

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

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
12 April 2001 (12.04.2001)

PCT

(10) International Publication Number
WO 01/25445 A1

- (51) International Patent Classification⁷: C12N 15/57, 9/64, C07K 16/40, C12N 5/10
- (21) International Application Number: PCT/US00/27029
- (22) International Filing Date:
29 September 2000 (29.09.2000)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
09/411,329 1 October 1999 (01.10.1999) US
- (71) Applicant: AMGEN INC. [US/US]; One Amgen Center Drive, Thousand Oaks, CA 91320-1799 (US).
- (72) Inventors: BOONE, Thomas, Charles; 3010 Deer Valley Avenue, Newbury Park, CA 91320 (US). LI, Huimin; 587F N. Ventu Park Road, No. 545, Newbury Park, CA 91320 (US). MANN, Michael, Benjamin; 1506 Rugby Circle, Thousand Oaks, CA 91360 (US).
- (74) Agents: ODRE, Steven, M. et al.; Amgen Inc., One Amgen Center Drive, M/S 27-4-A, Thousand Oaks, CA 91320-1799 (US).
- (81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.
- (84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).
- Published:
— With international search report.
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: FIBRINOLYTICALLY ACTIVE POLYPEPTIDE

MRFPSIFTAV LFAASSALAA PVNTTTEDET AOIPAEAVIG
 YSDLEGDFDV AVLPFSNSTN NGLLFINTTI ASIAAKEEGV
 SLEKREAEAS SIILESGNVN DYEVVYPRKV TPVPRGAVQP
 KYEDAMOYEF KVNSEPVVLH LEKNKGLFSE DYSETHYSPD
 GREITTYPLG EDHCYYHGRI ENDADSTASI SACNGLKGHF
 KLOGEMYLIE PLELSDSEAH AVYKYENVEK EDEAPKMCQV
 TONWESYEPI KKAFOFNLTK RSFPQRYVQL VIVADHRMNT
 KYNGDSDKIR QVWHQIVNTI NEIYRPLNIQ FTLVGLIWS
 NQDLITVTSV SHDTLASFGN WRETDLLRRQ RHDNAQLLTA
 IDFDGDTVGL AYVGGMCQIK HSTGVIQDHS AINLLVALTM
 AHELGHNLMG NHDGNQCHCG ANSCVMAAML SDQPSKLFSD
 CSKKDYQTFE TVNNPQCILN KP

(57) Abstract: A fibrinolytically active metalloproteinase polypeptide (called "novel acting thrombolytic") which is useful for blood clot lysis *in vivo* and methods and materials for its production by recombinant expression are described.

WO 01/25445 A1

- 1 -

FIBRINOLYTICALLY ACTIVE POLYPEPTIDE

5 Field of the Invention

 This invention relates to a fibrinolytically
active metalloproteinase of non-naturally occurring
sequence, to combinant methods for its manufacture, and
10 to its use in treating thrombosis *in vivo*.

Background of the Invention

 Fibrolase is an enzymatically active
15 polypeptide (specifically, a metalloproteinase)
composed of 203 amino acid residues that was originally
isolated by purification from the venom of the Southern
Copperhead snake; United States Patent No. 4,610,879,
issued September 9, 1986 (Markland et al.); and Guan et
20 al., Archives of Biochemistry and Biophysics, Volume
289, Number 2, pages 197-207 (1991). The enzyme
exhibits direct fibrinolytic activity with little or no
hemorrhagic activity, and it dissolves blood clots made
either from fibrinogen or from whole blood.

25 The amino acid sequence of fibrolase has also
been determined, with methods described for recombinant
production in yeast and use for the treatment of
thrombembolic conditions *in vivo*; Randolph et al.,
30 Protein Science, Cambridge University Press (1992),
pages 590-600, and European Patent Application No.
0 323 722 (Valenzuela et al.), published July 12, 1989.

- 2 -

Summary of the Invention

This invention provides a fibinolytic
5 metalloproteinase having the non-naturally occurring
linear array of amino acids depicted in SEQ ID NO: 1,
also referred to herein as "novel acting thrombolytic"
(or "NAT"). Also provided are nucleic acid molecules,
such as the one of SEQ ID NO: 2 and variants thereof
10 encoding NAT.

The term "mature" is used in its conventional
sense to refer to the biologically active polypeptide
which has been enzymatically processed *in situ* in the
15 host cell to cleave it from the prepro region.

Because of its fibrinolytic activity, NAT is
useful *in vivo* as a blood clot lysing agent to treat
thrombosis in a mammal (including rats, pigs and
20 humans).

The NAT polypeptide of this invention provides
advantages over naturally occurring fibrolase as a
therapeutic agent (i.e., the fibrinolytic polypeptide
25 found in snake venom). Native fibrolase is known to
contain several alternate N-termini: QQRFP, EQRFP and
ERFP (in which "E" designates a cyclized glutamine, or
pyroglutamic acid). More specifically, starting with
an N-terminus composed of QQRFP, the fibrolase molecule
30 undergoes degradation to result in two isoforms, having
N-terminal sequences of EQRFP and ERFP, respectively.
Recombinant fibrolase as produced in yeast typically
yields a mixture of all three of these forms and is
thus not homogeneous. See Loayza et al., Journal of
35 Chromatography, B 662, pages 227-243 (1994). Moreover,

- 3 -

the cyclized glutamine residue results in a "blocked" N-terminus which makes sequencing impossible.

In contrast, the recombinant NAT of this invention provides a single species: only one N-terminus is typically produced. The result is greater homogeneity of the end product compared to recombinant fibrolase, which is beneficial when medical applications are the intended end use.

Brief Description of the Figures

FIGURE 1 depicts in linear fashion the full amino acid sequence of NAT (SEQ ID NO: 3), consisting of the "prepro" region (underscored), which includes the "signal" peptide, and the mature polypeptide (non-underscored).

Detailed Description of the Invention

NAT may be produced by recombinant expression of the nucleic acid molecule of SEQ ID NO: 4, which encodes the full amino acid sequence of NAT (SEQ ID NO: 3), including the prepro region from nucleotides 1-783 and mature polypeptide from nucleotides 784-1386, in a suitable host. Following expression, the prepro region is enzymatically processed off in the host cell to yield the mature active polypeptide (SEQ ID NO: 1).

Preferably, NAT is produced recombinantly in yeast, as will be explained in greater detail further below.

The mature polypeptide (SEQ ID NO: 1) which is thus produced may or may not have an amino terminal

- 4 -

methionine, depending on the manner in which it is prepared. Typically, an amino terminal methionine residue will be present when the polypeptide is produced recombinantly in a non-secreting bacterial
5 (e.g., *E. coli*) strain as the host.

Besides the nucleic acid molecule of SEQ ID NO: 2, also utilizable are degenerate sequences thereof which encode the same polypeptide. The present
10 invention also embraces nucleic acid molecules that may encode additional amino acid residues flanking the 5' or 3' portions of the region encoding the mature polypeptide, such as sequences encoding alternative pre/pro regions (i.e., sequences responsible for
15 secretion of the polypeptide through cell membranes) in place of the "native" pre/pro region (i.e., found in naturally occurring fibrolase). The additional sequences may also be noncoding sequences, including regulatory sequences such as promoters of transcription
20 or translation, depending on the host cell.

NAT can be prepared using well known recombinant DNA technology methods, such as those set forth in Sambrook et al., Molecular Cloning: A
25 Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1989) and/or Ausubel et al., editors, Current Protocols in Molecular Biology, Green Publishers Inc. and Wiley and Sons, New York (1994). A DNA molecule encoding the polypeptide or truncated
30 version thereof may be obtained, for example, by screening a genomic or cDNA library, or by PCR amplification, to obtain a nucleic acid molecule encoding fibrolase, followed by replacement of the codons encoding the N-terminal amino acid residues QQR
35 with a codon for serine (S). Alternatively, a DNA molecule encoding NAT may be prepared by chemical

- 5 -

synthesis using methods well known to the skilled artisan, such as those described by Engels et al. in Angew. Chem. Intl. Ed., Volume 28, pages 716-734 (1989). Typically, the DNA will be several hundred
5 nucleotides in length. Nucleic acids larger than about one hundred nucleotides can be synthesized as several fragments using these same methods and the fragments can then be ligated together to form a nucleotide sequence of the desired length.

10

The DNA molecule is inserted into an appropriate expression vector for expression in a suitable host cell. The vector is selected to be functional in the particular host cell employed (i.e.,
15 the vector is compatible with the host cell machinery, such that expression of the DNA can occur). The polypeptide may be expressed in prokaryotic, yeast, insect (baculovirus systems) or eukaryotic host cells, although yeast is preferred as will be explained in
20 greater detail further below.

The vectors used in any of the host cells to express NAT may also contain a 5' flanking sequence (also referred to as a "promoter") and other expression
25 regulatory elements operatively linked to the DNA to be expressed, as well as enhancer(s), an origin of replication element, a transcriptional termination element, a complete intron sequence containing a donor and acceptor splice site, a signal peptide sequence, a
30 ribosome binding site element, a polyadenylation sequence, a polylinker region for inserting the nucleic acid encoding NAT, and a selectable marker element. Each of these elements is discussed below.

35 Optionally, the vector may also contain a "tag" sequence, i.e., an oligonucleotide sequence

- 6 -

located at the 5' or 3' end of the polypeptide-coding sequence that encodes polyHis (such as hexaHis) or another small immunogenic sequence (such as the c-myc or hemagglutinin epitope, for which antibodies, including monoclonal antibodies, are commercially available). This tag will be expressed along with NAT, and can serve as an affinity tag for purification of this polypeptide from the host cell. Optionally, the tag can subsequently be removed from the purified polypeptide by various means, for example, with use of a selective peptidase.

The 5' flanking sequence may be the native 5' flanking sequence, or it may be homologous (i.e., from the same species and/or strain as the host cell), heterologous (i.e., from a species other than the host cell species or strain), hybrid (i.e., a combination of 5' flanking sequences from more than one source), or synthetic. The source of the 5' flanking sequence may be any unicellular prokaryotic or eukaryotic organism, any vertebrate or invertebrate organism, or any plant, provided that the 5' flanking sequence is functional in, and can be activated by the host cell machinery.

The origin of replication element is typically a part of prokaryotic expression vectors purchased commercially and aids in the amplification of the vector in a host cell. Amplification of the vector to a certain copy number can, in some cases, be important for optimal expression of NAT. If the vector of choice does not contain an origin of replication site, one may be chemically synthesized based on a known sequence, and then ligated into the vector.

The transcription termination element is typically located 3' to the end of the polypeptide

- 7 -

coding sequence and serves to terminate transcription of the mRNA. Usually, the transcription termination element in prokaryotic cells is a G-C rich fragment followed by a poly T sequence. While the element is easily cloned from a library or even purchased commercially as part of a vector, it can also be readily synthesized using known methods for nucleic acid synthesis.

10 A selectable marker gene element encodes a protein necessary for the survival and growth of a host cell grown in a selective culture medium. Typical selection marker genes encode proteins that: (a) confer resistance to antibiotics or other toxins, for example, 15 ampicillin, tetracycline or kanamycin for prokaryotic host cells, zeocin for yeast host cells, and neomycin for mammalian host cells; (b) complement auxotrophic deficiencies of the cell; or (c) supply critical nutrients not available from complex media. Preferred 20 selectable markers for use in prokaryotic expression are the kanamycin resistance gene, the ampicillin resistance gene, and the tetracycline resistance gene.

The ribosome binding element, commonly called 25 the Shine-Dalgarno sequence (for prokaryotes) or the Kozak sequence (for eukaryotes), is necessary for the initiation of translation of mRNA. The element is typically located 3' to the promoter and 5' to the coding sequence of the polypeptide to be synthesized. 30 The Shine-Dalgarno sequence is varied but is typically a polypurine (i.e., having a high A-G content). Many Shine-Dalgarno sequences have been identified, each of which can be readily synthesized using methods set forth above and used in a prokaryotic vector. The 35 Kozak sequence typically includes sequences immediately before and after the initiating codon. A preferred

- 8 -

Kozak sequence is one that is associated with a high efficiency of initiation of translation at the AUG start codon.

5 In those cases where it is desirable for NAT polypeptide to be secreted from the host cell, a signal sequence may be used to direct the polypeptide out of the host cell where it is synthesized. Typically, the signal sequence is positioned in the coding region of
10 nucleic acid sequence, or directly at the 5' end of the coding region. Many signal sequences have been identified, and any of them that are functional in the selected host cell may be used here. Consequently, the signal sequence may be homologous or heterologous to
15 the polypeptide. Additionally, the signal sequence may be chemically synthesized using methods referred to above.

20 After the vector has been constructed and a nucleic acid has been inserted into the proper site of the vector, the completed vector may be inserted into a suitable host cell for amplification and/or polypeptide expression.

25 As mentioned, host cells may be prokaryotic (such as *E. coli*) or eukaryotic (such as a yeast cell, an insect cell, or a vertebrate cell). The host cell, whether it be yeast or some other host, when cultured under appropriate conditions can synthesize NAT, which
30 can subsequently be collected from the culture medium (if the host cell secretes it into the medium) or directly from the host cell producing it (if it is not secreted). After collection, NAT polypeptide can be purified using methods such as molecular sieve
35 chromatography, affinity chromatography, and the like.

- 9 -

Selection of the host cell will depend in large part on whether the manner in which the host cell is able to "fold" NAT into its native secondary and tertiary structure (e.g., proper orientation of
5 disulfide bridges, etc.) such that biologically active material is prepared by the cell. However, even where the host cell does not synthesize biologically active material, it may be "folded" after synthesis using appropriate chemical conditions, such as ones that are
10 known to those skilled in the art. In either case, proper folding can be inferred from the fact that biologically active material has been obtained.

Suitable host cells or cell lines may be
15 mammalian cells, such as Chinese hamster ovary cells (CHO) or 3T3 cells. The selection of suitable mammalian host cells and methods for transformation, culture, amplification, screening and product production and purification are known in the art.
20 Other suitable mammalian cell lines are the monkey COS-1 and COS-7 cell lines, and the CV-1 cell line. Further exemplary mammalian host cells include primate cell lines and rodent cell lines, including transformed cell lines. Normal diploid cells, cell strains derived
25 from *in vitro* culture of primary tissue, as well as primary explants, are also suitable. Candidate cells may be genotypically deficient in the selection gene, or may contain a dominantly acting selection gene. Still other suitable mammalian cell lines include but
30 are not limited to, HeLa, mouse L-929 cells, 3T3 lines derived from Swiss, Balb-c or NIH mice, BHK or HaK hamster cell lines.

Also useful as host cells are bacterial
35 cells. For example, the various strains of *E. coli* (e.g., HB101, DH5a, DH10, and MC1061) are well-known as

- 10 -

host cells in the field of biotechnology. Various strains of *B. subtilis*, *Pseudomonas* spp., other *Bacillus* spp., *Streptomyces* spp., and the like, may also be employed. Additionally, many strains of yeast
5 cells known to those skilled in the art are also available as host cells for expression of the polypeptide of the present invention. Also, where desired, insect cells may be utilized as host cells. See, for example, Miller et al., *Genetic Engineering*,
10 Volume 8, pages 277-298 (1986).

Insertion (also referred to as "transformation" or "transfection") of the vector into the selected host cell may be accomplished using such
15 methods as calcium phosphate, electroporation, microinjection, lipofection or the DEAE-dextran method. The method selected will in part be a function of the type of host cell to be used. These methods and other suitable methods are well known to the skilled artisan,
20 and are set forth, for example, in Sambrook et al., above.

The host cells containing the vector may be cultured using standard media well known to the skilled
25 artisan. The media will usually contain all nutrients necessary for the growth and survival of the cells. Suitable media for culturing *E. coli* cells are, for example, Luria Broth (LB) and/or Terrific Broth (TB). Suitable media for culturing eukaryotic cells are RPMI
30 1640, MEM, DMEM, all of which may be supplemented with serum and/or growth factors as required by the particular cell line being cultured. A suitable medium for insect cultures is Grace's medium supplemented with yeastolate, lactalbumin hydrolysate and/or fetal calf
35 serum, as necessary.

- 11 -

Typically, an antibiotic or other compound useful for selective growth of the transformed cells only is added as a supplement to the media. The compound to be used will be dictated by the selectable
5 marker element present on the plasmid with which the host cell was transformed. For example, where the selectable marker element is kanamycin resistance, the compound added to the culture medium will be kanamycin.

10 The amount of NAT produced in the host cell can be evaluated using standard methods known in the art. Such methods include, without limitation, Western blot analysis, SDS-polyacrylamide gel electrophoresis, non-denaturing gel electrophoresis, HPLC separation,
15 immunoprecipitation, and/or activity assays such as DNA binding gel shift assays.

If NAT is secreted from the host cells other than gram-negative bacteria, the majority will likely
20 be found in the cell culture medium. If NAT is secreted from gram-negative bacteria, it will to some degree be found in the periplasm. If NAT is not secreted, it will be present in the cytoplasm.

25 For intracellular NAT, the host cells are typically first disrupted mechanically. For NAT having a periplasmic location, either mechanical disruption or osmotic treatment can be used to release the periplasmic contents into a buffered solution. NAT
30 polypeptide is then isolated from this solution. Purification from solution can thereafter be accomplished using a variety of techniques. If NAT has been synthesized so that it contains a tag such as hexahistidine or other small peptide at either its
35 carboxyl or amino terminus, it may essentially be purified in a one-step process by passing the solution

- 12 -

through an affinity column where the column matrix has a high affinity for the tag or for the polypeptide directly (i.e., a monoclonal antibody). For example, polyhistidine binds with great affinity and specificity to nickel, thus an affinity column of nickel (such as the Qiagen nickel columns) can be used for purification. (See, for example, Ausubel et al., editors, Current Protocols in Molecular Biology, above).

10

Where, on the other hand, the polypeptide has no tag and no antibodies are available, other well known procedures for purification can be used. Such procedures include, without limitation, ion exchange chromatography, molecular sieve chromatography, reversed phase chromatography, HPLC, native gel electrophoresis in combination with gel elution, and preparative isoelectric focusing ("Isoprime" machine/technique, Hoefer Scientific). In some cases, two or more of these techniques may be combined to achieve increased purity.

15
20

Especially preferred for use in the production of NAT are yeast cells, and most advantageously those of the yeast genus known as *Pichia* (e.g., *Pichia pastoris*), because of the greater efficiency of refolding compared to, for instance, bacterial cells such as *E. coli*. Suitable recombinant methods of expression for this yeast strain are described in United States Patent Nos. 4,855,231 (Stroman et al.), 4,812,405 (Lair et al.), 4,818,700 (Cregg et al.), 4,885,242 (Cregg) and 4,837,148 (Cregg), the disclosures of which are incorporated herein by reference.

30
35

- 13 -

Notably, *Pichia* cells can also be used to express fibrolase with similar efficiency from DNA molecules encoding this metalloproteinase, and such a method constitutes an additional aspect of the present invention. Fibrolase is a known metalloproteinase which has been described in the scientific and patent literature; see Randolph et al., and European Patent Application No. 0 323 722, cited above. Typically, the fibrolase to be expressed will be of SEQ ID NO: 5, which is encoded by the cDNA molecule of SEQ ID NO: 6 (or variants thereof encoding the same amino acid sequence). The expression of fibrolase in such a system will typically involve a DNA molecule of SEQ ID NO: 7, which encodes "prepro" sequence (nucleotides 1-783) in addition to the "mature" polypeptide (nucleotides 784-1392).

Chemically modified versions of NAT in which the polypeptide is linked to a polymer or other molecule to form a derivative in order to modify properties are also included within the scope of the present invention. For human therapeutic purposes especially, it may be advantageous to derivatize NAT in such a manner by the attachment of one or more other chemical moieties to the polypeptide moiety. Such chemical moieties may be selected from among various water soluble polymers. The polymer should be water soluble so that the NAT polypeptide to which it is attached is miscible in an aqueous environment, such as a physiological environment. The water soluble polymer may be selected from the group consisting of, for example, polyethylene glycol, copolymers of ethylene glycol/propylene glycol, carboxymethylcellulose, dextran, polyvinyl alcohol, polyvinyl pyrrolidone, poly-1,3-dioxolane, poly-1,3,6-trioxane, ethylene/maleic anhydride copolymer, polyaminoacids

- 14 -

(either homopolymers or random or non-random copolymers (see further below regarding fusion molecules), and dextran or poly(n-vinyl pyrrolidone)polyethylene glycol, propylene glycol homopolymers, polypropylene
5 oxide/ethylene oxide co-polymers, polyoxyethylated polyols, polystyrenemaleate and polyvinyl alcohol. Polyethylene glycol propionaldenhyde may have advantages in manufacturing due to its stability in water.

10

The polymer may be of any molecular weight, and may be branched or unbranched. For polyethylene glycol, the preferred molecular weight is between about 2 kilodaltons (kDa) and about 100 kDa (the term "about"
15 indicating that in preparations of polyethylene glycol, some molecules will weigh more, some less, than the stated molecular weight) for ease in handling and manufacturing. Other sizes may be used, depending on the desired therapeutic profile (e.g., the duration of
20 sustained release desired, the effects, if any on biological activity, the ease in handling, the degree or lack of antigenicity and other known effects of the polyethylene glycol on a therapeutic protein).

25

The number of polymer molecules so attached may vary, and one skilled in the art will be able to ascertain the effect on function. One may mono-derivatize, or may provide for a di-, tri-, tetra- or some combination of derivatization, with the same or
30 different chemical moieties (e.g., polymers, such as different weights of polyethylene glycols). The proportion of polymer molecules to NAT polypeptide molecules will vary, as will their concentrations in the reaction mixture. In general, the optimum ratio
35 (in terms of efficiency of reaction in that there is no excess unreacted polypeptide or polymer) will be

- 15 -

determined by factors such as the desired degree of derivatization (e.g., mono, di-, tri-, etc.), the molecular weight of the polymer selected, whether the polymer is branched or unbranched, and the reaction
5 conditions.

The chemical moieties should be attached to NAT with consideration of effects on functional or antigenic domains of the polypeptide. There are a
10 number of attachment methods available to those skilled in the art. See, for example, EP 0 401 384 (coupling PEG to G-CSF), and Malik et al., Experimental Hematology, Volume 20, pages 1028-1035 (1992) (reporting the pegylation of GM-CSF using tresyl
15 chloride). By way of illustration, polyethylene glycol may be covalently bound through amino acid residues via a reactive group, such as, a free amino or carboxyl group. Reactive groups are those to which an activated polyethylene glycol molecule (or other chemical moiety)
20 may be bound. The amino acid residues having a free amino group may include lysine residues and the N-terminal amino acid residue. Those having a free carboxyl group may include aspartic acid residues, glutamic acid residues, and the C-terminal amino acid
25 residue. Sulfhydryl groups may also be used as a reactive group for attaching the polyethylene glycol molecule(s) (or other chemical moiety). Preferred for manufacturing purposes is attachment at an amino group, such as at the N-terminus or to a lysine group.
30 Attachment at residues important for receptor binding should be avoided if receptor binding is desired.

One may specifically desire N-terminally chemically modified derivatives. Using polyethylene
35 glycol as an illustration, one may select from a variety of polyethylene glycol molecules (by molecular

- 16 -

weight, branching, etc.), the proportion of polyethylene glycol molecules to polypeptide molecules in the reaction mixture, the type of pegylation reaction to be performed, and the method of obtaining the selected N-terminally pegylated NAT. The method of obtaining the N-terminally pegylated preparation (i.e., separating this moiety from other monopegylated moieties if necessary) may be by purification of the N-terminally pegylated material from a population of pegylated NAT molecules. Selective N-terminal chemical modification may be accomplished by reductive alkylation which exploits differential reactivity of different types of primary amino groups (lysine versus the N-terminal) available for derivatization. See PCT application WO 96/11953, published April 25, 1996. Under the appropriate reaction conditions, substantially selective derivatization of NAT at the N-terminus with a carbonyl group containing polymer is achieved. For example, one may selectively N-terminally pegylate NAT by performing the reaction at a pH which allows one to take advantage of the pK_a differences between the ϵ -amino group of the lysine residues and that of the α -amino group of the N-terminal residue of the polypeptide. By such selective derivatization, attachment of a polymer to a polypeptide is controlled: the conjugation with the polymer takes place predominantly at the N-terminus of the polypeptide and no significant modification of other reactive groups, such as lysine side chain amino groups, occurs. Using reductive alkylation, the polymer may be of the type described above, and should have a single reactive aldehyde for coupling to the polypeptide. Polyethylene glycol propionaldehyde, containing a single reactive aldehyde, may be used.

35

- 17 -

NAT or chemically modified derivatives in accordance with the invention may be formulated for in vivo administration, and most preferably via intra-thrombus (i.e., via localized delivery directly to the site of the clot in the blood vessel, e.g., as by catheter). Systemic delivery is normally not preferred due to the likelihood that innate α_2 macroglobulin in the general circulation may complex with NAT to prevent interaction with fibrin or fibrinogen, thus impairing clot lysis. However, there may be instances where larger amounts of NAT can be used which exceed the circulating levels of α_2 macroglobulin, thus enabling systemic administration and delivery. In general, encompassed within the invention are pharmaceutical compositions comprising effective amounts of NAT together with pharmaceutically acceptable diluents, preservatives, solubilizers, emulsifiers, adjuvants and/or carriers. By "effective amount" is meant amount sufficient to produce a measurable biological effect (i.e., a thrombolytically effective amount which effects lysis of the blood clot or clots being treated).

Typically, NAT will be in highly purified form, and any pharmaceutical composition being used as the delivery vehicle will normally be presterilized for use, such as by filtration through sterile filtration membranes.

One skilled in the art will be able to ascertain effective dosages by administration and observing the desired therapeutic effect. Particular effective doses within this range will depend on the particular disorder or condition being treated, as well as the age and general health of the recipient, and can be determined by standard clinical procedures. Where

- 18 -

possible, it will be desirable to determine the dose-response curve of the pharmaceutical composition first *in vitro*, as in bioassay systems, and then in useful animal model systems *in vivo* prior to testing in humans. The skilled practitioner, considering the therapeutic context, type of disorder under treatment, and other applicable factors, will be able to ascertain proper dosing without undue effort. Typically, a practitioner will administer the NAT composition until a dosage is reached that achieves the desired effect (i.e., lysis of the blood clot). The composition may be administered as a single dose, or as two or more doses (which may or may not contain the same amount of polypeptide) over time, or on a continuous basis.

NAT may also be used to generate antibodies in accordance with standard methods. The antibodies may be polyclonal, monoclonal, recombinant, chimeric, single-chain and/or bispecific, etc. To improve the likelihood of producing an immune response, the amino acid sequence of NAT can be analyzed to identify portions of the molecule that may be associated with increased immunogenicity. For example, the amino acid sequence may be subjected to computer analysis to identify surface epitopes, such as in accordance with the method of Hope and Woods, Proceedings of the National Academy of Science USA, Volume 78, pages 3824-3828 (1981).

Various procedures known in the art can be used for the production of polyclonal antibodies which recognize epitopes of NAT. For the production of antibody, various host animals can be immunized by injection with the polypeptide, including but not limited to rabbits, mice, rats, etc. Various adjuvants

- 19 -

may be used to increase the immunological response, depending on the host species, including but not limited to Freund's, mineral gels such as aluminum hydroxide (alum), surface active substances such as
5 lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanins, dinitrophenol, and potentially useful human adjuvants such as Bacille Calmette-Guerin and *Corynebacterium parvum*.

10

For the preparation of monoclonal antibodies directed toward NAT, any technique which provides for the production of antibody molecules by continuous cell lines in culture may be used. For example, the
15 hybridoma technique originally developed by Kohler and Milstein which is described in Nature, Volume 256, pages 495-497 (1975), as well as the trioma technique, the human B-cell hybridoma technique described by Kozbor et al. in Immunology Today, Volume 4, page 72
20 (1983), and the EBV-hybridoma technique to produce monoclonal antibodies described by Cole et al. in "Monoclonal Antibodies and Cancer Therapy", Alan R. Liss, Inc., pages 77-96 (1985), are all useful for preparation of monoclonal antibodies in accordance with
25 this invention.

The antibodies of this invention can be used therapeutically, such as to bind to and thereby neutralize or inhibit excess amounts of NAT *in vivo*
30 after administration. The antibodies can further be used for diagnostic purposes, such as in labeled form to detect the presence of NAT in a body fluid, tissue sample or other extract, in accordance with known diagnostic methods.

35

- 20 -

Description of Specific Embodiments

The invention is further illustrated in the following examples.

5

Example 1Derivation of NAT Sequence

- 10 An effective way to produce fibrolase is to express it initially as preprofibrolase in which cleavage by the protease kex-2 occurs at the junction of the "prepro" and "mature" regions to yield biologically active material ("mature" fibrolase).
- 15 From this design, the synthesis and processing of the preprofibrolase leads to secretion of mature fibrolase into the culture medium. The actual sequence at the cleaved junction is (...TKR↓QQRF...).
- 20 Kex-2 is an endoprotease that cleaves after two adjacent basic amino acids, in this case lysine(K)-arginine(R). Mature fibrolase expressed from DNA having the above mentioned sequence revealed that the expected N-terminal glutamine (Q) residue had in fact
- 25 undergone deamidation and cyclization to generate pyroglutamic acid (E). This chemical modification was deemed undesirable, since peptides with an N-terminal cyclized glutamine (pyroglutamic acid) residue fail to react in the Edman degradation procedure for amino acid
- 30 sequencing. Accordingly, both of the glutamine (Q) residues at the N-terminus in the sequence for mature fibrolase were deleted, resulting in an N-terminal arginine (R) residue. Since kex-2 cleaves after two adjacent basic amino acids as mentioned, it was
- 35 anticipated that the sequence (...KRRF...) would

- 21 -

present an ambiguous site for kex-2 cleavage. Accordingly, the N-terminal arginine (R) residue (shown underlined above) was replaced with a serine (S) residue to result in the sequence (...KRSF...). The
5 choice of serine was based on the need to introduce an amino acid which facilitates kex-2 cleavage when it occurs on the C-terminal side of the hydrolysis site. Rholam et al., European Journal of Biochemistry, Volume 227, pages 707-714 (1995).

10

As a result, the DNA sequence for preprofibrolase was modified by site-directed mutagenesis at the N-terminal coding region for mature fibrolase to substitute the codons for "QQR" with a
15 codon for "S", using a standard PCR protocol, thus resulting in preproNAT having the amino acid sequence of SEQ ID NO: 3. The oligonucleotides used to prime the PCR reactions are listed below, and their homology with the target sequence is also shown. Initially, two
20 PCR reactions were carried out using oligos 1 and 4 as one primer pair and oligos 2 and 3 as another primer pair, both with DNA of the parent gene as the template. The DNA products of these two reactions (601 and 815 nucleotides in length) were purified by agarose gel
25 electrophoresis and combined to serve as a template in a second round of PCR using oligos 1 and 2 as the primer pair. This final PCR product (1372 nucleotides in length) was cleaved with restriction endonucleases *XhoI* and *NotI*. The digest was deproteinized with
30 phenol/chloroform and DNA precipitated. A portion of the recovered DNA was ligated into the plasmid pPICZα (Invitrogen, Carlsbad, CA, Catalog No. VI95-20), which had been similarly cleaved with restriction endonucleases *XhoI* and *NotI*, enzymatically
35 dephosphorylated, and deproteinized with

- 22 -

phenol/chloroform. All subsequent steps were carried out according to the Invitrogen Pichia Expression Kit manual (Invitrogen Corp., Catalog No. K1710-01). The ligation reaction products were transformed into *E. coli* by electroporation and selected for survival on zeocin-containing solid media. The plasmid was isolated and the profibrolase region was confirmed by DNA sequencing. The plasmid was linearized by cleaving with restriction endonuclease *PmeI* and then transformed into *Pichia pastoris* GS115^{his⁺}. The GS115 strain is normally *his⁻*, so the *his⁺* genotype was restored by transformation with a DNA source carrying the wild type version of the *his4* gene. Alternatively, a *his⁺* strain can be obtained commercially from Invitrogen Corp. (X-33 cell line, Catalog No. C180-00). Integrants were selected as zeocin-resistant colonies. Candidate clones were induced in methanol-containing media, and the broth was assayed for NAT production on 4-20% PAGE, using Coomassie staining.

20

The oligonucleotides used for site-directed PCR mutagenesis were as follows:

Oligo 1 5'-TACTATTGCCAGCATTCGTGC-3' (SEQ ID NO: 8)

25

Oligo 2 5'-GCAAATGGCATTCTGACATCC-3' (SEQ ID NO: 9)

Oligo 3 (SEQ ID NO: 10)

5'-TCCAATTAACTTGACTAAGAGATCTTTCCCACAAAGATACGTAC-3'

30

Oligo 4 (SEQ ID NO: 11)

5'-GTACGTATCTTTGTGGGAAAGATCTCTTAGTCAAGTTTAATTGG-3'

The location of these oligonucleotides is shown below in relation to the double-stranded DNA sequence (SEQ ID NO: 12 coding or sense strand, SEQ ID

35

- 23 -

NO: 13 complementary or antisense strand) and corresponding amino acid sequence (SEQ ID NO: 14) of fibrolase (including the prepro region) being modified to create NAT. The N-terminal and C-terminal regions of mature fibrolase are indicated by underlining of terminal amino acid sequences (QQR^F and LNK^P). The N-terminal region (QQR^F) is the one being modified (to substitute S for QQR). For oligos 3 and 4, below, dashed lines are inserted to denote the location of the omitted codons encoding for residues QQ in the N-terminal region of fibrolase.

```
ATGAGATTCCTTCAATTTTACTGCTGTTTTATTTCGCAGCATCCTCCGCATTAGCTGCT
-----+-----+-----+-----+-----+-----+
15 TACTCTAAAGGAAGTTAAAAATGACGACAAAATAAGCGTCGTAGGAGGCGTAATCGACGA

M R F P S I F T A V L F A A S S A L A A

CCAGTCAACACTACAACAGAAGATGAAACGGCACAAATTCGGGCTGAAGCTGTCATCGGT
20 -----+-----+-----+-----+-----+-----+
GGTCAGTTGTGATGTTGTCTTCTACTTTGCCGTGTTTAAGGCCGACTTCGACAGTAGCCA

P V N T T T E D E T A Q I P A E A V I G

25 TACTCAGATTTAGAAGGGGATTTGATGTTGCTGTTTTGCCATTTTCCAACAGCACAAAT
-----+-----+-----+-----+-----+-----+
ATGAGTCTAAATCTTCCCCTAAAGCTACAACGACAAAACGGTAAAAGGTTGTCGTGTTTA

Y S D L E G D F D V A V L P F S N S T N
30

      Oligo 1      5'-TACTATTGCCAGCATTGCTGC-3'
AACGGGTTATGTTTATAAATACTACTATTGCCAGCATTGCTGCTAAAGAAGAAGGGGTA
-----+-----+-----+-----+-----+-----+
TTGCCCAATAACAAATATTTATGATGATAACGGTCGTAACGACGATTCTTCTTCCCCTAT
35

N G L L F I N T T I A S I A A K E E G V
```

WO 01/25445

PCT/US00/27029

- 24 -

XhoI
|
TCTCTCGAGAAAAGAGAGGCTGAAGCTTCTTCTATTATCTTGGAACTCTGGTAACGTTAAC
-----+-----+-----+-----+-----+-----+
5 AGAGAGCTCTTTTCTCTCCGACTTCGAAGAAGATAATAGAACCTTAGACCATTGCAATTG

S L E K R E A E A S S I I L E S G N V N

GATTACGAAGTTGTTTATCCAAGAAAGGTCACCTCCAGTTCCTAGGGGTGCTGTTCAACCA
10 -----+-----+-----+-----+-----+-----+
CTAATGCTTCAACAAATAGGTTCTTTCCAGTGAGGTCAAGGATCCCCACGACAAGTTGGT

D Y E V V Y P R K V T P V P R G A V Q P

15 AAGTACGAAGATGCCATGCAATACGAATTCAAGGTTAACAGTGAACCAGTTGTCTTGCAC
-----+-----+-----+-----+-----+-----+
TTCATGCTTCTACGGTACGTTATGCTTAAGTTCCAATTGTCACTTGGTCAACAGAACGTG

K Y E D A M Q Y E F K V N S E P V V L H
20 -----+-----+-----+-----+-----+-----+
TTGGAACAAAAACAAAGGTTTGTCTCTGAAGATTACTCTGAACTCATTACTCCCCAGAT
-----+-----+-----+-----+-----+-----+
AACCTTTTTTTGTTTCCAAACAAGAGACTTCTAATGAGACTTTGAGTAATGAGGGGTCTA

25 L E K N K G L F S E D Y S E T H Y S P D

GGTAGAGAAATTACTACTTACCCATTGGGTGAAGATCACTGTTACTACCATGGTAGAATC
-----+-----+-----+-----+-----+-----+
CCATCTCTTTAATGATGAATGGGTAACCCACTTCTAGTGACAATGATGGTACCATCTTAG
30 -----+-----+-----+-----+-----+-----+
G R E I T T Y P L G E D H C Y Y H G R I

GAAAACGATGCTGACTCCACTGCTTCTATCTCTGCTTGTAACGGTTTGAAGGGTCATTTT
-----+-----+-----+-----+-----+-----+
35 CTTTTGCTACGACTGAGGTGACGAAGATAGAGACGAACATTGCCAAACTTCCCAGTAAAG

E N D A D S T A S I S A C N G L K G H F

AAGTTGCAAGGTGAAATGTACTTGATTGAACCATTGGAATTGTCCGACTCTGAAGCCCAT
40 -----+-----+-----+-----+-----+-----+
TTCAACGTTCCACTTTACATGAACCTAAGTTGGAACCTTAACAGGCTGAGACTTCGGGTA

K L Q G E M Y L I E P L E L S D S E A H

- 25 -

GCTGTCTACAAGTACGAAAACGTCGAAAAGGAAGATGAAGCCCCAAAGATGTGTGGTGT
 -----+-----+-----+-----+-----+-----+-----+
 CGACAGATGTTTCATGCTTTTGCAGCTTTTCTCTTCTACTTCGGGGTTTCTACACACCACAA
 5 A V Y K Y E N V E K E D E A P K M C G V

Oligo 3 5'-TCCAATTAACTTGACTAAG
 ACCCAAACTGGGAATCATATGAACCAATCAAGAAGGCCTTCCAATTAACTTGACTAAG
 10 -----+-----+-----+-----+-----+-----+-----+
 TGGGTTTTGACCCCTTAGTATACTTGGTTAGTTCTTCCGGAAGGTTAATTTGAACTGATTC
 Oligo 4 3'-GGTTAATTTGAACTGATTC

T Q N W E S Y E P I K K A F Q L N L T K
 15 AGA-----TCTTTCCACAAAGATACGTAC-3'
 AGACAACAAAGATTCCACAAAGATACGTACAGCTGGTTATCGTTGCTGACCACCGTATG
 -----+-----+-----+-----+-----+-----+-----+
 TCTGTTGTTTCTAAGGGTGTCTTCTATGCATGTCGACCAATAGCAACGACTGGTGGCATACT
 20 TCT-----AGAAAGGGTGTCTTCTATGCATG-5'
 R Q Q R F P Q R Y V Q L V I V A D H R M

AACACTAAATACAACGGTGACTCTGACAAAATCCGTCAATGGGTGCACCAAATCGTCAAC
 25 -----+-----+-----+-----+-----+-----+-----+
 TTGTGATTTATGTTGCCACTGAGACTGTTTTAGGCAGTTACCCACGTGGTTTAGCAGTTG

N T K Y N G D S D K I R Q W V H Q I V N
 30 ACCATTAACGAAATCTACAGACCACTGAACATCCAATTCACTTTGGTTGGTTTGGAAATC
 -----+-----+-----+-----+-----+-----+-----+
 TGGTAATTGCTTTAGATGTCTGGTGACTTGTAGGTTAAGTGAAACCAACCAAACCTTTAG

T I N E I Y R P L N I Q F T L V G L E I
 35 TGGTCCAACCAAGATTTGATCACC GTTACTTCTGTATCCCACGACACTCTGGCATCCTTC
 -----+-----+-----+-----+-----+-----+-----+
 ACCAGGTGGTTCTAAACTAGTGGCAATGAAGACATAGGGTGCTGTGAGACCGTAGGAAG

40 W S N Q D L I T V T S V S H D T L A S F

- 26 -

GGTAACTGGCGTGAAACCGACCTGCTGCGTCGCCAACGTCATGATAACGCTCAACTGCTG
-----+-----+-----+-----+-----+-----+
CCATTGACCGCACTTTGGCTGGACGACGCAGCGGTTGCAGTACTATTGCGAGTTGACGAC

5 G N W R E T D L L R R Q R H D N A Q L L

ACCGCTATCGACTTCGACGGTGATACTGTTGGTCTGGCTTACGTTGGTGGCATGTGTCAA
-----+-----+-----+-----+-----+-----+
TGGCGATAGCTGAAGCTGCCACTATGACAACCAGACCGAATGCAACCACCGTACACAGTT

10 T A I D F D G D T V G L A Y V G G M C Q

CTGAAACATTCTACTGGTGTTATCCAGGACCACTCCGCTATTAACCTGCTGGTTGCTCTG
-----+-----+-----+-----+-----+-----+
15 GACTTTGTAAGATGACCACAATAGGTCCTGGTGAGGCGATAATTGGACGACCAACGAGAC

L K H S T G V I Q D H S A I N L L V A L

ACCATGGCACACGAACTGGGTCATAACCTGGGTATGAACCACGATGGCAACCAGTGTAC
-----+-----+-----+-----+-----+-----+
20 TGGTACCGTGTGCTTGACCCAGTATTGGACCCATACCTGGTGCTACCGTTGGTCACAGTG

T M A H E L G H N L G M N H D G N Q C H

25 TGCGGTGCAAACCTCCTGTGTTATGGCTGCTATGCTGTCCGATCAACCATCCAAACTGTTT
-----+-----+-----+-----+-----+-----+
ACGCCACGTTTGAGGACACAATACCGACGATACGACAGGCTAGTTGGTAGGTTTGACAAG

C G A N S C V M A A M L S D Q P S K L F

30 TCCGACTGCTCTAAGAAAGACTACCAGACCTTCCTGACCGTTAACAACCCGAGTGTATC
-----+-----+-----+-----+-----+-----+
AGGCTGACGAGATTCTTTCTGATGGTCTGGAAGGACTGGCAATTGTTGGGCGTCACATAG

35 S D C S K K D Y Q T F L T V N N P Q C I

NotI
|

CTGAACAAACCGTAAGCGGCCGCCAGCTTTCTAGAACAAAACCTCATCTCAGAAGAGGAT
-----+-----+-----+-----+-----+-----+
40 GACTTGTGTTGGCATTGCGCCGCGGTCGAAAGATCTTGTTTTTGAGTAGAGTCTTCTCCTA

L N K P * A A A S F L E Q K L I S E E D

- 27 -

CTGAATAGCGCCGTCGACCATCATCATCATCATCATTGAGTTTGTAGCCTTAGACATGAC
 -----+-----+-----+-----+-----+-----+-----+
 GACTTATCGCGGCAGCTGGTAGTAGTAGTAGTAGTAAC TCAAACATCGGAATCTGTACTG
 5 L N S A V D H H H H H H * V C S L R H D

 TGTTCCCTCAGTTCAAGTTGGGCACTTACGAGAAGACCGGTCTTGCTAGATTCTAATCAAG
 -----+-----+-----+-----+-----+-----+-----+
 10 ACAAGGAGTCAAGTTCAACCCGTGAATGCTCTTCTGGCCAGAACGATCTAAGATTAGTTC

 C S S V Q V G H L R E D R S C * I L I K

 AGGATGTCAGAATGCCATTTGCCTGAGAGATGCAGGCTTCATTTTGTACTTTTATT
 15 -----+-----+-----+-----+-----+-----+-----+
 TCCTACAGTCTTACGGTAAACGGACTCTCTACGTCCGAAGTAAAACTATGAAAAAATAA
 CCTACAGTCTTACGGTAAACG-5' Oligo 2

 R M S E C H L P E R C R L H F * Y F F I -
 20

Example 2

Expression of NAT in *Pichia pastoris*

25 When attempts were made to express the DNA
 for NAT in *E. coli*, very poor refolding and a
 requirement for dilute conditions reduced the
 purification efficiency. These and other
 30 considerations led to the usage of *Pichia pastoris*, a
 yeast species, as the host cell. A culture of selected
 clones of *Pichia pastoris* which had been transfected
 with prepro NAT cDNA (SEQ ID NO: 4) was inoculated into
 500 ml of the following inoculation growth medium:
 35

- 28 -

Per liter of batch medium

	Yeast extract	30.0 g
	Potassium phosphate dibasic	17.2 g
5	Glucose	20.0 g
	Biotin	0.004 g
	Water	to 1 liter
	Phosphoric acid, 85%	to adjust pH to 6.00

- 10 The transfected *P. pastoris* cells were incubated at 30°C in a shaker for about 30 to 32 hours. About 1% (w/v) of the resulting culture was used to inoculate a 10-liter fermentor. The fermentor contained sterilized basal salts and glucose (below).
- 15 Twelve milliliters per liter of PTM4 salts (PTM4 is a trace metals solution containing cupric sulfate pentahydrate, sodium iodide, manganese sulfate monohydrate, sodium molybdate dihydrate, boric acid, cobaltous chloride hexahydrate, zinc chloride, ferrous
- 20 sulfate heptahydrate, d-biotin, sulfuric acid and purified water) were added per liter of batch medium after fermentor sterilization. The fermentation growth temperature was 30°C. The fermentor pH was controlled with ammonium hydroxide and phosphoric acid at pH 6.00.
- 25 Zinc from the zinc salts added to the medium becomes incorporated into NAT as part of the metalloproteinase structure.

Basal salts per liter of batch medium

30	Phosphoric acid, 85%	26.7 ml
	Calcium sulfate	0.93 g
	Potassium sulfate	18.2 g
	Magnesium sulfate-7H ₂ O	14.9 g
35	Potassium hydroxide	4.13 g

- 29 -

Glucose	30.0 g
Water	to 1 liter

The batch culture was grown until the glucose
5 was completely consumed (17 to 20 hours). Then a fed-
batch phase was initiated. The fed-batch phase media
consisted of glucose and 12 ml of PTM4 salts per liter.
The induction feed consisted of glucose, 25% methanol,
and 12 ml of PTM4 salts per liter. At induction, the
10 temperature of the reactor was shifted to 20°C. The
induction phase lasted 60 to 75 hours. The conditioned
media were harvested and the cellular debris was
discarded.

15

Example 3Purification of NAT from *Pichia pastoris*

The yeast broth (conditioned media less
20 cellular debris) from Example 2 was clarified and the
pH and conductivity were adjusted to 6.5 and 10-20
mS/cm, respectively. The broth was loaded onto an
immobilized metal affinity resin that had been charged
with copper (Cu) and equilibrated with phosphate
25 buffered saline (PBS). The resin was washed with PBS
and eluted with an imidazole gradient (0-100 mM) in
PBS. Fractions containing "mature" NAT (SEQ ID NO: 1)
were pooled and diluted until the conductivity was less
than 1.5 mS/cm, pH 6.4. The diluted pool was loaded
30 onto an SP Sepharose resin (Amersham Pharmacia Biotech,
Inc., Piscataway, NJ) that had been equilibrated with
10 mM 2-(N-morpholino)ethanesulfonic acid (MES). The
column was washed with MES and eluted with a NaCl
gradient (0-500mM) in MES. Fractions containing NAT
35 were pooled and stored.

- 30 -

Example 4Thrombolysis in Acute Thrombosis of Rat
5 Carotid Artery; Comparison of NAT with Urokinase

To demonstrate that "mature" NAT (SEQ ID NO: 1) is biologically active and functionally unique, acute pharmacology studies were conducted in rats where focal injury to one of the carotid arteries was created by applying anodal current. This injury produces an occlusive thrombus which generally forms within fifteen minutes. Once the thrombus was formed, the artery was observed for a period of thirty minutes to assure that the carotid occlusion was stable. Then heparin and aspirin were administered intravenously to prevent further propagation of the thrombus. The animals were then treated with an intraarterial infusion of test material. Blood flow through the carotid artery was monitored during the delivery of test material so that successful clot lysis could be detected and the time at which clot lysis occurred could be noted. The percentage of experiments where clot lysis occurred was noted and group means were calculated for only those experiments where clot lysis was successful. As a measure of the hemorrhagic potential of the test material, any blood that was shed from the surgical site was collected with gauze swabs. The swabs were placed in a detergent solution to solubilize red blood cells and release hemoglobin, which was then quantified spectrophotometrically. Shed hemoglobin was used to calculate a volume of blood loss. Test data are reported in the Table below.

- 31 -

Table 1

Incidence of Clot Lysis, Time to Clot
Lysis and Surgical Blood Loss (Mean $\pm$ std. dev.)

5

	INCIDENCE OF LYSIS (%)	TIME TO LYSIS (min)	BLOOD LOSS (ml)
SALINE (n=6)	0% (0 of 6)	N/A	0.10 $\pm$ 0.23
Urokinase 25U/min (n=15)	33% (5 of 15)	55.3 $\pm$ 15.9	1.06 $\pm$ 1.59
Urokinase 250U/min (n=15)	86% (13 of 15)	33.5 $\pm$ 15.3	1.43 $\pm$ 1.45
NAT 2mg (n=14)	78% (11 of 14)	6.3 $\pm$ 5.8	0.96 $\pm$ 0.77

These studies establish that NAT is biologically active in an animal model of *in vivo* clot lysis. Further, clot lysis was achieved in a markedly reduced amount of time and with less blood loss from the surgical site, in comparison with urokinase. Thus, the activity profile of NAT can be distinguished from the plasminogen activator class of thrombolytic agents (represented by urokinase) in that clot lysis with NAT occurs more rapidly and with reduced hemorrhagic complications.

The fibrinolytic activity of NAT is comparable to that of fibrolase. In addition, as mentioned above the stability of the N-terminus of NAT results in a more

- 32 -

homogeneous end product upon recombinant expression,
which is a distinct advantage (i.e., the N-terminus
will not change over time resulting in a mixture of
different forms, thus making the polypeptide more
5 stable).

- 33 -

CLAIMS

1. A fibrinolytically active polypeptide
5 having the amino acid sequence of SEQ ID NO: 1.
2. The polypeptide of claim 1 which has
been made by recombinant expression in yeast.
- 10 3. A nucleic acid molecule encoding the
polypeptide of SEQ ID NO: 1.
4. The nucleic acid molecule of claim 3
which has the sequence of SEQ ID NO: 4.
15
5. An expression vector comprising
expression regulatory elements operatively linked to a
nucleic acid molecule according to claim 3.
- 20 6. An expression vector according to claim
5 in which the nucleic acid molecule has the nucleotide
sequence of SEQ ID NO: 4.
7. A host cell transformed or transfected
25 with the nucleic acid molecule of claim 3.
8. A transformed or transfected host cell
according to claim 7 which is a prokaryotic or
eukaryotic cell.
30
9. A transformed or transfected eukaryotic cell according
to claim 8 which is a yeast cell.
10. A transformed or transfected yeast cell
35 according to claim 9 which is *Pichia pastoris*.

- 34 -

11. A method for the production of a novel acting thrombolytic metalloproteinase encoded by the nucleic acid of claim 3, comprising culturing host cells containing a DNA molecule encoding said metalloproteinase operatively linked to regulatory elements for stimulating the expression thereof under conditions such that the DNA molecule is expressed, and isolating the expressed metalloproteinase from the host cell culture.

12. The method of claim 11 in which the DNA molecule is of SEQ ID NO: 4.

13. The method of claim 11 in which the host cells are yeast cells.

14. The method of claim 13 in which the yeast cells are of *Pichia pastoris*.

15. A method for the production of a metalloproteinase fibrinolytic agent selected from the group consisting of fibrolase and NAT, comprising culturing *Pichia* host cells containing a nucleic acid molecule of claim 3 DNA molecule encoding said metalloproteinase operatively linked to regulatory elements for stimulating the expression thereof under conditions such that the nucleic acid DNA molecule is expressed, and isolating the expressed metalloproteinase from the host cell culture.

16. The method of claim 15 in which the *Pichia* host cell is *Pichia pastoris*.

17. Use of a fibrinolytically active polypeptide of claim 1 for the manufacture of a medicament

18. An antibody against the polypeptide of claim 1.

- 35 -

19. The antibody of claim 18 which is a monoclonal antibody.

WO 01/25445

PCT/US00/27029

- 1 -

SEQUENCE LISTING

<110> Boone, Thomas C.
 Li, Huimin
 Mann, Michael B.

<120> FIBRINOLYTICALLY ACTIVE POLYPEPTIDE

<130> A-596

<140> 09/411,329

<141> 1999-10-01

<160> 20

<170> PatentIn Ver. 2.1

<210> 1

<211> 201

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: NAT (analog of
 fibrolase of Agkistrodon Contourtrix)

<400> 1

Ser Phe Pro Gln Arg Tyr Val Gln Leu Val Ile Val Ala Asp His Arg
 1 5 10 15

Met Asn Thr Lys Tyr Asn Gly Asp Ser Asp Lys Ile Arg Gln Trp Val
 20 25 30

His Gln Ile Val Asn Thr Ile Asn Glu Ile Tyr Arg Pro Leu Asn Ile
 35 40 45

Gln Phe Thr Leu Val Gly Leu Glu Ile Trp Ser Asn Gln Asp Leu Ile
 50 55 60

Thr Val Thr Ser Val Ser His Asp Thr Leu Ala Ser Phe Gly Asn Trp
 65 70 75 80

Arg Glu Thr Asp Leu Leu Arg Arg Gln Arg His Asp Asn Ala Gln Leu
 85 90 95

Leu Thr Ala Ile Asp Phe Asp Gly Asp Thr Val Gly Leu Ala Tyr Val
 100 105 110

Gly Gly Met Cys Gln Leu Lys His Ser Thr Gly Val Ile Gln Asp His
 115 120 125

Ser Ala Ile Asn Leu Leu Val Ala Leu Thr Met Ala His Glu Leu Gly
 130 135 140

His Asn Leu Gly Met Asn His Asp Gly Asn Gln Cys His Cys Gly Ala
 145 150 155 160

Asn Ser Cys Val Met Ala Ala Met Leu Ser Asp Gln Pro Ser Lys Leu
165 170 175

Phe Ser Asp Cys Ser Lys Lys Asp Tyr Gln Thr Phe Leu Thr Val Asn
180 185 190

Asn Pro Gln Cys Ile Leu Asn Lys Pro
195 200

```
<210> 2
<211> 603
<212> DNA
<213> Artificial Sequence
```

<220>
<223> Description of Artificial Sequence: Encodes NAT
(analog of fibrolase)

<400> 2						
tctttccac	aaagatacgt	acagctgggt	atcgttgcgt	accaccgtat	gaacactaaa	60
tacaacggtg	actctgacaa	aatccgtcaa	tgggtgcacc	aaatcgtcaa	caccattaac	120
gaaatctaca	gaccactgaa	catccaattc	actttggttg	gtttggaaaat	ctgggtccaac	180
caagatttga	tcacggttac	tctgttatcc	cacgacactc	tggcatcctt	cggtaactgg	240
cgtgaaaccg	acctgctgcg	tcgccaacgt	catgataacg	ctcaactgct	gaccgctatc	300
gacttcgacg	gtgatactgt	tggctctggct	tacgttggtg	gcatgtgtca	actgaaacat	360
tctactggtg	ttatccagga	ccactccgct	attaacctgc	tggttgctct	gaccatggca	420
cacgaactgg	gtcataacct	gggtatgaac	cacgatggca	accagtgtca	ctgcggtgca	480
aactcctgtg	ttatggctgc	tatgctgtcc	gatcaaccat	ccaaactggt	ctccgactgc	540
tctaagaaa	actaccagac	cttcttgacc	gttaacaacc	cgcagtgtat	cctgaacaaa	600
ccg						603

```
<210> 3
<211> 462
<212> PRT
<213> Artificial Sequence
```

<220>
<223> Description of Artificial Sequence: Native pro-NAT
(analog of fibrolase)

```
<400> 3
Met Arg Phe Pro Ser Ile Phe Thr Ala Val Leu Phe Ala Ala Ser Ser
  1             5             10             15
```

Ala Leu Ala Ala Pro Val Asn Thr Thr Thr Glu Asp Glu Thr Ala Gln
20 25 30

Ile Pro Ala Glu Ala Val Ile Gly Tyr Ser Asp Leu Glu Gly Asp Phe
35 40 45

Asp Val Ala Val Leu Pro Phe Ser Asn Ser Thr Asn Asn Gly Leu Leu
50 55 60

WO 01/25445

PCT/US00/27029

- 3 -

Phe Ile Asn Thr Thr Ile Ala Ser Ile Ala Ala Lys Glu Glu Gly Val
 65 70 75 80
 Ser Leu Glu Lys Arg Glu Ala Glu Ala Ser Ser Ile Ile Leu Glu Ser
 85 90 95
 Gly Asn Val Asn Asp Tyr Glu Val Val Tyr Pro Arg Lys Val Thr Pro
 100 105 110
 Val Pro Arg Gly Ala Val Gln Pro Lys Tyr Glu Asp Ala Met Gln Tyr
 115 120 125
 Glu Phe Lys Val Asn Ser Glu Pro Val Val Leu His Leu Glu Lys Asn
 130 135 140
 Lys Gly Leu Phe Ser Glu Asp Tyr Ser Glu Thr His Tyr Ser Pro Asp
 145 150 155 160
 Gly Arg Glu Ile Thr Thr Tyr Pro Leu Gly Glu Asp His Cys Tyr Tyr
 165 170 175
 His Gly Arg Ile Glu Asn Asp Ala Asp Ser Thr Ala Ser Ile Ser Ala
 180 185 190
 Cys Asn Gly Leu Lys Gly His Phe Lys Leu Gln Gly Glu Met Tyr Leu
 195 200 205
 Ile Glu Pro Leu Glu Leu Ser Asp Ser Glu Ala His Ala Val Tyr Lys
 210 215 220
 Tyr Glu Asn Val Glu Lys Glu Asp Glu Ala Pro Lys Met Cys Gly Val
 225 230 235 240
 Thr Gln Asn Trp Glu Ser Tyr Glu Pro Ile Lys Lys Ala Phe Gln Leu
 245 250 255
 Asn Leu Thr Lys Arg Ser Phe Pro Gln Arg Tyr Val Gln Leu Val Ile
 260 265 270
 Val Ala Asp His Arg Met Asn Thr Lys Tyr Asn Gly Asp Ser Asp Lys
 275 280 285
 Ile Arg Gln Trp Val His Gln Ile Val Asn Thr Ile Asn Glu Ile Tyr
 290 295 300
 Arg Pro Leu Asn Ile Gln Phe Thr Leu Val Gly Leu Glu Ile Trp Ser
 305 310 315 320
 Asn Gln Asp Leu Ile Thr Val Thr Ser Val Ser His Asp Thr Leu Ala
 325 330 335
 Ser Phe Gly Asn Trp Arg Glu Thr Asp Leu Leu Arg Arg Gln Arg His
 340 345 350
 Asp Asn Ala Gln Leu Leu Thr Ala Ile Asp Phe Asp Gly Asp Thr Val
 355 360 365

Gly Leu Ala Tyr Val Gly Gly Met Cys Gln Leu Lys His Ser Thr Gly
370 375 380

Val Ile Gln Asp His Ser Ala Ile Asn Leu Leu Val Ala Leu Thr Met
385 390 395 400

Ala His Glu Leu Gly His Asn Leu Gly Met Asn His Asp Gly Asn Gln
405 410 415

Cys His Cys Gly Ala Asn Ser Cys Val Met Ala Ala Met Leu Ser Asp
420 425 430

Gln Pro Ser Lys Leu Phe Ser Asp Cys Ser Lys Lys Asp Tyr Gln Thr
435 440 445

Phe Leu Thr Val Asn Asn Pro Gln Cys Ile Leu Asn Lys Pro
450 455 460

<210> 4

<211> 1386

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Encodes
pro-NAT (analog of fibrolase)

<400> 4

```

atgagatttc cttcaatttt tactgctggt ttattcgcag catcctccgc attagetgct 60
ccagtcaaca ctacaacaga agatgaaacg gcacaaattc cggctgaagc tgtcatcggt 120
tactcagatt tagaagggga ttctgatggt gctgttttgc cattttccaa cagcacaat 180
aactgggttat tgtttataaa tactactatt gccagcattg ctgctaaaga agaaggggta 240
tctctcgaga aaagagagcg tgaagcttct tctattatct tggaatctgg taacgttaac 300
gattacgaag ttgtttatcc aagaaaggtc actccagttc ctagggggtgc tgttcaacca 360
aagtacgaag atgccatgca atacgaattc aagggttaaca gtgaaccagt tgtcctgcac 420
ttggaaaaaa acaaagggtt gttctctgaa gattactctg aaactcatta ctccccagat 480
ggtagagaaa ttactactta cccattgggt gaagatcact gttactacca tggtagaatc 540
gaaaacgatg ctgactccac tgcttctatc tctgcttgta acggtttgaa gggtcatttc 600
aagttgcaag gtgaaatgta cttgattgaa ccattggaat tgtccgactc tgaagcccat 660
gctgtctaca agtacgaaaa cgtcgaaaaa gaagatgaag ccccaaagat gtgtgggtgt 720
acccaaaact gggaatcata tgaaccaatc aagaaggcct tccaattaaa cttgactaag 780
agatctttcc cacaaagata cgtacagctg gttatcgttg ctgaccaccg tatgaacact 840
aaatacaacg gtgactctga caaaatccgt caatgggtgc accaaatcgt caacaccatt 900
aacgaaatct acagaccact gaacatocaa ttcactttgg ttggtttgga aatctgggtcc 960
aaccaagatt tgatcacctg tacttctgta tcccacgaca ctctggcatc cttcggtaac 1020
tggcgtgaaa ccgacctgct gcgtcgccaa cgctcatgata acgctcaact gctgaccgct 1080
atcgacttcg acggtgatac tggtgggtctg gcttacgttg gtggcatgtg tcaactgaaa 1140
cattctactg gtgttatcca ggaccactcc gctattaacc tgctgggttc tctgacccatg 1200
gcacacgaac tgggtcataa cctgggtatg aaccacgatg gcaaccagtg tcaactgagg 1260
gcaaactcct gtgttatggc tgctatgctg tccgatcaac catccaaact gttctccgac 1320
tgctctaaga aagactacca gaccttcctg accgttaaca acccgcatgt tatcctgaac 1380
aaaccg                                     1386

```

WO 01/25445

PCT/US00/27029

- 5 -

<210> 5

<211> 203

<212> PRT

<213> Agkistrodon contortrix

<220>

<223> Native fibrolase of Agkistrodon contortrix

<400> 5

Gln Gln Arg Phe Pro Gln Arg Tyr Val Gln Leu Val Ile Val Ala Asp
 1 5 10 15

His Arg Met Asn Thr Lys Tyr Asn Gly Asp Ser Asp Lys Ile Arg Gln
 20 25 30

Trp Val His Gln Ile Val Asn Thr Ile Asn Glu Ile Tyr Arg Pro Leu
 35 40 45

Asn Ile Gln Phe Thr Leu Val Gly Leu Glu Ile Trp Ser Asn Gln Asp
 50 55 60

Leu Ile Thr Val Thr Ser Val Ser His Asp Thr Leu Ala Ser Phe Gly
 65 70 75 80

Asn Trp Arg Glu Thr Asp Leu Leu Arg Arg Gln Arg His Asp Asn Ala
 85 90 95

Gln Leu Leu Thr Ala Ile Asp Phe Asp Gly Asp Thr Val Gly Leu Ala
 100 105 110

Tyr Val Gly Gly Met Cys Gln Leu Lys His Ser Thr Gly Val Ile Gln
 115 120 125

Asp His Ser Ala Ile Asn Leu Leu Val Ala Leu Thr Met Ala His Glu
 130 135 140

Leu Gly His Asn Leu Gly Met Asn His Asp Gly Asn Gln Cys His Cys
 145 150 155 160

Gly Ala Asn Ser Cys Val Met Ala Ala Met Leu Ser Asp Gln Pro Ser
 165 170 175

Lys Leu Phe Ser Asp Cys Ser Lys Lys Asp Tyr Gln Thr Phe Leu Thr
 180 185 190

Val Asn Asn Pro Gln Cys Ile Leu Asn Lys Pro
 195 200

<210> 6

<211> 609

<212> DNA

<213> Agkistrodon contortrix

<220>

<223> Encodes native fibrolase of Agkistrodon contortrix

<400> 6

```

caacaaagat tcccacaaag atacgtacag ctgggttatcg ttgctgacca ccgtatgaac 60
actaaatata acggtgactc tgacaaaatc cgtcaatggg tgcaccaa atcgtaacacc 120
attaacgaaa tctacagacc actgaacatc caattcactt tgggttggtt ggaaatctgg 180
tccaaccaag atttgatcac cgttacttct gtatcccacg acactctggc atccttcggg 240
aactggcggt aaaccgacct gctgcgtcgc caacgctcatg ataacgctca actgctgacc 300
gctatcgact tcgacgggtga tactgttggt ctgggttacg ttgggtggcat gtgtcaactg 360
aaacattcta ctggtgttat ccaggaccac tccgctatta acctgctggt tgctctgacc 420
atggcacacg aactgggtca taacctgggt atgaaccacg atggcaacca gtgtcactgc 480
ggtgcaaact cctgtgttat ggctgctatg ctgtccgac aaccatccaa actgttctcc 540
gactgtctta agaaagacta ccagaccttc ctgaccgtta acaaccgcga gtgtatcctg 600
aacaaccg

```

609

<210> 7

<211> 1392

<212> DNA

<213> Agkistrodon contortrix

<220>

<223> Coding sequence of native pro-fibrolase of
Agkistrodon contortrix

<400> 7

```

atgagatttc cttcaatttt tactgctggt ttattcgcag catcctccgc attagctgct 60
ccagtcaaca ctacaacaga agatgaaacg gcacaaatc cggctgaagc tgtcatcggt 120
tactcagatt tagaagggga ttctgatggt gctgttttgc cattttccaa cagcacaat 180
aacgggttat tgtttataaa tactactatt gccagcattg ctgctaaaga agaaggggta 240
tctctcgaga aaagagaggc tgaagcttct tctattatct tggaatctgg taacgttaac 300
gattacgaag ttgtttatcc aagaaaggct actccagttc ctaggggtgc tgttcaacca 360
aagtacgaag atgccatgca atacgaatc aagggttaaca gtgaaccagt tgtcttgac 420
ttggaaaaaa acaaagggtt gtctctgaa gattactctg aaactcatta ctccccagat 480
ggtagagaaa ttactactta cccattgggt gaagatcact gttactacca tggtagaatc 540
gaaaacgatg ctgactccac tgcttctatc tctgcttgta acggtttgaa gggtcatttc 600
aagttgcaag gtgaaatgta cttgattgaa ccattggaat tgtccgactc tgaagcccat 660
gctgtctaca agtacgaaaa cgtcgaaaag gaagatgaag ccccaaagat gtgtgggtgt 720
acccaaaact gggaatcata tgaaccaatc aagaaggcct tccaattaaa cttgactaag 780
agacaacaaa gattcccaca aagatacgta cagctgggta tcgttgctga ccaccgtatg 840
aacactaaat acaacggtga ctctgacaaa atccgtcaat ggggtgacca aatcgtaaac 900
accattaacg aaatctacag accactgaac atccaattca ctttggttgg tttggaaatc 960
tggtccaacc aagatttgat caccgttact tctgtatccc acgacactct ggcaccttc 1020
ggtaactggc gtgaaaccga cctgctgctg cgccaacgtc atgataacgc tcaactgctg 1080
accgctatcg acttcgacgg tgatactggt ggtctggctt acgttggtgg catgtgtcaa 1140
ctgaaacatt ctactggtgt tatccaggac cactccgcta ttaacctgct ggttgcctcg 1200
accatggcac acgaactggg tcataacctg ggtatgaacc acgatggcaa ccagtgtcac 1260
tgcggtgcaa actcctgtgt tatggctgct atgctgtccg atcaaccatc caaactgttc 1320
tccgactgct ctaagaaaga ctaccagacc ttcctgaccg ttaacaaccc gcagtgtatc 1380
ctgaacaaac cg

```

1392

<210> 8

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:
Oligonucleotide

<400> 8

tactattgcc agcattgctg c

21

<210> 9

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:
Oligonucleotide

<400> 9

cctacagtct tacggtaaac g

21

<210> 10

<211> 45

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:
Oligonucleotide

<400> 10

tccaattaaa cttgactaag agatctttcc cacaaagata cgtac

45

<210> 11

<211> 44

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:
Oligonucleotide

<400> 11

ggttaatttg aactgattct ctagaaaggg tgtttctatg catg

44

<210> 12

<211> 1620

<212> DNA

<213> Agkistrodon contortrix

<220>

<221> misc_feature

<222> Complement((1)..(1620))

WO 01/25445

PCT/US00/27029

- 8 -

<223> Complementary (sense) strand of antisense strand
(See SEQ ID NO:13)

<220>

<223> Coding sequence of native pro-fibrolase of
Agkistrodon contortrix

<400> 12

```

atgagatttc cttcaatttt tactgctggt ttattcgcag catcctccgc attagctgct 60
ccagtcaaca ctacaacaga agatgaaacg gcacaaattc cggctgaagc tgtcatcggt 120
tactcagatt tagaagggga tttcgaatgt gctgttttgc cattttccaa cagcacaaat 180
aacgggttat tgtttataaa tactactatt gccagcattg ctgctaaaga agaaggggta 240
tctctcgaga aaagagagggc tgaagcttct tctattatct tggaatctgg taacgttaac 300
gattacgaag ttgtttatcc aagaaagggt actccagttc ctagggtggtc tgttcaacca 360
aagtacgaag atgccatgca atacgaattc aagggttaaca gtgaaccagt tgtcttgac 420
ttggaaaaaa acaaagggtt gttctctgaa gattactctg aaactcatta ctccccagat 480
ggtagagaaa ttactactta cccattgggt gaagatcact gttactacca tggtagaatc 540
gaaaacgatg ctgactccac tgcttctatc tctgcttgta acggtttgaa gggtcatttc 600
aagttgcaag gtgaaatgta cttgattgaa ccattggaat tgcccgactc tgaagcccat 660
gctgtctaca agtacgaaaa cgtcgaaaag gaagatgaag ccccaaagat gtgtgggtgt 720
acccaaaact gggaatcata tgaaccaatc aagaaggcct tccaattaaa cttgactaag 780
agacaacaaa gattcccaca aagatacgtc cagctgggta tcgttgctga ccaccgtatg 840
aacactaaat acaacgggtg ctctgacaaa atccgtcaat ggggtgcacca aatcgtaaac 900
accattaacg aaatctacag accactgaac atccaattca ctttgggttg tttggaaatc 960
tggtccaacc aagatttgat caccgttact tctgtatccc acgacactct ggcaccttc 1020
ggtaactggc gtgaaaccga cctgctgcgt cgccaacgtc atgataacgc tcaactgctg 1080
accgctatcg acttcgacgg tgatactggt ggtctggctt acgttggttg catgtgtcaa 1140
ctgaaacatt ctactggtgt tatccaggac cactccgcta ttaacctgct ggttgctctg 1200
accatggcac acgaactggg tcataacctg ggtatgaacc acgatggcaa ccagtgtcac 1260
tgcggtgcaa actcctgtgt tatggctgct atgtgtccg atcaaccatc caaactgttc 1320
tccgactgct ctaagaaaga ctaccagacc ttcccgaccg ttaacaaccg gcagtgtatc 1380
ctgaacaaac cgtaagcggc cgccagcttt ctagaacaaa aactcatctc agaagaggat 1440
ctgaatagcg ccgtcgacca tcatcatcat catcattgag tttgtagcct tagacatgac 1500
tgctcctcag ttcaagttgg gcacttacga gaagaccggt cttgctagat tctaatacaag 1560
aggatgtcag aatgccattt gcctgagaga tgcaggcttc atttttgata cttttttatt 1620

```

<210> 13

<211> 1620

<212> DNA

<213> Agkistrodon contortrix

<220>

<221> misc_feature

<222> Complement((1)..(1620))

<223> Complementary (antisense) strand of sense strand
(See SEQ ID NO: 12)

<220>

<223> Anti-coding sequence of native pro-fibrolase of
Agkistrodon contortrix

<400> 13

```

aataaaaaag tatcaaaaat gaagcctgca tctctcaggc aaatggcatt ctgacatcct 60
cttgattaga atctagcaag accggtcttc togtaatgac ccaacttgaa ctgaggaaca 120
gtcatgtcta aggctacaaa ctcaatgatg atgatgatga tggctcgacgg cgctatttcag 180

```

```

atcctcttct gagatgagtt tttgttctag aaagctggcg gccgcttacg gtttgttcag 240
gatacactgc gggttgttaa cggtcaggaa ggtctggttag tctttcttag agcagtcgga 300
gaacagtttg gatggttgat cggacagcat agcagccata acacaggagt ttgcaccgca 360
gtgacactgg ttgccatcgt ggttcatacc cagggttatga cccagttcgt gtgccatggg 420
cagagcaacc agcaggttaa tagcggagtg gtcctggata acaccagtag aatgtttcag 480
ttgacacatg ccaccaacgt aagccagacc aacagtatca ccgtcgaagt cgatagcggg 540
cagcagttga gcgttatcat gacgttggcg acgcagcagg tcggtttcac gccagttacc 600
gaaggatgcc agagtgtcgt gggatacaga agtaacgggtg atcaaactct gggtggacca 660
gatttccaaa ccaaccaaag tgaattggat gttcagtggt ctgtagattt cgttaatggg 720
gttgacgatt tgggtcaccc attgacggat tttgtcagag tcaccgttgt atttagtggt 780
catacgggtg tcagcaacga taaccagctg tacgtatctt tgtgggaatc tttgttgtct 840
cttagtcaag ttaattgga aggccttctt gattggttca tatgattccc agttttgggt 900
aacaccacac atctttgggg cttcatcttc cttttcgacg ttttcgtact tgtagacagc 960
atgggcttca gagtccgaca attccaatgg ttcaatcaag tacatttcac cttgcaactt 1020
gaaatgaccc ttcaaaccgt tacaagcaga gatagaagca gtggagtcag catcgttttc 1080
gattctacca tggtagtaac agtgatcttc acccaatggg taagtagtaa tttctctacc 1140
atctggggag taatgagttt cagagtaatc ttcagagAAC aaacctttgt ttttttccaa 1200
gtgcaagaca actggttcac tgtaaacctt gaattcgtat tgcattggcat cttcgtactt 1260
tggttgaaca gcacccctag gaactggagt gacctttctt ggataaacia cttcgtaatc 1320
gttaacgita ccagattcca agataataga agaagcttca gcctctcttt tctcgagaga 1380
tacccttctt tctttagcag caatgctggc aatagtagta tttataaaca ataaccggtt 1440
atttgtgctg ttggaaaatg gcaaaacagc aacatcgaaa tccccttcta aatctgagta 1500
accgatgaca gcttcagccg gaatttgtgc cgtttcatct tctgtttagt tgttgactgg 1560
agcagctaag gcggaggatg ctgcgaataa aacagcagta aaaattgaag gaaatctcat 1620

```

<210> 14

<211> 540

<212> PRT

<213> Agkistrodon contortrix

<220>

<223> Xaa in positions 465, 493, 516 and 536 correspond to stop codons.

<220>

<223> Native pro-fibrolase of Agkistrodon contortrix

<400> 14

```

Met Arg Phe Pro Ser Ile Phe Thr Ala Val Leu Phe Ala Ala Ser Ser
 1             5             10             15

```

```

Ala Leu Ala Ala Pro Val Asn Thr Thr Thr Glu Asp Glu Thr Ala Gln
      20             25             30

```

```

Ile Pro Ala Glu Ala Val Ile Gly Tyr Ser Asp Leu Glu Gly Asp Phe
 35             40             45

```

```

Asp Val Ala Val Leu Pro Phe Ser Asn Ser Thr Asn Asn Gly Leu Leu
 50             55             60

```

```

Phe Ile Asn Thr Thr Ile Ala Ser Ile Ala Ala Lys Glu Glu Gly Val
 65             70             75             80

```

```

Ser Leu Glu Lys Arg Glu Ala Glu Ala Ser Ser Ile Ile Leu Glu Ser
      85             90             95

```

Gly Asn Val Asn Asp Tyr Glu Val Val Tyr Pro Arg Lys Val Thr Pro
 100 105 110
 Val Pro Arg Gly Ala Val Gln Pro Lys Tyr Glu Asp Ala Met Gln Tyr
 115 120 125
 Glu Phe Lys Val Asn Ser Glu Pro Val Val Leu His Leu Glu Lys Asn
 130 135 140
 Lys Gly Leu Phe Ser Glu Asp Tyr Ser Glu Thr His Tyr Ser Pro Asp
 145 150 155 160
 Gly Arg Glu Ile Thr Thr Tyr Pro Leu Gly Glu Asp His Cys Tyr Tyr
 165 170 175
 His Gly Arg Ile Glu Asn Asp Ala Asp Ser Thr Ala Ser Ile Ser Ala
 180 185 190
 Cys Asn Gly Leu Lys Gly His Phe Lys Leu Gln Gly Glu Met Tyr Leu
 195 200 205
 Ile Glu Pro Leu Glu Leu Ser Asp Ser Glu Ala His Ala Val Tyr Lys
 210 215 220
 Tyr Glu Asn Val Glu Lys Glu Asp Glu Ala Pro Lys Met Cys Gly Val
 225 230 235 240
 Thr Gln Asn Trp Glu Ser Tyr Glu Pro Ile Lys Lys Ala Phe Gln Leu
 245 250 255
 Asn Leu Thr Lys Arg Gln Gln Arg Phe Pro Gln Arg Tyr Val Gln Leu
 260 265 270
 Val Ile Val Ala Asp His Arg Met Asn Thr Lys Tyr Asn Gly Asp Ser
 275 280 285
 Asp Lys Ile Arg Gln Trp Val His Gln Ile Val Asn Thr Ile Asn Glu
 290 295 300
 Ile Tyr Arg Pro Leu Asn Ile Gln Phe Thr Leu Val Gly Leu Glu Ile
 305 310 315 320
 Trp Ser Asn Gln Asp Leu Ile Thr Val Thr Ser Val Ser His Asp Thr
 325 330 335
 Leu Ala Ser Phe Gly Asn Trp Arg Glu Thr Asp Leu Leu Arg Arg Gln
 340 345 350
 Arg His Asp Asn Ala Gln Leu Leu Thr Ala Ile Asp Phe Asp Gly Asp
 355 360 365
 Thr Val Gly Leu Ala Tyr Val Gly Gly Met Cys Gln Leu Lys His Ser
 370 375 380

WO 01/25445

PCT/US00/27029

- 11 -

Thr Gly Val Ile Gln Asp His Ser Ala Ile Asn Leu Leu Val Ala Leu
385 390 395 400

Thr Met Ala His Glu Leu Gly His Asn Leu Gly Met Asn His Asp Gly
405 410 415

Asn Gln Cys His Cys Gly Ala Asn Ser Cys Val Met Ala Ala Met Leu
420 425 430

Ser Asp Gln Pro Ser Lys Leu Phe Ser Asp Cys Ser Lys Lys Asp Tyr
435 440 445

Gln Thr Phe Leu Thr Val Asn Asn Pro Gln Cys Ile Leu Asn Lys Pro
450 455 460

Xaa Ala Ala Ala Ser Phe Leu Glu Gln Lys Leu Ile Ser Glu Glu Asp
465 470 475 480

Leu Asn Ser Ala Val Asp His His His His His His Xaa Val Cys Ser
485 490 495

Leu Arg His Asp Cys Ser Ser Val Gln Val Gly His Leu Arg Glu Asp
500 505 510

Arg Ser Cys Xaa Ile Leu Ile Lys Arg Met Ser Glu Cys His Leu Pro
515 520 525

Glu Arg Cys Arg Leu His Phe Xaa Tyr Phe Phe Ile
530 535 540

<210> 15

<211> 203

<212> PRT

<213> Agkistrodon contortrix

<220>

<223> Native fibrolase of Agkistrodon contortrix

<220>

<223> Xaa in position 1 represents the chemical compound, pyroglutamic acid. This chemical compound is referred to in the specification as the letter "E".

<400> 15

Xaa Gln Arg Phe Pro Gln Arg Tyr Val Gln Leu Val Ile Val Ala Asp
1 5 10 15

His Arg Met Asn Thr Lys Tyr Asn Gly Asp Ser Asp Lys Ile Arg Gln
20 25 30

Trp Val His Gln Ile Val Asn Thr Ile Asn Glu Ile Tyr Arg Pro Leu
35 40 45

WO 01/25445

PCT/US00/27029

- 12 -

```

Asn Ile Gln Phe Thr Leu Val Gly Leu Glu Ile Trp Ser Asn Gln Asp
  50          55          60
Leu Ile Thr Val Thr Ser Val Ser His Asp Thr Leu Ala Ser Phe Gly
  65          70          75          80
Asn Trp Arg Glu Thr Asp Leu Leu Arg Arg Gln Arg His Asp Asn Ala
          85          90          95
Gln Leu Leu Thr Ala Ile Asp Phe Asp Gly Asp Thr Val Gly Leu Ala
          100          105          110
Tyr Val Gly Gly Met Cys Gln Leu Lys His Ser Thr Gly Val Ile Gln
          115          120          125
Asp His Ser Ala Ile Asn Leu Leu Val Ala Leu Thr Met Ala His Glu
          130          135          140
Leu Gly His Asn Leu Gly Met Asn His Asp Gly Asn Gln Cys His Cys
          145          150          155          160
Gly Ala Asn Ser Cys Val Met Ala Ala Met Leu Ser Asp Gln Pro Ser
          165          170          175
Lys Leu Phe Ser Asp Cys Ser Lys Lys Asp Tyr Gln Thr Phe Leu Thr
          180          185          190
Val Asn Asn Pro Gln Cys Ile Leu Asn Lys Pro
          195          200

```

<210> 16

<211> 202

<212> PRT

<213> Agkistrodon contortrix

<220>

<223> Native fibrolase of Agkistrodon contortrix

<220>

<223> Xaa in position 1 represents the chemical compound, pyroglutamic acid. This chemical compound is referred to in the specification as the letter "E".

<400> 16

```

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

```

WO 01/25445

PCT/US00/27029

- 13 -

Ile Gln Phe Thr Leu Val Gly Leu Glu Ile Trp Ser Asn Gln Asp Leu -
 50 55 60
 Ile Thr Val Thr Ser Val Ser His Asp Thr Leu Ala Ser Phe Gly Asn
 65 70 75 80
 Trp Arg Glu Thr Asp Leu Leu Arg Arg Gln Arg His Asp Asn Ala Gln
 85 90 95
 Leu Leu Thr Ala Ile Asp Phe Asp Gly Asp Thr Val Gly Leu Ala Tyr
 100 105 110
 Val Gly Gly Met Cys Gln Leu Lys His Ser Thr Gly Val Ile Gln Asp
 115 120 125
 His Ser Ala Ile Asn Leu Leu Val Ala Leu Thr Met Ala His Glu Leu
 130 135 140
 Gly His Asn Leu Gly Met Asn His Asp Gly Asn Gln Cys His Cys Gly
 145 150 155 160
 Ala Asn Ser Cys Val Met Ala Ala Met Leu Ser Asp Gln Pro Ser Lys
 165 170 175
 Leu Phe Ser Asp Cys Ser Lys Lys Asp Tyr Gln Thr Phe Leu Thr Val
 180 185 190
 Asn Asn Pro Gln Cys Ile Leu Asn Lys Pro
 195 200

<210> 17

<211> 538

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Analog form of
 native pro-fibrolase of Agkistrodon contortrix

<400> 17

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

WO 01/25445

PCT/US00/27029

- 14 -

Ser Leu Glu Lys Arg Glu Ala Glu Ala Ser Ser Ile Ile Leu Glu Ser
 85 90 95

Gly Asn Val Asn Asp Tyr Glu Val Val Tyr Pro Arg Lys Val Thr Pro
 100 105 110

Val Pro Arg Gly Ala Val Gln Pro Lys Tyr Glu Asp Ala Met Gln Tyr
 115 120 125

Glu Phe Lys Val Asn Ser Glu Pro Val Val Leu His Leu Glu Lys Asn
 130 135 140

Lys Gly Leu Phe Ser Glu Asp Tyr Ser Glu Thr His Tyr Ser Pro Asp
 145 150 155 160

Gly Arg Glu Ile Thr Thr Tyr Pro Leu Gly Glu Asp His Cys Tyr Tyr
 165 170 175

His Gly Arg Ile Glu Asn Asp Ala Asp Ser Thr Ala Ser Ile Ser Ala
 180 185 190

Cys Asn Gly Leu Lys Gly His Phe Lys Leu Gln Gly Glu Met Tyr Leu
 195 200 205

Ile Glu Pro Leu Glu Leu Ser Asp Ser Glu Ala His Ala Val Tyr Lys
 210 215 220

Tyr Glu Asn Val Glu Lys Glu Asp Glu Ala Pro Lys Met Cys Gly Val
 225 230 235 240

Thr Gln Asn Trp Glu Ser Tyr Glu Pro Ile Lys Lys Ala Phe Gln Leu
 245 250 255

Asn Leu Thr Lys Arg Ser Phe Pro Gln Arg Tyr Val Gln Leu Val Ile
 260 265 270

Val Ala Asp His Arg Met Asn Thr Lys Tyr Asn Gly Asp Ser Asp Lys
 275 280 285

Ile Arg Gln Trp Val His Gln Ile Val Asn Thr Ile Asn Glu Ile Tyr
 290 295 300

Arg Pro Leu Asn Ile Gln Phe Thr Leu Val Gly Leu Glu Ile Trp Ser
 305 310 315 320

Asn Gln Asp Leu Ile Thr Val Thr Ser Val Ser His Asp Thr Leu Ala
 325 330 335

Ser Phe Gly Asn Trp Arg Glu Thr Asp Leu Leu Arg Arg Gln Arg His
 340 345 350

Asp Asn Ala Gln Leu Leu Thr Ala Ile Asp Phe Asp Gly Asp Thr Val
 355 360 365

Gly Leu Ala Tyr Val Gly Gly Met Cys Gln Leu Lys His Ser Thr Gly
 370 375 380

Val Ile Gln Asp His Ser Ala Ile Asn Leu Leu Val Ala Leu Thr Met
 385 390 395 400

Ala His Glu Leu Gly His Asn Leu Gly Met Asn His Asp Gly Asn Gln
 405 410 415

Cys His Cys Gly Ala Asn Ser Cys Val Met Ala Ala Met Leu Ser Asp
 420 425 430

Gln Pro Ser Lys Leu Phe Ser Asp Cys Ser Lys Lys Asp Tyr Gln Thr
 435 440 445

Phe Leu Thr Val Asn Asn Pro Gln Cys Ile Leu Asn Lys Pro Xaa Ala
 450 455 460

Ala Ala Ser Phe Leu Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu Asn
 465 470 475 480

Ser Ala Val Asp His His His His His His Xaa Val Cys Ser Leu Arg
 485 490 495

His Asp Cys Ser Ser Val Gln Val Gly His Leu Arg Glu Asp Arg Ser
 500 505 510

Cys Xaa Ile Leu Ile Lys Arg Met Ser Glu Cys His Leu Pro Glu Arg
 515 520 525

Cys Arg Leu His Phe Xaa Tyr Phe Phe Ile
 530 535

<210> 18

<211> 602

<212> DNA

<213> Agkistrodon contortrix

<220>

<223> Fragment of fibrolase of Agkistrodon contortrix

<400> 18

tactattgcc agcattgctg ctaaagaaga aggggtatct ctcgagaaaa gagaggctga 60
 agcttcttct attatcttgg aatctggtaa cgtaacgat tacgaagttg tttatccaag 120
 aaaggctact ccagttccta ggggtgctgt tcaaccaaag tacgaagatg ccatgcaata 180
 cgaattcaag gttaacagtg aaccagttgt cttgcacttg gaaaaaaca aagggtttgtt 240
 ctctgaagat tactctgaaa ctcttactc cccagatggg agagaaatta ctacttacc 300
 attgggtgaa gatcactgtt actaccatgg tagaatcgaa aacgatgctg actccactgc 360
 ttctatctct gcttgtaacg gtttgaaggg tcatttcaag ttgcaagggtg aaatgtactt 420
 gattgaacca ttggaattgt ccgactctga agcccatgct gtctacaagt acgaaaacgt 480
 cgaaaaggaa gatgaagccc caaagatgtg tgggtgttacc caaaactggg aatcatatga 540
 accaatcaag aaggccttcc aattaaactt gactaagaga tctttccac aaagatacgt 600
 ac 602

<210> 19

<211> 816

WO 01/25445

PCT/US00/27029

- 16 -

<212> DNA

<213> Agkistrodon contortrix

<220>

<223> Fragment of fibrolase of Agkistrodon contortrix

<400> 19

```

tccaattaaa cttgactaag agatctttcc cacaaagata cgtacagctg gttatcgttg 60
ctgaccaccg tatgaacact aaatacaacg gtgactctga caaaatccgt caatgggtgc 120
accaaactcg caacaccatt aacgaaatct acagaccact gaacatccaa ttcactttgg 180
ttggtttgga aatctgggtc aaccaagatt tgatcacctg tacttctgta tcccacgaca 240
ctctggcacc cttcggtaac tggcgtgaaa ccgacctgct gcgtcgccaa cgtcatgata 300
acgctcaact gctgaccgct atcgacttcg acggtgatac tgttggtctg gcttacgttg 360
gtggcatgtg tcaactgaaa cattctactg gtgttatcca ggaccactcc gctattaacc 420
tgctgggttg tctgaccatg gcacacgaac tgggtcataa cctgggtatg aaccacgatg 480
gcaaccagtg tcaactgcgt gcaaaactct gtgttatggc tgctatgctg tccgatcaac 540
catccaaact gttctccgac tgctctaaga aagactacca gaccttctct accgttaaca 600
accgcagtg tctctgaac aaaccgtaag cggccgcccag ctttctagaa caaaaactca 660
tctcagaaga ggatctgaat agcgccgctg accatcatca tcatcatcat tgagtttgta 720
gccttagaca tgactgttcc tcagttcaag ttgggcactt acgagaagac cggctctgct 780
agattctaat caagaggatg tcagaatgcc atttgc 816

```

<210> 20

<211> 1373

<212> DNA

<213> Agkistrodon contortrix

<220>

<223> Fragment of fibrolase of Agkistrodon contortrix

<400> 20

```

tactattgcc agcattgctg ctaaagaaga aggggtatct ctcgagaaaa gagaggctga 60
agcttcttct attatcttgg aatctggtaa cgtaacgat tacgaagttg tttatccaag 120
aaaggtcact ccagttccta ggggtgctgt tcaaccaaag tacgaagatg ccagtcaata 180
cgaattcaag gttaacagtg aaccagttgt cttgcacttg gaaaaaaca aaggtttggt 240
ctctgaagat tactctgaaa ctcttactc cccagatggg agagaaatta ctacttacc 300
attgggtgaa gatcactgtt actaccatgg tagaatcgaa aacgatgctg actccactgc 360
ttctatctct gcttgtaacg gtttgaaggg tcatttcaag ttgcaagggt aaatgtactt 420
gattgaacca ttggaattgt ccgactctga agcccctgct gtctacaagt acgaaaacgt 480
cgaaaaggaa gatgaagccc caaagatgtg tgggtgttacc caaaactggg aatcatatga 540
accaatcaag aaggccttcc aattaaactt gactaagaga tctttcccac aaagatacgt 600
acagctgggt atcggtgctg accaccgtat gaacactaaa tacaacggtg actctgacaa 660
aatccgtcaa tgggtgcacc aaatcgtaaa caccattaac gaaatctaca gaccactgaa 720
catccaattc actttgggtg gtttggaaat ctggtccaac caagatttga tcaccgttac 780
ttctgtatcc caccgactc tggcatcctt cggttaactg cgtaaaaccg acctgctgcg 840
tcgccaacgt catgataacg ctcaactgct gaccgctatc gacttcgacg gtgatactgt 900
tggtctggct tacgttggtg gcatgtgtca actgaaacat tctactgggt ttatccagga 960
ccactccgct attaacctgc tgggtgctct gaccatggca cacgaactgg gtcataacct 1020
gggtatgaac cagatggca accagtgtca ctgcggtgca aactcctgtg ttatggctgc 1080
tatgtgtcc gatcaaccat ccaaactgtt ctccgactgc tctaagaaag actaccagac 1140
cttctgacc gttaacaacc cgcagtgtat cctgaacaaa ccgtaagcgg ccgccagctt 1200
tctagaacaa aaactcatct cagaagagga tctgaatagc gccgtcgacc atcatcatca 1260
tcatcattga gtttgtagcc ttagacatga ctgttctca gttcaagttg ggcacttacg 1320
agaagaccgg tcttgctaga ttctaataca gaggatgtca gaatgccatt tgc 1373

```