

РЕПУБЛИКА БЪЛГАРИЯ



(19) BG

(11) 106577 A
7(51) C 12 N 15/57

ЗАЯВКА ЗА ПАТЕНТ

ЗА

ИЗОБРЕТЕНИЕ

C 12 N 9/64

C 07 K 16/40

C 12 N 5/10

ПАТЕНТНО ВЕДОМСТВО

(21) Регистров № 106577

(22) Заявено на 04.04.2002

(24) Начало на действие
на патента от:

Приоритетни данни

(31) 411329 (32) 01.10.99 (33) US

(41) Публикувана заявка в
бюлетин № 4 на 30.04.2003

(45) Отпечатано на

(46) Публикувано в бюлетин №
на

(56) Информационни източници:

(62) Разделена заявка от рег. №

(71) Заявител(и):

AMGEN INC.

THOUSAND OAKS, CA (US)

(72) Изобретател(и):

Thomas Charles Boone

Huimin Li

Newbury Park, CA

Michael Benjamin Mann

Thousand Oaks, CA (US)

(74) Представител по индустриална
собственост:

Любомир Илиев Косев, 1000 София,
ул. "Княз Борис I" 61 А, ет. 2, офис 8

(86) № и дата на РСТ заявка:

PCT/US00/27029, 29.09.2000

(87) № и дата на РСТ публикация:

WO01/25445, 12.04.2001

(54) ФИБРИНОЛИТИЧНО АКТИВЕН ПОЛИПЕПТИД

(57) Изобретението се отнася до фибринолитично активен полипептид металопротеиназа, който е тромболитик с ново действие, полезен за разтваряне на кръвни съсиреци *in vivo*, и до методи и материали за получаването му чрез рекомбинантна експресия.

19 претенции

BG 106577 A

ФИБРИНОЛИТИЧНО АКТИВЕН ПОЛИПЕПТИД

Област на изобретението

Това изобретение се отнася до фибринолитично активната металопротеиназа на последователност, която не се среща в естествено състояние и до комбинираните методи за нейното получаване и използване за лечение на тромбоза в жив организъм.

История на изобретението

Фибролазата е ензимно-активен полипептид (особено металопротеиназата), състояща се от 203 аминокиселинни остатъци, която първоначално беше изолирана чрез пречистване на змийска отрова от Южен Копърхед; Патент на Съединените щати No 4,610,879, издаден на 09 септември 1986 г. (*Markland* и гр.)

и *Guan* и др., Архиви по Биохимия и Биофизика, том 289. брой 2, стр. 197-207 (1991 г.). Ензимът проявява директна фибринолитична активност с малка или никаква хеморагична активност, който разтваря кръвни съсиреци, образувани от фибриноген или от кръвта.

Аминокиселинната последователност на фибролазата е била определена с методи, описани за рекомбинантно получаване на дрожди и използването ѝ за лечение на тромбемболично състояние в жив организъм; *Randolf* и други, Наука за протеините, Издателство на Университета в Кембридж (1992 г.), страници 590-600, и заявка за европейски патент No 0 323 722 (*Valenzuela* и др.), публикувана на 12 юли 1989 г.

Резюме на изобретението

Това изобретение предоставя фибринолитична метало-протеиназа с несъществуваща в природата линейна проба от аминокиселини, посочени в SEQ ID No: 1, спомената тук като "тромболитик с ново действие" (или NAT). Също така са дадени молекули на нуклеиновата киселина, като тази в SEQ ID No: 2 и варианти на същата, кодиращи NAT.

Терминът "зрял" се използва в неговия обичаен смисъл, за означаване на биологично-активния полипептид, който е бил обработен ензимно на място в клетка-домакин, за да се откъсне от областта *prepro*.

Поради неговата фибринолитична активност, NAT е полезен *ин vivo* като разтварящ агент на кръвни съсиреци, за лечение на тромбоза в бозайник, (включително плъхове, свине и хора).

Полипептидът от това изобретение има предимства пред естествено срещащата се фибролаза, като терапевтичен агент (например, фибринолитичният полипептид, намиращ се в змийската отрова). Естествената фибролаза съдържа няколко алтернативни N-терминала: QQRFF, EQRFP и ERFP, (в които "Е" означава цикличен глутамин, или пироглутаминова киселина). По-точно, като се започне с N-терминала на QQRFP, молекулата на фибролазата претърпява разпадане до две изоформи, съответно с N-терминални последователности на EQRFP и ERFP. Рекомбинантната фибролаза, получена в дрожди, обикновено произвежда смес от всички тези три форми и следователно не е хомогенна. Вижте Loayza и други, Журнал по хроматография, В 662, страници 227-243 (1994 г.). Освен това, цикличният глутаминов остатък в "блокирания" N-терминал, прави невъзможно образуването на последователност.

Обратно, рекомбинантният NAT от това изобретение дава отделни видове: само N-терминалът се получава по обикновен начин. Резултатът е по-голяма хомогенност на крайния продукт в сравнение с рекомбинантната фибролаза, която е полезна когато се предвиждат медицински приложения за пациентите.

Кратко описание на фигурите

Фигура 1 изобразява линейния вид на пълната аминокиселинна последователност на NAT (SEQ ID No:3), съставена от област "prepro", (подчертана), която съдържа "сигналния" пептид и зрелия полипептид (неподчертан).

Подробно описание на изобретението

NAT може да се получи от рекомбинантната експресия на молекулата на нуклеиновата киселина от SEQ ID No: 4, която кодира пълната аминокиселинна последователност на NAT (SEQ ID No: 3), включително областта *prepro* от нуклеотиди 1 -783 и зрелия полипептид от нуклеотиди 784-1386, в подходящ приемник. Следвайки експресията, областта *prepro* е преработена ензимно в клетката-домакин, за да се получи зрял активен полипептид (SEQ ID No: 1).

За предпочитане е NAT да се получи рекомбинатно в дрожди, както ще бъде обяснено по-подробно по-долу.

Зрелият полипептид (SEQ ID No: 1), получен по този начин, може да има или не аминокиселинен терминален метионин, в зависимост от начина, по който е изготвен. Обикновено аминокиселинният терминален метионинов остатък ще присъства когато полипептидът се получава рекомбинатно в

несекреторен бактериален щам като домакин (*например Ешерихия коли*).

Освен молекулата на нуклеиновата киселина от SEQ ID No: 2, също се използват и изродени последователности на същата, които кодират същия полипептид. Настоящото изобретение разглежда също така молекули на нуклеиновата киселина, които могат да кодират допълнителни аминокиселинни остатъци, разположени странично на 5' или 3' части на областта, кодираща зрелия полипептид, като последователностите, кодиращи *pre/pro* области (например, последователности, отговорни за отделянето на полипептида през клетъчните мембрани) на мястото на "естествената" *pre/pro* област, (т.е. намираща се в естествено срещащата се фибролаза). Допълнителните последователности могат също така да бъдат некодирани последователности, включително регулиращи последователности, като например, ускорители на транскрипция или трансляция, в зависимост от клетката-домакин.

NAT може да се приготви като се използват добре познатите методи на рекомбинантната ДНК технология, като например изложените в *Sambrook* и други, Молекулно клониране, лабораторен Наръчник, Издателство на Лабораторията в Колд Спринг Харбър, Ню Йорк (1989 г.) и/или *Ausubel* и други издатели, Текущи протоколи в Молекулната биология, Издатели *Green Inc.* и *Wiley* и Синове, Ню-Йорк (1994 г.). ДНК молекула, кодираща

полипептида или скъсените версии на същата могат да се получат, например чрез екраниране на геном или проба *cDNA*, или чрез усилване на *PCR*, за да се получи молекула на нуклеинова киселина, кодираща фибролаза, последвана от замяна на кодоните, кодиращи N-терминалните аминокиселинни остатъци QQR, с кодон за серин (S). Обратно, ДНК молекулата, кодираща NAT, може да бъде получена чрез химичен синтез, използвайки методи, добре познати на специалистите в тази област, като например, описаните от Енгелс и други в *Angew*, Международно издателство по химия, том 28, страници 716-734 (1989 г.). Обикновено, ДНК представлява няколко стотици нуклеотиди по дължината. Нуклеинови киселини, по-големи от около сто нуклеотида, могат да се синтезират като няколко фрагмента, използвайки същите тези методи; след това фрагментите могат да се свържат заедно, за да образуват нуклеотидна последователност с желаната дължина.

ДНК молекула се внася в съответния експресиен носител за експресия в подходяща клетка-домакин. Носителят се избира да бъде функционален в използваната клетка-домакин (например, носителът е съвместим с клетката-домакин, в която може да възникне експресия на ДНК). Полипептидът може да е експресиран в прокариотна клетка, дрожди, насекомо (бакуловирусни системи) или евкариотни клетки-домакини, въпреки, че се предпочитат дрожди, както ще бъде обяснено по-нататък.

Носителите, използвани във всяка от клетките-домакини за експресия на NAT могат да съдържат също 5' странични последователности (наричани още "промотори") и други експресивни регулаторни елементи, оперативно свързани с ДНК за експресия, както и ускорител(и), началото на репликиращия елемент, транскрибиращ терминален елемент, пълна интронова последователност, съдържаща донорна и акцепторна съединителна страна, сигнална пептидна последователност, рибозомен свързващ страничен елемент, полиагенилираща последователност, полилинкерна област за внасяне на нуклеинова киселина, кодираща NAT и елемент на избираемия маркер. Всеки от тези елементи се обсъжда по-нататък.

По избор, носителят може да съдържа също така "крайна" последователност, т.е. олигонуклеотидна последователност, разположена в края 5' или 3' на полипептидно-кодиращата последователност полихис (като например хексахис) или друга малка имуногенна последователност (като например, с-мик или хемаглутининов епитоп, за който антителата, включително моноклоналните антитела, се намират в търговската мрежа.

Тази крайна последователност ще бъде експресирана заедно с NAT, и може да служи като афинитетен край за очистване на този полипептид от клетката-домакин. По избор, крайната последователност може впоследствие да бъде отстранена от пречистения полипептид по различни начини, например, като се използва избрана пептидаза.

Страничната последователност 5' може да бъде естествената 5' или хомоложната последователност (т.е. от същите видове и/или щам като клетката-домакин), хетероложната (т.е. от видовете, различни от тези на клетката-домакин или щам), хибриди (т.е. комбинацията от 5' странична последователност с повече от един източник), или синтетична. Източникът на страничната последователност 5' може да бъде всеки едноклетъчен прокариотен или еукариотен организъм, всяко гръбначно или безгръбначно животно или всяко растение, при условие, че страничната последователност 5' е функционална и може да се активира от механизма на клетката-домакин.

Началото на репликиращия елемент обикновено е част от прокариотни експресивни носители, закупени в търговската мрежа и подпомага усилването на носителя в клетката-домакин.

Усилването на носителя до определен номер на копията може в някои случаи да бъде важно за оптималната експресия на NAT. Ако избраният носител не съдържа началото на, репликиращата страна, такъв може да бъде синтезиран химически, въз основа на известна последователност, и след това да бъде свързан в носителя.

Терминалният транскрибиращ елемент обикновено е разположен в 3', към края на кодиращата полипептидна последователност и служи да прекрати копирането (транскрипцията) на *mRNA*. Обикновено, транскрибиращият

терминален елемент в прокариотни клетки е G-C богат фрагмент, последван от поли-T последователност. Докато елементът се клонира лесно от проба или е закупен в търговската мрежа като част от носителя, той също така може да бъде синтезиран лесно, чрез използване на известните методи за синтез на нуклеинова киселина.

Избираем маркер елемент на ген кодира протеин, необходим за оцеляването и растежа на клетката-домакин, развиваща се в хранителната среда на подбраната култура. Обикновените избирателни гени маркери кодират протеините, които:

(а) оказват съпротивление на антибиотиците или на другите токсини, например на ампицилин, тетрациклин или канамицин за прокариотни клетки-домакини, зеоцин за клетки-домакини за грозди, и неомицин за клетки-домакини от бозайници;

(b) допълващи недостига от оксотропин в клетката; или

(c) доставя важни хранителни вещества, които не се намират в комплексната среда. Предпочетените избирателни маркери за използване в прокариотната експресия са резистентни гени на канамицин, ампицилин и тетрациклин.

Рибозомният свързващ елемент, обикновено *Shine-Dalgarno* последователност (за прокариоти) или последователност на *Kozak* (за еукариоти), е необходим за започване на транслацията на mRNA. Елементът обикновено е разположен в 3' към

промоторите 5' и към кодиращата последователност на полипептида, който ще се синтезира. Последователността *Shine-Dalgarno* е различна, но обикновено е полипурин (например, с високо съдържание на A-G). Много последователности *Shine-Dalgarno* са били идентифицирани, всяка от които може да бъде синтезирана лесно чрез използване на методите, изложени по-горе и използвани в прокариотния носител. Последователността на *Kozak* обикновено съдържа последователности непосредствено преди и след започващия кодон. Предпочитаната последователност на *Kozak* е онази, която се свързва с висока ефективност при започване на трансляция при AUG начален кодон.

В онези случаи, когато е желателно полипептидът NAT да бъде отделен от клетката-домакин, може да се използва сигнална последователност, за да насочи полипептида навън от клетката-домакин, където е синтезирана. Обикновено сигналната последователност е разположена в кодиращата област на последователността на нуклеиновата киселина, или непосредствено в края на 5' на кодиращата област. Много сигнални последователности са били идентифицирани и всяка от тях, която е функционална в избраната клетка-домакин, може да бъде използвана тук. Следователно, сигналната последователност може да бъде хомоложна или хетероложна на полипептида. Освен това, сигналната последователност може да бъде синтезирана химически с помощта на методи, описани по-горе.

След като носителът е бил изграден и нуклеиновата киселина е внесена в съответното място на носителя, целият носител може да се вложи в подходяща клетка-домакин за усилване и/или експресия на полипептида.

Като беше споменато, клетки-домакини могат да бъдат прокариотни (като например *Ешерихия коли*) или еукариотни (като например, клетка от грозд), клетка от насекомо или клетка на гръбначно животно). Клетката-домакин, било от грозд или от някой друг домакин, когато се култивира при подходящи условия, може да синтезира NAT, който впоследствие може да бъде събран от хранителната среда (ако клетката-домакин я отделя в средата) или непосредствено от клетката-домакин, която го произвежда (ако той не се отделя). След събирането му, полипептидът NAT може да се пречисти чрез използване на методи като хроматографския метод на молекулното сито, афинитетната хроматография и други.

Подборът на клетка-домакин ще зависи в голяма степен от начина, по който клетката-домакин може да "нагъне" NAT в своята естествена вторична и третична структура (например, съответна ориентация на бисулфидните мостове и др.), така че от нея да се получи биологично активен материал. Обаче, дори когато клетката-домакин не синтезира биологично активен материал, той може да бъде "нагънат" след синтеза, като се използват съответни химически условия, като тези, известни на

специалистите в тази област. Във всеки от случаите, правилното нагъване може да се предвиди от факта, че трябва да се получи биологично активен материал.

Подходящи клетки-домакини или клетъчни линии могат да бъдат клетки от бозайници, например клетки от яйчниците на китайския хамстер (CHO) или клетки 3Т3. Изборът на подходящи клетки-домакини от бозайници и методите за преобразуване, култура, усилване, екраниране, изготвяне на продукта и очистване са известни на специалистите в тази област. Други подходящи клетъчни линии от бозайници са маймунските COS-7 и CV-1 клетъчни линии. Следващи примерни клетки-домакини от бозайници са клетъчни линии на примати и гризачи, включително трансформирани клетъчни линии. Нормалните диплоидни клетки, клетъчните щамове, получени от култура на първична тъкан *ин-витро*, както и от първични външни култури, също така са подходящи. Клетки кандидати могат да са генотипно недостатъчни в избрания ген, или могат да съдържат доминиращо действие на избрания ген. И други подходящи клетъчни линии от бозайници включват, но не са ограничени до *Hela*, миши клетки L-929, 3Т3 линии, получени от Swiss, Balb-с или NIH мишки, ВНК или HaK клетъчни линии от хамстер.

Също така полезни като клетки-домакини са бактериалните клетки. Например, различните щамове на *Ешерихия коли* (като

HB101, DH5a, DH10, и MC1061) са добре познати като клетки-домакини в областта на биотехнологията. Различни щамове на *B. subtilis*, *Pseudomonas spp.*, други *Bacillus spp.*, *Streptomyces spp.* и подобни, също могат да бъдат използвани. Освен това, много клетки на щамове на грожди, известни на специалистите в тази област, също могат да бъдат клетки-домакини за експресия на полипептида от това изобретение. При желание, клетки от насекомо също така могат да се използват като клетки-домакини. Вижте, например. Милър и др., Генно инженерство, том 8, страници 277-298 (1986 г.).

Внасянето на носител в избраната клетка-домакин (известно също така като "преобразуване" или "трансфекция") може да се осъществи чрез използване на методи като калциев фосфат, електропорация, микроинжектиране, липофекция или метод DEAE - декстран. Избраният метод ще зависи отчасти от вида на клетката-домакин, която ще се използва. Тези и други подходящи методи са добре познати на специалистите в тази област и са разгледани по-горе, например в *Sambrook* и други.

Клетките-домакини, съдържащи носителя, могат да се изследват като култури, като се използват стандартни среди, добре познати на специалистите. Средите обикновено съдържат всички необходими хранителни вещества за развитието и оцеляването на клетките. Подходящи среди за клетките на културата *Ешерихия коли* са например *Luria Broth (LB)* и / или

Terrific Broth (TB). Подходящи среди за развитие на евкариотни клетки са RPMI 1640, MEM, DMEM, които могат да бъдат добавени към серум и / или фактори на растежа, както се изисква от клетъчната линия, която ще бъде култура. Подходяща среда за култури на насекоми е средата на *Grace* с добавена субстанция от дрожди, лакталбумин хидролизат и / или серум от ембрион на теле, ако е необходимо.

Обикновено, антибиотик или друго полезно съединение за избирателно развитие на трансформирани клетки се добавя просто към средата. Съединението, което ще се използва, се определя от избирателния елемент маркер, намиращ се в плазмид, с който е трансформирана клетката-домакин. Например, когато избирателният елемент маркер е резистентен канамицин, добавеното съединение към средата на културата ще бъде канамицин.

Количеството на получения NAT в клетката-домакин може да се определи като се използват известни стандартни методи. Тези методи включват, без ограничение, западен анализ за петна, SDS - полиакриламидна гел електрофореза, HPLC разделяне, имуноутаяване, и / или активни проби като ДНК гел свързващи проби.

Ако NAT се отделя в клетки-домакини, различни от грам-негативните бактерии, вероятно е по-голямата част да се намери

в средата на клетъчната култура. Ако NAT се отделя от грам-негативни бактерии, той може да бъде открит до известна степен в периплазмата. Ако NAT не се отделя, той ще се намира в цитоплазмата.

За вътрешно-клетъчен NAT, обикновено клетките-домакини първо се разкъсват механично. За NAT с периплазмено разположение може да се използва механично разрушаване или осмотично третиране, за да се освободи периплазменото съдържание в буфериран разтвор. След това полипептидът NAT се изолира от този разтвор. Може да се извърши очистване от разтвора с разнообразни техники. Ако NAT е бил синтезиран така, че да има крайна група като хектохистидин или друг малък пептид в двата си терминала - карбоксилен или амино, той може да пречистен значително в едностепен процес чрез прекарване на разтвора през афинитетна колона, където матрицата на колоната силен афинитет към крайната група или директно към полипептида (т.е. моноклонално антитяло). Например, полихистидинът се свързва с голям афинитет и специфичност към никел, така че за очистване може да се използва колона от никел (като никелова колона *Quiagen*). (Вижте, например, *Ausbel* и гр. редактори, Текущи протоколи в молекулярната биология, по-горе).

От друга страна, когато полипептидът няма крайна група и няма налични антитела, могат да се използват други добре известни процедури за очистване. Такива процедури включват, без

ограничение, йонно-обменна хроматография, хроматография на молекулното сито, хроматография на обратната фаза, HPLC, електрофореза на природен гел в комбинация с отмиване на гела и подготвително изоелектрично фокусиране ("Изопрайм" механизъм / методика, научен Хьофер). В някои случаи, две или повече методики могат да се комбинират, за да се постигне по-висока степен на чистота.

Особено предпочитани за използване при получаването на NAT са клетките от дрожди и най-вече на дрождевия генус, известен още и като *Pichia* (например *Pichia pastoris*) поради голямата ефективност при нагъване, в сравнение с бактериалните клетки на *Ешерихия коли*. Подходящи рекомбинантни методи за експресия на този дрождев щам са описани в Патенти на САЩ No 4,855,231 (*Stroman* и гр.), 4,812,405 (*Lair* и гр.), 4,818,700 (*Cregg* и гр.), 4,885,242 (*Cregg*) и 4,837,148 (*Cregg*), описанията на които са включени тук за справка.

За отбелязване е, че клетките на *Pichia* могат да се използват също за експресия на фибролаза с подобна ефективност на молекулите на ДНК, кодиращи тази металопротеиназа и такъв метод представлява допълнителен аспект на настоящото изобретение. Фибролазата е известна металопротеиназа, която е описана в научната и патентна литература; вижте *Randolf* и гр. и Заявка за Европейски Патент No 0 323 722, цитирана по-горе. Обикновено, фибролазата за експресиране е от последовател-

ността SEQ ID No: 5, която се кодира от молекула на cDNA от SEQ ID No: 6 (или варианти на същата, кодираща същата аминокиселинна последователност). Експресирането на фибролазата в такава система обикновено включва молекула ДНК от SEQ ID No: 7, която кодира "репро" последователност (нуклеотиди 1-783), в допълнение към "зрелия" полипептид (нуклеотиди 784-1392).

Химически модифицираните версии на NAT, в който полипептидът е свързан с полимер или с друга молекула, за да формира производна за изменение на свойствата, също са включени в обхвата на настоящото изобретение. Особено за терапевтични цели на хора, може би е предимство NAT да се получава чрез присъединяване на една или повече групи към тази на полипептида. Такива химически групи могат да се изберат измежду различни водоразтворими полимери. Полимерът би трябвало да е водоразтворим, така че полипептидът NAT, към който е прикачен, да може да се смеси с водната среда, например с физиологичен разтвор. Водоразтворимият полимер може да се избере от групата, съдържаща например полиетиленов гликол, ко-полимери на етиленов гликол / пропиленов гликол, карбоксиметилцелулоза, декстран, поливинилов пиролидон, поли-1,3-диоксан, поли-1,3,6-триоксан, етилен / малеинов анхидрид ко-полимер, полиаминокиселини (било хомополимери или случайни или неслучайни ко-полимери (вижте по-долу относно сливането на молекули), и декстран или поли(п-винилов пиролидон)поли-

етиленов гликол, хомополимери на пропиленов гликол, ко-полимери на полипропиленов окис / етиленов окис, полиоксиетилирани полиоли, полистирен-малеат и поливинилов алкохол. Полиетиленовият гликол пропионалдехид може да има преимущества в производството поради неговата устойчивост във вода.

Полимерът може да има произволно молекулно тегло, и може да е разклонен или не. За полиетиленовия гликол предпочитаното молекулно тегло е приблизително между 2 килодалтона (kDa) и 100 kDa (терминът “приблизително” означава, че при получаването на полиетиленовия гликол някои молекули тежат повече, а някои по-малко от означеното молекулно тегло) за лесна манипулация и производство. Могат да се използват и другите числа, в зависимост от желания терапевтичен профил (например, продължителността на желаното поддържано отделяне, влиянието, ако има такова, върху биологичната активност, лесното манипулиране, степенята или липсата на антигенност и други известни влияния на полиетиленовия гликол върху терапевтичния протеин).

Броят на прикрепените полимерни молекули може да варира и специалистът в тази област може да установи влиянието върху функцията. Възможно е да се направи моноderivатизация, но също и ди-, три-, тетра- или някоя комбинация на derivатизацията, със същите или с различни химични групи (например, полимери с

различно тегло на полиетиленовите гликоли). Пропорцията на полимерните молекули към полипептидните молекули на NAT ще се променя, както и техните концентрации в реакционната смес.. По принцип, оптималното съотношение (чрез ефективността на реакцията, по отношение на нереагирал полипептид или полимер) ще бъде определена от фактори, като например желаната степен на дериватизация (например, моно, di-, tri-, и т.н.), молекулното тегло на избрания полимер, дали полимерът е разклонен или не и условията на реакцията.

Химичните групи би трябвало да се прикачват към NAT, като се има предвид въздействието върху функционалните или антигенни области на полипептида. Има известен брой методи за присъединяване, известни на специалистите в тази област. Вижте, например, EP 0 401 384 (свързването на PEG към G-CSF) и Малик и др. , Експериментална хематология, том 20, страници 1028-1035 (1992 г.)(относно пегилацията на GM-CSF с използване на трезилов хлорид). За илюстрация, полиетиленовият гликол може да се свърже ковалентно чрез аминокиселинни остатъци към реактивна група, като например свободната амина или карбоксилна група. Реактивните групи са онези, към които може да се свърже активираната молекула на полиетиленовия гликол (или друга химична група). Аминокиселинните остатъци, имащи свободна аминогрупа, могат да съдържат лизинови остатъци и N-терминален аминокиселинен остатък. Онези, които имат свободна карбоксилна група, могат да съдържат остатъци на

аспарагинова киселина, остатъци на глутаминова киселина, и C-терминален аминокиселинен остатък. Като реактивна група могат да бъдат използвани сулфидрилни групи за прикачване към молекулата (-те) на полиетиленовия гликол (или друга химическа група). Предпочитано за целите на производството е присъединяване към амино групата, като например, към N-терминала или към лизиновата група. Трябва да се избягва присъединяване към остатъци, които са важни за свързването на рецептор, ако това свързване на рецептора е желано.

В специални случаи може да се желаят N-терминално химически модифицирани деривати. Използвайки полиетиленовия гликол като илюстрация, специалистът може да избере от разнообразието на молекули на полиетиленов гликол (по молекулно тегло, разклонения и др.) пропорцията на молекулите на полиетиленов гликол към молекулите на полипептида в реакционната смес, вида на пегилацията, която трябва да бъде изпълнена и метода на получаване на избрания N-терминален пегилиран NAT. Методът на получаване на N-терминален пегилиран препарат (например, отделяйки тази група от другите монопегилирани групи, ако е необходимо) може да бъде чрез пречистване на N-терминален пегилиран материал от популацията на пегилираните NAT молекули. Избраната N-терминална химична модификация може да се извърши чрез редуктивна алкилация, която използва диференциалната реактивност на различните видове първични amino групи (лизин

срещу N-терминала), налични за дериватизация. Вижте РСТ заявка WO 96 / 11953, публикувано на 25 април, 1996 г. При подходящи условия за реакция е постигната значителна избирателна дериватизация на NAT в N-терминал с карбонилна група, съдържаща полимер. Например, може селективно да се пегулира NAT при N-терминала като се извърши реакцията при рН, което позволява да се вземе предимство от разликите рКа между ϵ -амино група на лизинови остатъци и тази на α -амино група на N-терминалния остатък на полипептида. С помощта на такава избирателна дериватизация се контролира присъединяването на полимер към полипептида: конюгацията с полимера става преимуществено в N-терминала на полипептида и не се появява никаква значителна модификация на другите реактивни групи, като например, аминокислотите на лизиновата страна на веригата. Използвайки редуктивна алкилация, полимерът може да бъде от описания по-горе вид и би трябвало да има единичен реактивен алдехид за свързване с полипептида. Може да се използва полиетиленов гликол пропионалалдехид, съдържащ единичен реактивен алдехид.

NAT или химически модифицирани производни съгласно изобретението могат да се формулират за приемане *ин vivo*; най-за предпочитане е през интратромба (т.е. чрез локализирано внасяне директно към стената на съсирека в кръвоносния съд, т. е. чрез катетър). Обикновено не се предпочита системното внасяне, поради вероятността собственият $\alpha 2$ макроглобулин в

общото кръвообращение да взаимодейства с NAT и да предотврати взаимодействието с фибрина или фибриногена, влошавайки по този начин разтварянето на съсирека. Има случаи, обаче, когато могат да се използват по-големи количества NAT, които превишават циркулационните нива на α_2 макроглобулина, като по този начин се улеснява системното назначаване и внасяне на медикамента в организма на пациента. Изобщо, фармацевтични състави, обхванати от изобретението, съдържат ефективни количества от NAT, заедно с фармацевтично приемливи разредители, консерванти, разтворители, емулгатори, спомагателни агенти и /или носители. "Ефективно количество" означава това количество, което е достатъчно да произведе измерим биологичен ефект (например, тромболитично ефективно количество, което влияе на съсирека в кръвта или при лечението на съсиреците).

Обикновено, NAT ще бъде в много пречистена форма и всеки фармацевтичен състав, който се използва като носител, по правило ще бъде стерилизиран за употреба, например чрез филтриране през стерилни филтриращи мембрани.

Специалистът в тази област ще може да установи ефективните дози чрез лекарско назначение и наблюдение на терапевтичния ефект. Определените ефективни дози в този обхват ще зависят от лекуваното заболяване или увреждане, както и от възрастта и общото здравно състояние на

пациента и могат да се установят чрез стандартните клинични процедури. Когато е възможно, желателно е да се определи кривата "доза - реакция" на фармацевтичния състав първо *ин витро*, както в проба на биосистема и след това в полезна животинска система 1 *ин vivo* преди да се изпитва у хора. Умелият практик обмисля терапевтичния контекст, вида на лекуваното заболяване и други приложими фактори, и може да определи правилната доза без излишно усилие. Обикновено практикът предписва състава NAT до получаването на доза с желания ефект (т.е., разтваряне на кръвния съсирек). Съставът може да се предписва като единична, двойна или многократна доза, (която може да съдържа или не същото количество полипептид) за определено време, или на непрекъсната база.

NAT може да бъде използван също така да генерира антитела, съгласно стандартните методи. Антителата могат да бъдат поликлонални, моноклонални, рекомбинантни, химерични, с единична и/или с двойна верига, и т.н. За да се подобри приликата на получаване на имунна реакция, аминокиселинната последователност на NAT може да се анализира, за да се идентифицират частите на молекулата, което може да се свърже с увеличената имуногенност. Например, аминокиселинната последователност може да се подложи на компютърен анализ, за да се определят повърхностните епитопи, както е според метода на Хоуп и Уудс, Процедури на Националната Академия на науките на САЩ, том 78, страници 3824-3828 (1981 г.).

Различни процедури, известни на науката, могат да се използват за получаване на поликлонални антитела, които разпознават епитопите на NAT. За получаване на антитяло, различни животни-домакини могат да се имунизират с инжекция с полипептида, включително, но не ограничено до зайци, мишки, плъхове и др. Различни спомагателни агенти могат да се използват, за да се увеличи имунологичната реакция в зависимост от видовете - домакини, включително, но не ограничено до минерални гелове на Фройнд, като алуминиев хидроокис (стипца), повърхностно-активни вещества, като например лизолецитин, плуронови полиоли, полианиони, пептиди, маслени емулсии, лимфет хемоцианини, динитрофенол, и потенциално полезните човешки помощни агенти като *Bacille Calmette-Guerin* и *Corynebacterium parvum*.

За изготвянето на моноклонални антитела, насочени към NAT, може да се използва всеки метод, който дава възможност за получаване на молекули на антитяло чрез непрекъснати клетъчни линии в културата. Например, методът хибридома, разработен първоначално от Кьолер и Милщайн, описан в Природа, том 256, страници 495-497 (1975 г.), както и методът триома, метод на човешка В-клетъчна хибридома, описана от Kozbor и други в Имунологията днес, том 4, страница 72 (1983 г.), и методът EBV-хибридома за получаване на моноклонални антитела, описани от Cole и други в "Многоклонални антитела и ракова терапия", Alan R. Liss, Inc., страници 77-96 (1985 г.) са полезни за получаването на

моноклонални антитела, съгласно това изобретение.

Антителата на това изобретение могат да се използват терапевтично, след като бъдат предписани, например за свързване към излишни количества NAT, неутрализирането и инхибирането им по този начин *ин vivo*.

Антителата могат да бъдат използвани и за диагностични цели, например като образецът за откриване наличие на NAT в телесните течностите, образци от тъкани или други екстракти, съгласно известните диагностични методи.

Описание на специфични приложения

Изобретението се илюстрира от следните примери.

Пример 1

Деривация на последователността NAT

Ефективен начин за получаване на фибролаза е първоначално да се експресира като препрофибролаза, в която се извършва свързването на протеаза *kex-2* при връзката на "грего" и "зрелите" области за получаване на биологично активен материал ("зряла" фибролаза). От тази разработка, синтезът и преработката на препрофибролазата водят до отделянето на зряла фибролаза в хранителната среда на културата. Действителната последователност в мястото на разцепване е (...TKR↓QQRF ...).

Kex-2 е ендопротеаза, която се разцепва впоследствие на две съседни главни аминокиселини, в този случай на лизин(K) - аргинин (R). Зрялата фибролаза, експресирана от ДНК и имаща гореспоменатата последователност, разкрива, че очакваният остатък N-терминален глутамин (Q) всъщност е претърпял дезамидация и циклизация, за да произведе пироглутаминова киселина (E). Тази химическа модификация бе считана за нежелателна, тъй като остатъкът от пептиди с N-терминален циклизиран глутамин (пироглутаминова киселина) не успява да

реагира при процедурата на разпадането *Edman* за образуване на аминокиселинната последователност. Съответно, и двата остатъка на глутамин (Q) в N-терминала на последователността на зряла фибролаза са били заличени, което е довело до появата на остатък N-терминален аргинин (R). Понеже *kex-2* се разцепва след две съседни главни аминокиселини, както е споменато, очакваше се, че последователността (...KRRF...) би представила неясна страна за деление *kex-2*. Съответно, остатъкът на N-терминалния аргинин (R) (посочен по-горе) бе заменен с остатък на серин (S), за да се получи в резултат последователността (...KRSF..). Изборът на серин се основаваше на необходимостта от въвеждане на аминокиселина, която улеснява отцепването на *kex-2*, когато това става в C-терминалната страна на хидролизната стена *Rholam* и др (Европейски журнал по биохимия, том 227, страници 707-714 (1995 г.).

В резултат последователността на ДНК за препрофилазата бе модифицирана от странично - ориентирана мутагенеза при N-терминалната кодираща област на зряла фибролаза, за да замести кодоните *QQR" с кодона за "S", използвайки стандартния протокол PCR, което довежда до *preproNAT*, имащ аминокиселинната последователност от SEQ ID No: 3. Олигонуклеотидите, използвани в първичните реакции PCR са изброени по-долу и е показана също така тяхната целева последователност. Отначало бяха осъществени две PCR реакции с използване на олиго 1 и 4 като първична двойка и олиго 2 и 3

камо друга първична двойка, и двете с ДНК от родителския ген като темплейт. Продуктите ДНК от тези две реакции (601 и 815 нуклеотиди в дължината) бяха пречистени с електрофореза на агароза гел и бяха комбинирани, за да служат като темплейт за втория рунд на PCR, като се използва олигон 1 и 2 като първична двойка. Този краен PCR продукт (1372 нуклеотиди в дължината) беше отцепен с ограничителните ендонуклеази - *XhoI* и *NotI*, ензиматично дефосфорилизирани, и депротеинизирани с фенол/хлороформ. Всички следващи стъпки бяха проведени съгласно набора указания на Invitrogen Pichia Expression (Invitrogen Corp., Каталог No K1710-01). Продуктите от свързващата реакция бяха трансформирани в *Escherichia coli* с електропорация и бяха избрани за оцеляване в твърда среда, съдържаща зеоцин. Плазмидът беше изолиран и беше потвърдена областта на профибролазата с ДНК последователност. Плазмидът беше линеализиран чрез отцепване с ограничителната ендонуклеаза *PmeI* и след това трансформиран в *Pichia pastoris* GS115his⁺. Щамът GS115 обикновено е his⁻, така че his⁺ геномът беше възстановен чрез трансформация с източник на ДНК с дивия тип версия на гена *kis4*. Обратно, щамът his⁺ може да се получи от търговската мрежа - от Invitrogen Corp. (X-33 клетъчна линия, Каталог No C180-00). Интегриращи агенти бяха избрани като зеоцин-съпротивителни колонии. Бяха индуцирани канديدат клонове в среди, съдържащи метанол, и хранителната среда беше изпробвана за получаване на NAT при 4-20% PAGE, използвайки оцветяване Coomassie.

Олигонуклеотидите, използвани за странично-ориентирана PCR мутагенеза бяха следните:

Oligo 1 5'-ТАСТАТТGCCAGCATTGCTGC-3' (SEQ ID No: 8)

Oligo2 5'-GCAAATGGCATTCTGACATCC-3' (SEQ ID No: 9)

Oligo 3 5'-TCCAATTAAACTTGACTAAGAGA
TCTTTCCCACAAAGATACGTAC-3' (SEQ ID No: 10)

Oligo 4 5' - GTACGTATCTTTGTGGGAAAGA
TCTCTTAGTCAA GTTTAATTGG-3' (SEQ ID No: 11)

Разположението на тези олигонуклеотиди е показано по-долу във връзка с двойно - усуканата ДНК последователност SEQ ID No: 12, кодираща или имаща смисъл верига, SEQ ID No: 13 допълнителна или антисмислова верига и съответстващата аминокиселинна последователност (SEQ ID No 14) на фибролазата (включително *prepro*-областта) е модифицирана, за да създаде NAT. N-терминалната и C-терминалната области на зрялата фибролаза са посочени, като са подчертани терминалните аминокиселинни последователности (QQRF и LNKP). N-терминалната област QQRF е тази, която е модифицирана да замени QQR с S). За олиго 3 и 4 по-долу са вмъкнати линии с тирета, за да обозначат разположението на изпуснатите кодони, кодиращи остатъците QQ в N-терминалната област на фибролазата.

ATGAGATTTCCTTCAATTTTACTGCTGTTTTATTTCGCAGCATCCTCCGCATTAGCTGCT
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----
 TACTCTAAAGGAAGTTAAAAATGACGACAAAATAAGCGTCGTAGGAGGCGTAAT
 CGACGA

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

CCAGTCAACACTACAACAGAAGATGAAACGGCACAAATTCCGGCTGAAGCTGTCATCGGT
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----
 GGTCAAGTTGTGATGTTGTCTTCTACTTTGCCGTGTTTAAGGCCGACTTCGACAGTAGCCA

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

TACTCAGATTTAGAAGGGGATTTTCGATGTTGCTGTTTTGCCATTTTCCAACAGCACAAAT
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----
 ATGAGTCTAAATCTTCCCCTAAAGCTACAACGACAAAACGGTAAAAGGTTGTCGTGT
 TTA

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

Oligo 1 5'-TACTATTGCCAGCATTGCTGC-3'
 AACGGGTTATTGTTTATAAATACTACTATTGCCAGCATTGCTGCTAAAGAAGAAGGGGTA
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----
 TTGCCCAATAACAAATATTTATGATGATAACGGTCGTAACGACGATTCTTCTTCCCCAT

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

XhoI

TCTCTCGAGAAAAGAGAGGGCTGAAGCTTCTTCTATTATCTTGGAATCTGGTAACGTTAAC
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----
 AGAGAGCTCTTTTCTCTCCGACTTCGAAGAAGATAATAGAACCTTAGACCATTGCAATTG

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

GATTACGAAGTTGTTTATCCAAGAAAGGTCACTCCAGTTCCTAGGGGTGCTGTTCAACCA
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----
 CTAATGCTTCAACAAATAGGTTCTTCCAGTGAGGTCAAGGATCCCCACGACAAGTTGGT

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

AAGTACGAAGATGCCATGCAATACGAATTCAAGGTTAACAGTGAACCAGTTGTCTTGAC
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----
 TTCATGCTTCTACGGTACGTTATGCTTAAGTTCCAATTGTCACCTGGTCAACAGAACGTG

K Y E D A M Q Y E F K V N S E P V V L H

TTGGAAAAAACAAGGTTTGTCTCTGAAGATTACTCTGAAACTCATTACTCCCCAGAT
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----
 AACCTTTTTTTGTTTCCAAACAAGAGACTTCTAATGAGACTTTGAGTAATGAGGGGTCTA

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

GGTAGAGAAATTACTACTTACCCATTGGGTGAAGATCACTGTTACTACCATGGTAGAATC
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----
 CCATCTCTTTAATGATGAATGGGTAACCCACTTCTAGTGACAATGATGGTACCATCTTAG

G R E I T T Y P L G E D H C Y Y H G R I

GAAAACGATGCTGACTCCACTGCTTCTATCTCTGCTTGTAACGGTTTGAAGGGTCATTTCT
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----
 CTTTGTCTACGACTGAGGTGACGAAGATAGAGACGAACATTGCCAAACTTCCCAGT
 AAAG

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

AAGTTGCAAGGTGAAATGTACTTGATTGAACCATTGGAATTGTCCGACTCTGAAGCCCAT
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----
 TTCAACGTTCCACTTTACATGAACTAAGTGGTAACCTTAACAGGCTGAGACTTCGGGTA

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

GCTGTCTACAAGTACGAAAACGTCGAAAAGGAAGATGAAGCCCCAAAGATGTGTGGTGTT
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----
 CGACAGATGTTTCATGCTTTTGCAGCTTTTCTTCTACTTCGGGGTTTCTACACACCACAA

A V Y K Y E N V E K E D E A P K M C G V

Oligo 3

5'- TCCAATTAACTTGACTAAGACCCAAAAGTGGGAATCATATGAACCAATCAAGAAGG
 CCTTCCAATTAACTTGACTAAG

-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----
 TGGGTTTTGACCCTTAGTATACTTGGTTAGTTCTTCCGGAAGGTTAATTTGAACTGATTC
 Oligo 4 3' - GGTTAATTTGAACTGATTC

T Q N W E S Y E P I K K A F Q L N L T K

AGA -----TCTTTCCCAACAAAGATACGTAC-3'
 AGACAACAAAGATTCCCAACAAAGATACGTACAGCTGGTTATCGTTGCTGACCACCGTATG
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----

TCTGTTGTTTCTAAGGGTGTTTCTATGCATGTCGACCAATAGCAACGACTGGTGGCATACT
 TCT-----AGAAAGGGTGTTTCTATGCATG-5'

R O O R E P Q R Y V Q L V I V A D H R M

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

N T K Y N G D S D K I R Q W V H Q I V N

ACCATTAACGAAATCTACAGACCACTGAACATCCAATTCACCTTTGGTTGGTTTGGAAATC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----
TGGTAATTGCTTTAGATGTCTGGTGACTTGTAGGTTAAGTGAAACCAACCAAACCTTTAG

T I N E I Y R P L N I Q F T L V G L E I

TGGTCCAACCAAGATTTGATCACCGTTACTTCTGTATCCCACGACACTCTGGCATCCTTC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----
ACCAGGTTGGTTCTAAACTAGTGGCAATGAAGACATAGGGTGCTGTGAGACCGTAGGAAG

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

GGTAACTGGCGTGAAACCGACCTGCTGCGTCGCCAACGTCATGATAACGCTCAACTGCTG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----
CCATTGACCGCACTTTGGCTGGACGACGCGAGCGGTTGCAGTACTATTGCGAGTTGACGAC

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

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

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

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

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

ACCATGGCACACGAACTGGGTCTATAACCTGGGTATGAACCACGATGGCAACCAGTGTCAC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----
TGGTACCGTGTGCTTGACCCAGTATTGGACCCATACTTGGTGCTACCGTTGGTCACAGTG

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

TGCGGTGCAAACCTCTGTGTTATGGCTGCTATGCTGTCCGATCAACCATCCAAACTGTTTC
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----
 ACGCCACGTTTGAGGACACAATACCGACGATACGACAGGCTAGTTGGTAGGTTTGACAAG

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

TCCGACTGCTCTAAGAAAGACTACCAGACCTTCCTGACCGTTAACAACCCGCAAGTGTATC
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----
 AGGCTGACGAGATTCTTTCTGATGGTCTGGAAGGACTGGCAATTGTTGGGCGTTCACATAG

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

NotI

I

CTGAACAAACCGTAAGCGGCCGCCAGCTTTCTAGAACA AAAACTCATCTCAGAAGAGGAT
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----
 GACTTGTTTGGCATTTCGCCGGCGGTGAAAGATCTTGTTTTTGAGTAGAGTCTTCTCCTA

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

CTGAATAGCGCCGTCGACCATCATCATCATCATCATTGAGTTTGTAGCCTTAGACATGAC
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----
 GACTTATCGCGGCAGCTGGTAGTAGTAGTAGTAGTA ACTCAAACATCGGAATCTGTACTG

L N S A V D H H H H H * V C S L R H D

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

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

AGGATGTCAGAATGCCATTTGCCTGAGAGATGCAGGCTTCATTTTTGATACTTTTTTATT
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----
 TCCTACAGTCTTACGGTAAACGGACTCTCTACGTCCGAAGTAAAAACTATGAAAAAATA
 A CCTACAGTCTTACGGTAAACG-5' Oligo 2

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

Пример 2

Експресия на NAT в *Pichia pastoris*

Когато бяха правени опитите за експресия на ДНК за NAT в *Ешерихия коли*, много слабо се преназвине и изискването за

условия на разреждането намаляват ефективността на очистването. Тези и други предпоставки доведоха до използването на *Pichia pastoris*, вид грозди, като клетка-домакин. Културата на избрани клонове на *Pichia pastoris*, която беше прехвърлена с ретро NAT cDNA (SEQ ID No: 4) бе добавена в 500 ml на следващата хранителна среда:

На литър от партидата от хранителната среда

Екстракт от грозди	30.0 g
Калиев фосфат двуосновен	17.2 g
Глюкоза	20.0 g
Биотин	0.004 g
Вода	около 1 литър
Фосфорна киселина, 85%	за нагласяване на pH = 6.00

Прехвърлените клетки на *P. pastoris* бяха инкубирани при 30 °C в размесител около 30 до 32 часа. Около 1% (w/v) от получената култура бе използван на прибавяне към 10-литров ферментатор. Той съдържа стерилизирани основни соли и глюкоза (по-долу). Дванадесет милилитра на литър от соли РТМ4 (РТМ4 е разтвор на трасиращи метали, съдържащ меден сулфат пентахидрат, натриев йодид, манганов сулфат монохидрат, натриев молибдат дихидрат, борна киселина, кобалтов хлорид хексахидрат, цинков хлорид, железен сулфат хептахидрат, d-биотин, сярна киселина, пречистена вода) бяха добавени към литър от партидата на хранителната среда след стерилизация във ферментатора. Ферментационната температура бе 30° C. pH на ферментатора

беше контролирано с амониев хидроокис и фосфорна киселина до стойност 6.00. Цинк от цинковите соли, добавен към хранителната среда се включва в NAT като част от структурата на металопротеиназата.

Основни соли от партидата на хранителната среда

Фосфорна киселина, 85%	26.7 ml
Калциев сулфат	0.93 g
Калиев сулфат	18.2 g
Магнезиев сулфат - 7H ₂ O	14.9 g
Калиев хидроокис	4.13 g
Глюкоза	30.0 g
Вода	на 1 литър

Партидата на културата бе оставена да расте до пълното изконсумиране на глюкозата (от 17 до 20 часа). Тогава започна подхранващата фаза. Хранителната среда съдържаше глюкоза и 12 ml PTM4 соли на литър. Индуцираното подхранване съдържаше глюкоза, 25% метанол и 12 ml PTM4 соли на литър. При това подхранване температурата на реактора бе установена на 20 °C. Индуцираната фаза продължи 60 - 75 часа. Третираната среда бе обработена и останалият клетъчен материал бе изхвърлен.

Пример 3

Пречистване на NAT от *Pichia pastoris*

Субстратът от грозди (обработената среда без остатъчния клетъчен материал) от Пример 2 беше избистрен и рН и проводи-

мостта бяха нагласени съответно на 6.5 и 10-20 mS/cm. Хранителната среда беше изсипана върху неподвижна смола с метален афинитет, заредена с мед (Cu) и уравновесена с фосфатен буферизиран солен физиологичен разтвор (PBS). Смолата беше промита с (PBS) и разрежена със имидазол с градиент (0-100 mM) в (PBS). Части, съдържащи "зрял" NAT (SEQ ID No:1) бяха поставени в резервоар и бяха разреждани докато проводимостта стана под 1.5 mS/cm, а pH = 6.4. Разрежданият състав бе изсипан върху смола SP Sepharose (Amersham Pharmacia Biotech, Inc., Пискамауей, Ню Джърси), която е била изравнена с 10 mM 2-(N-morpholino) етансулфонова киселина (MES). Колоната бе промита с MES и разрежена с NaCl с градиент (0-500 mM) в MES. Частите, съдържащи NAT, бяха събрани и съхранени.

Пример 4

Тромболиза при остра форма на тромбоза у плъх

Сънна артерия: Сравнение на NAT с урокиназа

За да се демонстрира, че "зрелият" NAT (SEQ ID No: 1) е биологично активен и уникален функционално, бяха проведени старателни фармакологични изследвания с плъхове, на които бяха причинени огнищни рани чрез аноден ток на една от сънните артерии. Тази рана създава оклузен тромб, който обикновено се образува за петнадесет минути. След образуването на тромба, артерията бе наблюдавана в продължение на тридесет минути, за да се уверим, че каротидната оклузия е стабилна. Тогава бяха

предписани интравенозно хепарин и аспирин, за да се предотврати по-нататъшното движение на тромба. След това животните бяха лекувани с вътрешно-артериална инфузия на тестовия материал. Кръвният поток през сънната артерия бе следен по време на внасянето на тестовия материал, така че можеше да се установи сполучливото разтваряне на съсирека и времето, когато е започнало разтварянето на съсирека. Процентът на експериментите, когато започваше разтваряне на съсирека, бе отбелязан и общите средства бяха изчислявани само за онези експерименти, където разтварянето на съсирека бе сполучливо. Като мярка на хеморагичния потенциал на тестовия материал, цялата отделена кръв при хирургичната намеса бе събрана с марлени тампони. Пробите с тампони бяха поставени в разтвор на детергент, за разтваряне на червените кръвни клетки и за отделяне на хемоглобина, количество на който след това беше определено спектрофотометрично. Събраният хемоглобин бе използван за изчисляване обема на изгубената кръв. Данните от теста са показани в таблицата по-долу.

Тзи проучвания установяват, че NAT е биологично активен в животински модел при разтваряне на съсирек *ин vivo*. Освен това разтварянето на съсирек е постигнато при забележимо малко време и с по-малка кръвна загуба от мястото на хирургическата намеса, в сравнение с урокиназата. По този начин, профилът на активност на NAT може да се различи от плазминогенния активаторен клас на тромболитичните агенти (представени от

урокиназата), при която разтварянето на съсирек с NAT става по-бързо и с понижени хеморагични усложнения.

Таблица 1

Начало на разтваряне на съсирека, време на разтварянето му и загуба на кръв при хирургическата намеса (Средни и стандартни отклонения).

	Начало на разтварянето (%)	Време на разтварянето (минути)	Загуба на кръв (ml)
Физиологичен разтвор (n=6)	0 % (0 от 6)	N/A	0.10 ± 0.23
Урокиназа 25 U/min (n=15)	33% (5 от 15)	55.3 ± 15.9	1.06 ± 1.59
Урокиназа 250 U/min (n=15)	86 % (13 от 15)	33.5 ± 15.3	1.43 ± 1.45
NAT 2 mg (n=14)	78% (11 от 14)	6.3 ± 5.8	0.96 ± 0.77

Фибринолитичната активност на NAT е сравнима с тази на фибролазата. Освен това, както беше споменато по-горе, стабилността на N-терминала на NAT има в резултат по-хомогенен краен продукт при рекомбинантна експресия, което е явно преимущество (например, N-терминалът няма да се промени след време, в резултат на смес от различни форми, така правейки полипептида по-стабилен).

ПРЕТЕНЦИИ

1. Фибринолитично активен полипептид, който има аминокиселинна последователност SEQ ID No: 1.
2. Полипептидът от Претенция 1, който е бил получен от рекомбинатна експресия в дрожди.
3. Молекула на нуклеинова киселина, кодираща полипептида от SEQ ID No: 1.
4. Молекулата на нуклеиновата киселина от Претенция 3, която има последователност SEQ ID No: 4.
5. Експресивен носител, съдържащ експресивни регулиращи елементи, оперативно свързан с молекула на нуклеинова киселина според Претенция 3.
6. Експресивен носител според претенция 5, в който молекулата на нуклеиновата киселина има нуклеотидна последователност SEQ ID No: 4.
7. Клетка-домакин, трансформирана или преобразувана молекула на нуклеиновата киселина от Претенция 3.

8. Трансформирана или преобразувана клетка-домакин според Претенция 7, която е прокариотна или евкариотна клетка.
9. Трансформирана или преобразувана прокариотна клетка според Претенция 8, която е клетка от дрожди.
10. Трансформирана или преобразувана клетка от дрожди според Претенция 9, която е *Pichia pastoris*.
11. Методът за получаване на NAT, включващ клетки-домакини от култура, съдържащ ДНК молекула, кодираща металопротеиназата, оперативно свързана с регулиращите елементи за стимулиране експресията на същата при условие, че молекулата на ДНК е експресирана, и изолиране на експресираната металопротеиназа от клетката-домакин на културата.
12. Методът от Претенция 11, в която молекулата на ДНК е от SEQ ID No: 4.
13. Методът от Претенция 11, в която клетките-домакини са клетки от дрожди.

14. Методът от Претенция 13, в който грождевите клетки са на *Pichia pastoris*.
15. Метод за получаване на металопротеиназа фибринолитичен агент, избран от групата, състояща се от фибролаза и NAT, съдържаща клетки-домакини от културата *Pichia*, съдържащи ДНК молекула, кодираща въпросната металопротеиназа, свързана оперативно с регулиращите елементи за стимулиране експресията на същата при условие, че молекулата на ДНК е експресирана и изолиране на експресираната металопротеиназа, от клетката-домакин на културата.
16. Методът от Претенция 15, в който клетката-домакин *Pichia* е *Pichia pastoris*.
17. Методът за лечението на тромбоза в бозайник, включващ предписване на тромболитично ефективно количество NAT за бозайника.
18. Антитяло срещу полипептида на Претенция 1.
19. Антитялото от Претенция 17, което е многокласно антитяло.

SEQUENCE LISTING

<110> Amgen Inc.

<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 atcgttgctg accaccgtat gaacactaaa 60
tacaacggtg actctgacaa aatccgtcaa tgggtgcacc aaatcgtcaa caccattaac 120
gaaatctaca gaccactgaa catccaattc actttgggtg gtttggaat ctggtccaac 180
caagatttga tcaccgttac ttctgtatcc cagcacctc tggcatcctt cggtaactgg 240
cgtgaaaccg acctgctgcg tcgccaacgt catgataacg ctcaactgct gaccgctatc 300
gacttcgacg gtgatactgt tggctctggc tacgttggtg gcatgtgtca actgaaacat 360
tctactgggtg ttatccagga ccactccgct attaacctgc tggttgctct gaccatggca 420
cacgaactgg gtcataacct gggtatgaac cacgatggca accagtgtca ctgcggtgca 480
aactcctgtg ttatggctgc tatgctgtcc gatcaaccat ccaaactgtt ctccgactgc 540
tctaagaaag actaccagac cttcctgacc gtaacaacc 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
```

```
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

<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 attagctgct 60
 ccagtcaaca ctacaacaga agatgaaacg gcacaaattc 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 ctagggggtgc tgttcaacca 360
 aagtacgaag atgccatgca atacgaattc aagggttaaca gtgaaccagt tgtcttgac 420
 ttggaaaaaa acaaagggtt gttctctgaa gattactctg aaactcatta ctcccagat 480
 ggtagagaaa ttactactta cccattgggt gaagatcact gttactacca tggtagaatc 540
 gaaaacgatg ctgactccac tgcttctatc tctgcttgta acgggttgaa gggtcatttc 600
 aagttgcaag gtgaaatgta cttgattgaa ccattggaat tgtccgactc tgaagcccat 660
 gctgtctaca agtacgaaaa cgtcgaaaag gaagatgaag ccccaaagat gtgtgggtgt 720
 acccaaaact gggaatcata tgaaccaatc aagaaggcct tccaattaaa cttgactaag 780
 agatctttcc cacaagata cgtacagctg gttatcgttg ctgaccaccg tatgaacact 840
 aaatacaacg gtgactctga caaatccgt caatgggtgc accaaatcgt caacaccatt 900
 aacgaaatct acagaccact gaacatccaa ttcactttgg ttgggttgga aatctgggtc 960
 aaccaagatt tgatcacctg tacttctgta tcccacgaca ctctggcatc cttcggtaac 1020
 tggcgtgaaa ccgacctgct gcgtcgccaa cgctcatgata acgctcaact gctgaccgct 1080
 atcgacttcg acgggtgatac tggttggtctg gcttacgttg gtggcatgtg tcaactgaaa 1140
 cattctactg gtgttatcca ggaccactcc gctattaacc tgctgggtgc tctgaccatg 1200
 gcacacgaac tgggtcataa cctgggtatg aaccacgatg gcaaccagtg tcaactgcgg 1260
 gcaaactcct gtgttatggc tgctatgctg tccgatcaac catccaaact gttctccgac 1320
 tgctctaaga aagactacca gaccttcctg accgttaaca acccgcagtg tatcctgaac 1380
 aaaccg 1386

<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
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
 actaaataca acggtgactc tgacaaaatc cgtcaatggg tgcaccaa at cgtcaacacc 120
 attaacgaaa tctacagacc actgaacatc caattcactt tgggttggtt ggaaatctgg 180
 tccaaccaag atttgatcac cgttacttct gtatcccacg acactctggc atccttcggg 240
 aactggcggtg aaaccgacct gctgcgtcgc caacgctcatg ataacgctca actgctgacc 300
 gctatcgact tcgacgggtga tactgttggt ctggcttacg ttggtggcat gtgtcaactg 360
 aaacattcta ctgggtgttat ccaggaccac tccgctatta acctgctggg tgctctgacc 420
 atggcacacg aactgggtca taacctgggt atgaaccacg atggcaacca gtgtcactgc 480
 ggtgcaaact cctgtgttat ggctgctatg ctgtccgatc aaccatccaa actgttctcc 540
 gactgctcta agaaagacta ccagaccttc ctgaccgtta acaaccgcga gtgtatcctg 600
 aacaaaccg 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 gcacaaattc cggctgaagc tgtcatcggt 120

tactcagatt	tagaagggga	tttcgatggt	gctgttttgc	cattttccaa	cagcacaaat	180
aacgggttat	tgtttataaa	tactactatt	gccagcattg	ctgctaaaga	agaaggggta	240
tctctcgaga	aaagagaggc	tgaagcttct	tctattatct	tggaatctgg	taacgttaac	300
gattacgaag	ttgtttatcc	aagaaaggte	actccagttc	ctaggggtgc	tggtcaacca	360
aagtacgaag	atgccatgca	atacgaatcc	aagggttaaca	gtgaaccagt	tgtcttgcac	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	cccaaagat	gtgtgggtgt	720
acccaaaact	gggaatcata	tgaaccaatc	aagaaggcct	tccaattaaa	cttgactaag	780
agacaacaaa	gattcccaca	aagatacgta	cagctgggta	tcgttgctga	ccaccgtatg	840
aacactaaat	acaacgggtga	ctctgacaaa	atccgtcaat	gggtgcacca	aatcgtcaac	900
accattaacg	aaatctacag	accactgaac	atccaattca	ctttgggtgg	tttggaatc	960
tggtccaacc	aagatttgat	caccgttact	tctgtatccc	acgacactct	ggcatccttc	1020
ggtaactggc	gtgaaaccga	cctgctgcgt	cgccaacgtc	atgataacgc	tcaactgctg	1080
accgctatcg	acttcgacgg	tgatactggt	ggctgggctt	acgttggtgg	catgtgtcaa	1140
ctgaaacatt	ctactgggtg	tatccaggac	cactccgcta	ttaacctgct	ggttgctctg	1200
accatggcac	acgaactggg	tcataacctg	ggatgaacc	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 caciaagata 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))
<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 ttatttcgag catcctccgc attagctgct 60
ccagtcaaca ctacaacaga agatgaaacg gcacaaattc cggctgaagc tgtcatcggt 120
tactcagatt tagaagggga ttctgatggt gctgttttgc cattttccaa cagcacaaat 180
aacgggttat tgtttataaa tactactatt gccagcattg ctgtataaga agaaggggta 240
tctctcgaga aaagagagggc tgaagcttct tctattatct tggaatctgg taacgttaac 300
gattacgaag ttgtttatcc aagaaagggtc actccagttc ctaggggtgc tgttcaacca 360
aagtacgaag atgccatgca atacgaattc aagggttaaca gtgaaccagt tgtcttgac 420
ttggaaaaaa acaaagggtt gttctctgaa gattactctg aaactcatta ctcccagat 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
agacaacaaa gattcccaca aagatacgta cagctgggta tcgttgctga ccaccgtatg 840
aacactaat acaacgggtga ctctgacaaa atccgtcaat ggggtgcacca aatcgtcaac 900
accattaacg aaatctacag accactgaac atccaattca ctttggttgg tttggaaatc 960
tggteccaacc aagatttggat caccgttact tctgtatccc acgacactct ggcacacctc 1020
ggtaactggc gtgaaaccga cctgctgcgt cgccaacgtc atgataacgc tcaactgctg 1080
accgctatcg acttcgacgg tgatactggt ggtctggctt acgttgggtg catgtgtcaa 1140
ctgaaacatt ctactggtgt tatccaggac cactccgcta ttaacctgct gggtgctctg 1200
accatggcac acgaactggg tcataaacctg ggtatgaacc acgatggcaa ccagtgtcac 1260
tgcgggtgcaa actcctgtgt tatggctgct atgctgtccg atcaaccatc caaactgttc 1320
tccgactgct ctaagaaaga ctaccagacc ttcctgaccg ttaacaacct gcagtgtatc 1380
ctgaacaaac cgtaagcggc cgccagcttt ctagaacaaa aactcatctc agaagaggat 1440
ctgaatagcg ccgtcgacca tcatcatcat catcattgag tttgtagcct tagacatgac 1500
tgttcctcag ttcaagttgg gcacttacga gaagaccggc 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 tcgtaagtgc ccaacttgaa ctgaggaaca 120
gtcatgtcta aggcatacaa ctcaatgatg atgatgatga tggtcgacgg cgctattcag 180
atcctcttct gagatgagtt tttgttctag aaagctggcg gccgcttacg gtttgttcag 240
gatacactgc gggttgttaa cggtcaggaa ggtctggtag tctttcttag agcagtcgga 300
gaacagtttg gatggttgat cggacagcat agcagccata acacaggagt ttgcaccgca 360
gtgacactgg ttgccatcgt ggttcatacc caggttatga cccagttcgt gtgccatggg 420
cagagcaacc agcagggttaa tagcggagtg gtcctggata acaccagtag aatgtttcag 480
ttgacacatg ccaccaacgt aagccagacc aacagtatca ccgtcgaagt cgatagcggt 540
cagcagttga gcgttateat gacgttgagg acgcagcagg tcggtttcac gccagttacc 600
gaaggatgcc agagtgtcgt gggatacaga agtaacgggt atcaaatctt ggttggaaca 660
gatttccaaa ccaaccaaag tgaattggat gttcagtggt ctgtagattt cgtaaatggg 720
gttgacgatt tgggtgcacc attgacggat tttgtcagag tcaccgttgt atttagtgtt 780
catacgggtg tcagcaacga taaccagctg tacgtatctt tgtgggaatc tttgttgtct 840
cttagtcaag ttttaattgga aggccttctt gattgggtca tatgattccc agttttgggt 900
aacaccacac atctttgggg cttcatcttc cttttcgacg ttttcgtact tgtagacagc 960
atgggcttca gagtcggaca 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 gcacccttag gaactggagt gaccttctt ggataaaca cttcgtaatc 1320
gttaacgtta ccagattcca agataataga agaagcttca gcctctctt tctcgagaga 1380
tacccttctt tcttttagcag caatgctggc aatagtagta tttataaaca ataaccggtt 1440
atttgtgctg ttggaaaatg gcaaacagc aacatcgaaa tccccttcta aatctgagta 1500
accgatgaca gtttcagccg gaatttgtgc cgtttcatct tctgtttagt tgttgactgg 1560
agcagctaata gcgaggagtg 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
 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

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
 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
 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 cgtaaagat tacgaagttg tttatccaag 120
aaaggctact ccagttccta ggggtgctgt tcaaccaaag tacgaagatg ccatgcaata 180
cgaattcaag gttaacagtg aaccagtgtg cttgcacttg gaaaaaaca aagggttgtt 240
ctctgaagat tactctgaaa ctctactctc 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 tgggtgtacc caaaactggg aatcatatga 540
accaatcaag aaggccttcc aattaaactt gactaagaga tctttccac aaagatacgt 600
ac 602

<210> 19
<211> 816
<212> DNA
<213> Agkistrodon contortrix

<220>
<223> Fragment of fibrolase of Agkistrodon contortrix

<400> 19
tccaattaaa cttgactaag agatctttcc cacaaagata cgtacagctg gttatcggtg 60
ctgaccaccg tatgaacact aaatacaacg gtgactctga caaaatccgt caatgggtgc 120
accaaatcgt caacaccatt aacgaaatct acagaccact gaacatccaa ttcacttttg 180
ttgggtttgga aatctggtec aaccaagatt tgatcaccgt tacttctgta tcccacgaca 240
ctctggcatc cttcggtaac tggcgtgaaa ccgacctgct gcgtcgccaa cgatcatgata 300
acgtcaact gctgaccgct atcgacttcg acggtgatac tgggtgtctg gcttacgttg 360
gtggcatgtg tcaactgaaa cattctactg gtgttatcca ggaccactcc gctattaacc 420
tgctgggttg tctgaccatg gcacacgaac tgggtcataa cctgggtatg aaccacgatg 480
gcaaccagtg tcaactgcgt gcaaactcct gtgttatggc tgctatgctg tccgatcaac 540
catccaaact gttctccgac tgctctaaga aagactacca gaccttcctg accgttaaca 600
acccgcagtg tatcctgaac aaaccgtaag cggccgccag ctttctagaa caaaaactca 660
tctcagaaga ggatctgaat agcgccgtcg accatcatca tcatcatcat tgagtttgta 720
gccttagaca tgactgttcc tcagttcaag ttgggcactt acgagaagac cgtcttctgct 780

<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 cgttaacgat tacgaagttg tttatccaag 120
aaaggtcact ccagttccta ggggtgctgt tcaaccaaag tacgaagatg ccatgcaata 180
cgaattcaag gttaacagtg aaccagttgt cttgcacttg gaaaaaaaca aagggtttgtt 240
ctctgaagat tactctgaaa ctcttactc cccagatggt 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
acagctgggt atcgttgctg accaccgtat gaacactaaa tacaacggtg actctgacaa 660
aatccgtcaa tgggtgcacc aaatcgtaa caccattaac gaaatctaca gaccactgaa 720
catccaattc actttgggtg gtttggaat ctggtccaac caagatttga tcaccgttac 780
ttctgtatcc cagcactc tggcatcctt cggtaaactg cgtaaaaccg acctgctgcg 840
tcgccaacgt catgataacg ctcaactgct gaccgctatc gacttcgacg gtgatactgt 900
tgggtctggct tacgttgggt gcatgtgtca actgaaacat tctactggtg 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 gttttagtgc ttagacatga ctgttctca gttcaagttg ggcacttacg 1320
agaagaccgg ttttgc taga ttctaataca gaggatgtca gaatgccatt tgc 1373
```