



(86) Date de dépôt PCT/PCT Filing Date: 2002/02/08

(87) Date publication PCT/PCT Publication Date: 2002/08/29

(45) Date de délivrance/Issue Date: 2012/01/24

(85) Entrée phase nationale/National Entry: 2003/08/20

(86) N° demande PCT/PCT Application No.: EP 2002/001306

(87) N° publication PCT/PCT Publication No.: 2002/066628

(30) Priorité/Priority: 2001/02/20 (DE101 08 100.6)

(51) Cl.Int./Int.Cl. *C12N 15/62* (2006.01),
C07K 1/30 (2006.01), *C07K 1/36* (2006.01),
C07K 14/62 (2006.01), *C07K 14/815* (2006.01),
C12N 15/81 (2006.01), *C12P 21/02* (2006.01)

(72) Inventeur/Inventor:
HABERMANN, PAUL, DE

(73) Propriétaire/Owner:
SANOFI-AVENTIS DEUTSCHLAND GMBH, DE

(74) Agent: BERESKIN & PARR LLP/S.E.N.C.R.L.,S.R.L.

(54) Titre : UTILISATION DE PEPTIDES APTES A L'HYPERSECRETION DANS DES PROCEDES PERMETTANT LEUR PRODUCTION ET L'AMELIORATION EN PARALLELE DES FORMES EXPORTEES D'UN OU DE PLUSIEURS POLYPEPTIDES D'INTERET

(54) Title: SUPERSECRETABLE PEPTIDES, PROCESSES FOR THEIR PRODUCTION, AND PARALLEL IMPROVEMENT OF THE SECRETED FORM OF ONE OR MORE OTHER POLYPEPTIDES

(57) Abrégé/Abstract:

The invention relates to an expression cassette of the form Px-Sx-Bn-(ZR)-transport peptide-(Z₁Z₂)-protein(Y)-(Z₁Z₂)-protein (Y_m)-T; wherein the expression cassette codes for a transport peptide linked via a sequence Z₁Z₂ to a second protein which in turn is linked via Z₁Z₂ to a protein Y1 which either corresponds to Y or can be different from Y and the transport peptide improves the rate of secretion and Y and/or Y_m where Px is any promoter DNA sequence selected in such a way that optimal yields of the protein of interest become obtainable; Sx is any DNA which, accordingly, encodes any signal or leader sequence which allows optimal yields; B_n is 1-15 genetically encodable amino acids or a chemical bond; Z is the codon of an amino acid selected from the group comprising Lys and Arg; Z₁ is the codon of an amino acid selected from the group comprising Lys and Arg; Z₂ is the codon of an amino acid selected from the group comprising Lys and Arg; protein Y_m is a DNA sequence encoding any protein which can be produced and secreted by yeast (m=1-5) or is a chemical bond (m=0); R is an arginine codon; transport peptide is a DNA sequence encoding a peptide which can be transported efficiently and which can pass membranes; protein Y is a DNA sequence encoding any protein which can be produced and secreted by yeast and whose biological activity, when Y_m is not a chemical bond, is not impaired by a basic dipeptide extension or allows degradation of the extension by carboxypeptidases; T is an untranslated DNA sequence advantageous to the expression, and to processes for the production and use thereof.



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

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
29 August 2002 (29.08.2002)

PCT

(10) International Publication Number
WO 02/066628 A3

(51) International Patent Classification⁷: **C12N 15/62**,
C07K 14/815, C12N 15/81, C07K 1/30, 1/36, 14/62

GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZM, ZW.

(21) International Application Number: PCT/EP02/01306

(22) International Filing Date: 8 February 2002 (08.02.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
101 08 100.6 20 February 2001 (20.02.2001) DE

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

- with international search report
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments

(71) Applicant: **AVENTIS PHARMA DEUTSCHLAND GMBH** [DE/DE]; Brüningstrasse 50, 65929 Frankfurt (DE).

(72) Inventor: **HABERMANN, Paul**; Rossertstrasse 35, 65817 Eppstein (DE).

(88) Date of publication of the international search report:
14 November 2002

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH,

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: SUPERSECRETABLE PEPTIDES, PROCESSES FOR THEIR PRODUCTION, AND PARALLEL IMPROVEMENT OF THE SECRETED FORM OF ONE OR MORE OTHER POLYPEPTIDES

(57) Abstract: The invention relates to an expression cassette of the form Px-Sx-Bn-(ZR)-transport peptide-(Z₁Z₂)-protein(Y)-(Z₁Z₂)-protein (Y_m)-T; wherein the expression cassette codes for a transport peptide linked via a sequence Z₁Z₂ to a second protein which in turn is linked via Z₁Z₂ to a protein Y1 which either corresponds to Y or can be different from Y and the transport peptide improves the rate of secretion and Y and/or Y_m where Px is any promoter DNA sequence selected in such a way that optimal yields of the protein of interest become obtainable; Sx is any DNA which, accordingly, encodes any signal or leader sequence which allows optimal yields; B_n is 1-15 genetically encodable amino acids or a chemical bond; Z is the codon of an amino acid selected from the group comprising Lys and Arg; Z₁ is the codon of an amino acid selected from the group comprising Lys and Arg; Z₂ is the codon of an amino acid selected from the group comprising Lys and Arg; protein Y_m is a DNA sequence encoding any protein which can be produced and secreted by yeast (m=1-5) or is a chemical bond (m=0); R is an arginine codon; transport peptide is a DNA sequence encoding a peptide which can be transported efficiently and which can pass membranes; protein Y is a DNA sequence encoding any protein which can be produced and secreted by yeast and whose biological activity, when Y_m is not a chemical bond, is not impaired by a basic dipeptide extension or allows degradation of the extension by carboxypeptidases; T is an untranslated DNA sequence advantageous to the expression, and to processes for the production and use thereof.



WO 02/066628 A3

Description

Use of supersecretable peptides in processes for their production and for parallel improvement of the exported forms of one or more other polypeptides of interest

5

In view of economic viability, processes for the production of pharmaceutically relevant proteins must lead to biologically active products of the highest possible purity. The expression of such relevant proteins in yeasts is widely used here. The production of proteins such as insulin, GM-CSF (Leukine[®]) and hirudin (Refludan[®]) is an example
10 of the successful development of genetic engineering processes which are based on the synthesis of the particular protein or precursors thereof in yeast. Generally, yeasts can directly synthesize particularly hirudins with good yields which are on the gram scale when using *Hansenula polymorpha* (Weydemann et al., Appl. Microbiol Biotechnol. 44: 377 –385, 1995) or *Pichia pastoris* (Rosenfeld et al., Protein Expr. Purif
15 :4 , 476 –82, 1996).

EP-A 0 324 712 describes the hirudin derivative (Refludan[®]) whose N-terminal amino acid is leucine and its constitutive expression in *Saccharomyces cerevisiae* strain Y79. EP-A 0 347 781 describes a mini-proinsulin and, by way of example, its expression in
20 bakers' yeast. Refludan[®] and insulin are produced by carrying out two separate expressions.

Surprisingly, we have now found that hirudin derivatives and mini-proinsulin derivatives can be obtained from a common precursor protein by fusing the precursor protein to a
25 signal or leader sequence, which is recognized by yeasts as a secretion signal, via a basic dipeptide, preferably Lys – Arg, and likewise introducing between the N-terminal hirudin derivative and the mini-proinsulin derivative a cleavage site which is recognized by a yeast endoprotease. Here too, preference is given to a basic dipeptide, for example Lys - Arg. After expression, a hirudin derivative extended by Lys - Arg and the
30 mini-proinsulin derivative starting with the first amino acid of the insulin B chain are found in the supernatant. Surprisingly, we have found here that the yield of mini-proinsulin markedly improves compared with the yield achievable by direct signal –

mini-proinsulin expression, whereas the yield of the hirudin derivative remains nearly the same. Surprisingly, hirudin thus acts as a kind of enhancer peptide with respect to the yield of mini-proinsulin.

5 Peptides which can act as enhancer proteins are usually those which are relatively small and which are secreted naturally in large amounts over a short period, for example from glandular tissue. Peptides of this type, which include, for example, snake venom or eglin C or TAP (tick anticoagulant peptide), are distinguished by extremely good export compatibility. The invention relates to such proteins, too.

10

Another advantage may result from the hirudin derivative having equal or better pharmaceutical properties compared with hirudin which is already used in pharmaceuticals. In this case, it becomes possible to produce two or even more pharmaceuticals from one and the same fermentation. As a consequence, less
15 fermentation capacity is required. This is directly beneficial to production costs.

However, the production of a plurality of products is optional. The amount needed of Refludan, for example, is less than that of insulin, and this may result in processes in which one of the pharmaceutically interesting substances is discarded.

20

To improve the yield, it is possible, as suggested in patent application EP-A 0 200 655, to place a short peptide sequence via Lys- Arg N-terminally in front of the hirudin derivative, as a linker to the signal or leader sequence. It is also obvious to the skilled worker that the choice of signal or leader sequence directly affects the yield of
25 the protein of interest. The selection of such a sequence is the subject of further optimizations. The sequence located at the 3' end of the expression cassette, too, directly affects the yield by influencing mRNA stability. Here too, it is obvious to the skilled worker that said sequence can be optimized for each protein of interest to be expressed. This is also true for the choice of a suitable promoter, which can be
30 inducible or constitutively active. The choice of vector system and host system is equally important for the yield. Thus, instead of bakers' yeast which has been used by way of example, it is also possible to use the yeasts *Pichia pastoris*, *Hansenula*

polymorpha or *K. lactis* together with vectors or expression cassettes, which have in each case been optimized for the different physiology.

Another advantage of processes allowing secretion into the medium is the simpler
 5 protein-chemical workup of the protein of interest. Surprisingly, we have found that mini-proinsulin can be concentrated in the presence of hirudin by filtration through membranes having an exclusion limit for molecules with a molecular weight of greater than 10 kDa. The molecules are found almost exclusively in the retentate. It is obvious to the skilled worker that the development of novel separation techniques and new
 10 combinations of process steps always make it possible to improve purification processes. This is directly beneficial to the yield and therefore to production costs.

The invention thus relates to a DNA-molecule (alternative term: expression cassette) of the form:

15

$P_x-S_x-B_n-(ZR)-\text{transport peptide}-(Z_1Z_2)-\text{protein}(Y)-(Z_1Z_2)-\text{protein}(Y_m)-T$;

wherein the expression cassette codes for a transport peptide linked via a sequence Z_1Z_2 to a second protein which in turn is linked via Z_1Z_2 to a protein Y_1 which either corresponds to Y or can be different from Y and the transport peptide improves the
 20 rate of secretion of Y and/or Y_m , where:

P_x is any promoter DNA sequence selected in such a way that optimal yields of the protein of interest become obtainable;

S_x is any DNA which, accordingly, encodes any signal or leader sequence which
 25 allows optimal yields;

B_n is 1-15 genetically encoded amino acids or a chemical bond;

Z is the codon of an amino acid selected from the group comprising Lys and Arg;

Z_1 is the codon of an amino acid selected from the group comprising Lys and Arg;

Z_2 is the codon of an amino acid selected from the group comprising Lys and Arg;

30 R is an Arg codon;

transport peptide is a DNA sequence encoding a peptide which can be transported efficiently and which can pass membranes, such as hirudin or a hirudin derivative, for example;

protein Y is a DNA sequence encoding any protein which can be produced and
5 secreted by yeast;

protein Y_m is a DNA sequence encoding any protein which can be produced and secreted by yeast ($m = 1-5$) or is a chemical bond ($m = 0$);

T is an untranslated DNA sequence advantageous to the expression.

10 Another embodiment of the invention is a fusion protein encoded by any of the above-mentioned DNA molecules.

A further embodiment of the invention is a multicopy vector and a plasmid comprising the above-mentioned DNA-molecule.

15

An additional embodiment of the invention is a host cell comprising the above-mentioned DNA-molecule, or the above-mentioned multicopy vector or the above-mentioned plasmid, as a part of its chromosome, as a part of a mini-chromosome, or extra-chromosomally, wherein preferentially said host cell is a yeast, in particular
20 selected from the group comprising of *S. cerevisiae*, *K. lactis*, *H. polymorpha* and *P. pastoris*.

Another embodiment of the invention is a process of fermenting the above-mentioned proteins, in which

25

(a) the above-mentioned DNA-molecule, the above-mentioned multicopy vector, or the above-mentioned plasmid is expressed in an above-mentioned host cell, and

(b) the expressed proteins are isolated from the supernatant of the cell culture,
30 wherein in particular after completion of fermentation, the pH is adjusted to 2,5-3,5 in order to precipitate non-desired proteins and the expressed proteins are isolated from the supernatant of the precipitation.

Another embodiment of the invention is the above mentioned process, in which process after separating the fermentation supernatant from the host cells, the host cells are repeatedly cultured in fresh medium, and the released fusion protein is isolated from each supernatant obtained during cultivation.

5

Another embodiment of the invention is the above mentioned process, wherein a process step for concentrating the expressed protein in the supernatant after precipitation is selected from a group comprising microfiltration, hydrophobic interaction chromatography and ion exchange chromatography.

10

An additional embodiment of the invention is a process for preparing insulin, in which

- (a) proinsulin is expressed as protein (Y) of the above-mentioned expression cassette in the above-mentioned process;
- 15 (b) the proinsulin of step (a) is isolated and treated with trypsin and carboxypeptidase B; and
- (c) insulin is isolated from the reaction mixture of step (b),

in particular, wherein the transport peptide is hirudin or a hirudin derivative which is
20 destroyed or biologically inactivated after step (a) or (b).

A further embodiment of the invention is a protein wherein the protein is a hirudin-derivative with two basic amino acid residues at its C-terminal end.

25 Leeches of the type *Hirudo* have developed, for example, various isoforms of the thrombin inhibitor hirudin. Hirudin has been optimized for pharmaceutical requirements by artificial variation of the molecule, for example exchange of the N-terminal amino acid (e.g. EP 0 324 712). The invention includes the use of hirudin and hirudin variants. Particular embodiments of the invention use one of the natural hirudin
30 isoforms (the natural isoforms are together denoted "hirudin"). A natural isoform is, for example, Val-Val-hirudin or Ile-Thr-hirudin. Other embodiments of the invention use a variant of a natural hirudin isoform. A variant is derived from a natural hirudin isoform

but contains, for example, additional amino acids and/or amino acid deletions and/or amino acid exchanges compared with the natural isoform. A hirudin variant may contain alternating peptide segments of natural hirudin isoforms and new amino acids. Hirudin variants are known and are described, for example, in DE 3 430 556. Hirudin
5 variants are commercially available in the form of proteins (Calbiochem Biochemicals, cat. no. 377-853, -950, -960). The term "hirudin derivative" denotes sequences which are at least 40% homologous to natural hirudin.

The expression cassette is preferably introduced into yeasts such as *S. cerevisiae*, *K. lactis*, *H. polymorpha* or *P. pastoris*. Said expression cassette may have one or more
10 copies stably integrated into the particular yeast genome or may be present extrachromosomally on a multicopy vector. It is obvious to the skilled worker that this technique is also applicable to other systems such as animal cell culture or plant cells. This is also a subject of the invention.

15

The expression system described below serves as an example. It is obvious to the skilled worker that, in order to introduce the expression cassette into said selected system, the appropriate recombinant DNA constructs must be made depending on the type of host system selected. Accordingly, industrial fermentation can be optimized in
20 relation to the selected host/vector system. Accordingly, the examples are nonrestrictive.

Example 1: Construction of a yeast expression plasmid encoding hirudin (Refludan) – Lys-Arg – mini-proinsulin

25

Starting materials are the plasmids pK152 (PCT/EP00/08537), pSW3 (EP-A 0 347 781) and the recombinant yeast plasmid derivative coding for bovine interleukin 2 (Price et al. Gene 55, 1987). The yeast plasmid is distinguished by the fact that it carries the α factor leader sequence under the control of the yeast ADH2
30 promoter. This sequence is followed by the bovine interleukin 2 cDNA sequence which is connected via a KpnI restriction enzyme recognition site and which contains, after manipulation, an NcoI restriction enzyme recognition site in the untranslated 3' end

which is unique in the vector. Thus, the cDNA sequence can readily be removed from the plasmid via KpnI/NcoI cleavage. Since good expression yields have been reported, it can be assumed that the remaining 3' interleukin 2 sequence (as T) has a stabilizing effect on the mRNA and thus need not be deleted or replaced by a yeast terminator sequence . Plasmid pK152 carries the DNA sequence coding for Leu–hirudin (Refludan) and plasmid pSW3 carries the DNA sequence for mini-proinsulin. The gene sequence which is to encode hirudin – Lys Arg – mini-proinsulin is first prepared by means of PCR technology. For this purpose, 4 primers are prepared with the aid of the Expedite[™] DNA synthesis system:

10

i. hir_insfkr (SEQ ID NO: 1, encoded protein segment: SEQ ID NO: 2)

I P E E Y L Q K R F V N Q H L C

5' - ATCCCTGAGGAATACCTTCAGAAAGCGATTTGTTAACCAACACTTGTGTGG-3 '

59 60 61 62 63 64 65 B1 B2 B3 B4 B5 B6 B7

15

ii. hir_insrevkr (SEQ ID NO: 3)

5' - CCTCACAAGTG TTGGTTAACA AATCGCTTCT GAAGGTATTC CTCAGGGAT-3 '

20 iii. hirf1 (SEQ ID NO: 4, encoded protein segment: SEQ ID NO: 5)

L T Y T D C

5' - TTTTTTTGGATCCTTTGGATAAAAGACTTACGTATACTGACTGCAC

25 iv. insnco1rev (SEQ ID NO: 6)

5' - TTTTTTCCAT GGGTCGACTATCAG

Primer hir_insfkr describes the junction between codons for the terminal amino acids of
30 hirudin (59 – 65) and the insulin sequence B1 – B7 via the Lys – Arg linker. Primer
hir_insrevkr is 100% complementary thereto. Primer hirf1 codes for the start of the
hirudin gene extended to the KpnI cleavage site as described in EP-A 0 324 712.

Primer insncoirev marks the 3' end of the synthetic mini-proinsulin according to EP-A 0 347 781. Two standard polymerase chain reactions are carried out using the primer pairs hirf1/hir_insrevkr with DNA of plasmid pK152 as template and hir_insfrkr/insncoirev with DNA of plasmid pSW3 as template. The reactions are carried out in 100 µl of PCR buffer with, in each case, 200 nmol of primer, 1 µl of polymerase and 100 ng of vector. Step 1 is a 2-minute incubation at 95°C. This is then followed by 25 cycles of 30'' at 95°C, 30'' at 55°C and 30'' at 72°C. The last cycle is followed by an incubation at 72 °C for 3 minutes, and the reaction is subsequently stopped.

10 Since the primers hir_insrevkr and hir_insfrkr are 100% complementary, the DNA products of the two products overlap according to said sequence so that in a third reaction, using the products of the first two reactions as templates and the primers hirf1 and insncoirev, a DNA fragment is formed, which encodes hirudin and mini-proinsulin separated by Lys – Arg. The PCR fragment is digested by the enzymes KpnI und NcoI and then, in a T4 ligase reaction, inserted into the pαADH2 vector opened by KpnI / NcoI. In analogy to example 7 of EP-A 0 347 781, competent E. coli strain MM294 cells are then transformed with the ligation mixture. Plasmid DNA is then isolated from two clones for characterization by means of DNA sequence analysis. After confirmation of the inserted DNA sequence, DNA of a plasmid preparation is used to transform cells of bakers' yeast strain Y79, according to said example. However, when using the pαADH2 vector, introduction of the vector is followed by selecting for complementation of the trp1-1 mutation, in contrast to said example. For another control, plasmid DNA is reisolated from yeast transformants and analyzed by means of restriction analysis. The expression vector constructed is denoted pADH2Hir_KR_Ins. Expression is carried out according to example 4.

Example 2: Construction of a yeast expression plasmid encoding hirudin (Refludan) – Lys-Arg – insulin B chain – Lys-Arg – insulin A chain

30

Patent application EP-A 0 195 691 describes proinsulin derivatives which can contain the dipeptide XY, where X and Y each correspond to either Lys or Arg, as linker

between the B and A chains of insulin. The following example describes the preparation of an expression vector for proinsulin derivatives of this kind. A DNA sequence which codes for a proinsulin derivative of the form B chain – Lys-Arg – A chain is selected by way of example and synthesized accordingly.

5

The synthesis of the gene segment is carried out according to example 1. The oligonucleotide sequences used are hirf1 and insncoirev. The oligonucleotides B_KR_Af1 and B_KR_Arev1 are synthesized de novo.

10 B_KR_Af1 has the sequence (SEQ ID NO: 7)

5' CTTCTACACTCC**AAAGACGAAACGCGGTATCG**-3'

and B_KR_Arev1 has the sequence (SEQ ID NO: 8)

15

5' CAACATTGTTCAACGATACCGCGTTTCGTCTTT-3'

The part shown in bold type of the two primers depicted indicates the partially overlapping sequence. Both primers pair exactly with the sequence of the mini-proinsulin gene of EP-A 0 347 781, apart from the 6 underlined nucleotides. The underlined part corresponds to codons for Lys and Arg. DNA of the plasmid pADH2Hir_KR_Ins constructed according to example 1 serves as template in the PCR.

As described in example 1, two polymerase chain reactions are carried out using the primer pairs hirf1 / B_KR_Arev and insncoirev / B_KR_Af1. The template in each case is DNA of the plasmid pADH2Hir_KR_Ins constructed in example 1. The products of both reactions serve as template in a third PCR using the primer pair hirf1 and insnco1. The reaction product from PCR 3 is cleaved with NcoI/ Sall and inserted into the opened pαADH2 vector. After sequence and restriction analysis, the correct plasmid is referred to as pADHHirKR_B_KR_A.

30

Example 3: Construction of a yeast plasmid coding for hirudin – Lys-Arg – simian proinsulin

Patent application EP-A 489 780 describes a plasmid, pINT90d, which contains cDNA of simian proinsulin (Wetekam et al., Gene 19, p.179-183, 1982). DNA of said plasmid and DNA of plasmid pK152 serve as templates. The primer hirf1 described in example 1 is used and three further primers are synthesized.

Primer insncorev reversely binds to the 3' region of the insulin gene cloned in pINT90d and has the sequence:

5'-TTTTTCCATGGTCATGTTTGACAGCTTATCAT-3' (SEQ ID NO: 9)

The underlined sequence indicates the recognition site for the restriction enzyme NcoI.

Primer hir_insfkr has the sequence:

5'-ATCCCTGAGG AATACCTTCA **GAAGCGATTT** GTGAACCAGC ACCTGTGCGG C-3'
(SEQ ID NO: 10)

Here, the nucleotides in bold type indicate the Lys-Arg linker between hirudin and proinsulin.

Primer hir_insrevkr is completely complementary to primer hir_inskr and has the sequence:

5' -GCCGCACAGG TGCTGGTTCA CAAAT**CGCTT** CTGAAGGTAT TCCTCAGGGA T-3'
(SEQ ID NO: 11)

Corresponding to example 1, two polymerase chain reactions are carried out. The primer pair hirf1 / hir_insrevkr is reacted with DNA of plasmid pK152 and the primer pair hir_insfkr / insncorev is reacted with DNA of plasmid pINT91d. As described in

example 1, the products of both reactions serve as template in a third PCR using the primer pair hirf1 / insncorev. The DNA product of this reaction includes the sequence for hirudin_Lys-Arg_proinsulin. It is subsequently cleaved with the enzymes NcoI and KpnI and, corresponding to example 1, inserted into the plasmid p α ADH2. Accordingly
 5 expression vector for any natural proinsulin derivatives may be constructed.

Example 4: Expression of the recombinant products

The expression is divided into two phases. Firstly, a preculture is cultivated in yeast
 10 minimal medium. The medium has the following composition per 1 l:

	6.7 g	-	yeast nitrogen base (without amino acids)
	5.0 g	-	casamino acids (vitamin-free)
	0.008%	-	adenine
	0.008%	-	uracil
15	2%	-	glucose

The main or expression culture is inoculated with an aliquot of the preculture.

The main culture medium contains per liter:

20

	10 g	-	yeast extract
	20 g	-	peptone
	0.008%	-	adenine
	0.008%	-	uracil
25	4%	-	glucose

Using the media described, expression is carried out in a shaken flask in the following way: 0.3 ml of a preculture which has been cultivated overnight is diluted with 80 ml of prewarmed medium and incubated with vigorous shaking at 30°C for approx. 24 h. In
 30 each case, 1 ml of the culture produced in this way is then centrifuged, after determining the optical density, and, after removing the cells, the supernatant is lyophilized and analyzed by means of SDS –PAGE. The biologically active hirudin

content is determined by carrying out a thrombin inhibition assay.

An alternative fermentation protocol provides for the cells to be removed by filtration or careful centrifugation. While isolating the protein of interest from the medium, the cells
5 are provided with fresh prewarmed main culture medium containing alcohol and not more than 0.5% glucose as carbon source, and thus fermentation is continued without interruption. This step can be repeated up to 5 times.

10 Example 5: Thrombin inhibition test

The hirudin concentration is determined according to the method of Griebach et al. (Thrombosis Research 37, pp. 347 –350 , 1985). For this purpose, specific amounts of a Refludan standard are included in the measurements in order to establish a
15 calibration curve from which the yield in mg/l can be determined directly.

Example 6: Cloning and expression of the hirudin – Lys-Arg – mini-proinsulin fusion protein in the *P. pastoris* system

20

Invitrogen® sells a cloning and expression kit for preparing recombinant proteins using *P. pastoris* as host system. For this, a detailed technical protocol regarding preparation and subsequent expression of a *P. pastoris* system for the production of a desired recombinant protein is provided so that only the construction of the expression vector
25 encoding the desired protein has to be described when following said protocols. The EasySelect™ Pichia expression kit (catalog no. K1740-01) is used.

The pPICZαA vector is part of the kit. Opening the vector by the restriction enzymes XhoI and SacII makes it possible to append, similar to example 1, a protein of interest
30 to the alpha factor leader sequence and to test for secretion into the supernatant.

Cloning requires two primers. Primer pichia_H_If1 (SEQ ID NO: 12) has the sequence:

5' - TTTTTTCTCGAGAAAAGA CTTACGTATACTGAC - 3'

XhoI

Hir₁ Hir₂ etc.

Primer pichia_H_Irev2 (SEQ ID NO: 13) has the sequence:

5

5' - TTTTGGCGCCGAATTCACCTATTAGTTACAGTAGTTTCC - 3'

SacII

EcoRI

A21

The template used is DNA of plasmid pADH2Hir_KR_Ins. A standard PCR with both
 10 primers produces a DNA product which contains the sequence hirudin – Lys-Arg –
 mini-proinsulin extended by the XhoI and SacII integration sites. If the DNA product is
 cleaved appropriately and the fragment is isolated, said fragment can be inserted into
 the opened vector DNA in a T4 DNA ligase reaction. In deviation from the
 manufacturer's protocol, E. coli strain MM294, described in example 1, is transformed
 15 with the ligation mixture and recombinant colonies are screened for on zeocine
 selection plates. Plasmid DNA is reisolated from clones and then characterized by
 means of restriction and DNA sequence analysis. Using the plasmid constructed in this
 way, a P. pastoris expression clone for production of the peptides is then prepared,
 following the manufacturer's instructions.

20

Example 7: Purification of mini-proinsulin and hirudin

The purification requires separation of the two proteins at an early stage. After
 25 completion of the expression, the medium is analyzed by means of analytical RP-
 HPLC. In contrast to most other polypeptides found in the supernatant due to either
 spontaneous lysis of yeast cells or secretion, the two proteins, hirudin and mini-
 proinsulin, are not precipitated at pH 2.5–3. The culture medium is therefore acidified
 appropriately and then, after completion of the precipitation, the precipitate and the
 30 cells are removed by centrifugation. After centrifugation, the medium is adjusted to pH
 3.5–7 and the two components hirudin and mini-proinsulin are separated from one
 another by means of hydrophobic interaction chromatography, for example by using a

chromatography column filled with Diaion HP20[®] material. Hirudin can then be isolated from the hirudin-containing fractions according to EP-A 0 549 915 and insulin can be isolated from the mini-proinsulin-containing fractions according to EP-A 0 347 781.

5

Example 8: Preparation of insulin from mini-proinsulin

At the end of the expression period, the culture medium is adjusted to pH 6.8 and
10 trypsin is then added with stirring so that a final concentration of 4–8 mg per liter is established. After incubation for approx. 4 hours, the fermentation broth treated in this way is adjusted to pH 2.5–3. After 1–6 hours of precipitation, the precipitate is removed. The mono-Arg-insulin formed is then isolated via ion exchange chromatography, by way of example on S – Sepharose[®] in a buffer of 50 mM lactic
15 acid and 30% isopropanol (pH 3.5). Elution is carried out by means of an NaCl gradient of 0.05–0.5 M salt. The product-containing fractions are diluted 1:1 with H₂O and then ZnCl₂ is added, so that a 0.1% strength ZnCl₂ solution is formed. Mono-Arg-insulin precipitates at pH 6.8 and by way of example is converted to insulin according to EP-A 0 324 712.

20

Example 9: Preparation of insulin from mini-proinsulin after filtration

At the end of the expression period, cells and supernatant components are removed
25 by precipitation at pH 2.5 to 3. Then the medium is concentrated via filtration through membranes having an exclusion limit of 10 kDa. Like the hirudin derivative, mini-proinsulin is found quantitatively in the retentate and can then be processed to insulin according to example 8.

15

SEQUENCE LISTING

<110> Aventis Pharma Deutschland GmbH

<120> Use of supersecretable peptides in processes for their
production and for parallel improvement of the exported
forms of one or more other polypeptides of interest

<130> 9982-766

<140> 2,438,935

<141> 2002-02-08

<150> 10108100.6

<151> 2001-02-20

<160> 13

<170> PatentIn Ver. 2.1

<210> 1

<211> 50

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:hir_insfkr

<400> 1

atccctgagg aataccttca gaagcgattt gttaaccaac acttgtgtgg 50

<210> 2

<211> 16

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:protein
hir_insfkr

<400> 2

Ile	Pro	Glu	Glu	Tyr	Leu	Gln	Lys	Arg	Phe	Val	Asn	Gln	His	Leu	Cys
1					5				10					15	

<210> 3

<211> 50

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:hir_insrevkr

<400> 3

cctcacaagt gttggttaac aaatcgcttc tgaaggtatt cctcagggat 50

<210> 4

<211> 46

16

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:hirf1

<400> 4

tttttttgggaa tccttttggat aaaagactta cgtataactga ctgcac

46

<210> 5

<211> 6

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:protein hirf1

<400> 5

Leu Thr Tyr Thr Asp Cys

1

5

<210> 6

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:insncolrev

<400> 6

ttttttcccat gggtcgacta tcag

24

<210> 7

<211> 32

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:B_KR_Af1

<400> 7

cttctacact ccaaagacga aacgcggtat cg

32

<210> 8

<211> 33

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:B_KR_Arev1

<400> 8

caacattggt caacgatacc gcgttttcgtc ttt

33

<210> 9

<211> 33
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence:insncorev

<400> 9
 ttttttccat ggatcatgttt gacagcttat cat 33

<210> 10
 <211> 51
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence:hir_insfkr

<400> 10
 atccctgagg aataccttca gaagcgattt gtgaaccagc acctgtgcgg c 51

<210> 11
 <211> 51
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence:hir_insrevkr

<400> 11
 gccgcacagg tgctgggttca caaatcgctt ctgaaggtat tcctcaggga t 51

<210> 12
 <211> 34
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence:pichia_H_lf1

<400> 12
 tttttttctc gagaaaagac ttacgtatac tgac 34

<210> 13
 <211> 40
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence:pichia_H_lrev2

<400> 13
 tttttggcgc cgaattcact attagttaca gtagttttcc 40

What is claimed is:

1. A nucleic acid which comprises:

P_x-S_x-(ZR)-nucleic acid encoding a transport peptide-(Z₁Z₂)-nucleic acid encoding a protein(Y)-T; wherein:

P_x comprises a promoter;

S_x comprises a nucleic acid encoding a signal or leader;

Z is a codon for lysine or arginine;

R is an arginine codon;

transport peptide comprises hirudin or lepirudin;

Z₁ is a codon for lysine or arginine;

Z₂ is a codon for lysine or arginine;

protein(Y) comprises mini-proinsulin or proinsulin; and

T is a terminator.

2. The nucleic acid according to claim 1 wherein the hirudin is Val-Val-hirudin or Ile-Thr-hirudin.

3. A protein encoded by the nucleic acid of claim 1 or claim 2.

4. A multicopy vector comprising the nucleic acid of claim 1 or claim 2.

5. A plasmid comprising the nucleic acid of claim 1 or claim 2.

6. A host cell comprising the nucleic acid of claim 1 or claim 2, a multicopy vector of claim 4, or a plasmid of claim 5 as a part of the host cell chromosome, as a part of a mini-chromosome, or extra-chromosomally.

7. The host cell of claim 6, wherein the host cell is a yeast.

8. The host cell of claim 7, wherein the yeast is selected from *Saccharomyces cerevisiae*, *Kluyveromyces fragilis*, *Hansenula polymorpha*, and *Pichia pastoris*.

9. A process of fermenting the protein according to claim 3 comprising:
- (a) culturing the host cell of any one of claims 6-8; and
 - (b) isolating the expressed protein from the supernatant of the cell culture of (a).
10. The process of claim 9, further comprising adjusting the pH to 2.5 to 3.5 in order to precipitate non-desired proteins.
11. The process of claim 10, further comprising repeatedly culturing the host cell in fresh medium after separating the supernatant from the host cell; and isolating the expressed protein from each supernatant obtained.
12. The process of any one of claims 9-11, wherein a process step for concentrating the expressed protein in the supernatant after precipitation is selected from microfiltration, hydrophobic interaction chromatography and ion exchange chromatography.
13. A process for the preparation of insulin comprising:
- (a) expressing proinsulin as protein(Y) from the nucleic acid of claim 1 or 2 using the process of any one of claims 9-12;
 - (b) isolating the proinsulin of (a) and treating with trypsin and carboxypeptidase B; and
 - (c) isolating insulin from the reaction mixture of (b).
14. The process of claim 13, wherein the transport peptide, hirudin or lepirudin, is destroyed or biologically inactivated after (a) or (b).