

METHOD FOR PRODUCTION OF RECOMBINANT HUMAN THROMBIN '644

TECHNICAL FIELD

The present application relates to a method for producing recombinant human thrombin
5 from recombinant prothrombin using recombinant ecarin.

BACKGROUND OF THE INVENTION

Thrombin is a key enzyme in the coagulation cascade. By thrombin mediated proteolytic digestion of fibrinogen into fibrin monomer, a cascade reaction leading to clot formation is
10 started. Clot formation is the first step in wound healing. In addition thrombin is a chemo attractant to cells involved in wound healing, and, the fibrin network formed act as a scaffold for collagen-producing fibroblasts, increases phagocytosis, promotes angiogenesis and binds growth factors thus further supporting the healing process. The rate of clot formation is dependent on the concentration of thrombin and fibrinogen. Because of the
15 important function in clot formation thrombin has been utilised in a number of products intended for haemostasis and/or as tissue sealants or "glues", both as stand-alone products (i.e. Thrombin-JMI) or in combination with fibrin or other compounds (i.e. Tisseel, Hemaseel, Crosseal). The potential fields of use are numerous; skin grafting, neuro surgery, cardiac surgery, toracic surgery, vascular surgery, oncologic surgery, plastic
20 surgery, ophthalmologic surgery, orthopedic surgery, trauma surgery, head and neck surgery, gynecologic and urologic surgery, gastrointestinal surgery, dental surgery, drug delivery, tissue engineering and dental cavity haemostasis.

So far the thrombin in approved thrombin-containing products on the market is derived
25 either from human or bovine plasma. Using plasma derived protein confers several disadvantages as limited availability and safety concerns such as risk for transmission of viruses and prions and the risk of triggering autoantibody formation (bovine products). Cases where antibody formation due to bovine thrombin exposure has lead to significant bleeding disorders are known.

In vivo thrombin is obtained from activation of prothrombin through the coagulation cascade. Activation through the coagulation cascade is dependent on the presence of a functional GLA-domain containing 8-10 glutamic residues converted to gamma-carboxyglutamate. *In vitro*, also incomplete gamma-carboxylated prothrombin can be converted to thrombin by the use of prothrombin activators such as ecarin. Ecarin, a snake venom derived from the Kenyan viper *Echis carinatus* is a procoagulant, a protease which cleaves human prothrombin between residues Arg₃₂₀-Ile₃₂₁ to generate meizothrombin. Further autocatalytic processing results in the formation of meizothrombin desF1 and then alpha-thrombin, which is the mature active form of thrombin.

An ideal commercial thrombin manufacturing process would use a recombinant thrombin precursor and a recombinant protease produced at high productivity without addition of animal-derived components. Further requirements would be robust performance, convenience and low cost.

A big obstacle for efficient recombinant human thrombin (rh-thrombin) has been to obtain high yields of prothrombin. Although extensive efforts have been spent, obtaining high yields of prothrombin under conditions suitable for production of biologicals has long remained a challenge. Yonemura et al. (J Biochem 135:577-582, 2004) have used recombinant GLA-domain-less prethrombin digested with recombinant ecarin to generate recombinant human thrombin. The productivity of prethrombin at process scale was 150-200 mg/L, which is a modest productivity for commercial scale production. Recombinant production of ecarin has also been described in WO 01/04146. In this publication generation of rh-thrombin is exemplified by conversion of recombinant prothrombin produced in COS cells by a recombinant ecarin produced from CHO cells. However, the exemplified methods are not suitable for large-scale production and animal-derived components are used.

Recombinant ecarin is produced as a prepro-protein that needs to be activated. Problems to efficiently activate the r-ecarin are described in both publications and the suggested activation procedures are far from optimal.

Thus there is a need for improved methods to obtain recombinant human thrombin. During our efforts to obtain improved productivity of gamma-carboxylated human prothrombin we made the surprising discovery that co-expression with gamma-glutamyl carboxylase (GGCX) vastly improved also the productivity of incompletely carboxylated prothrombin
5 (see WO2005038019).

The present invention describes a process to efficiently produce human thrombin from recombinant prothrombin obtained by the expression method as described in WO2005038019. Recombinant carboxylated or incompletely carboxylated prothrombin
10 combined with recombinant ecarin has not previously been used for manufacturing of recombinant thrombin. Further, the procedure for activating recombinant ecarin is new. The methods described would be suitable for large scale rh-thrombin manufacturing without the addition of animal-derived components.

15 SUMMARY OF THE INVENTION

According to a first aspect of the invention, a method is provided for producing recombinant human thrombin from recombinant prothrombin using recombinant ecarin having the sequence SEQ ID NO 2 or a homologue thereof.

According to a another aspect, a pharmaceutical composition is provided comprising a
20 recombinant thrombin according to said method, in combination with pharmaceutically acceptable carriers, vehicles and/or adjuvants.

According to further aspect, an isolated DNA sequence is provided coding for recombinant ecarin according to SEQ ID NO 2 or a homologue thereof, having at least 80 % identity to SEQ ID NO 2.

25 According to another aspect, a vector is provided comprising an isolated DNA sequence coding for recombinant ecarin according to SEQ ID NO 2 or a homologue thereof, having at least 80 % identity to SEQ ID NO 2.

According to yet another aspect, a cell line is provided comprising a vector comprising an isolated DNA sequence coding for recombinant ecarin according to SEQ ID NO 2 or a homologue thereof, having at least 80 % identity to SEQ ID NO 2.

5 BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1. FII+GGCX construct

Figure 2. Ecarin construct

Figure 3. Example of a process outline for thrombin manufacturing

10 Figure 4. Nucleotide sequence alignment of recombinant ecarin used in the present invention and wild type ecarin

Figure 5. Amino acid alignment of recombinant ecarin used in the present invention and wild type ecarin

Figure 6. Graph showing the activation of recombinant ecarin during cell death over time

15 Figure 7. Activation of recombinant ecarin in cell cultures over time, assayed by SDS-PAGE

Figure 8. Chromatogram from CIEX purification of rh-thrombin.

Figure 9. Non-reduced SDS-PAGE analyses of fractions obtained by CIEX purification

Detailed description of the invention

20 The invention consists in one part of a cell line derived by stable transfection with a vector (Figure 1) encoding human prothrombin (FII) associated by suitable control sequences and human gamma-glutamyl carboxylase (GGCX) associated by suitable control sequences. Control sequences should be chosen so that prothrombin expression is in excess of the GGCX expression by at least a factor of 10. The host cell is preferably a eukaryotic cell.

25 Typical host cells include, but are not limited to insect cells, yeast cells, and mammalian cells. Mammalian cells are particularly preferred. Suitable mammalian cells lines include, but are not limited to, CHO, HEK, NS0, 293, Per C.6, BHK and COS cells, and derivatives thereof. In one embodiment the host cell is the mammalian cell line CHO-S. The obtained prothrombin producing cell line is grown under culture conditions optimised for high yield

30 of prothrombin disregarding gamma-carboxylation. Vitamin K may or may not be added to the growth medium.

It will be appreciated that the invention is not restricted to a particular prothrombin or gamma-glutamyl carboxylase or protein encoding sequence of one of these proteins to be co-expressed. Moreover, and in particular with respect to blood coagulation factors, numerous mutant forms of the proteins have been disclosed in the art. The present invention is equally applicable to prothrombin and gamma-glutamyl carboxylase mutant forms, including naturally occurring allelic variants, of the proteins as it is to wild-type sequence. In one embodiment the invention can be undertaken with any wild-type protein or one with at least 90%, preferably at least 95% sequence identity thereto. In another embodiment, sequences listed in Table 1 can be used.

Table 1

PROTEIN	CDNA EMBL ACC#	SPLICE VARIANTS (PROTEIN)	MUTATIONS	GENE EMBL ACC#
Glutamate gamma carboxylase	BC013979	2; BC013979; AF253530	1 SNP (EMBL# U65896); 2 SNPs (OMIM# 137167)	U65896
Prothrombin	V00595	1; V00595	approx. 100 SNP's (EMBL# AF478696)	AF478696

Each of these proteins, including their nucleic acid and amino acid sequences, are well known. Table 2 identifies representative sequences of wild-type and mutant forms of the various proteins that can be used in the present invention.

The term "gamma-glutamyl carboxylase" or "GGCX", as used herein, refers to a vitamin K dependent enzyme that catalyses carboxylation of glutamic acid residues.

20

GGCX enzymes are widely distributed, and have been cloned from many different species such as the beluga whale *Delphinapterus leucas*, the toadfish *Opsanus tau*, chicken (*Gallus gallus*), hagfish (*Myxine glutinosa*), horseshoe crab (*Limulus polyphemus*), and the cone snail *Conus textile* (Begley et al., 2000, *ibid*; Bandyopadhyay et al. 2002, *ibid*). The carboxylase from conus snail is similar to bovine carboxylase and has been expressed in

25

COS cells (Czerwicz et al. 2002, *ibid*). Additional proteins similar to GGCX can be found in insects and prokaryotes such as *Anopheles gambiae*, *Drosophila melanogaster* and *Leptospira* with NCBI accession numbers: gi 31217234, gi 21298685, gi 24216281, gi 24197548 and (Bandyopadhyay et al., 2002, *ibid*), respectively. The carboxylase enzyme displays remarkable evolutionary conservation. Several of the non-human enzymes have shown, or may be predicted to have, activity similar to that of the human GGCX we have used, and may therefore be used as an alternative to the human enzyme.

Table 2 identifies representative sequences of predicted proteins homologous to human GGXC (sorted after species origin) that can be used in the present invention.

Table 2

Species	Data base accession #/ID
<i>Homo sapiens</i> (man)	NM_000821.2 HUMGLUCARB HUMHGCA BC004422 HSU65896 AF253530.1
<i>Papio hamadryas</i> (red baboon)	AC116665.1
<i>Delphinapterus leucas</i> (white whale)	AF278713
<i>Bos taurus</i> (bovine)	NM_174066.2 BOVCARBOXG BOVBGCA
<i>Ovis aries</i> (domestic sheep)	AF312035
<i>Rattus norvegicus</i> (brown rat)	NM_031756.1 AF065387
<i>Mus musculus</i> (mouse)	NM_019802.1 AF087938
<i>Opsanus tau</i> (bony fishes)	AF278714.1
<i>Conus textile</i> (molluscs)	AY0044904.1

	AF382823.2
<i>Conus imperialis</i> (molluscs)	AF448234.1
<i>Conus episcopatus</i> (molluscs)	AF448233.1
<i>Conus omaria</i> (molluscs)	AF448235.1
<i>Drosophila melanogaster</i> (fruit fly)	NM_079161.2
<i>Anopheles gambiae</i> (mosquito)	XM_316389.1
<i>Secale cereale</i> (monocots)	SCE314767
<i>Triticum aestivum</i> (common wheat)	AF280606.1
<i>Triticum urartu</i> (monocots)	AY245579.1
<i>Hordeum vulgare</i> (barley)	BLYHORDCA
<i>Leptospira interrogans</i> (spirochetes)	AE011514.1
<i>Streptomyces coelicolor</i> (high GC Gram+ bacteria)	SCO939109 SCO939124 AF425987.1
<i>Streptomyces lividans</i> (high GC Gram+ bacteria)	SLU22894
<i>Streptomyces virginiae</i> (high GC Gram+ bacteria)	SVSNBDE
<i>Micrococcus luteus</i> (high GC Gram+ bacteria)	MLSPCOPER
<i>Chlamydomonas reinhardtii</i> (green algae)	AF479588.1
<i>Dictyostelium discoideum</i> (slime mold)	AC115612.2
<i>Coturnix coturnix</i> (birds)	AF364329.1
<i>Bradyrhizobium japonicum</i> (α -protoeobacteria)	AP005937.1
<i>Rhodobacter sphaeroides</i> (α -proteobacteria)	RSY14197
<i>Sinorhizobium meliloti</i> (α -proteobacteria)	RME603647 AF119834
<i>Mesorhizobium loti</i> (α -proteobacteria)	AP003014.2
<i>Chromobacterium violaceum</i> (β -proteobacteria)	AE016910.1 AE016918.1
<i>Pseudomonas aeruginosa</i> (γ -proteobacteria)	AE004613.1 AF165882
<i>Xanthomonas axonopodis</i> (γ -proteobacteria)	AE011706.1
<i>Human herpesvirus 8</i>	KSU52064

	KSU75698
	AF305694
	AF360120
	AF192756

Each of the above-identified GGCX proteins and GGCX proteins from other species can be used as the carboxylase enzyme in the present invention.

5

A second part of the invention is a cell line stably transfected with a polynucleotide encoding ecarin and associated control elements (Figure 2). The ecarin encoding sequence may be optimised for expression in mammalian cells, but is not limited to such sequences. In one embodiment of the invention the sequence according to SEQ ID NO 2 or a
10 homologue thereof is used to express ecarin. A homologue of SEQ ID NO 2 coding for ecarin may have at least 80%, 85%, 90%, 95%, 97%, 98% or 99% identity to the sequence SEQ ID NO 2. The host cell is preferably a eukaryotic cell. Typical host cells include, but are not limited to insect cells, yeast cells, and mammalian cells. Mammalian cells are particularly preferred. Suitable mammalian cells lines include, but are not limited to, CHO,
15 HEK, NS0, 293, Per C.6, BHK and COS cells, and derivatives thereof. In one embodiment the host cell is the mammalian cell line CHO-S.

In one embodiment prothrombin and ecarin are produced from cells originating from the same parent cell line. This cell line origin may be, but is not limited to, Chinese Hamster
20 Ovary cells (CHO) including derivatives and NS0 (myeloma BALB/c mouse) including derivatives. The purpose of using the same cell line background is to facilitate purification and evaluation of purity of the thrombin product.

In another embodiment ecarin and prothrombin are produced from different host cell line;
25 i.e. CHO and NS0, respectively.

In one aspect of the invention use of recombinant ecarin is preferred as this facilitates detection of non-thrombin product derived components during the thrombin generation

process and in the final thrombin product. In a second aspect recombinant ecarin is preferred due to reduced risk for exposure to allergenic or toxic components that may be present in ecarin derived from snake venom. In a third aspect ecarin from snake venom is not preferred due to batch variation and limited batch size of ecarin preparations.

5

The crude prothrombin and the crude ecarin are mixed and incubated under conditions that allow formation of thrombin, such as described in Example 3. Generated thrombin is then purified by methods described in Example 4 or by other methods known by persons skilled in the art. Alternatively prothrombin and/or ecarin can first be purified by methods known
10 in the art and then mixed to obtain thrombin. Thrombin is then purified from non-product components.

An example of a suitable thrombin manufacturing process is outlined in Figure 3.

A method is provided for producing recombinant human thrombin from recombinant
15 prothrombin using recombinant ecarin having the sequence SEQ ID NO 2 or a homologue thereof. The recombinant ecarin can be expressed and secreted by a cell containing the gene comprising the nucleotide sequence SEQ ID NO 2 or a homologue thereof in CHO-S cells, which ecarin has an amino acid sequence equal to that of wild type ecarin.

In the above method the recombinant protrombin is subjected to recombinant ecarin, which
20 recombinant ecarin can be isolated in active form after extra-cellular expression by CHO-S cells, said cells being left to apoptosis/necrosis for a time sufficient to activate said ecarin, whereupon a human recombinant thrombin is isolated.

The recombinant prothrombin can be produced by a cell-line comprising a prothrombin expressing gene having a nucleotide sequence comprising the sequence SEQ. ID. NO. 1 or
25 an homologue thereof. A homologue of SEQ ID NO 1 coding for prothrombin may have at least 80%, 85%, 90%, 95%, 97%, 98% or 99% identity to the sequence SEQ ID NO 1. The recombinant prothrombin can be a mixture of fully carboxylated prothrombin and incompletely carboxylated prothrombin. In one embodiment, the recombinant prothrombin is a fully carboxylated prothrombin and in another embodiment, the recombinant
30 prothrombin is an incompletely carboxylated prothrombin.

A further aspect of the invention relates to the recombinant thrombin obtained by the method according to the invention. A pharmaceutical composition can be designed comprising the recombinant thrombin obtained by the method according to the invention, in combination with pharmaceutically acceptable carriers, vehicles and/or adjuvants. The pharmaceutical composition can be in an applicable form.

In one embodiment thrombin produced by the described method can be used in the manufacturing of tissue sealants ("glues") in combination with other proteins, i.e. fibrin originating from recombinant cells, transgenic animals or human plasma. In another embodiment thrombin produced by the described method can be used as a stand-alone product, freeze dried as single active component or in combination with a non-protein matrix, or, in solution as single active component or in combination with other active components.

Suitable mix-in components would be, but is not limited to, collagen, chitin, degradable polymers, cellulose, recombinant coagulation factors and fibrinogen from transgenic or recombinant sources.

The potential fields of use for the tissue sealants ("glues") are numerous; skin grafting, neuro surgery, cardiac surgery, toracic surgery, vascular surgery, oncologic surgery, plastic surgery, ophthalmologic surgery, orthopedic surgery, trauma surgery, head and neck surgery, gynecologic and urologic surgery, gastrointestinal surgery, dental surgery, drug delivery, tissue engineering and dental cavity haemostasis.

A further aspect of the invention relates to a method for obtaining coagulation by administering a therapeutically effective amount of a recombinant human thrombin obtained using the method according to the invention to a patient.

Another aspect of the present invention is an isolated DNA sequence according SEQ ID NO 2 or homologues thereof coding for a recombinant ecarin. A homologue of SEQ ID NO 2 coding for ecarin may have at least 80%, 85%, 90%, 95%, 97%, 98% or 99% identity to the sequence SEQ ID NO 2. SEQ ID NO 2 is a designed sequence that has been optimised

for optimal expression. The sequence is particularly suited for expression in mammalian cell systems.

According to another aspect a vector comprising SEQ ID NO 2 or a homologue thereof is provided. Said vector can be designed to overexpress SEQ ID NO 2 or a homologue thereof and is operably linked to expression control sequences permitting expression of ecarin encoded by SEQ ID NO 2 or a homologue thereof. According to a third aspect a host cell comprising said vector is provided that is capable of expressing ecarin encoded by SEQ ID NO 2 or a homologue thereof. This host cell is preferably a eukaryotic cell. Typical host cells include, but are not limited to insect cells, yeast cells, and mammalian cells. Mammalian cells are particularly preferred. Suitable mammalian cells lines include, but are not limited to, CHO, HEK, NS0, 293, Per C.6, BHK and COS cells, and derivatives thereof. In one embodiment the host cell is the mammalian cell line CHO-S.

According to another embodiment of the present invention a polypeptide comprising an amino acid sequence encoded by SEQ ID NO: 2 or a homologue thereof and obtained by the method described in Example 2.

The sequence identity between two sequences can be determined by pair-wise computer alignment analysis, using programs such as, BestFit, PILEUP, Gap or FrameAlign. The preferred alignment tool is BestFit. In practise, when searching for similar/identical sequences to the query search, from within a sequence database, it is generally necessary to perform an initial identification of similar sequences using suitable algorithms such as Blast, Blast2, NCBI Blast2, WashU Blast2, FastA, or Fasta3, and a scoring matrix such as Blosum 62. Such algorithms endeavour to closely approximate the "gold-standard" alignment algorithm of Smith-Waterman. Thus, the preferred software/search engine program for use in assessing similarity, i.e., how two primary polypeptide sequences line up is Smith-Waterman. Identity refers to direct matches, similarity allows for conservative substitutions.

Experimental section

The invention will be further described by means of the following examples which shall not be interpreted as limiting the scope of the appended claims.

5 Example 1**High yield production of recombinant human prothrombin in CHO cells.**

The P1E2 cell line containing the construct PN32 shown in Figure 1 having the nucleotide sequence SEQ ID NO: 1, was grown in a fermentor according to the method described in WO2005038019, using a protein and animal component free growth medium in order to
10 produce prothrombin for use in thrombin manufacturing. The cells were grown either by batch or perfusion culture methods (Table 1) and the amount of prothrombin produced was measured by an ecarin assay. This ecarin assay was performed essentially as the Chromogenix assay (Mölndal, Sweden) using purified plasma-derived human prothrombin (Haematologic Technologies Inc., Vermont, USA) as standard.

15

Table 1. Examples of yield of prothrombin in experimental fermentor runs

Experiment ID	Culture method & time	Viable cells (million cells/mL)	Prothrombin mg/L
CC2LC (272-8)	Batch, 238 h	5.9	281
CC2LD (272-8)	Batch, 238 h	6.2	276
326-11B	Perfusion, 259 h	18	722

The fermentor experiments showed that both batch and perfusion culture methods can be used to produce prothrombin suitable for production of recombinant thrombin (Table 1).
20 The share of fully carboxylated prothrombin obtained in these fermentor runs was about 55-87 %, the rest being incompletely carboxylated prothrombin.

Example 2

Production of recombinant ecarin in CHO cells.

An ecarin encoding sequence having the nucleotide sequence SEQ ID NO: 2 optimised for
5 expression in mammalian cells was synthesized and cloned into the Invitrogen vector
pCDNA 3.1+ (Figure 2). An alignment of the nucleotide sequence of the recombinant
ecarin used in the present invention to the sequence of wild type ecarin (GI:717090) is seen
in Figure 4. As can be seen in Figure 5 this recombinant ecarin is 100 % homologous to the
amino acid sequence for wild type ecarin. This construct, AZ ecarin (SEQ ID NO. 3), was
10 used to stably transfect CHO-S cells (Invitrogen). Ecarin is secreted by the host cell to the
extra-cellular space, and in order to screen for ecarin producing clones, culture supernatant
samples were removed and mixed with recombinant human prothrombin (rhFII) to a final
concentration of 1 mg rhFII /L in assay buffer (50 mM Tris-HCl, pH 7.4 containing 0.1%
BSA). This mix was incubated 20-40 minutes at 37°C. The thrombin generated by the
15 action of ecarin present in the sample was then detected by adding a 1-2 mM solution of
the chromogenic thrombin substrate S-2238 (Chromogenix, Mölndal). Colour development
was monitored and stopped when suitable using 20% acetic acid. To estimate the activity
of the recombinant ecarin produced, snake venom derived ecarin with a declared activity
was purchased from Sigma and used as standard. The best producing cell line obtained
20 produced up to 7000 U ecarin per litre culture in lab scale shaker cultures grown in animal
component free medium.

Activation of recombinant ecarin

The above method produces the recombinant ecarin as a pro protein. Thus, activation by
25 removal of the pro-part is necessary for optimal activity. To our surprise, we found that
activation was most conveniently obtained by continued incubation of the culture for at
least 7 days after the death of the ecarin producing cells (Figure 6). The culture medium
used was CD-CHO supplemented with HT-supplement, non-essential amino acids and
Glutamax I (as recommended by Invitrogen for CHO-S), and growth conditions were
30 shaker bottles at 37°C in an atmosphere containing 5% carbon dioxide. Culture samples

were assayed for activity as described above. As can be seen from Figure 6, the activity of recombinant ecarin increased during the activation period.

Samples from culture supernatants were also separated by SDS-PAGE and blotted to
5 nitrocellulose membranes. Labelling of the membrane was performed with polyclonal rabbit
serum directed towards the mature part of ecarin expressed as inclusion bodies in *E. coli*.
“M” indicates the molecular weight marker and numbers refer to day of sample collection.
As can be seen from Figure 7 the recombinant ecarin remains stable for more than a week
after the death of the cells. Activation of ecarin may also take place at lower temperatures,
10 for instance as low as room temperature, but will then require longer times for activation.
Ecarin will remain stable for severable months in room temperature in the presence of dead
host cells. The activity will increase gradually until it levels out. A decrease in activity has
not been observed except in the presence of bacterial infections or high temperatures.
Efforts to use trypsin for activation of ecarin were made, but were not successful.

15

Example 3

Conversion of prothrombin to thrombin by ecarin.

The ecarin protease converts prothrombin to meizothrombin, an intermediate form of
thrombin that has thrombin catalytic activity. Further processing into thrombin is achieved
20 by auto-catalyses. To determine the estimated amount of ecarin culture needed for
converting prothrombin into thrombin, we performed a series of test digests. Different
amounts of ecarin-containing culture supernatants as obtained in Example 2, were mixed
with 1 mg/ml prothrombin (as obtained in example 1) in PBS buffer (Cambrex). Incubation
of the mixtures was done at 37°C for 1-3 hours. Samples were then analysed by SDS-
25 PAGE to identify the amount of recombinant ecarin needed for complete conversion of
prothrombin into thrombin. By this procedure we found that the recombinant ecarin was
very potent; one litre of ecarin culture supernatant at 7000 U/L is capable of complete
conversion of 64 grams of prothrombin into thrombin in less than 3 hours at 37°C.
Normally recombinantly produced prothrombin has to be purified in order to separate
30 fully-carboxylated prothrombin from the incompletely carboxylated prothrombin. However
this is not necessary for the present invention as the recombinant ecarin is able to

efficiently activate both the fully carboxylated and the incompletely carboxylated prothrombins.

Example 4

5 Purification of thrombin

Thrombin obtained by the procedure described in example 3 was purified by cation-exchange chromatography (CIEX) using ÄKTA-FPLC (GE Healthcare) and an SP-Sephacrose HP column (GE Healthcare) equilibrated with 25 mM sodium-phosphate buffer, pH 6.5. Ecarin-digested prothrombin prepared as in example 3 was adjusted to pH 6.2 and
10 a conductivity of approximately 8 mS/cm. Thrombin was eluted with a 1M sodium chloride gradient in column equilibration buffer over 20 column volumes (Figure 8). Selected fractions were analysed by SDS-PAGE (Figure 9). Thrombin activity was confirmed by incubation with the chromogenic thrombin substrate S-2238 (Chromogenix, Mölndal).

15

Example 5

Analyses of rh-thrombin obtained.

To further analyse the obtained thrombin, kinetic parameters were determined using the chromogenic thrombin substrate S-2366 (Chromogenix). Activity was estimated by
20 titration with hirudin. The rh-thrombin was for all parameters; Activity, K_{kat} and V_{max} , similar to plasma-derived human α -thrombin from Haematologic Technologies Inc. (Vermont, USA).

Purified thrombin was also subjected to N-terminal sequencing: Reduced thrombin heavy
25 and light polypeptide chains were separated by SDS-PAGE and blotted to Immobilon P membrane (Millipor). The excised bands were sequenced by the Edman degradation method. Heavy chain N-terminal first five amino acids were confirmed to be IVEGS, and the light chain five N-terminal amino acids were TFGS as expected.

CLAIMS

1. A method for producing recombinant human thrombin comprising: contacting recombinant prothrombin with recombinant active ecarin; wherein said recombinant active ecarin is obtained from host cells that secrete recombinant ecarin into the extracellular space of a cell culture medium, wherein said medium is separated from cell debris and contacted with said recombinant prothrombin with no additional purification steps.
2. The method according to claim 1, wherein said recombinant active ecarin is produced by a method comprising:
 - (a) transfecting a host cell with recombinant prepro-ecarin;
 - (b) incubating said host cell in growth medium to generate a cell culture, wherein the cells in said culture express and secrete recombinant prepro-ecarin to the extracellular space;
 - (c) continuing incubation of the cell culture from step (b) such that substantially all of the cells have ceased to be viable;
 - (d) continuing incubation of the cell culture from step (c) for an activation period sufficient to produce recombinant active ecarin in the medium.
3. The method according to claim 1 or 2, wherein said recombinant ecarin is encoded by a nucleic acid having the sequence of or having 80%, 85%, 90%, 95%, 97% or 99% identity to SEQ ID NO 2.
4. A method according to any of the preceding claims, wherein the recombinant prothrombin and/or the recombinant ecarin is expressed in a mammalian host cell.
5. The method of claim 4, wherein said recombinant prothrombin is encoded by a nucleic acid having the sequence of SEQ ID NO 1.
6. A method according to any of the preceding claims, wherein the recombinant prothrombin is produced by a host cell that co-expresses recombinant gamma-glutamyl carboxylase (GGCX).
7. A method according to any of the preceding claims, wherein the recombinant prothrombin is a mixture of fully carboxylated prothrombin and incompletely carboxylated prothrombin.
8. A method according to any of the preceding claims, wherein the recombinant prothrombin is a fully carboxylated prothrombin.

9. A method according to any of the preceding claims, wherein the recombinant prothrombin is an incompletely carboxylated prothrombin.
10. A recombinant human thrombin obtained by the method of any of the preceding claims.
11. A pharmaceutical composition comprising a recombinant thrombin according to claim 10, in combination with pharmaceutically acceptable carriers, vehicles and/or adjuvants.

SEQUENCE LISTING

<110> AstraZeneca AB
<120> METHOD FOR PRODUCTION OF RECOMBINANT HUMAN THROMBIN `644
<130> 102644
<140> PCT/SE2008/050836
<141> 2008-07-04
<150> 60/948207
<151> 2007-07-06
<160> 3
<170> PatentIn version 3.4
<210> 1
<211> 7280
<212> DNA
<213> Synthetic
<400> 1
gacggatcgg gagatctccc gatcccctat ggtcgactct cagtacaatc tgctctgatg 60
ccgcatagtt aagccagtat ctgctccctg ctgtgtgtgt ggaggtcgct gagtagtgcg 120
cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc 180
ttagggtttag gcgttttgcg ctgcttcgcy atgtacgggc cagatatacg cgttgacatt 240
gattattgac tagttattaa tagtaatcaa ttacggggtc attagtcat agcccatata 300
tggagttccg cgttacataa cttacggtaa atggcccgc tggctgaccg cccaacgacc 360
cccgccatt gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc 420
attgacgtca atgggtggac tatttacggt aaactgcca cttggcagta catcaagtgt 480
atcatatgcc aagtacgcc cctattgacg tcaatgacgg taaatggccc gcctggcatt 540
atgccagta catgacctta tgggacttct ctacttggca gtacatctac gtattagtca 600
tcgctattac catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg 660
actcacgggg atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc 720
aaaatcaacg ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcy 780
gtaggcgtgt acggtgggag gtctatataa cgagagctct ctggctaact agagaaccca 840
ctgcttactg gcttatcgaa attaatacga ctactatag ggagacccaa gctggctagc 900
gtttaaactt aagcttggtg ccgagctcgg atccactagt ccagtgtggt ggaattctgc 960
agatatccac catgatccag atcctgctgg tgatcatctg cctggccgtg tccccctacc 1020
agggtgctc catcatcctg ggcagcggca acgtgaacga ctacgaggtg gtgtaccccc 1080
agaaggtgac cgccctgcc aagggcgccg tgcagcagcc cgagcagaaa tacgaggacg 1140
ccatgcagta cgagttcgag gtgaaggcg agcccggtgt gctgcacctg gagaagaaca 1200
aggagctgtt cagcgaggac tacagcgaga cccactacag cagcgacgac agggagatca 1260
ccaccaaccc cagcgtggag gaccactgct actaccacgg ccggatccag aacgacgccg 1320
agagcaccgc cagcatcagc gcctgtaatg gcctgaaggc cacttcaag ctgagaggcg 1380
agacctactt catcgagccc ctgaagatcc ccgacagcga ggcccacgcc gtgtacaagt 1440
acgagaacat cgagaacgag gacgaggccc ctaagatgtg tggcgtgacc caggacaact 1500
gggagagcga cgagcccatc aagaaaaccc tgggcctgat cgtgcccccc cacgagagaa 1560
agttcgagaa gaagttcatc gaactggtgg tcgtggtgga ccacagcatg gtgaccaagt 1620

acaacaacga cagcaccgcc atcaggacct ggatctacga gatgctgaac accgtgaacg 1680
agatctacct gcccttcaac atcagagtgg cctgggtggg cctggagttc tgggtgaacg 1740
gcgacctgat caacgtgacc agcaccgccg acgacaccct gcacagcttc ggcgagtga 1800
gagccagcga cctgctgaac cggaagagac acgatcacgc ccagctgctg accaatgtga 1860
ccctggacca ctccaccctg ggcattcacct tcgtgtacgg catgtgtgaag agcgaccgga 1920
gcgtggagct gatcctggac tacagcaaca tcaccttcaa catggcctac atcatcgccc 1980
acgagatggg ccacagcctg ggcattgctgc acgacaccaa gttctgtacc tgtggcgcca 2040
agccctgtat catgttcggc aaggagagca tccctcccc taaggagttc agcagctgct 2100
cctacgacca gtacaataag tacctgctga agtacaaccc caagtgtatc ctggaccccc 2160
ccctgagaaa ggacatcgcc agccctgccg tgtgtggcaa tgagatctgg gaggagggcg 2220
aggagtgtga ctgtggcagc ccagccgact gtagaaaccc ctgctgtgat gccgccacct 2280
gtaagctgaa gcctggcgcc gagtgtggca acggcgagtg ctgtgacaag tgtaagatcc 2340
ggaaggccgg caccgagtgt agaccgccca gggacgattg tgacgtggcc gagcactgta 2400
ccggccagag cgccgagtgc ccagaaacg agttccagag gaacggccag ccttgccctga 2460
acaacagcgg ctactgctac aacggcgact gccccatcat gctgaaccag tgtatcgccc 2520
tgttcagccc cagcgccacc gtggcccagg acagctgctt ccagagaaac ctgcagggca 2580
gctactacgg ctactgtacc aaggagatcg gctactacgg aaagaggttc ccctgtgccc 2640
ctcaggacgt gaagtgtggc aggtgtact gcctggacaa ctcttcaag aaaaacatga 2700
ggtgtaagaa cgactacagc tacgccgacg agaacaaggg catcgtggag cccggcacca 2760
agtgtgagga cggcaaagtg tgtatcaacc ggaagtgtgt ggacgtgaac accgcctact 2820
gatgagcggc cgctcgagtc tagagggccc gtttaaacc gctgatcagc ctcgactgtg 2880
ccttctagtt gccagccatc tgttgtttgc cctcccccg tgccttctt gacctggaa 2940
ggtgccactc cactgtcct ttcctaataa aatgaggaaa ttgcatcgca ttgtctgagt 3000
aggtgtcatt ctattctggg ggggtgggtg gggcaggaca gcaaggggga ggattgggaa 3060
gacaatagca ggcattgctg ggatgcggtg ggtctatgg cttctgaggc ggaaagaacc 3120
agctggggct ctagggggta tccccacgcg ccctgtagcg gcgcattaag cgcggggggt 3180
gtgggtggta cgcgcagcgt gaccgctaca cttgccagcg ccctagcgcc cgctccttct 3240
gctttcttcc ctctcttct cgcacgttc gccggcttct cccgtcaagc tctaaatcgg 3300
ggcatccctt taggggttccg atttagtgct ttacggcacc tcgaccccaa aaaacttgat 3360
tagggtgatg gttcacgtag tgggccatcg ccctgataga cggtttttcg ccctttgacg 3420
ttggagtcca cgttctttaa tagtggactc ttgttccaaa ctggaacaac actcaacct 3480
atctcggctc attcttttga ttataagggt attttgggga tttcgcccta ttggttaaaa 3540
aatgagctga ttttaacaaa atttaacgcg aattaattct gtggaatgtg tgtcagttag 3600
ggtgtggaaa gtccccaggc tccccaggca ggcagaagta tgcaaagcat gcattctaat 3660
tagtcagcaa ccagggtgtg aaagtcccca ggctccccag caggcagaag tatgcaaagc 3720
atgcattctc attagtcagc aaccatagtc ccgcccctaa ctccgcccct cccgccccta 3780
actccgcca gttccgcca ttctccgcc catggctgac taattttttt tatttatgca 3840
gaggccgagg ccgcctctgc ctctgagcta ttccagaagt agtgaggagg ctttttttga 3900

03-09-2008

ggcctaggct tttgcaaaaa gctcccggga gcttgatat ccattttcgg atctgatcaa 3960
gagacaggat gaggatcggt tcgcatgatt gaacaagatg gattgcacgc aggtttctccg 4020
gccgcttggg tggagaggct attcggctat gactgggcac aacagacaat cggtgctct 4080
gatgccgcgc tgttcgggct gtcagcgcag gggcgcccg ttctttttgt caagaccgac 4140
ctgtccgggtg cctgaatga actgcaggac gaggcagcgc ggctatcgtg gctggccacg 4200
acgggcgttc cttgcgcagc tgtgctcgac gttgtcactg aagcgggaag ggactggctg 4260
ctattgggcy aagtgcggg gcaggatctc ctgtcatctc accttgctcc tgccgagaaa 4320
gtatccatca tggctgatgc aatgcggcgg ctgcatacgc ttgatccggc tacctgcca 4380
ttcgaccacc aagcgaaca tcgcatcgag cgagcacgta ctcgatgga agccggtctt 4440
gtcgatcagg atgatctgga cgaagagcat caggggctcg cgccagccga actgttcgcc 4500
aggctcaagg cgcgcagtc cgacggcgag gatctcgtcg tgacctatgg cgatgcctgc 4560
ttgccgaata tcatggtgga aaatggccgc tttctggat tcatcgactg tggccggctg 4620
ggtgtggcgg accgctatca ggacatagcg ttggctaccc gtgatattgc tgaagagctt 4680
ggcggcgaat gggctgaccg ctctcctcgtg ctttacggta tcgccgctcc cgattcgag 4740
cgcatcgctt tctatcgctt tcttgacgag ttcttctgag cgggactctg gggttcgaaa 4800
tgaccgacca agcgcgcgc aacctgccat cagcagatct cgattccacc gccgccttct 4860
atgaaagggt gggcttcgga atcgtttcc gggacgcggc ctggatgatc ctccagcgcg 4920
gggatctcat gctggagttc ttcgcccacc ccaacttgtt tattgcagct tataatggtt 4980
acaaataaag caatagcatc acaaatttca caaataaagc atttttttca ctgcattcta 5040
gttgtggttt gtccaaactc atcaatgtat cttatcatgt ctgtataccg tcgacctcta 5100
gctagagctt ggcgtaatca tggctcatagc tgtttcctgt gtgaaattgt tatccgctca 5160
caattccaca caacatacga gccggaagca taaagtgtaa agcctggggt gcctaatagag 5220
tgagctaact cacattaatt gcgttgcgct cactgcccgc tttccagtcg ggaaacctgt 5280
cgtgccagct gcattaatga atcgccaac gcgcggggag aggcggtttg cgtattgggc 5340
gctcttcgc ttcctcgctc actgactcgc tgcgctcgtt cgttcggtg cggcgagcgg 5400
tatcagctca ctcaaaggcg gtaatacggc tatccacaga atcaggggat aacgcaggaa 5460
agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc gcgttgctgg 5520
cgtttttcca taggtccgc cccctgacg agcatcaca aaatcgacgc tcaagtcaga 5580
ggtggcgaaa cccgacagga ctataaagat accaggcgtt tccccctgga agctccctcg 5640
tgcgctctcc tgttcgacc ctgccgctta ccggatacct gtccgccttt ctcccttcgg 5700
gaagcgtggc gctttctcaa tgctcacgct gtaggtatct cagttcgggtg taggtcgctt 5760
gctccaagct gggctgtgtg cacgaacccc ccgttcagcc cgaccgctgc gccttatccg 5820
gtaactatcg tcttgagtc aacccggtaa gacacgactt atcgccactg gcagcagcca 5880
ctggtaacag gattagcaga gcgaggtatg taggcggtgc tacagagttc ttgaagtgg 5940
ggcctaacta cggctacact agaaggacag tatttggtat ctgcgctctg ctgaagccag 6000
ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa acaaaccacc gctggtagcg 6060
gtggtttttt tgtttgcaag cagcagatta cgcgcagaaa aaaaggatct caagaagatc 6120
ctttgatctt ttctacggg tctgacgctc agtggaaaga aaactcacgt taagggattt 6180
tggtcatgag attatcaaaa aggatcttca cctagatcct tttaaattaa aaatgaagtt 6240

ttaaataaat ctaaagtata tatgagtaaa cttgggtctga cagttaccaa tgcttaataca 6300
 gtgaggcacc tatctcagcg atctgtctat ttcgttcac catagttgcc tgactccccg 6360
 tcgtgtagat aactacgata cgggagggct taccatctgg cccagtgct gcaatgatac 6420
 cgcgagaccc acgtccaccg gctccagatt tatcagcaat aaaccagcca gccggaagg 6480
 ccgagcgcag aagtgggtcct gcaactttat ccgcctccat ccagtctatt aattgttgcc 6540
 gggaaagctag agtaagtagt tcgccagtta atagtttgcg caacgttggt gccattgcta 6600
 caggcatcgt ggtgtcacgc tcgtcgtttg gtatggcttc attcagctcc ggttcccaac 6660
 gatcaaggcg agttacatga tccccatgt tgtgcaaaaa agcggttagc tccttcggtc 6720
 ctccgatcgt tgtcagaagt aagttggcgg cagtgttacc actcatgggt atggcagcac 6780
 tgcataattc tcttactgtc atgccatccg taagatgctt ttctgtgact ggtgagtact 6840
 caaccaagtc attctgagaa tagtgtatgc ggcgaccgag ttgctcttgc ccggcgctca 6900
 tacgggataa taccgcgcca catagcagaa ctttaaaagt gctcatcatt ggaaaacgtt 6960
 cttcggggcg aaaactctca aggatcttac cgtgttgag atccagttcg atgtaacca 7020
 ctctgtcacc caactgatct tcagcatctt ttactttcac cagcgtttct ggggtgagca 7080
 aaacaggaag gcaaaatgcc gcaaaaaagg gaataaggcg gacacggaaa tgttgaatac 7140
 tcatactctt cctttttcaa tattattgaa gcatttatca gggttattgt ctcatgagcg 7200
 gatacatatt tgaatgtatt tagaaaaata aacaaatagg ggttcgcgc acatttcccc 7260
 gaaaagtgcc acctgacgtc 7280

<210> 2
 <211> 1851
 <212> DNA
 <213> Synthetic

<400> 2
 atgatccaga tctgtctggt gatcatctgc ctggccgtgt tcccctacca gggctgctcc 60
 atcatcctgg gcagcggcaa cgtgaacgac tacgaggtgg tgtaccccca gaaggtgacc 120
 gccctgcccc agggcgccgt gcagcagccc gagcagaaat acgaggacgc catgcagtac 180
 gagttcgagg tgaaggcgga gccgtgggtg ctgcacctgg agaagaacaa ggagctgttc 240
 agcgaggact acagcgagac ccactacagc agcgacgaca gggagatcac caccaacccc 300
 agcgtggagg accactgcta ctaccacggc cggatccaga acgacgccga gagcaccgcc 360
 agcatcagcg cctgtaatgg cctgaaggcg cacttcaagc tgagaggcga gacctacttc 420
 atcgagcccc tgaagatccc cgacagcgag gccacgccg tgtacaagta cgagaacatc 480
 gagaacgagg acgaggcccc taagatgtgt ggcgtgacct aggacaactg ggagagcgac 540
 gagcccatca agaaaaccct gggcctgac gtgccccccc acgagagaaa gtctcgagaag 600
 aagttcatcg aactgggtgg cgtgggtggac cacagcatgg tgaccaagta caacaacgac 660
 agcaccgcca tcaggacctg gatctacgag atgctgaaca ccgtgaacga gatctacctg 720
 ccttcaaca tcagagtggc cctgggtggc ctggagttct ggtgtaacgg cgacctgac 780
 aacgtgacca gcaccgccga cgacaccctg cacagcttcg gcgagtggag agccagcgac 840
 ctgctgaacc ggaagagaca cgatcacgcc cagctgctga ccaatgtgac cctggaccac 900
 tccaccctgg gcacacactt cgtgtacggc atgtgtaaga gcgaccggag cgtggagctg 960
 atcctggact acagcaacat cacctcaac atggcctaca tcatcgcccc cgagatgggc 1020

cacagcctgg gcatgctgca cgacaccaag ttctgtacct gtggcgccaa gccctgtatc 1080
 atgttcggca aggagagcat cctccccc aaggagttca gcagctgtc ctacgaccag 1140
 tacaataagt acctgctgaa gtacaacccc aagtgtatcc tggaccccc cctgagaaag 1200
 gacatcgcca gccctgccgt gtgtggcaat gagatctggg aggagggcga ggagtgtgac 1260
 tgtggcagcc cagccgactg tagaaacccc tgtgtgatg ccgccacctg taagctgaag 1320
 cctggcgccg agtgtggcaa cggcgagtgc tgtgacaagt gtaagatccg gaaggccggc 1380
 accgagtgtg gacccgccag ggacgattgt gacgtggccg agcactgtac cggccagagc 1440
 gccgagtgcc ccagaaacga gttccagagg aacggccagc cttgcctgaa caacagcggc 1500
 tactgtaca acggcgactg ccccatcatg ctgaaccagt gtatcgccct gttcagcccc 1560
 agcgccaccg tggcccagga cagctgtctc cagagaaacc tgcagggcag ctactacggc 1620
 tactgtacca aggagatcgg ctactacgga aagaggttcc cctgtgcccc tcaggacgtg 1680
 aagtgtggca ggctgtactg cctggacaac tccttcaaga aaaacatgag gtgtaagaac 1740
 gactacagct acgccgacga gaacaagggc atcgtggagc ccggcaccaa gtgtgaggac 1800
 ggcaaaagtgt gtatcaaccg gaagtgtgtg gacgtgaaca ccgcctactg a 1851

<210> 3
 <211> 10139
 <212> DNA
 <213> Synthetic

<400> 3
 gacggatcgg gagatctccc gatccccat ggtgcactct cagtacaatc tgctctgatg 60
 ccgcatagtt aagccagtat ctgctccctg cttgtgtgtt ggaggtcgct gagtagtgcg 120
 cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc 180
 ttagggttag gcgttttgcg ctgcttcgag atgtacgggc cagatatacg cgttgacatt 240
 gattattgac tagttattaa tagtaataca ttacggggtc attagttcat agcccatata 300
 tggagtcccg cgttacataa cttacggtaa atggcccgc tggctgaccg cccaacgacc 360
 cccgccatt gacgtcaata atgacgtatg tcccatagat aacgccaata gggactttcc 420
 attgacgtca atgggtggag tatttacggt aaactgcca cttggcagta catcaagtgt 480
 atcatatgcc aagtacgcc cctattgacg tcaatgacgg taaatggccc gcctggcatt 540
 atgccagta catgacctta tgggactttc ctacttgga gtacatctac gtattagtca 600
 tcgctattac catgggtgat cggttttggc agtacatcaa tgggcgtgga tagcggtttg 660
 actcacgggg atttccaagt ctccaccca ttgacgtcaa tgggagtttg ttttggcacc 720
 aaaatcaacg ggactttcca aaatgtcgta acaactccgc ccattgacg caaatggcg 780
 gtaggcgtgt acggtgggag gtctatataa gcagagctct ctggctaact agagaacca 840
 ctgcttactg gcttatcgaa attaatacga ctactatag ggagaccaa gctggctagc 900
 gtttaaactt aagcttggtg ccgagctcgg atccactagt ccagtgtggt ggaattgccc 960
 ttattcctca gtgaccaggg agctgacaca ctatggcgca cgtccgaggc ttgcagctgc 1020
 ctggctgcct ggccctggct gccctgtgta gccttggtga cagccagcat gtgttcctgg 1080
 ctctcagca agcacggtcg ctgctccagc gggccggcg agccaacacc ttcttgagg 1140
 aggtgcgcaa gggcaacctg gagcgagagt gcgtggagga gacgtgcagc tacgaggagg 1200
 ccttcgaggc tctggagtcc tccacggcta cggatgtgtt ctgggccaag tacacagctt 1260

gtgagacagc gaggacgcct cgagataagc ttgctgcatg tctggaaggt aactgtgctg 1320
agggctctggg tacgaactac cgagggcatg tgaacatcac ccggtcaggc attgagtgcc 1380
agctatggag gagtgcctac ccacataagc ctgaaatcaa ctccactacc catcctgggg 1440
ccgacctaca ggagaatttc tgccgcaacc ccgacagcag caccacggga ccctggtgct 1500
acactacaga cccaccctg aggaggcagg aatgcagcat ccctgtctgt ggccaggatc 1560
aagtcactgt agcgatgact ccacgctccg aaggctccag tgtgaatctg tcacctccat 1620
tgagagcagt gtccctgat cgggggcagc agtaccaggg gcgcctggcg gtgaccacac 1680
atgggctccc ctgcctggcc tggggccagc cacaggccaa ggccctgagc aagcaccagg 1740
acttcaactc agctgtgcag ctggtggaga acttctgccg caaccacagc ggggatgagg 1800
agggcgtgtg gtctatgtg gccgggaagc ctggcgactt tgggtactgc gacctcaact 1860
attgtgagga ggccgtggag gaggagacag gagatgggct ggatgaggac tcagacaggg 1920
ccatcgaagg gcgtaccgcc accagtgagt accagacttt cttcaatccg aggacctttg 1980
gctcgggaga ggcagactgt gggctgcgac ctctgttcga gaagaagtcg ctggaggaca 2040
aaaccgaaag agagctcctg gaatcctaca tcgacggcg cttgtggag ggctcggatg 2100
cagagatcgg catgtcacct tggcaggtga tgcttttccg gaagagtccc caggagctgc 2160
tgtgtggggc cagcctcatc agtgaccgct gggctctcac cgcgcgccac tgctcctgt 2220
accgcacctg ggacaagaac ttcaccgaga atgaccttct ggtgcgcatt ggcaagcact 2280
cccgccaccg gtacgagcga aacattgaaa agatatccat gttggaaaag atctacatcc 2340
accccaggta caactggcgg gagaacctgg accgggacat tgccctgatg aagctgaaga 2400
agcctgttgc cttcagtgc tacattcacc ctgtgtgtct gcccgacagg gagacggcag 2460
ccagcttctc ccaggctgga tacaaggggc gggtgacagg ctggggcaac ctgaaggaga 2520
cgtggacagc caacgttggg aaggggcagc ccagtgtcct gcaggtgggt aacctgccc 2580
ttgtggagcg gccggtctgc aaggactcca cccgatccg catcactgac aacatgttct 2640
gtgctggtta caagcctgat gaagggaac gaggggatgc ctgtgaaggt gacagtgggg 2700
gacctttgt catgaagagc cctttaaca accgctggta tcaaatgggc atcgtctcat 2760
ggggtgaagg ctgtgaccgg gatgggaaat atggcttcta cacacatgtg ttccgcctga 2820
agaagtggat acagaaggtc attgatcagt ttggagagta gaagggaat tctgcagata 2880
tccagcacag tggcgccgc tcgagtctag agggcccggt taaaccgct gatcagctc 2940
gactgtgcct tctagtgtcc agccatctgt tgtttgcccc tccccgtgc cttccttgac 3000
cctggaaggt gccactcca ctgtccttcc taataaaaat gaggaaattg catcgcatg 3060
tctgagtagg tgctattcta ttctggggg tggggtggg caggacagca agggggagga 3120
ttgggaagac aatagcaggc atgctgggga tgcggtgggc tctatggctt ctgaggcgg 3180
aagaaccagc tggggtctta ggggtatcc ccacgcgcc tgtagcggcg cattaagcgc 3240
ggcggtgtg gtggttacgc gcagcgtgac cgctacactt gccagcgccc tagcgccgc 3300
tcctttcgtc ttcttccctt cctttctcgc cacgttcgcc ggctttcccc gtcaagctct 3360
aaatcgggg ctccttttag ggttccgatt tagtgctta cggcacctc gaccccaaaa 3420
aacttgatta gggctgtgga atgtgtgtca gttagggtgt ggaaagtccc caggctcccc 3480
agcaggcaga agtatgcaa gcatgcact caattagtca gcaaccagg gtggaaagtc 3540

cccaggctcc	ccagcaggca	gaagtatgca	aagcatgcat	ctcaattagt	cagcaaccat	3600
agtcccgccc	ctaaactccg	ccatcccgcc	cctaactccg	cccagttccg	cccattctcc	3660
gccccatggc	tgactaattt	tttttattta	tgcagaggcc	gaggccgcct	cggcctctga	3720
gctattccag	aagtagtgag	gaggcttttt	tggaggccta	ggcttttgca	aaaagctctc	3780
tggctaacta	gagaaccac	tgcttactgg	cttatcgaaa	ttaatacgac	tcactatagg	3840
gagaccaag	ctggctagcg	tttaaactta	agcttggtac	cgagctcgga	tccactagtc	3900
cagtgtgggtg	gaattgccct	ttccgcagag	caatggcggg	gtctgccggg	tccgcgcgga	3960
cctcgcccag	ctcagataaa	gtacagaaa	acaaggctga	actgatctca	gggcccaggc	4020
aggacagccg	aataggga	ctcttgggtt	ttgagtggac	agatttgtcc	agttggcgga	4080
ggctgggtgac	cctgctgaat	cgaccaacgg	accctgcaag	cttagctgtc	tttcgttttc	4140
tttttgggtt	cttgatggtg	ctagacattc	cccaggagcg	ggggctcagc	tctctggacc	4200
ggaaataacct	tgatgggctg	gatgtgtgcc	gcttcccctt	gctggatgcc	ctacgccac	4260
tgccacttga	ctggatgtat	cttgtctaca	ccatcatgtt	tctgggggca	ctgggcatga	4320
tgctgggcct	gtgtaccgg	ataagctgtg	tgttattcct	gctgccatac	tggtatgtgt	4380
ttctcctgga	caagacatca	tggaacaacc	actcctatct	gtatgggttg	ttggcctttc	4440
agctaacatt	catggatgca	aaccactact	ggctctgtga	cggtctgctg	aatgcccata	4500
ggaggaatgc	ccacgtgccc	ctttggaact	atgcagtgtc	ccgtggccag	atcttcattg	4560
tgtacttcat	tgccgggtgtg	aaaaagctgg	atgcagactg	ggttgaaggc	tattccatgg	4620
aatattttgtc	ccggcactgg	ctcttcagtc	ccttcaaact	gctgttgtct	gaggagctga	4680
ctagcctgct	ggctgtgcac	tgggggtggc	tgctgcttga	cctctcagct	ggtttcctgc	4740
tcttttttga	tgtctcaaga	tccattggcc	tggtctttgt	gtcctacttc	cactgcatga	4800
attcccagct	tttcagcatt	ggtatgttct	cctacgtcat	gctggccagc	agccctctct	4860
tctgtctccc	tgagtggcct	cggaaagctgg	tgctcctactg	ccccgaagg	ttgcaacaac	4920
tgttgccct	caaggcagcc	cctcagccca	gtgtttcctg	tgtgtataag	aggagccggg	4980
gcaaaaagtgg	ccagaagcca	gggctgcgcc	atcagctggg	agctgccttc	accctgctct	5040
acctcctgga	gcagctattc	ctgccctatt	ctcattttct	caccagggc	tataacaact	5100
ggacaaatgg	gctgtatggc	tattcctggg	acatgatgg	gcactccgc	tcccaccagc	5160
acgtgaagat	cacctaccgt	gatggccgca	ctggcgaaact	gggctacctt	aacctgggg	5220
tatttacaca	gagtcggcga	tggaaaggatc	atgcagacat	gctgaagcaa	tatgccactt	5280
gcctgagccg	cctgcttccc	aagtataatg	tcactgagcc	ccagatctac	tttgatattt	5340
gggtctccat	caatgaccgc	ttccagcaga	ggatttttga	ccctcgtgtg	gacatcgtgc	5400
aggccgcttg	gtcacccttt	cagcgcacat	cctgggtgca	accactcttg	atggacctgt	5460
ctccctggag	ggccaagtta	caggaaatca	agagcagcct	agacaaccac	actgaggtgg	5520
tcttcattgc	agatttccct	ggactgcact	tggagaattt	tgtgagtga	gacctgggca	5580
acactagcat	ccagctgctg	cagggggaag	tgactgtgga	gcttgtggca	gaacagaaga	5640
accagactct	tcgagaggga	gaaaaaatgc	agttgcctgc	tggtgagtac	cataaggtgt	5700
atacgacatc	acctagccct	tcttgcctaca	tgtacgtcta	tgtcaacact	acagagcttg	5760
cactggagca	agacctggca	tatctgcaag	aattaaagga	aaaggtggag	aatggaagtg	5820
aaacagggcc	tctaccccca	gagctgcagc	ctctgttggga	aggggaagta	aaagggggcc	5880

ctgagccaac acctctggtt cagacctttc ttagacgcca acaaaggctc caggagattg 5940
aacgccggcg aaatactcct ttccatgagc gattcttcgc cttcttggtg cgaaagctct 6000
atgtctttcg ccgcagcttc ctgatgactt gtatctcact tcgaaatctg atattaggcc 6060
gtccttcctt ggagcagctg gcccaggagg tgacttatgc aaacttgaga ccctttgagg 6120
cagttggaga actgaatccc tcaaacacgg attcttcaca ttctaactct cctgagtcaa 6180
atcctgatcc tgtccactca gagttctgaa gggggccaga tgttgaagg gcaattcgag 6240
tctagagggc ccgcctgat agacggtttt tcgcccttg acgttgaggt ccacgttctt 6300
taatagtgga ctcttgttcc aaactggaac aacactcaac cctatctcgg tctattcttt 6360
tgatttataa gggattttgc cgatttcggc ctattgggta aaaaatgagc tgatttaaca 6420
aaaatttaac gcgaattaat tctgtggaat gtgtgtcagt tagggtgtgg aaagtcccca 6480
ggctccccag caggcagaag tatgcaaagc atgcatctca attagtcagc aaccagggtg 6540
ggaaagtccc caggctcccc agcaggcaga agtatgcaa gcacatctct caattagtca 6600
gcaaccatag tcccgcctt aactccgcc atcccgcctt taactccgcc cagttccgcc 6660
cattctccgc cccatggctg actaattttt tttatttatg cagaggccga ggccgcctct 6720
gcctctgagc tattccagaa gtatgagga ggcttttttg gaggcctagg cttttgcaa 6780
aagctcccg gagcttgat atccattttc ggatctgac aagagacagg atgaggatcg 6840
tttcgcatga ttgaacaaga tggattgcac gcaggttctc cggccgcttg ggtggagagg 6900
ctattcggct atgactgggc acaacagaca atcggctgct ctgatgccgc cgtgttccgg 6960
ctgtcagcgc aggggcgcct gggtcttttt gtcaagaccg acctgtccgg tgccctgaat 7020
gaactgcagg acgaggcagc gcggctatcg tggtggcca cgacggcgt tccttgcca 7080
gctgtgctcg acgttgac tgaagcggga agggactggc tgctattggg cgaagtgcgc 7140
gggcaggatc tcctgtcacc tcacctgct cctgccgaga aagtatccat catggctgat 7200
gcaatgcggc ggctgcatac gcttgatccg gctacctgcc cattcgacca ccaagcga 7260
catcgcatcg agcagcacg tactcgatg gaagccggtc ttgtcgatca ggatgatctg 7320
gacgaagagc atcaggggct cgcgccagc gaactgttcg ccaggctcaa ggcgcgcag 7380
cccgcggcg aggatctcgt cgtgacctat ggcatgcct gcttgccgaa tatcatggtg 7440
gaaaatggcc gcttttcttg attcatcgac tgtggccggc tgggtgtggc ggaccgctat 7500
caggacatag cgttggctac ccgtgatatt gctgaagagc ttggcggcga atgggctgac 7560
cgcttctctg tgctttacgg tatcgccgt cccgatccgc agcgcacgc cttctatcgc 7620
cttcttgacg agttcttctg agcgggactc tggggttcga aatgaccgac caagcgacgc 7680
ccaacctgcc atcacgagat ttcgattcca ccgcgcctt ctatgaaagg ttgggcttcg 7740
gaatcgtttt ccgggacgcc ggctggatga tcctccagcg cggggatctc atgctggagt 7800
tcttcgcca ccccaacttg tttattgcag cttataatgg ttacaaataa agcaatagca 7860
tcacaaattt cacaaataaa gcattttttt cactgcattc tagttgtggt ttgtccaaac 7920
tcacaaatgt atcttatcat gtctgtatc cgctgcacct tagctagagc ttggcgtaat 7980
catggtcata gctgtttcct gtgtgaaatt gttatccgct cacaattcca cacaacatac 8040
gagccggaag cataaagtgt aaagcctggg gtgcctaagt agtgagctaa ctcacattaa 8100
ttgcgttgcg ctcaactgcc gctttccagt cgggaaacct gtcgtgccag ctgcattaat 8160

gaatcggcc	acgcgcggg	agaggcggt	tgcgtattg	gcgctcttc	gcttctctgc	8220
tcactgactc	gctgcgctcg	gtcgcttcggc	tcggcgagc	ggtatcagct	cactcaaagg	8280
cggtaatacg	gttateccaca	gaatcagggg	ataacgcagg	aaagaacatg	tgagcaaaa	8340
gccagcaaaa	ggccaggaac	cgtaaaaagg	ccgcgttgct	ggcgtttttc	cataggctcc	8400
gccccctga	cgagcatcac	aaaaatcgac	gctcaagtca	gagggtggcg	aacccgacag	8460
gactataaag	ataccaggcg	tttccccctg	gaagctccct	cgtgcgctct	cctgttccga	8520
ccctgcgct	taccggatac	ctgtccgctt	ttctcccttc	gggaagcgtg	gcgctttctc	8580
atagctcacg	ctgtaggtat	ctcagttcgg	tgtaggtcgt	tcgctccaag	ctgggctgtg	8640
tgcacgaacc	ccccgttcag	cccgaaccgt	gcgccttctc	cggtaactat	cgtcttgagt	8700
ccaacccgg	aagacacgac	ttatcgccac	tggcagcagc	cactggtaac	aggattagca	8760
gagcgaggta	tgtaggcggt	gctacagagt	tcttgaagt	gtggcctaac	tacggctaca	8820
ctagaagaac	agtatttgg	atctgcgctc	tgtgaagcc	agttaccttc	ggaaaaagag	8880
ttggtagctc	ttgatccggc	aaacaaacca	ccgctggtag	cggttttttt	gtttgcaagc	8940
agcagattac	gcgcagaaaa	aaaggatctc	aagaagatcc	tttgatcttt	tctacggggt	9000
ctgacgctca	gtggaacgaa	aactcacgtt	aagggatttt	ggcatgaga	ttatcaaaaa	9060
ggatcttcac	ctagatcctt	ttaaattaaa	aatgaagttt	taaatcaatc	taaagtatat	9120
atgagtaaac	ttggtctgac	agttaccaat	gcttaatcag	tgaggcacct	atctcagcga	9180
tctgtctatt	tcgttcatcc	atagttgctt	gactccccgt	cgtgtagata	actacgatac	9240
gggagggtt	accatctggc	cccagtgctg	caatgatacc	gcgagacca	cgctcaccgg	9300
ctccagattt	atcagcaata	aaccagccag	ccggaagggc	cgagcgaga	agtggctctg	9360
caactttatc	cgcctccatc	cagtcctatta	attgttgccg	ggaagctaga	gtaagtagtt	9420
cgccagttaa	tagtttgctc	aacgttggtg	ccattgctac	aggcatcgtg	gtgtcacgct	9480
cgtcgtttgg	tatggcttca	ttcagctccg	gttcccaacg	atcaaggcga	gttacatgat	9540
ccccatggt	gtgcaaaaaa	gcggttagct	ccttcgggtc	tccgatcggt	gtcagaagta	9600
agttggcgc	agtgttatca	ctcatgggta	tggcagcact	gcataattct	cttactgtca	9660
tgccatccgt	aagatgcttt	tctgtgactg	gtgagtactc	aaccaagtca	ttctgagaat	9720
agtgtatg	gcgaccgagt	tgctcttgcc	cgcggtcaat	acgggataat	accgcgccac	9780
atagcagaac	tttaaaagt	ctcatcattg	gaaaacgttc	ttcggggcga	aaactctcaa	9840
ggatcttacc	gctgttgaga	tccagttcga	tgtaaccac	tcgtgcaccc	aactgatctt	9900
cagcatcttt	tactttcacc	agcgtttctg	ggtgagcaaa	aacaggaagg	caaatgccg	9960
caaaaaagg	aataagggcg	acacggaaat	gttgaatact	catactcttc	ctttttcaat	10020
attattgaag	cattttatcag	ggttattgtc	tcatgagcgg	atacatattt	gaatgtattt	10080
agaaaaataa	acaaatagg	gttccgcgca	catttccccg	aaaagtgcc	cctgacgtc	10139