be obtained, it is necessary to synthesize the P2₁₋₆₃-Val-Gly-OH molecule in the form of a pre-protein. More particularly, an amino acid sequence named "leader peptide", responsible for an efficient secretion must be present at the NH₂ end of P2₁₋₆₃-Val-Gly-OH¹³⁻¹⁴. This extra sequence is then cleaved off, in vivo, during secretion, by a specific leader peptidase, yielding the correct mature sequence¹⁵.

30 In the insect cells expression system, we used the leader peptide of the Vesicular Stomatitis Virus (VSV) G protein, as will be described in detail in Example 5. In the case of E. coli many examples of secretion systems have been described in the literature¹⁶⁻¹⁷. Among them, we have selected the system based on the secretion signal of the Outer Membrane Protein of E. coli (Omp A) previously

35

- 23 -

published¹⁸.

In order to prepare a suitable source of the P2₁₋₆₃-Val-Gly-OH gene for the expression vectors used in the Examples, 5 six synthetic complementary oligonucleotides were synthesized using an automated DNA synthesizer (Applied Biosystems) and their sequence is shown in Figure 3.

We designed the first two complementary oligonucleotides coding for the OmpA leader peptide preceded by the OmpA 10 Shine-Dalgarno sequence known to be responsible for an efficient translation of the messenger RNA¹⁹.

Their sequence, shown in Fig. 3 as oligos 1 and 2, includes also the beginning of the P2₁₋₆₃-Val-Gly-OH gene coding for the first 10 amino acids.

15 Following enzymatic phosphorylation the six oligos were assembled using DNA

ligase and the resulting double-strand sequence was inserted in the M13 phage vector mp18, (Yanisch-Perron et al., Gene 33, 103-119, 1985) obtaining the recombinant plasmid OMP-P2VG 20 which is shown in Figure 4. In order to enable insertion of the P2₁₋₆₃-Val-Gly-OH gene in the M13 vector, HindIII and BamHI sites were also added in the synthetic oligos. The correct nucleotide sequence has been verified by the Sanger method carried out on the single strand phage DNA²⁰.

25

Example 4: Expression and secretion of P2₁₋₆₃-Val-Gly-OH from E. Coli cells

From OMP-P2VG the P2₁₋₆₃-Val-Gly-OH gene can be excised as a HindIII-BamHI fragment which codes for the OmpA Shine-Dalgarno and leader peptide followed by the P2₁₋₆₃-Val-Gly-OH 30 coding sequence. This restriction fragment is now ready to be inserted in an appropriate expression vector. Several expression systems

could, theoretically, be employed to obtain high level 35 production of heterologous proteins in bacteria. The system based on the promoter P_{mp} has been used with success in our

- 24 -

laboratory in the past¹⁹. Again, even in the case of the selected promoter, the levels of expression of a given polypeptide cannot be predicted.

5 Plasmid pFC33, shown in Figure 5, has already been described in the literature¹⁹. It carries the resistance to the antibiotic ampicillin and the bacterial promoter P_{up} which drives expression of proapolipoprotein A1. Following digestion of pFC33 with HindIII and BamHI, the large HindIII-BamHI
10 BamHI fragment, carrying the antibiotic resistance gene and the promoter, was isolated and joined to the HindIII-BamHI fragment from OMP-P2VG coding for the P₂₁₋₆₃-Val-Gly-OH gene. The details of this construction are shown in Figure 5. We isolated a new plasmid, named pFC-P2VG, which is the final
15 plasmid for the production of P₂₁₋₆₃-Val-Gly-OH in E. coli.

 An object of the present invention is the use of E. coli strains of the type B for the expression and secretion to the periplasm of P₂₁₋₆₃-Val-Gly-OH and the other anti-thrombin analogues of the invention. Indeed, we have found that
20 insertion of plasmid pFC-P2VG in type B strains of the bacterium E. coli brings high level production of P₂₁₋₆₃-Val-Gly-OH. Interestingly, different strain types of E. coli do not work as efficiently and it seems, therefore, that the host strain type is crucial for the successful production of
25 P₂₁₋₆₃-Val-Gly-OH.

 Several type B strains of E. coli are available and can be used for the production of P₂₁₋₆₃-Val-Gly-OH. Preferred strains are ATCC 12407, ATCC 11303, NCTC 10537. Below is an example of transformation of strain NCTC 10537 with plasmid
30 pFC-P2VG and subsequent cultivation of the transformant. Competent cells of strain NCTC 10537 were prepared using the calcium chloride procedure of Mandel and Higa²¹. Approximately 200μl of a preparation of these cells at 1 x 10⁹ cells per milliliter were transformed with 2μl of plasmid
35 DNA (approximate concentration 5μg/ml). Transformants were selected on plates of L-agar containing 100μg/ml ampicillin.

- 25 -

Two small colonies were streaked with wooden tooth picks (each as three streaks about 1 cm long) onto L-agar containing the same antibiotic.

5 After 12 hours incubation at 37°C, portions of the streaks were tested for P_{2,63}-Val-Gly-OH protein production by inoculation onto 10 ml of LB medium (containing ampicillin at a concentration of 150µg/ml) and incubated overnight at 37°C. The following day the cultures were diluted 1:100 in M9

10 medium, containing the same concentration of ampicillin, and incubated for 6 hours at 37°C.

20 ml of such culture were centrifuged at 12000xg, 4°C, for 10 minutes. The bacterial pellet was resuspended in 2 ml of 33 mM HCl Tris pH 8; an equal volume of a second solution

15 33 mM EDTA, 40% sucrose was then added and the total mixture was incubated under mild shaking conditions at 37°C for 10 minutes. Following centrifugation, the permeabilized cells were resuspended in 2 ml of cold water and left for 10 minutes in ice. The resulting supernatant was isolated by

20 centrifugation and represents the periplasmic fraction of the bacterial cell.

Using a chromogenic assay that is based on the inhibition of the thrombin ability to cleave a synthetic substrate S-2238²², we have measured the presence of anti-

25 thrombin activity in the periplasmic fraction of P_{2,63}-Val-Gly-OH producing cells but not in control periplasmic fractions.

With the similar approach we have also constructed a new expression/secretion plasmid for P_{2,63}-Val-Gly-OH where the

30 promoter P_{lpp/lac}¹⁷ is present instead of the promoter P_{lpp}. This different plasmid, named pOMPA-P2VG, is shown in Figure 6. Following insertion of this plasmid in *E. coli* strains of the type B, high levels of active P_{2,63}-val-Gly-OH were also obtained. As starting plasmid for the construction of pOMPA-

35 P2VG we used the plasmid pIN-III-ompA3 described by Ghrayb et al²³. Conditions for cultivation and induction of expression

- 26 -

with isopropyl- β -D-thiogalactopyranoside (IPTG) were as previously described²³.

5 Example 5: Expression and secretion of protein P2₁₋₆₃-Val-Gly-OH from insect cells

To obtain secretion of protein P2₁₋₆₃-val-Gly-OH from recombinant insect cells we had to join the P2₁₋₆₃-Val-Gly-OH coding sequence to a leader peptide that is efficiently
10 recognized by these cells. We have used the leader peptide of the Vesicular Stomatitis Virus (VSV) G protein²⁴. Similarly to what is described above, a synthetic DNA sequence coding for the VSVG protein leader peptide followed by the beginning of the P2₁₋₆₃-Val-Gly-OH gene has been
15 prepared and the nucleotide sequence is given in Figure 7. Also in this case we provided convenient restriction sites (HindIII, BamHI and BalI) to allow joining to the rest of the P2₁₋₆₃-Val-Gly-OH gene and to the expression vector.

The synthetic HindIII-BalI fragment was joined to
20 apurified BalI-BamHI fragment from OMP-P2VG carrying the P2₁₋₆₃-Val-Gly-OH gene and inserted in M13mp18 previously cut with HindIII and BamHI. This construction which yielded a new plasmid named VSV-P2VG is schematically shown in Figure 8. From VSV-P2VG we have excised a BamHI-BamHI DNA fragment
25 carrying the P2₁₋₆₃-Val-Gly-OH gene fused to the VSV leader peptide which was then inserted into the vector pACYM1²⁵, as shown in Figure 9. The resulting plasmid was named pAc-P2VG.

To obtain expression in insect cells, the VSV-P2VG coding sequence must be transferred to the baculovirus genome
30 under the transcriptional control of the polyhedrin promoter. For this purpose, we co-transfected insect cells with a wild-type baculovirus DNA and with the transfer vector pAc-P2VG. As insect cells, Spodoptera frugiperda cells were chosen as host cells. Experimental details are as follows:

35 S. frugiperda cells were transfected with a mixture of infectious AcNPV DNA and plasmid DNA representing the

- 27 -

individual recombinant transfer vectors by a modification of the procedure described by Summers et al²⁶. One microgram of viral DNA was mixed with 25-100 µg of plasmid DNA and
5 precipitated with (final concentrations) 0.125 M calcium chloride in the presence of 20 mM HEPES buffer, pH 7.5, 1 mM disodium hydrogen orthophosphate, 5mM potassium chloride, 140 mM sodium chloride and 10 mM glucose (total volume 1ml).

The DNA suspension was inoculated onto a monolayer of
10 10⁶ S. frugiperda cells in a 35-mm tissue culture dish, allowed to adsorb to the cells for 1 h at room temperature, then replaced with 1 ml of medium. After incubation at 28°C for 3 days the supernatant fluids were harvested and used to produce plaques in S. frugiperda cell monolayers. Plaques
15 containing recombinant virus were identified by their lack of polyhedra when examined by light microscopy. Virus from such plaques was recovered and after further plaque purification was used to produce polyhedrin-negative virus stocks.

The above procedure allowed us to isolate a recombinant
20 baculovirus whose genome carried the P2₁₋₆₃-Val-Gly-OH gene under control of the polyhedrin promoter and of the VSV G protein leader peptide. We used this virus to infect S. frugiperda cells according to well-established procedures²⁶, at a multiplicity of infection of 10. Infected cells were
25 then cultivated in spinner culture or in monolayers in the presence of 10% foetal calf serum according to published methods²⁶. In both conditions, the S-2238 chromogenic assay showed the presence of an anti-thrombin activity in the culture supernatants of the infected cells.

30

Example 6: Expression of protein P2₁₋₆₃-Val-Gly-OH in the cytoplasm of insect cells

Protein P2₁₋₆₃-Val-Gly-OH could also be produced and accumulated in the cytoplasm of S. frugiperda cells. This
35 approach generally gives a better yield of heterologous

- 28 -

proteins since it utilizes the expression signals of polyhedrin which is a non-secreted viral protein.

Our approach to obtain large quantities of recombinant protein P2₁₋₆₃-Val-Gly-OH is based on the expression of a fusion polypeptide where the first 18 amino acids of polyhedrin are joined in frame to the 65 amino acids of P2₁₋₆₃-Val-Gly-OH. The presence of the NH₂ end sequence of polyhedrin allows high level expression²⁷. In addition, between the polyhedrin portion and the P2₁₋₆₃-Val-Gly-OH sequence we put a methionine residue which allows the release of the P2₁₋₆₃-Val-Gly-OH moiety by treatment of the hybrid protein with CNBr.

Similarly to the previous approaches, we prepared a synthetic DNA fragment which could allow the joining of the BalI-BamHI fragment from OMP-P2VG to an appropriate transfer vector. The new synthetic piece, shown in Figure 10, includes also BamHI and BalI sites for subsequent manipulations.

A different transfer vector, pAcFT1, carrying the nucleotide sequence coding for the first 18 amino acids of polyhedrin has been obtained (Figure 11). Briefly, the EcoRV-BamHI fragment of pAcYM1²⁵ has been replaced by a synthetic oligonucleotide containing the polyhedrin gene sequence from nucleotide -92 to nucleotide +55. A convenient BamHI site is present after this sequence and it has been used for insertion of the complete P2₁₋₆₃-Val-Gly-OH coding sequence according to a scheme illustrated in Figure 12. Through this construction, we obtained a new plasmid, named pAcFT1-P2VG, which has been used to transfer the hybrid gene to the baculovirus genome.

The recombinant baculovirus was obtained as described in Example 5. Infection of S. frugiperda cells was carried out according to standard procedures²⁶. Cultivation of infected insect cells leads to the cytoplasmic accumulation of the fusion protein. This hybrid protein was the source of

- 29 -

recombinant protein P2₁₋₆₃-Val-Gly-OH. Several methods are available from the literature which can be used to cleave the hybrid with CNBr²⁸⁻²⁹. The application of the method of Olson
5 et al has allowed us to obtain the correct polypeptidic sequence of P2₁₋₆₃-Val-Gly-OH. This molecule displayed anti-thrombin activity.

Example 7

Expression and secretion in E. coli and insect cells of
10 the analogue P1₁₋₆₃-Val-Gly-OH can be achieved following procedures analogous to those described above under examples 3 to 6. For this purpose, six synthetic complementary oligonucleotides were prepared and their sequence is shown in Figure 13. Oligos 1,2,5 and 6 are identical to those used for
15 the genetic construction of Figure 3, since the differences in the sequence of P1₁₋₆₃-Val-Gly-OH and P2₁₋₆₃-Val-Gly-OH are in the central part of the molecule and the corresponding codons are located in oligos 3 and 4.

20 Example 8

The analogue P1₁₋₆₃-Val-Gly-OH is converted to the corresponding amidated product P1₁₋₆₃-Val-NH₂ by incubation with a partially purified protease-free PAM enzyme preparation, as, for
example, obtained from rat thyroid medullary carcinoma³⁰ or
25 from the conditioned medium of cells derived from the same tumour³¹.

Alternatively, recombinant DNA techniques can also be applied to prepare an amidating enzyme.

The enzymatic amidating reaction has been carried out in an
30 aqueous buffer, for example TSE buffer, supplemented with copper ions, ascorbate, catalase, potassium iodide, SDS and Tween 20.

The progress of the amidation reaction can be monitored by the method of Corbett and Corbett³² which detects the
35 formation of glyoxylic acid after derivatization with nitrosobenzene.

- 30 -

The amidation product P_{1-63} -Val-NH₂ can be purified according to standard techniques well known to the skilled in the art.

Example 9

5 Biological activity of $P_{2,1-63}$ -Val-Gly-OH

Antithrombin activity

The antithrombin activity of the analogues $P_{2,1-63}$ -Val-Gly-OH of the present invention was determined on the basis of the rapid and stoichiometric reaction of the analogue peptides
10 with thrombin.

This activity was measured quantitatively by means of titration of a standard solution of thrombin³³.

Thrombin activity was calibrated with the International Standard Preparation of Thrombin (c 70/157) obtained from the
15 National Institute for Biological Standards and Control (London, U.K.) and expressed in National Institute of Health Units (N.I.H. units).

The thrombin neutralizing activity of the samples was expressed as antithrombin units (ATU); one ATU is the amount
20 of test compound which neutralizes an NIH unit of thrombin. Recombinant hirudin (HV1 variant; rHV1) and recombinant P2 protein (rP2) were considered as reference compounds.

The test was performed as follows:

0.2 ml of a standard solution of 0.05% human fibrinogen (Kabi Vitrum, Sweden) in Tris HCl buffer pH 7.4 were incubated at
25 37°C in the presence of 0.01 to 0.1 ml of the solutions of $P_{2,1-63}$ -Val-Gly-OH.

Thereafter, aliquots of 0.005 ml (0.5 NIH units) of a 100 NIH/ml thrombin standard solution were added progressively
30 each minute and mixed gently.

The end point of the titration was considered to be reached when a fibrin clot was formed within one minute.

Indeed, clot formation only occurred when a sufficient amount of thrombin was added to the reaction mixture, the added
35 thrombin being able to neutralize the total amount of the analogue present in the mixture.

- 31 -

The antithrombin activity of the analogue P2₁₋₆₃-Val-Gly-OH has been expressed in ATU/mg protein (see table 1).

Protein content was determined by aminoacid analysis.

5 Chromogenic substrate assay

The principle of the method is based on the inhibitory activity of P2₁₋₆₃-Val-Gly-OH on the reaction between thrombin and its specific chromogenic substrate S-2238³⁴.

10 Recombinant hirudin HV1 and recombinant P2 protein were used as references.

The test was performed as follows:

630 μ l of a 0.5 mM solution of S-2238 in 0.05 M Tris buffer pH 8.3 containing 0.1N NaCl and 100 KIU/ml aprotinin were pre-incubated for 10 min. at 25°C. After this pre-incubation
15 period, 70 μ l of a 2 IU/ml thrombin solution in 0.025 M sodium phosphate buffer pH 6.6 containing 0.3 M NaCl and 0.5% bovine albumin were added.

For the controls, 70 μ l of buffer were used.

Immediately after the addition of thrombin, the absorbance
20 variations were recorded spectrophotometrically at a wave length of 405 nm.

Recordings were made kinetically each minute over a total period of 6 minutes.

Thereafter, 20 μ l of the solutions of the tested analogue
25 were added and absorbance variations were recorded again as described before.

Final concentrations of the test compounds were 1.25, 2.5, 7.5, 10 and 15 ng/ml in 0.025 M sodium phosphate buffer pH 6.6 containing 0.3 M NaCl.

30 The absorbance variations per minute (Δ Abs/min) were determined respectively before and after the addition of the sample and the mean Δ Abs/min. were calculated.

The inhibition of the reaction between thrombin and the chromogenic substrate S-2238 was calculated as the ratio
35 (expressed as percentage) between the mean Δ Abs/min. before and after the addition of the analogue.

- 32 -

A standard curve has been established on the basis of the percent inhibition versus the concentration of the sample. Final results were expressed as final concentrations of the test compounds inducing 50% inhibition of the reaction between thrombin and its chromogenic substrate.

The thrombin time

The antithrombin activity of the analogue P2₁₋₆₃-Val-Gly-OH has been determined by measuring the thrombin time in human plasma.

Recombinant hirudin HV1 (rHV1) and recombinant P2 protein (rP2) were used as references.

The test was performed as follows:

Mixtures of normal human citrated plasma containing increasing final concentrations of the tested analogue or the reference compounds were placed in an automatic coagulometer (ACL 300 research, Instrumentation Laboratory, Italy), and human thrombin (Fibrindex, Ortho Diagnostics, Italy: 5 IU/ml final concentration) was added.

The clotting times were recorded and plotted against the final concentration of each sample.

Final results were expressed as the final concentration of each compound required to double the normal value of the thrombin time (ratio=2).

In all the test performed, the antithrombin activity of P2₁₋₆₃-Val-Gly-OH resulted considerably superior to that of the reference compounds.

Table 1 reports the data obtained with the analogue P2₁₋₆₃-Val-Gly-OH in the biological assays discussed above.

30

35

- 33 -

TABLE 1

5	Compound	ATU/mg	S-2238 IC50 in ng/ml	Thrombin Time in µg/ml
	rHV1	11.000	5.47	ratio=2
10	rP2	11.900	5.43	at 0.132 µg/ml ratio=2
15	P2 ₁₋₆₃ -Val-	20.000	0.6	at 0.131 µg/ml ratio=2
	Gly-OH			at 0.047 µg/ml

20

25

30

35

References

- 1) Markwardt, F. 1970, Methods in Enzymology, 19, p. 924
- 2) Markwardt, F. 1985, Biomed. Biochim. Acta. 44, p. 1007
- 5 3) Markwardt, F. Hauptmann, J., Nowak, G., Klessen, C., and
Walsmann, P. 1982. Thromb. Haemostasis 47, p. 226
- 4) A.F. Bradbury et al., 1982, Nature, 289: 686-688
- 5) G.S. Wand et al., 1985, Neuroendocrinol, 41, 482-489
- 6) B. EIPPER et al. 1983, Proc.Natl.Acad.Sci: USA 80, 5144-
10 5148
- 7) Breddam, K. et al. Int. J. Pept. Prot. Res.: 37, 153-
160, 1991;
- 8) Henriksen, D.B. et al. J. Am. Chem. Soc.: 114, 1876-
1877, 1992
- 15 9) Lundblad, R.L., 1971, Biochemistry, 10: 2501-2506
- 10) Dupont D., Keim P., Chui A., Bello R. and Wilson K.,
Derivatizer-Analyser User Bulletin No. 1, Applied Biosystems
Inc., 1987
- 11) Grosjeans H. and Fiers W. 1982. Gene, 18, p. 199
- 20 12) Maniatis T., Fritsch E.F. and Sambrook J. 1982. Cold
Spring Harbor, NY
- 13) Blobel G. and Dobberstein B. 1975. J. Cell Biology, 67,
p. 83
- 14) Pages J.M. 1983, Biochimie, 65, p. 531
- 25 15) Wolfe P.B. 1983. J. Biol. Chem. 258, p. 12073
- 16) Talmadge K., Stahl S. and Gilbert W. 1980. Proc. Natl.
Acad. Sci. USA, 77, p. 3369
- 17) Oka T., Sakamoto S., Miyoshi K., Fuwa T., Yoda K.,
Yamasaki M., Tamura G. and Miyake K. 1985. Proc. Natl. Acad.
30 Sci. USA, 82, p. 7212
- 18) Henning V., Royer H.D., Teather R.M., Hindennach I. and
Hollenberg C.P. 1979. Proc. Natl. Acad. Sci. USA, 76,
p. 4360
- 19) Isacchi A., Sarmientos P., Lorenzetti R. and Soria M.
35 1989, Gene 81, p. 129
- 20) Sanger, F., Nicklen, S., and Coulson, A.R. 1977, Proc.

- 35 -

- Natl. Acad. Sci. USA 74, p. 5463.
- 21) Mandel M. and Higa A.J. 1970. J. Mol. Biology, 53, p.154
- 22) Krstenansky, J.K., and Mao, S.J.T. 1987. FEBS Lett. 211,
5 p. 10
- 23) Ghrayeb J., Kimura H., Takahara M., Hsiung H., Masui Y.
and Inouye M. 1984. EMBO Journal 3, p. 2437
- 24) Bailey, M.J., McLeod, D.A., Kang, C., and Bishop, D.H.L.
1989. Virology 169, p. 323
- 10 25) Matsuura, Y., Possee, R.D., Overton, H.A. and Bishop.
D.H.L. 1987. J. Gen. Virol. 68, p. 1233
- 26) Summers, M.D., and Smith, G.E. 1987, Texas Agricultural
Experiment Station Bulletin No. 1555
- 27) Luckow, V.A. and Summers, M.D. 1988, Virology, 167, p.56
- 15 28) Gross E. 1967. Methods in Enzymology, 11, p. 238
- 29) Olson H., Lind P., Pohl G., Henrichson C., Mutt V.,
Jornvall H., Josephson S., Uhlen M. and Lake M. 1987,
Peptides, 9, p. 301
- 30) N.M. Mehta et al. 1988, Arch.Biochem.Biophys. 261-44-45
- 20 31) J.P. Gilligan et al. 1989, Endocrinol. 124, 2729-2736.
- 32) J. Org.Chem. 1980, 45, 2834-2839
- 33) Lundblad R.L., et al., Methods in Enzymology, 45, 156-
161, 1976).
- 34) Wiman et al., BBA, 579, 142-154, 1979

25

- 36 -

SEQUENCE LISTING

(1) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 64 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Hirudinaria manillensis*

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

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

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 64 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Hirudinaria manillensis*

- 37 -

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

```

Val Ser Tyr Thr Asp Cys Thr Glu Ser Gly Gln Asn Tyr Cys Leu Cys
1           5           10           15

Val Gly Ser Asn Val Cys Gly Glu Gly Lys Asn Cys Gln Leu Ser Ser
          20           25           30

Ser Gly Asn Gln Cys Val His Gly Glu Gly Thr Pro Lys Pro Lys Ser
      35           40           45

Gln Thr Glu Gly Asp Phe Glu Glu Ile Pro Asp Glu Asp Ile Leu Asn
    50           55           60

```

(3) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 42 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Hirudinaria manillensis*

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

```

Val Ser Tyr Thr Asp Cys Thr Glu Ser Gly Gln Asn Tyr Cys Leu Cys
1           5           10           15

Val Gly Ser Asn Val Cys Gly Glu Gly Lys Asn Cys Gln Leu Ser Ser
          20           25           30

Ser Gly Asn Gln Cys Val His Gly Glu Gly
      35           40

```

- 38 -

(4) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 63 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: E. coli

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION: 1..63
- (D) OTHER INFORMATION: /function= "leader peptide"
/standard_name= "OmpA leader"

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

ATGAAAAAGA CAGCTATCGC GATTGCAGTG GCACTGGCTG GTTTCGCTAC
CGTAGCGCAG GCC 63

50

(5) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 48 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: vesicular stomatitis virus

- 39 -

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION: 1..48

(D) OTHER INFORMATION: /function= "leader peptide"
/standard_name= "VSV G protein leader"

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

ATGAAGTGCC TTTTGTACTT AGCCTTTTTA TTCATTGGGG TGAATTGC

48

(6) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 13 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Hirudinaria manillensis*

(ix) FEATURE:

(B) LOCATION: 1..13

(D) OTHER INFORMATION: /note= "This sequence is amino
acids 1..13 of seq id no:1"

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

Val	Ser	Tyr	Thr	Asp	Cys	Thr	Glu	Ser	Gly	Gln	Asn	Tyr
1				5					10			

- 40 -

(7) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 13 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Hirudinaria manillensis*

(ix) FEATURE:

- (B) LOCATION: 1..13
- (D) OTHER INFORMATION: /note= "This sequence is amino acids 14..26 of seq id no:1"

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

Cys	Leu	Cys	Val	Gly	Gly	Asn	Leu	Cys	Gly	Gly	Gly	Lys
1				5					10			

(8) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Hirudinaria manillensis*

- 41 -

(ix) FEATURE:

(B) LOCATION: 1..10

(D) OTHER INFORMATION: /note= "This sequence is amino acids 27..36 of seq id no:1"

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

His	Cys	Glu	Met	Asp	Gly	Ser	Gly	Asn	Lys
1				5				10	

(9) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 11 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Hirudinaria manillensis

(ix) FEATURE:

(B) LOCATION: 1..11

(D) OTHER INFORMATION: /note= "This sequence is amino acids 37..47 of seq id no:1"

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

Cys	Val	Asp	Gly	Glu	Gly	Thr	Pro	Lys	Pro	Lys
1				5				10		

- 42 -

(10) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Hirudinaria manillensis*

(ix) FEATURE:

- (B) LOCATION: 1..11
- (D) OTHER INFORMATION: /note= "This sequence is amino acids 37..47 of seq id no:1"

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

Cys	Val	Asp	Gly	Glu	Gly	Xaa	Pro	Lys	Pro	Lys
1				5					10	

(11) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Hirudinaria manillensis*

(ix) FEATURE:

- (A) NAME/KEY: Region
- (B) LOCATION: 1..16

- 43 -

(D) OTHER INFORMATION: /note= "amino acids 48..64 of seq
id no:1"

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

1	Ser	Gln	Thr	Glu	Gly	Asp	Phe	Glu	Glu	Ile	Pro	Asp	Glu	Asp	Ile	Leu
				5				10							15	

Asn

(12) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 13 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Hirudinaria manillensis

(ix) FEATURE:

(B) LOCATION: 1..13

(D) OTHER INFORMATION: /note= "This sequence is amino
acids 1..13 of seq id no:2"

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

1	Val	Ser	Tyr	Thr	Asp	Cys	Thr	Glu	Ser	Gly	Gln	Asn	Tyr
				5				10					

(13) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 13 amino acids

- 44 -

- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Hirudinaria manillensis*

(ix) FEATURE:

(B) LOCATION: 1..13

(D) OTHER INFORMATION: /note= "This sequence is amino acids 14..26 of seq id no:2"

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

1	Cys	Leu	Cys	Val	Gly	Ser	Asn	Val	Cys	Gly	Glu	Gly	Lys
			5						10				

(14) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Hirudinaria manillensis*

(ix) FEATURE:

(A) NAME/KEY: Region

(B) LOCATION: 1..21

(D) OTHER INFORMATION: /note= "This sequence is amino acids 27..47 of seq id no:2"

- 45 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Asn Cys Gln Leu Ser Ser Ser Gly Asn Gln Cys Val His Gly Glu Gly
1 5 10 15
Xaa Pro Lys Pro Lys
20

(15) INFORMATION FOR SEQ ID NO:15:

(i) **SEQUENCE CHARACTERISTICS:**

- (A) LENGTH: 17 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Hirudinaria manillensis*

(ix) **FEATURE:**

- (A) NAME/KEY: Region
(B) LOCATION: 1..21
(D) OTHER INFORMATION: /note= "This sequence is amino acids 48..64 of seq id no:2"

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

Ser Gln Thr Glu Gly Asp Phe Glu Glu Ile Pro Asp Glu Asp Ile Leu
1 5 10 15
Asn

- 46 -

(16) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 198 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Hirudinaria manillensis*

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

GTTTCTTACA CCGACTGCAC CGAATCTGGC CAGAACTACT GCCTGTGCGT	50
TGGTTCTAAC GTTTGCGGTG AAGGTAAAAA CTGCCAGCTG TCTTCTTCTG	100
GTAACCAGTG CGTTCACGGT GAAGGTACCC CGAAACCGAA ATCTCAGACT	150
GAAGGTGACT TCGAAGAAAT TCCGGACGAA GACATCCTGG TTGGTTAG	198

(17) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 198 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Hirudinaria manillensis*

- 47 -

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

GTTCCTTACA CCGACTGCAC CGAATCTGGC CAGAACTACT GCCTGTGCGT 50
TGGTGGTAAC CTGTGCGGTG GTGGTAAACA CTGCGAAATG GATGGTTCTG 100
GTAACAAATG CGTTGATGGT GAAGGTACCC CGAAACCGAA ATCTGAGACT 150
GAAGGTGACT TCGAAGAAAT TCCGGACGAA GACATCCTGG TTGGTTAG 198

(18) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 116 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(vii) IMMEDIATE SOURCE:

- (B) CLONE: oligo 1

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

AGCTTTGATA ACGAGGCGCA AAAAATGAAA AAGACAGCTA TCGCGATTGC 50
AGTGGCACTG GCTGGTTTCG CTACCGTAGC GCAGGCCGTT TCTTACACCG 100
ACTGCACCGA ATCTGG 116

(19) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 117 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

- 48 -

(ii) MOLECULE TYPE: DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Hirudinaria manillensis*

(vii) IMMEDIATE SOURCE:

(B) CLONE: oligo 2

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

AACTATTGCT CCGCGTTTTT TACTTTTTCT GTCGATAGCG CTAACGTCAC 50

CGTGACCGAC CAAAGCGGTG GCATCGCGTC CGGCAAAGAA TGTGGCTGAC 100

GTGGCTTAGA CCGGTCT

117

(20) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 88 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Hirudinaria manillensis*

(vii) IMMEDIATE SOURCE:

(B) CLONE: oligo 3

- 49 -

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

CCAGAACTAC TGCCTGTGCG TTGGTTCTAA CGTTTGCGGT GAAGGTAAAA 50
ACTGCCAGCT GTCTTCTTCT GGTAACCACT GCGTTCAC 88

(21) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 87 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Hirudinaria manillensis*

(vii) IMMEDIATE SOURCE:

- (B) CLONE: oligo 4

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

TGATGACGGA CACGCAACCA AGATTGCAAA CGCCACTTCC ATTTTGTACG 50
GTCGACAGAA GAAGACCATT GGTCACGCAA GTGCCAC 87

(22) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 85 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

- 50 -

(ii) MOLECULE TYPE: DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Hirudinaria manillensis*

(vii) IMMEDIATE SOURCE:

(B) CLONE: oligo 5

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

GGTGAAGGTA CCCCGAAACC GAAATCTCAG ACTGAAGGTG ACTTCGAAGA 50
AATTCGGGAC GAAGACATCC TGGTTGGTTA GTAAG 85

(23) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 85 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Hirudinaria manillensis*

(vii) IMMEDIATE SOURCE:

(B) CLONE: oligo 6

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

TTCCATGGGG CTTTGGCTTT AGAGTCTGAC TTCCACTGAA GCTTCTTTAA 50
GGCCTGCTTC TGTAGGACCA ACCAATCATT CCTAG 85

- 51 -

(24) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 91 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: vesicular stomatitis virus

(vii) IMMEDIATE SOURCE:

- (B) CLONE: oligo 1 VSV

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

AGCTTGGATC CACTATGAAG TGCCTTTTGT ACTTAGCCTT TTTATTCATT 50

GGGGTGAATT GCGTTTCTTA CACCGACTGC ACCGAATCTG G 91

(25) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 87 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: vesicular stomatitis virus

(vii) IMMEDIATE SOURCE:

- (B) CLONE: oligo 2 VSV

- 52 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

CCAGATTCGG TGCAGTCGGT GTAAGAAACG CAATTGACCC CAATGAATAA 50
AAAGGCTAAG TACAAAAGGC ACTTCATAGT GGATCCA 87

(26) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 37 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Hirudinaria manillensis*

(vii) IMMEDIATE SOURCE:

- (B) CLONE: oligo 1 fusion

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

GATCCATGGT TTCTTACACC GACTGCACCG AATCTGG 37

(27) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Hirudinaria manillensis*

- 53 -

(vii) IMMEDIATE SOURCE:

(B) CLONE: oligo 2 fusion

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

CCAGATTCGG TGCAGTCGGT GTAAGAAACC ATG

33

(28) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 150 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: baculovirus

(vii) IMMEDIATE SOURCE:

(B) CLONE: oligo 1 polyhedrin

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

ATCATGGAGA TAATTAAAAT GATAACCATC TCGCAAATAA ATAAGTATTT 50

TACTGTTTTTC GTAACAGTTT TGTAATAAAA AAACCTATAA ATATGCCGGA 100

TTATTCATAC CGTCCCACCA TCGGGCGTAC CTACGTGTAC GACAACACCG 150

(29) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 154 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

- 54 -

(ii) MOLECULE TYPE: DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: baculovirus

(vii) IMMEDIATE SOURCE:

(B) CLONE: oligo 2 polyhedrin

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

GATCCGGTGT TGTCGTACAC GTAGGTACGC CCGATGGTGG GACGGTATGA	50
ATAATCCGGC ATATTTATAG GTTTTTTTAT TACAAACTG TTACGAAAAC	100
AGTAGAATAC TTATTTATTT GCGAGATGGT TATCATTTTA ATTATCTCCA	150
TGAT	154

(30) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 88 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Hirudinaria manillensis

(vii) IMMEDIATE SOURCE:

(B) CLONE: oligo 3

- 55 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

CCAGAACTAC TGCCTGTGCG TTGGTGGTAA CCTGTGCGGT GGTGGTAAAC 50
ACTGCGAAAT GGATGGTTCT GGTAACAAAT GCGTTGAT 88

(31) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 87 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Hirudinaria manillensis*

(vii) IMMEDIATE SOURCE:

(B) CLONE: oligo 4

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

TGATGACGGA CACGCAACCA CCATTGGACA CGCCACCACC ATTTGTGACG 50
CTTTACCTAC CAAGACCATT GTTTACGCAA CTACCAC 87

32) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

- 56 -

(ii) MOLECULE TYPE: peptide

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Hirudinaria manillensis

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

Met	Phe	Ser	Leu	Lys	Leu	Phe	Val	Val	Phe	Leu	Ala	Val	Cys	Ile	Cys
1				5				10						15	
Val Ser Gln Ala															
20															

(33) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 204 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Hirudinaria manillensis

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

GTTTCTTACA CCGACTGCAC CGAATCTGGC CAGAACTACT GCCTGTGCGT	50
TGGTTCTAAC GTTTGCGGTG AAGGTAAAAA CTGCCAGCTG TCTTCTTCTG	100
GTAACCAGTG CGTTCACGGT GAAGGTACCC CGAAACCGAA ATCTCAGACT	150
GAAGGTGACT TCGAAGAAAT TCCGGACGAA GACATCCTGA ACGGTGCTTA	200
GTAA	204

- 57 -

CLAIMS

1. A polypeptide comprising a sequence represented by formula (I)



wherein P_{1-63} is the amino acid sequence from position 1 to position 63 of the polypeptide P1 (SEQ ID NO: 1) or P2 (SEQ ID NO: 2), Xaa is any amino acid residue and Z is -OH, -NH₂, -Gly-OH or a dipeptide consisting of:

- a glycine or a polar amino acid residue, followed by an amino acid selected from the group consisting of: alanine, glycine, valine, leucine, isoleucine, proline, cysteine, methionine, serine, threonine, lysine, arginine, histidine, asparagine, glutamine, aspartic acid and glutamic acid, with the proviso that when Z is -OH, -NH₂ or -Gly-OH, Xaa can not be asparagine and a pharmaceutically acceptable salt thereof.

2. A polypeptide according to claim 1 in which the amino acid sequence represented by formula (I) is preceded by a Met residue.

3. A polypeptide according to claim 1 in which the amino acid sequence represented by formula (I) is preceded by all or part of a leader sequence.

4. A polypeptide according to claim 3, in which all or part of the following leader sequence is present:

- 58 -

Met Phe Ser Leu Lys Leu Phe Val Val Phe Leu Ala Val Cys Ile Cys
Val Ser Gln Ala.

5. A polypeptide according to claim 1, which is a fusion protein in which the amino acid sequence represented by formula (I) is fused to a carrier sequence.

6. A process for the preparation of a polypeptide as defined in any one of the preceding claims or a pharmaceutically acceptable salt thereof, which process comprises:

(a) providing a host, transformed with an expression vector comprising a DNA sequence encoding a said polypeptide, under such conditions that the said polypeptide is expressed;

(b) isolating the said polypeptide thus obtaining a polypeptide as defined in any one of the precedings claims wherein Z is -OH or Gly-OH or a pharmaceutically acceptable salt thereof; and

(c) if desired converting the polypeptide as defined in any one of any one of the preceding claims wherein Z is -OH or Gly-OH into the corresponding C-terminal amidated derivative by way of an amidating enzyme, so obtaining a polypeptide as defined in any one of the preceding claims wherein Z is NH_2 or a pharmaceutically acceptable salt thereof.

- 59 -

7. A process according to claim 6, wherein the host is a bacterium, yeast, mammalian cell line, insect cell line or animal.

8. A process according to claim 7, wherein the host is a strain of E. coli type B or a Spodoptera frugiperda cell line.

9. A process for the preparation of a polypeptide as defined in any one of claims 1 to 5 or a pharmaceutically acceptable salt thereof, which process comprises:

- (a) chemically synthesising the said polypeptide; and
- (b) isolating the said polypeptide thus obtained or a pharmaceutically acceptable salt thereof.

10. A process according to claim 9, wherein step (a) is effected by solid-phase synthesis.

11. A pharmaceutical composition comprising a pharmaceutically acceptable carrier or diluent and, as active ingredient, a polypeptide as defined in any one of claims 1 to 5 or a pharmaceutically acceptable salt thereof.

12. An expression vector comprising a DNA sequence encoding a polypeptide as defined in any one of claims 1 to 5 wherein Z is -OH, Gly-OH or a dipeptide consisting of:

- a glycine or a polar amino acid residue, followed by an amino acid selected from the group consisting of: alanine, glycine, valine, leucine, isoleucine, proline, cysteine,

- 60 -

methionine, serine, threonine, lysine, arginine, histidine, asparagine, glutamine, aspartic acid and glutamic acid.

13. A vector according to claim 12, which is a plasmid.

14. A vector according to claim 12, which is a virus.

15. A vector according to claim 14, wherein the virus is a recombinant baculovirus in which the polyhedrin promoter is operably linked to the said DNA sequence.

16. A vector according to any one of claims 12 to 15, wherein the said DNA sequence further encodes a leader peptide capable of directing secretion of the said polypeptide from cells in which the said polypeptide is expressed.

17. A vector according to any one of claims 12 to 16, wherein the said DNA sequence encodes a fusion protein which is cleavable to release a polypeptide of formula (I) as defined in claim 1.

18. A host transformed with a compatible expression vector according to any one of claims 12 to 17.

19. A host according to claim 18, which is a bacterium, yeast, mammalian cell line, insect cell line or animal.

20. A host according to claim 19, which is a strain of E. coli type B or a Spodoptera frugiperda cell line.

21. A DNA sequence encoding a polypeptide as defined in any one of claims 1 to 5 wherein Z is -OH, Gly-OH or a dipeptide consisting of:

- 61 -

- a glycine or a polar amino acid residue, followed by an amino acid selected from the group consisting of: alanine, glycine, valine, leucine, isoleucine, proline, cysteine, methionine, serine, threonine, lysine, arginine, histidine, asparagine, glutamine, aspartic acid and glutamic acid.

22. A DNA sequence according to claim 21, which is a synthetic DNA sequence.

23. A DNA sequence according to claim 21, which is a mutagenized cDNA.

24. A DNA sequence according to any one of claims 21 to 23, which further encodes a leader peptide capable of directing secretion of the said polypeptide from cells in which the said polypeptide is expressed.

25. A DNA sequence according to any one of claims 21 to 24, which encodes a fusion protein which is cleavable to release a polypeptide of formula (I) as defined in claim 1.

26. A process for the preparation of a host in which a polypeptide as defined in any one of claims 1 to 5 wherein Z is - OH, Gly-OH or a dipeptide consisting of:

- a glycine or a polar amino acid residue, followed by an amino acid selected from the group consisting of: alanine, glycine, valine, leucine, isoleucine, proline, cysteine, methionine, serine, threonine, lysine, arginine, histidine, asparagine,

- 62 -

glutamine, aspartic acid and glutamic acid is able to be expressed, which process comprises transforming a host with a compatible expression vector according to any one of claims 12 to 17.

27. A process according to claim 26, wherein the said expression vector has been prepared by:

(a) chemically synthesising a DNA sequence encoding the said polypeptide; and

(b) inserting the said DNA sequence into an expression vector.

28. A process according to claim 26, wherein the said expression vector has been prepared by:

(a) producing and isolating a cDNA encoding the polypeptide P1 (SEQ ID NO: 1) or P2 (SEQ ID NO: 2) from mRNA from a leech of the species Hirudinaria manillensis;

(b) mutagenising the isolated cDNA so as to obtain a nucleotide sequence encoding a polypeptide as defined in any one of claims 1 to 5 wherein Z is -OH, Gly-OH or a dipeptide consisting of:

- a glycine or a polar amino acid residue, followed by an amino acid selected from the group consisting of: alanine, glycine, valine, leucine, isoleucine, proline, cysteine, methionine, serine, threonine, lysine, arginine, histidine, asparagine, glutamine, aspartic acid and glutamic acid; and

- 63 -

(c) inserting the said nucleotide sequence into an expression vector.

29. A DNA sequence according to claim 21, which consists essentially of the sequence depicted in SEQ ID NO: 16.

30. A DNA sequence according to claim 21, which consists essentially of the sequence depicted in SEQ ID NO: 17.

31. A DNA sequence according to claim 21, which consists essentially of the sequence depicted in SEQ ID NO: 33.

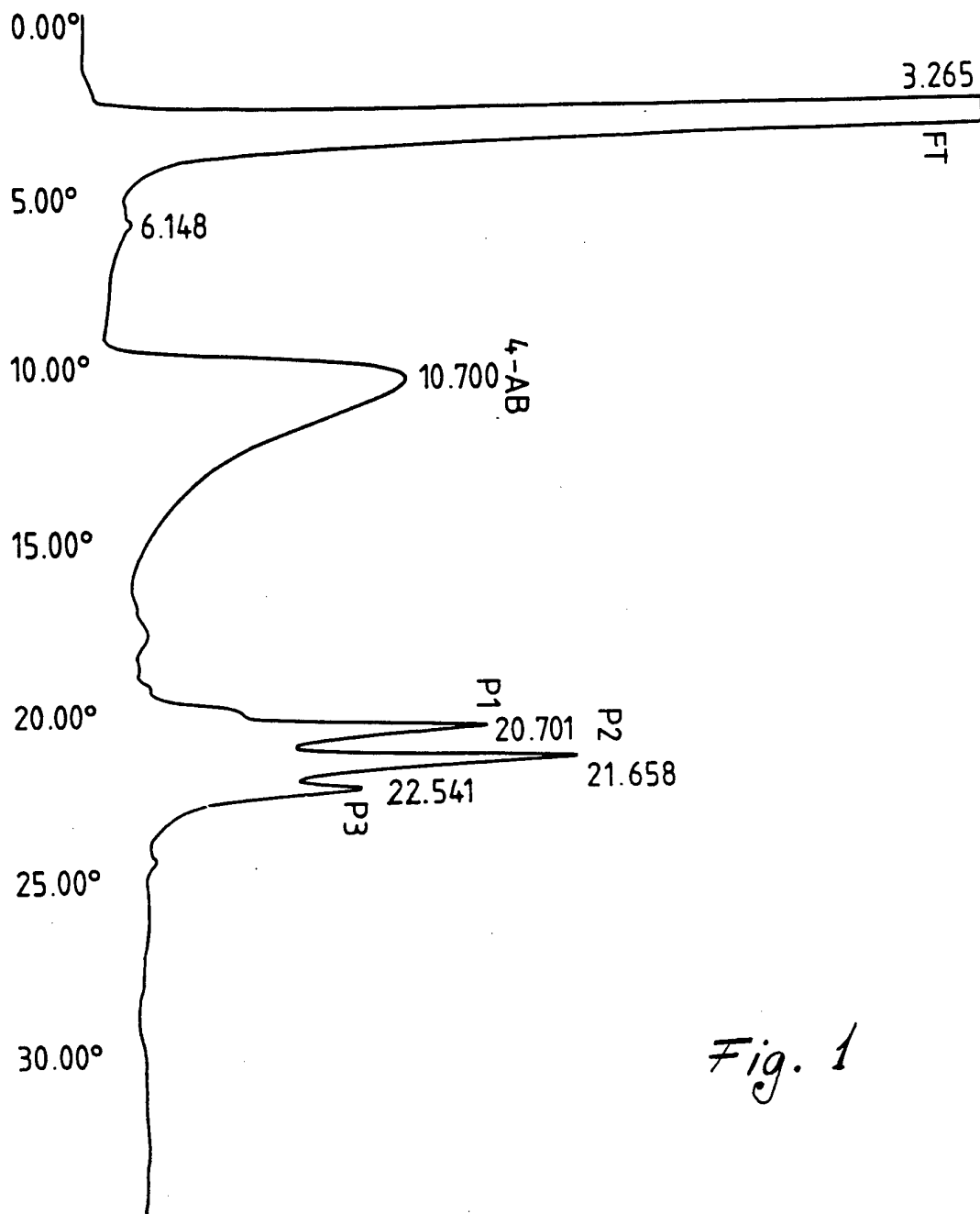


Fig. 1

CAL. METHOD		00	PA	PB	
		SF			
		.10000010+03	.10000010+01	.10000010+01	
NO.	NAME	RT	A OR H	MK	CONC
1		3.265	15519340	M	54.4024
2		6.148	363941		1.2757
3		10.708	7082058	M	24.8258
4		20.701	2114406	M	7.4119
5		21.658	2308858	M	8.0936
6		22.541	1138332		3.9903
TOTAL			28526936		100.0000

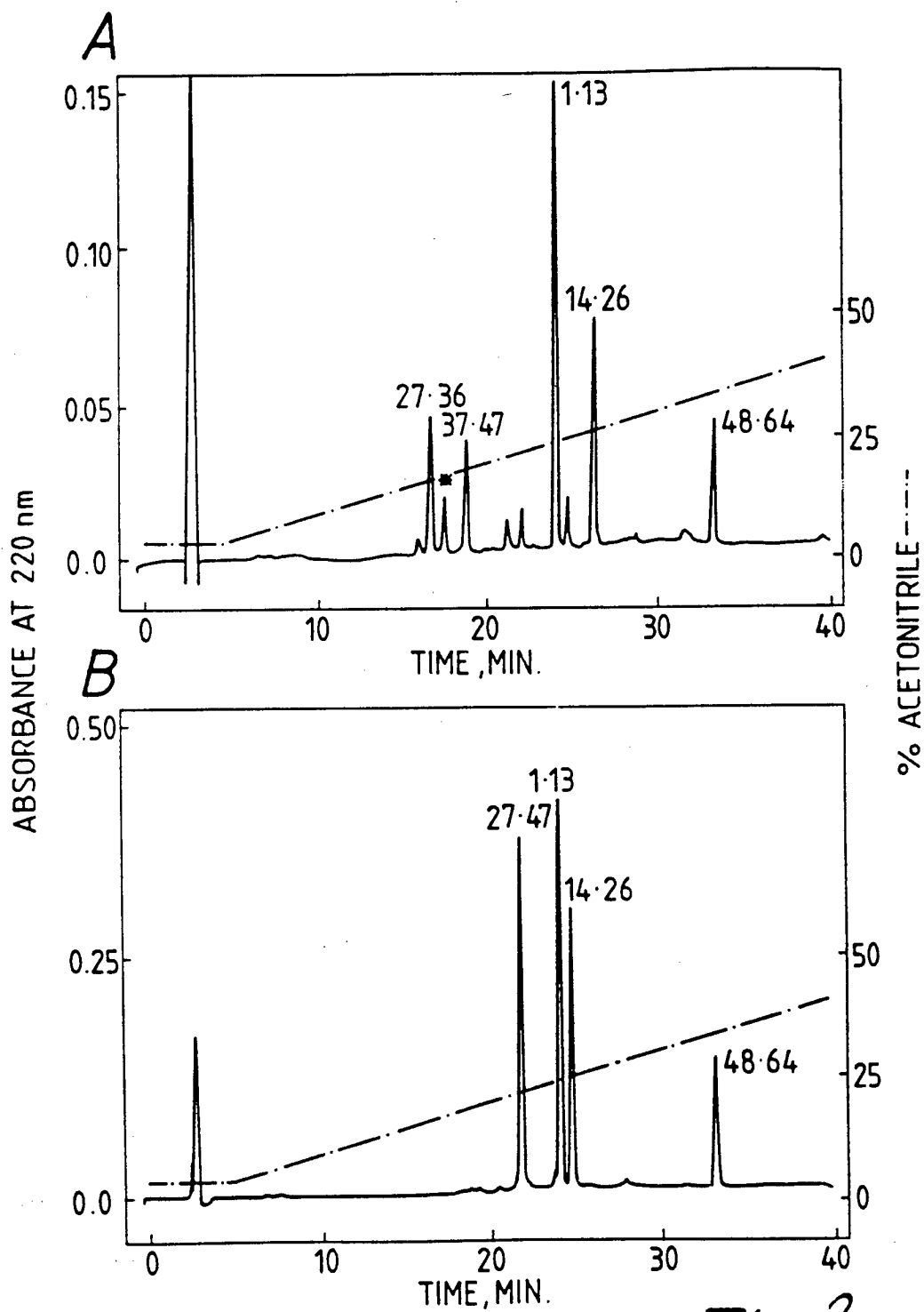


Fig. 2

HIND III
 5' AGCTTTGATAACGAGCGCGCAAAAATGAAAAAGACAGCTATCGCGATTGCGATTGCGACTGGCTGGTTTCGCTACCGTAGCGGAGGCC
 3' AACTATTGCTCCGCGTTTTTTACTTTTTCGATAGCGCTAACGTCACCGTGACCGACCAAAAGCGGTGGCATCGCGTCCGG
 OL/IG0 1 (SEQ ID NO: 18)
 OL/IG0 2 (SEQ ID NO: 19)

BAL
 GTTTCCTACAOOGACTGCACOGAATCTGG 3' 5' CCAGAACTACTGOCIGTGGTTCCTAAGTTTGGGTGAAGGTAAAACTGCCAGCTGCTCTCT
 CAAAGAATGTGGCTGACGTGGCTTAGAOCGGTCT 5' 3' TGATCAOOGGACAOCCAACCAAGATTGCCAAOCCACTTCCATTTTGTGACGGTGGACAGAAGA
 OL/IG0 3 (SEQ ID NO: 20)
 OL/IG0 4 (SEQ ID NO: 21)

VAL
 TCTGGTAACAGTGGTTTAC 3' 5' GGTGAAGGTACCCCGAACCAGAAATCTCAGACTGAAGGTGACTTGGAAAGAAATTCOCCGACGAAGACATCCTGGTT
 AGAOCATTGGTCAOCCAGTCCAC 5' 3' TTCCATGGGGCTTTGGCTTTAGAGTCTGACTTCCACTGAGCTTCTTTAAGGCCCTGCTTCTGTAGGACCAA
 OL/IG0 5 (SEQ ID NO: 22)
 OL/IG0 6 (SEQ ID NO: 23)

GLY BAM HI
 GGTTAGTAAG 3'
 CCAATCATTCCTAG 5'

Fig. 3

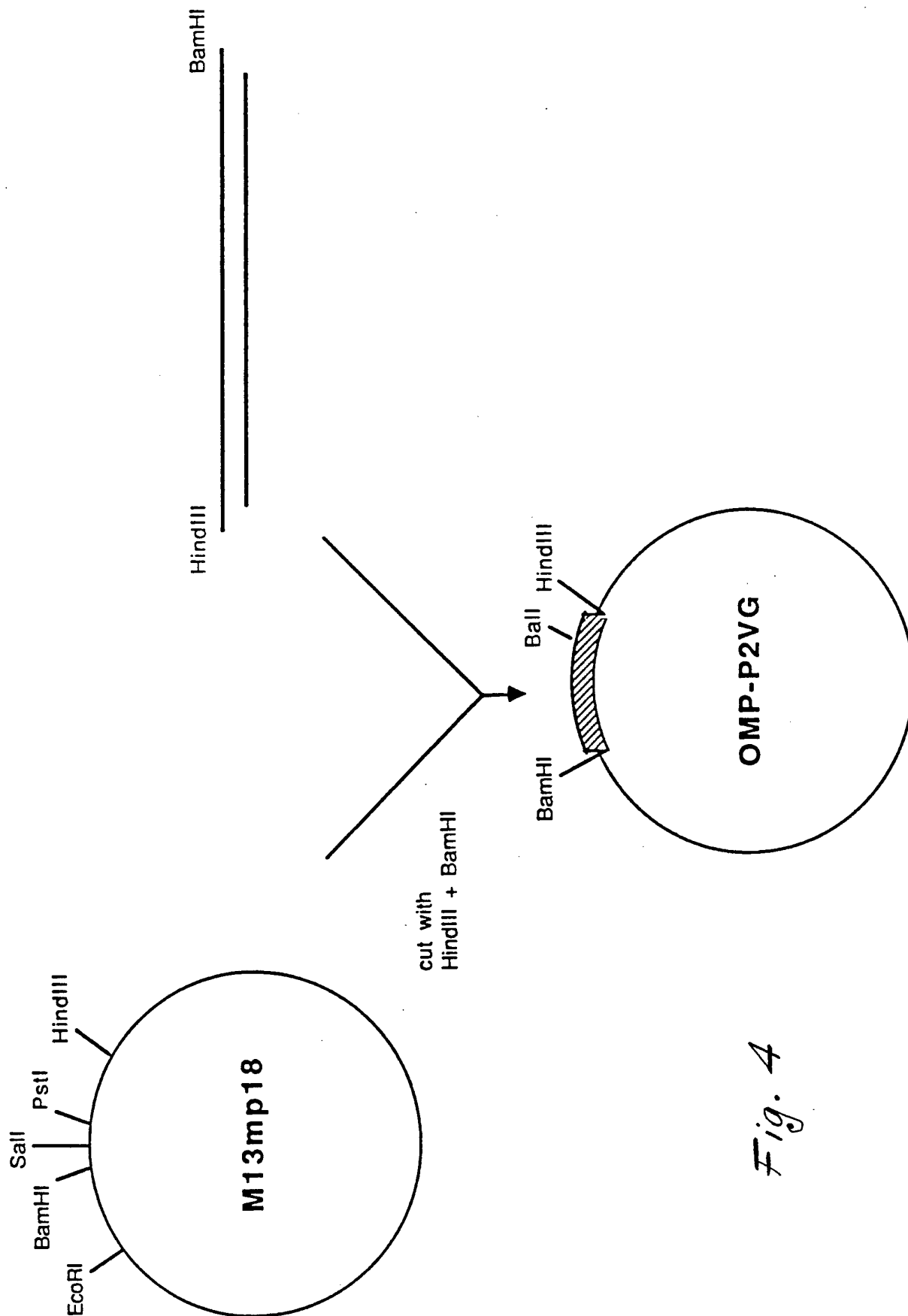
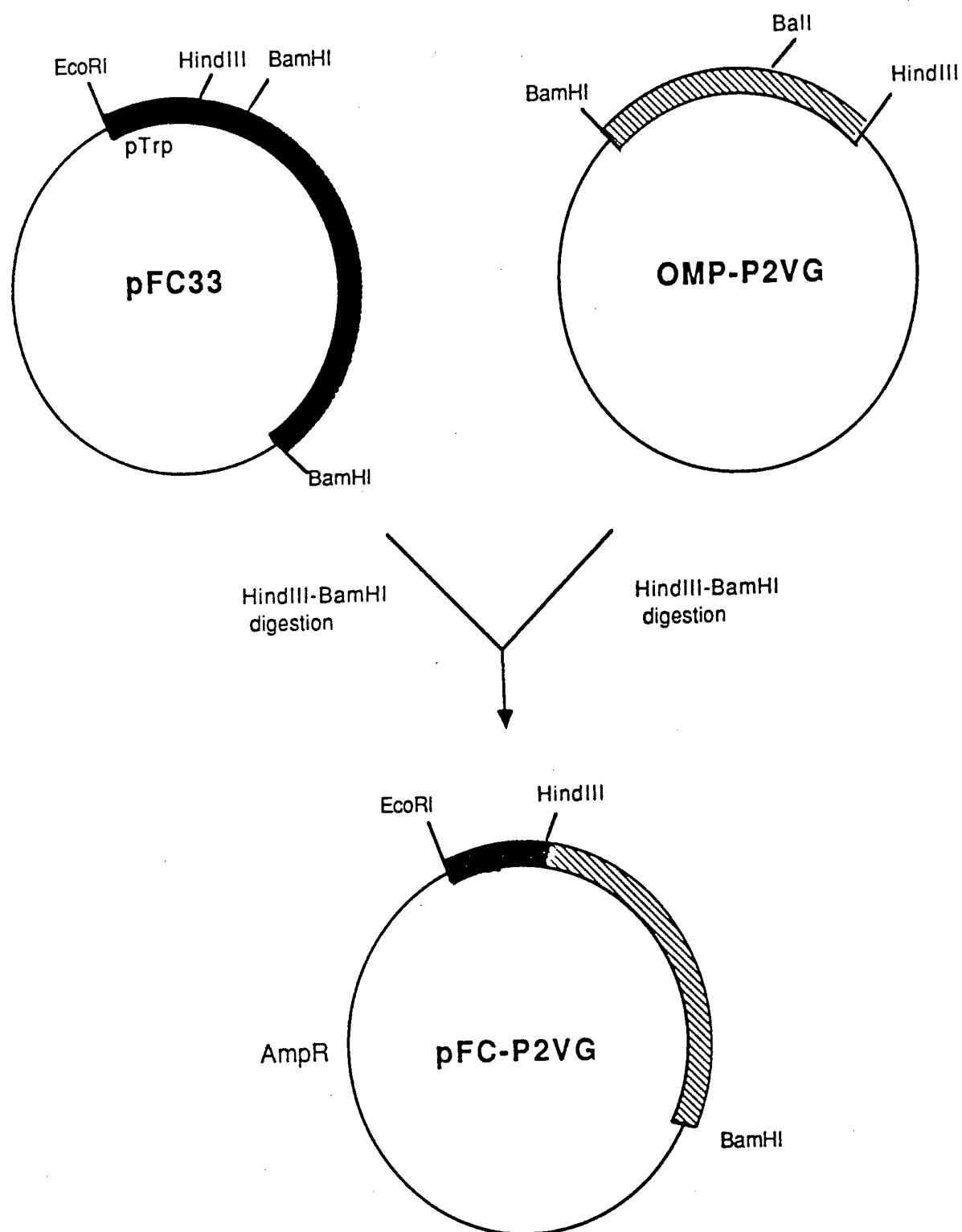
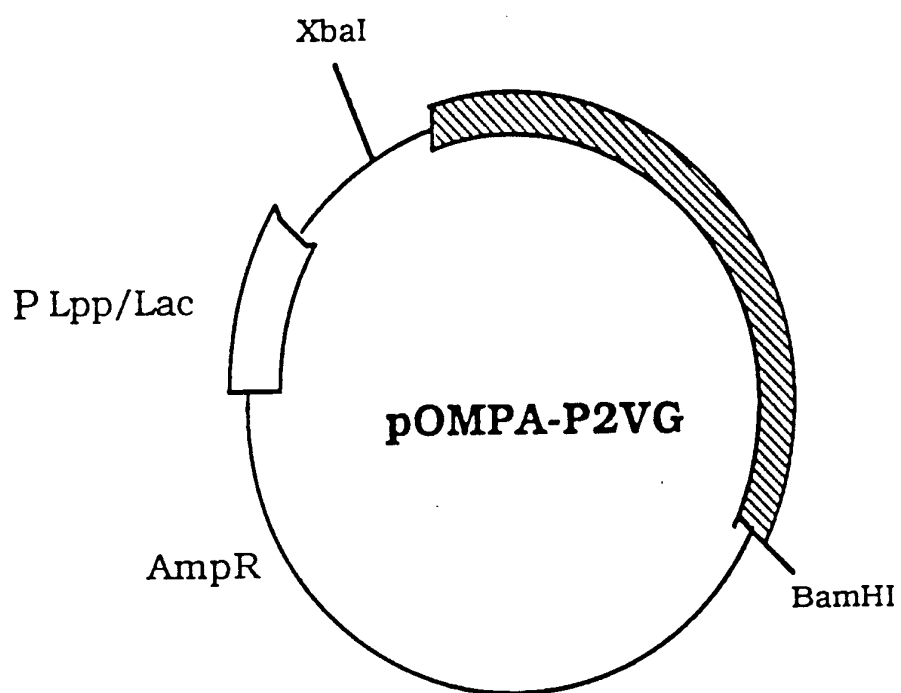


Fig. 4

*Fig. 5*

*Fig. 6*

oligo 1 VSV (SEQ ID NO: 24)

5' AGCTTGGATC CACTATGAAG TGCCTTTTGT ACTTAGCCTT TTTATTCATT GGGGTGAATT
GCGTTTCTTA CACCGACTGC ACCGAATCTG G 3'

oligo 2 VSV (SEQ ID NO: 25)

5' CCAGATTCGG TGCAGTCGGT GTAAGAAACG CAATTGACCC CAATGAATAA AAAGGCTAAG
TACAAAAGGC ACTTCATAGT GGATCCA 3'

ASSEMBLING:

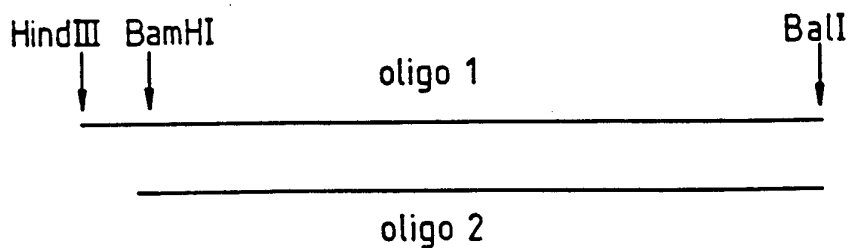


Fig. 7

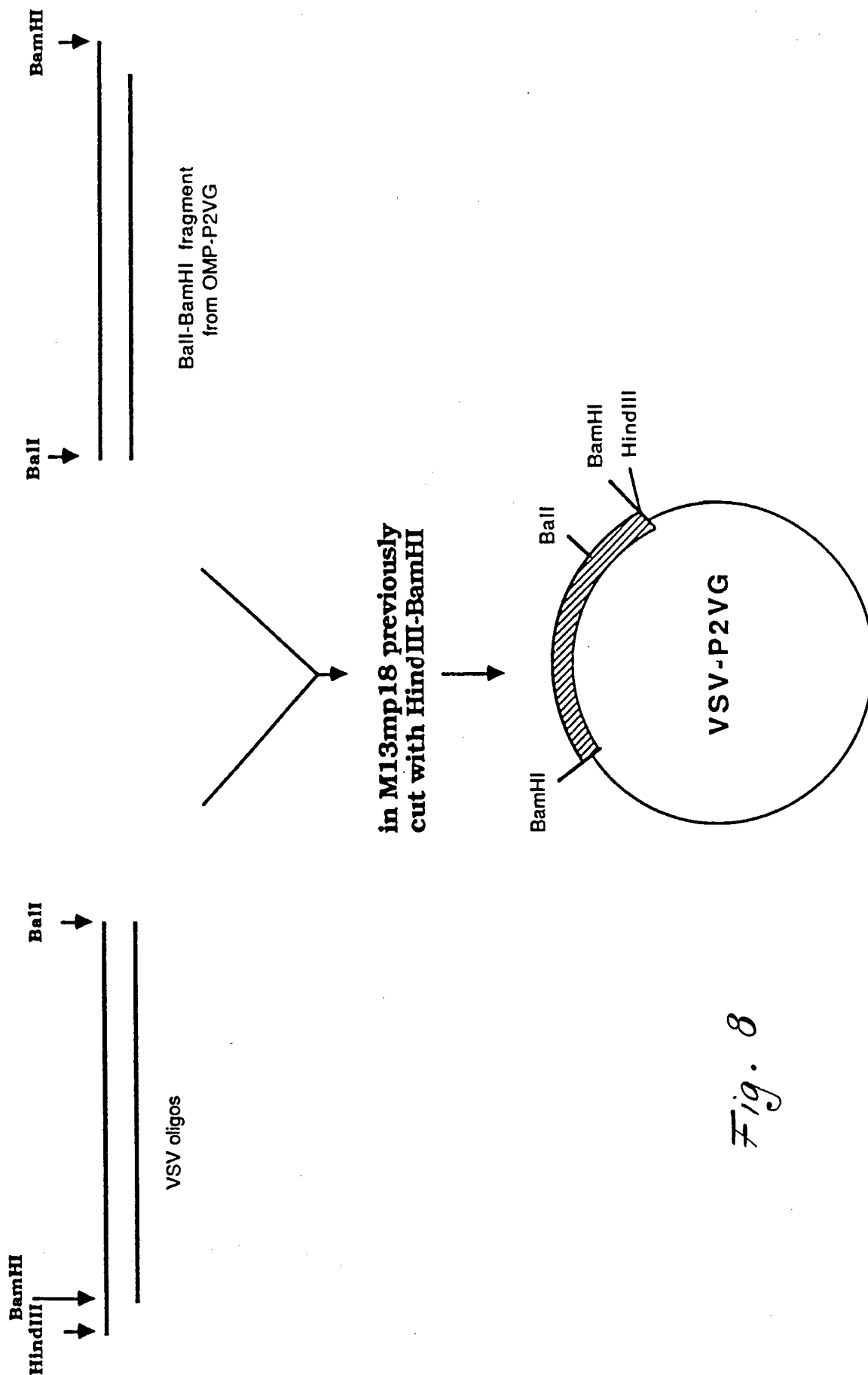


Fig. 8

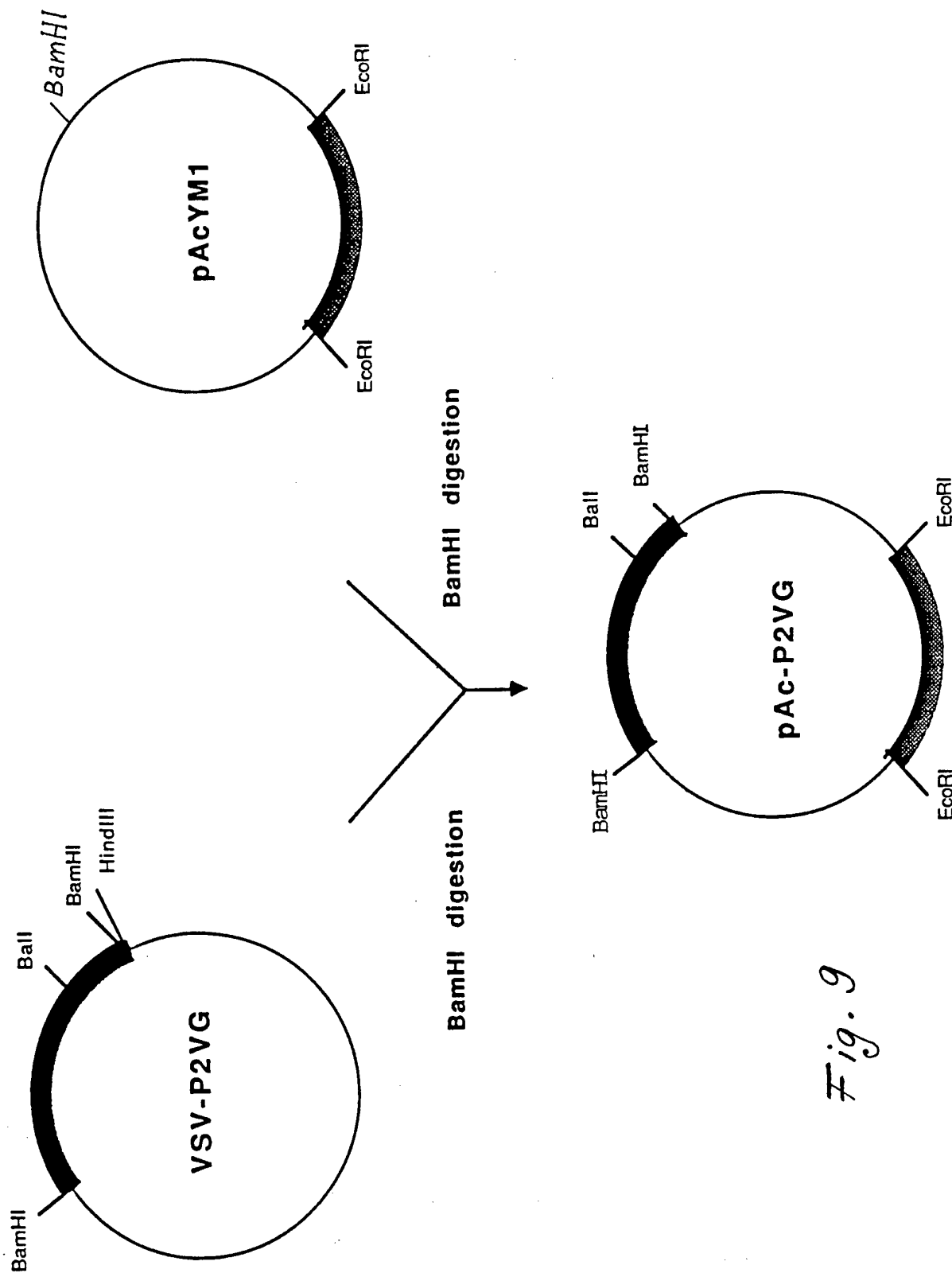


Fig. 9

oligo1 fusion (SEQ ID NO: 26)

5' GATCCATGGT TTCTTACACC GACTGCACCG AATCTGG 3'

oligo 2 fusion (SEQ ID NO: 27)

5' CCAGATTCGG TGCAGTCGGT GTAAGAAACC ATG 3'

ASSEMBLING:

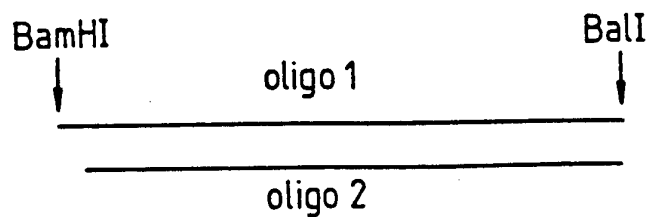


Fig. 10

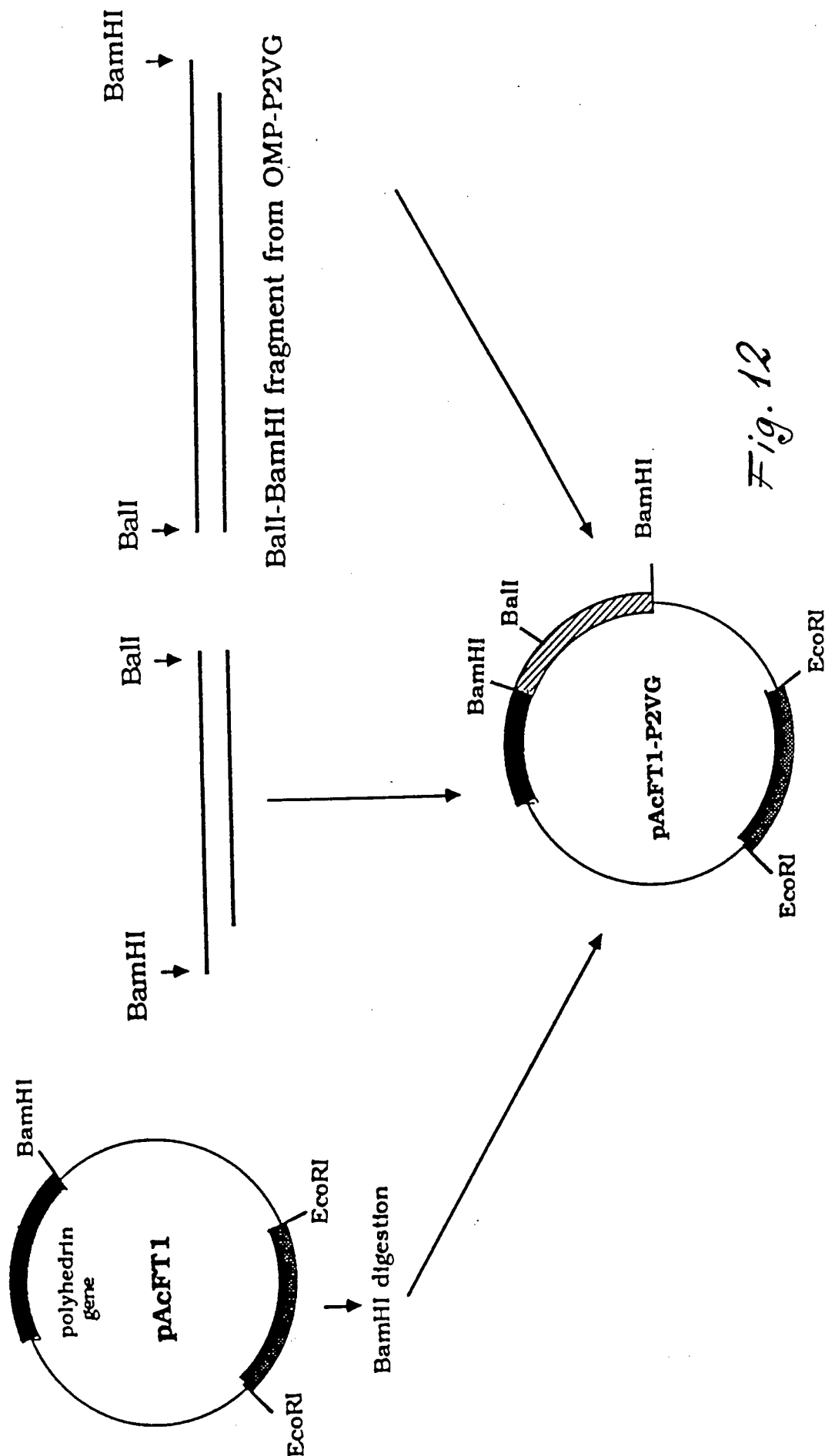


Fig. 12

HindIII
 5' AGCTTTGATAACGAGCGCAAAATGAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGTAGCGCAGGCC
 3' AACTATTGCTCCGCGTTTTTTACTTTTCTGCTAGCGCTAACGTACCGTGACCGACCAAGCGGTGGCATCGCGTCCGG
oligo 1 (SEQ ID NO: 18)
oligo 2 (SEQ ID NO: 19)

BalI
 GTTTCCTACACCGACTCCACCGAATCTGG 3' | CCAGAACTACTGCTGTGGTTGGTTAAGTAAACACTGCGAAATGGATGGT
 CAAAGAATGTCGCTGAAGTGGCTTAGACCGGTCT 5' | 3' TGATGAOAGACAAGCAACCACTTGGACAAGCCACCACTTTGTGACCGCTTTAOCCTAACA
oligo 3 (SEQ ID NO: 30)
oligo 4 (SEQ ID NO: 31)

oligo 5 (SEQ ID NO: 22)
 TCTGGTAACAATCCGTTCAT 3' | 5' GGTAAGGTACCCCGAAACCGAAATCTCAGACTGAAGGTGACTTCGAAAGAAATTCGGACGAAGACATCCTGGTT
 AGAOCATTTGTTAOCGAACCTACCAC 5' | 3' TTCCATGGGGCTTTGGCTTTAGAGTCTGACTTCCTCCACTGAAGCTTCTTTAAGGCTGCTTCGTAGGACCA
oligo 6 (SEQ ID NO: 23)

Gly **BamHI**
 GGTTAGTAAG 3'
 CCAATCATTCCTAG 5'

Fig. 13

INTERNATIONAL SEARCH REPORT

International Application No

PCT/EP 94/01979

A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 C12N15/15 C07K14/815 C07K19/00 C12N15/62 A61K38/58
 C12N15/70 C12N15/85 C12N1/21 C12N5/16 //(C12N1/21,
 C12R1:19)

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 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 practical, 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.
X	EUR. J. BIOCHEM. (1993), CODEN: EJBCAI; ISSN: 0014-2956, vol. 214, no.1, May 1993 pages 295-304, SCACHERI, EMANUELA ET AL 'Novel hirudin variants from the leech Hirudinaria manillensis. Amino acid sequence, cDNA cloning and genomic organization' see the whole document ---	1-8, 12- 31
X	EP, A, 0 501 821 (ERBA CARLO SPA) 2 September 1992 cited in the application see the whole document ---	1-31
X	EP, A, 0 520 502 (ERBA CARLO SPA) 30 December 1992 see page 4, line 17 - line 51 ---	1, 2, 6
-/--		

☒ Further documents are listed in the continuation of box C.☒ Patent family members are listed in 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

11 November 1994

Date of mailing of the international search report

29 -11- 1994

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
 NL - 2280 HV Rijswijk
 Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
 Fax: (+31-70) 340-3016

Authorized officer

Fuhr, C

INTERNATIONAL SEARCH REPORT

Internat	Application No
PCT/EP 94/01979	

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	EP,A,0 373 767 (BIOPHARM LTD) 20 June 1990 see claims; examples -----	1,11

INTERNATIONAL SEARCH REPORT

 Internati Application No
 PCT/EP 94/01979

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP-A-0501821	02-09-92	AU-B- 643876	25-11-93
		AU-A- 1122192	03-09-92
		CN-A- 1066272	18-11-92
		HU-A- 64592	28-01-94
		JP-A- 5247090	24-09-93
		US-A- 5356875	18-10-94

EP-A-0520502	30-12-92	AU-A- 2193792	25-01-93
		WO-A- 9300363	07-01-93
		EP-A- 0600901	15-06-94

EP-A-0373767	20-06-90	AT-T- 111484	15-09-94
		AU-B- 636857	13-05-93
		AU-A- 4522589	28-05-90
		CA-A- 2002924	11-05-90
		DE-D- 68918246	20-10-94
		WO-A- 9005143	17-05-90
		HU-A- 62016	29-03-93
		JP-T- 4502907	28-05-92
