



159159

WO 01/21662

1

PCT/EP00/08537

Description

Signal sequences for preparing Leu-hirudin by secretion by *E. coli* into the culture medium

5

The leech-derived product Refludan[®] shows good therapeutic properties in clinical trials (The Lancet, Vol. 353, p. 429 – 438). This leads to the conclusion that larger amounts of the product are likely to be required in future. The biological active ingredient in the product is [Leu¹, Thr²]-63-desulfatohirudin which is described in European patent 0 324 712 and is called "Leu-hirudin" for short hereinafter.

10

European patent 0 448 093 describes a process for preparing hirudin. The preferred embodiment of the patent comprises a hirudin whose N-terminal amino acid consists of alanine. Fusion of this hirudin to the signal sequence of α -cyclodextrin glycosyltransferase (CGTase) and transformation of an expression vector encoding this fusion protein, as described in the patent, into an *E. coli* secretor mutant make it possible to prepare Ala-hirudin with crude yields of more than 2 grams per liter. European patent 0 549 915 describes variants of Ala-hirudin with improved stability. Preparation of these variants using the *E. coli* secretor system results in yields of several grams per liter. The yields are thus distinctly higher than described by Dodt *et al.* for the hirudin variant HV1 (FEBS LETTERS vol. 202 373–377 1986). A negligible increase in the yield compared therewith is described in US patent 5,573,929 by expressing the expression cassette via a pUC vector in a known manner in place of the pBR322-derived vector of Dodt *et al.* Bender *et al.* (Appl. Microbiol Biotechnol 34, p. 203 –207 1990) describe the secretion of Thr-hirudin, which is described in European patent 0 171 024, by *Streptomyces lividans*. However, the yields compared with the yields mentioned in European patents 0 448 093 and 0 549 915 are once again distinctly less. This also applies to expression in *E. coli* B as found by P. de Taxis du Poet *et al.* for secretion of the hirudin variant HV1 via the signal sequence Ompa of *E. coli*. The authors find yields of 300 mg/ l hirudin in the periplasm and about 40 mg/l in the cell supernatant. The expression in insect cell systems, which is also described in the article, is low (400 μ g/l).

15

20

25

30

35

Yields which can be achieved with the yeast expression systems *Hansenula polymorpha* or *Pichia pastoris* come closest to the yields



G00002

described in European patents 0 448 093 and 0 549 915, in contrast to the levels achieved with *S. cerevisiae*.

5 Rosenfeld *et al.* (Protein Expression and Purification 8 , 476 – 482, 1996) describe the expression and secretion of hirudin by the yeast *Pichia pastoris*. Yields of about 1.5 g/l of culture broth are achieved in this case. A similar order of magnitude can be achieved with the yeast *Hansenula polymorpha* (Appl. Microbiol. Biotechnol. 44, 377 – 385 1995). However, a considerable disadvantage of such expression systems is that the fermentation times are distinctly longer than for the *E. coli* system. It would thus be
10 advantageous if Leu-hirudin could, like Ala-hirudin, be prepared by secretion by *E. coli*.

However, this is not possible with the system described in European patent
15 0 448 093. For this reason, it is proposed in the patent to extend the Leu-hirudin sequence by the tripeptide Ala–Thr–Arg to produce a pre-Leu-hirudin which is finally converted after reaction with trypsin into the native active ingredient Leu-hirudin. Following this proposal in a shaken flask experiment results in crude yields which are distinctly worse than described
20 for Ala-hirudin. Thus, at first sight, no distinct advantage is clearly evident compared with later yeast expression systems.

The object on which the present invention was based was accordingly to prepare a fusion protein in which the combination of signal sequence and
25 Leu-hirudin permits direct processing to Leu-hirudin and subsequent secretion of native Leu-hirudin in high yields by *E. coli*. This is prerequisite for developing a process having an advantageous effect on the costs of production of Refludan both in the fermentation and in the subsequent purification owing to the improved initial hirudin concentration.

30 It has now been found, surprisingly, that signal sequences which permit direct secretion of Leu-hirudin by *E. coli* exist, and that in this case the secretion is in fact observed to be more efficient than described in European patent 0 448 093. It is thus possible to develop a process which makes large amounts of Leu-hirudin available without great expenditure.
35 The invention relates thereto.

In order to find advantageous signal sequences, a method of PCR-assisted signal sequence screening is introduced. This method uses the DNA encoding the protein of interest as template and a defined reverse PCR

primer, and variable forward primers which permit the synthesis of a DNA section which encodes a signal sequence coupled to a gene of interest. The reaction proceeds as shown in the scheme depicted in Figure 1. It is clear to the skilled worker that the number of reaction steps may vary according to the length of the signal sequence to be synthesized. Short signal sequences can be prepared with one reaction step, and longer sequences with two, three or more reactions. In addition, the number of reactions also depends on the apparatus used to synthesize the oligonucleotides used as primers. The signal peptide gene fusion synthesized in this way can then be cleaved specifically with the enzymes recognizing restriction sites 1 and 2 and be inserted into a correspondingly opened expression vector. The system becomes of general significance when hirudin is chosen as the gene of interest. Variable selection of the N-terminal amino acid of hirudin is moreover possible. Although this has a certain effect on the binding of hirudin to thrombin (change in the binding constant), the inhibitory effect of hirudin in relation to the thrombin activity remains measurable.

The patent EP-B1 0 448 093 describes the secretion of hirudin into the culture supernatant. The hirudin concentration therein can be determined directly via the well-known thrombin inhibition assay. The hirudin concentration is a direct measure of the efficiency of secretion and thus of the elimination of the signal sequence. The patent describes, however, that, for example, hirudin starting with the amino acid leucine cannot be efficiently released into the supernatant via the signal sequence of the CGTase. It is now possible, using the method described above, to search for signal sequences which effectively permit this. It is now possible in a similar way to investigate the secretion of hirudins starting with one of the other 19 amino acids. This results in each case in a spectrum of signal sequences which permit in a model way the efficient processing of the carboxy-terminal amino acid of the signal peptide and the peptide residue attaching thereto. It is thus possible to make a preselection of signal peptides for efficient secretion of any desired protein into the periplasm and thus increase the chances of developing an advantageous process for preparing a protein. The invention likewise relates thereto. The process can be expedited or automated by shaking the transformation mixture of ligand mixture and competent cells as liquid culture in a selection medium overnight and, the next day, inoculating with an aliquot of the cells as described in example 11 medium which contains inducer to carry out the

induction, but centrifuging most of the culture and freezing out the cell pellet. If hirudin activity is found on expression, the corresponding expression plasmid can be re-isolated from the cells, linearized and separated by gel electrophoresis from any autoligation products. The linear
 5 plasmid DNA is then religated and transformed anew into the host strain. It is then possible for individual colonies to be isolated and tested for their expression efficiency. It is possible to proceed in this case in such a way that the process meets the criteria of pharmaceutical approval.

10 A further advantage of the procedure is that it is easy to investigate different variants of a signal peptide, as arise in the course of evolution by exchange of amino acids between individual species, side by side for their ability to secrete a hirudin efficiently.

15 The process is also advantageous by comparison with the use of computer programs as described by Nielsen *et al.* (Protein Engineering 10, 1-6, 1997), with the aid of which it is possible to predict cleavage sites between signal sequence and a protein of interest. However, it is found that the predictions made therewith are not correct in every case, so that advan-
 20 tageous combinations may easily be overlooked. In addition, there is no relation between the prediction of correct processing and the yield which can actually be achieved.

One aspect of the invention is a hirudin precursor comprising a signal
 25 sequence selected from the group comprising the signal sequences of the outer membrane protein of *Serratia marcescens*, the oprF protein of *Pseudomonas fluorescens*, the lamb B protein of *Escherichia coli* (encoded by the lambda receptor (lamB) gene), and the fumarate reductase of *Shewanella putrefaciens*, preferably selected from the group comprising the
 30 signal sequence of the outer membrane protein of *Serratia marcescens* and the fumarate reductase of *Shewanella putrefaciens*, onto which there is C-terminal attachment of the sequence of Leu-hirudin.

Another aspect of the invention is a process for preparing Leu-hirudin, in
 35 which a hirudin precursor as described above occurs as intermediate, wherein

- (a) an expression plasmid comprising a DNA sequence coding for the hirudin precursor is prepared;

- (b) the expression plasmid from (a) is expressed in a suitable *E. coli* cell;
- 5 (c) the hirudin precursor is secreted from *E. coli* and simultaneously processed; and
- (d) the Leu-hirudin is isolated directly from the culture medium.

10 Likewise an aspect of the invention is the use of a hirudin precursor as described above for preparing Leu-hirudin, preferably in a process as described above.

15 A further aspect of the invention is a process for finding a suitable signal peptide for secretory expression of any desired protein in *E. coli*, wherein

- (a) hirudin or a hirudin derivative which has an antithrombotic effect and which has a defined amino acid aa_x at its N terminus which is connected N-terminally to a signal peptide to be tested is expressed in *E. coli*;
- 20 (b) the expression rate is determined by measuring the hirudin activity in the culture supernatant;
- (c) steps (a) and (b) are repeated with various signal peptides;
- (d) a suitable signal peptide is selected by comparing the expression rates represented by the hirudin activities found in step (b).

25

Likewise an aspect of the invention is the use of hirudin or a hirudin derivative which has an antithrombotic effect and which has a defined amino acid aa_x at its N terminus for finding a signal peptide which makes it possible to secrete efficiently a precursor protein consisting of the signal peptide and any other desired protein with the N-terminal amino acid aa_x with simultaneous elimination of the signal peptide from *E. coli*, in particular where aa_x is leucine.

30

A further aspect of the invention is a process for preparing any desired protein by secretory expression in *E. coli*, wherein

35

- (a) a suitable signal peptide is found by the process for finding a suitable signal peptide;

- (b) a nucleic acid construct coding for a precursor protein consisting of the suitable signal peptide from (a) and the desired protein is expressed in *E. coli*; and
- (c) the desired protein is isolated from the culture supernatant,

5

in particular where the N-terminal amino acid of the desired protein is leucine, and the expression takes place via a nucleic acid construct in which the sequence comprising the signal peptide codes for a signal peptide selected from the group comprising the outer membrane protein of

10 *Serratia marcescens*, the oprF protein of *Pseudomonas fluorescens*, the lamb B protein of *Escherichia coli*, and the fumarate reductase of *Shewanella putrefaciens*.

15 The synthesis of signal sequences which permit efficient synthesis and secretion of Leu-hirudin is to be described by way of example. Likewise described is the synthesis of other signal sequences which did not lead to the objective or gave worse results in relation to the yield. The examples are intended in this connection to explain the concept of the invention on the basis of the selection of signal sequences on the basis of Leu-hirudin

20 but not to be restricted thereto.

The processes described can be used for the purification of Refludan; this is described, for example, in example 11.

25 Example 1: Synthesis of a fusion gene coding for a fusion protein consisting of Leu-hirudin and the signal sequence of the outer membrane protein from *Serratia marcescens*

The expression plasmid used is the vector pJF118 which is described in

30 European patent 0 468 539, in Figure 1, because this is identical in its basic structure to the vector pCM7053 described in European patent 0 448 093.

The template used is the plasmid pK152 which is mentioned in Example 1

35 in European patent 0 448 093 and which harbors the hirudin sequence corresponding to European patent 0 171 024.

The membrane protein has been described by Braun, G. and Cole, S.T: (Mol. Gen. Genet. 195, 321-328 , 1984).

To synthesize the required DNA section, three oligonucleotide sequences are prepared.

- 5 Oligonucleotide hirrev has the sequence:

5' TTTTTTAAAG CTTGGGCTGC AGGTC 3' [SEQ ID NO: 1]
HindIII

- 10 The primer hybridizes with the region 227 –210 bp of the hirudin gene depicted in Table 1.

Primer smompaf1 has the sequence:

- 15 5'-TGGCACTGGC AGGTTTCGCT ACCGTAGCGC AAGCCcttac gtatactgac tgca – 3'
[SEQ ID NO: 2]

- 20 The primer hybridizes with nucleotides 1-19 of the hirudin sequence depicted in Table 1. The hybridizing part of the primer sequence is symbolized by small letters. The remainder of the sequence hybridizes with the region 229 bp - 263 bp of the sequence published by Braun, G. and Cole, S.T. (Mol. Gen. Genet. 195, 321-328, 1984).

Primer smompaf2 has the sequence:

- 25 5'- tttttgaat tcATGAAAAA GACAGCTATC GCATTAGCAG TGGCACTGGC AGGTTTC –3'
[SEQ ID NO: 3]

- 30 The primer sequence hybridizes from the 13 bp position onwards with the 201 bp – 245 bp sequence published by Braun and Cole and thus overlaps with the primer sequence smompaf2. The 1 - 12 position of the primer contains a recognition site for the restriction enzyme *EcoRI* and, adjoining, 6 T nucleotides in order to make recognition by the enzyme possible.

- 35 In a standard PCR (such as, for example, 94°C: 10'', 50°C: 30'', 72°C: 45'', 25 cycles) with DNA of the plasmid pK152, which harbors the sequence described in Table 1, as template, and the primers hirrev and smompaf1, the hirudin sequence is extended by the bacterial partial signal sequence. The reaction product is then reacted in a second PCR as template with the primers hirrev and smompaf2 under the same conditions. The reaction

product is a DNA fragment which codes for a fusion protein which consists of the hirudin sequence extended by the desired signal sequence. At the 5' end there is the recognition site for the restriction enzyme *EcoRI* and at the 3' end there is the recognition site for the enzyme *HindIII*.

5

The reaction product from the second PCR is reacted in a double-digestion mixture with the two restriction enzymes and is inserted as *EcoRI/HindIII* fragment into the vector DNA, which has been opened with the two enzymes, in a T4 DNA ligase reaction. Competent cells of the *E. coli* strain Mc1061 or the secretor mutant WCM100 are transformed with the ligation mixture and grown under selection pressure on ampicillin-containing plates. The next morning, expression as described in Example 6 is then compared with Ala-hirudin expression using the *E. coli* strain WCM100 / pCM7053. It is found that the expression obtained is about 1.5 times better than in the comparative test.

15

Example 2: Synthesis of the fusion protein of Leu-hirudin and the signal sequence of the *oprF* gene product from *Pseudomonas fluorescens*

Construction takes place in accordance with the scheme described in Example 1 with the exception that, in place of the primers smompaf1 / f2, there is use of two new primers which, in terms of their specificity for the hirudin gene and the sequence for recognition by the restriction enzyme *EcoRI*, have the same characteristics as the smompa primers but code for the required signal sequence of the *oprF* gene (De, E. *et al.*: FEMS Microbial. Lett. 127, 267 –272 , 1995).

25

Primer pfuf1 has the sequence:

30 5'GGTTCTCTTA TTGCCGCTAC TTCTTTCGGC GTTCTGGCAc ttacgtatac tgactgca 3'0
[SEQ ID NO: 4]

Primer pfuf2 has the sequence:

35 5'tttttgaat tcatgAAAAA CACCTTGGGC TTGGCCATTG GTTCTCTTAT TGCCGC 3'
[SEQ ID NO: 5]

In this case, the primer pfuf1 is used in accordance with Example 1 in PCR1 and primer pfuf2 is used correspondingly in PCR2. The expression is carried out by comparison with Ala-hirudin expression using the *E. coli*

strain WCM100 / pCM7053. It is found that the expression obtained is about 1.1 times better than in the comparative test. After fractionation by gel electrophoresis in the SDS-PAGE system, the hirudin band is isolated and the N-terminal sequence of the hirudin is determined. It is found that the sequence is completely intact and starts with the amino acid leucine. This result is surprising because the program for identifying the putative signal peptidase recognition site predicts an extension of the hirudin by valine.

10 Example 3: Synthesis of the fusion protein of Leu-hirudin and the signal sequence of the lamB gene product from *E. coli*

Construction takes place in accordance with the scheme described in Example 1 with the exception that, in place of the primers smompaf1 / f2, there is use of two new primers which, in terms of their specificity for the hirudin gene and the sequence for recognition by the restriction enzyme *EcoRI*, have the same characteristics as the smompa primers but code for the required signal sequence of the lamB gene (Clement, J.M. and Hofnung, M.: Cell 27, 507 –514, 1981).

20

Primer lambbf1 has the sequence:

5' GTTGCCGTCG CAGCGGGCGT AATGTCTGCT CAGGCAATGG CTcttacgta tactgactgc a 3'
[SEQ ID NO: 6]

25

Primer lambbf2 has the sequence:

5'tttttgaat tcATGATGAT TACTCTGCGC AAACCTCCTC TGGCGGTTGC CGTCGCAGC 3'
[SEQ ID NO: 7]

30

In this case, the primer lambbf11 is used in accordance with Example 1 in PCR1 and the primer lambbf2 is correspondingly used in PCR2. The expression is carried out by comparison with the Ala-hirudin expression using the *E. coli* strain WCM100 / pCM7053. It is found that the expression obtained is at the same level as in the comparative test. After fractionation by gel electrophoresis in the SDS-PAGE system, the hirudin band is isolated, and the N-terminal sequence of the hirudin is determined. It is found that the sequence is completely intact and starts with the amino acid leucine. This result is surprising because the program for identifying the

35

putative signal peptidase recognition site does not predict correct processing of hirudin.

5 Example 4: Synthesis of the fusion protein of Leu-hirudin and the signal sequence of the precursor of fumarate reductase flavoprotein subunit from *Shewanella putrefaciens*

10 Construction takes place in accordance with the scheme described in Example 1 with the exception that, in place of the primers smompaf1 / f2, there is use of two new primers which, in terms of their specificity for the hirudin gene and the sequence for recognition by the restriction enzyme *EcoRI*, have the same characteristics as the smompa primers but code for the required signal sequence from *Shewanella putrefaciens* (Pealing S.L. et al.: Biochemistry 31, 12132 – 12140, 1992). Since the publication describes
15 only the protein sequence, the amino acid sequence is translated in accordance with the codon tables into a DNA sequence so that the sequence which emerges for the

20 primer spfccf1 is as follows:

5' CTACCCTGAT GGGTACCGCT GGTCTGATGG GTACCGCTGT TGCTcttaag tatactgact gca 3'
[SEQ ID NO: 8]

25 Primer spfccf2 has the sequence:

5'tttttgaat tcATGAAAAA AATGAACCTG GCTGTTTGCA TCGCTACCCT GATGGGTACC 3'
[SEQ ID NO: 9]

30 In this case, the primer spfccf1 is used in accordance with Example 1 in PCR1 and primer spfccf2 is used correspondingly in PCR2. The expression is carried out by comparison with Ala-hirudin expression using the *E. coli* strain WCM100 / pCM7053. It is found that the expression obtained is about 1.5 times better than in the comparative test. After fractionation by gel electrophoresis in the SDS-PAGE system, the hirudin band is isolated
35 and the N-terminal sequence of the hirudin is determined. It is found that the sequence is completely intact and starts with the amino acid leucine. This result is surprising because the program for identifying the putative signal peptidase recognition site predicts processing on the carboxyl side of cysteine in position 6 of the hirudin sequence.

Example 5: Synthesis of the fusion protein of Leu-hirudin and the signal sequence of the β -lactamase precursor from pBR322

- 5 Construction takes place in accordance with the scheme described in Example 1 with the exception that, in place of the primers smompaf1 / f2, there is use of two new primers which, in terms of their specificity for the hirudin gene and the sequence for recognition by the restriction enzyme *EcoRI*, have the same characteristics as the smompa primers but code for the required signal sequence of the β -lactamase precursor protein (Sutcliffe J.G.; Cold Spring Harbor Symp. Quant. Biol. 43:77-90 (1978)).

Primer blattf1 has the following sequence:

- 15 5' CTGATCCCGT TCTTTGCAGC GTTCTGCCTG CCGGTTTTTCG CGcttacgta tactgactgc a 3'
[SEQ ID NO: 10]

Primer blattf2 has the sequence:

- 20 5' tttttgaat tcATGTCCAT CCAGCACTTC CGCGTCGCCC TGATCCCGTT CTTTGC 3'
[SEQ ID NO: 11]

- In this case, the primer blattf1 is used in accordance with Example 1 in PCR1 and primer blattf2 is used correspondingly in PCR2. The expression is carried out by comparison with Ala-hirudin expression using the *E. coli* strain WCM100 / pCM7053. It is found that the expression yield obtained is only 50%-90% of the yield obtained in the comparative test. After fractionation by gel electrophoresis in the SDS-PAGE system, the hirudin band is isolated and the N-terminal sequence of the hirudin is determined.
- 25 It is found that the sequence is completely intact and starts with the amino acid leucine. This result was predicted by the program for identifying the putative signal peptide recognition site.
- 30

- Example 6: Synthesis of the fusion gene of Leu-hirudin and the signal sequence of the precursor of alkaline phosphatase from *E. coli*
- 35

Construction takes place in accordance with the scheme described in Example 1 with the exception that, in place of the primers smompaf1 / f2, there is use of two new primers which, in terms of their specificity for the

hirudin gene and their sequence for recognition by the restriction enzyme *EcoRI*, have the same characteristics as the smompa primers but code for the required signal sequence of the alkaline phosphatase protein from *E. coli* (Shuttleworth, H., Taylor, J. and Minton, N. Nucleic Acids Res. 14 (21), 8689 (1986)).

Primer linkphoaf1 has the following sequence:

5' GCTGCCGCTG CTGTTACCCC CGGTTACCAA AGCGcttacg tatactgact gca 3'
[SEQ ID NO.: 12]

Primer linkphoaf2 has the sequence:

5' tttttgAAT TCATGAAACA GTCGACCATC GCGCTGGCGC TGCTGCCGCT GCTGTTC 3'
[SEQ ID NO.: 13]

The two primers are optimized in terms of the codon choice for *E. coli*, i.e. they do not correspond entirely to the natural sequence of the starting gene.

20

In this case, the primer linkphoaf1 is used in accordance with Example 1 in PCR1 and primer linkphoaf2 is used correspondingly in PCR2. The expression is carried out by comparison with Ala-hirudin expression using the *E. coli* strain WCM100 / pCM7053. It is found that the expression yield obtained is only a fraction of the yield obtained in the comparative test. After fractionation by gel electrophoresis in the SDS-PAGE system, the hirudin band is isolated and the N-terminal sequence of the hirudin is determined. It is found that the sequence is completely intact and starts with the amino acid leucine. This result was predicted by the program for identifying the putative signal peptidase recognition site. However, the poor yield is surprising.

30

Example 7: Synthesis of the fusion gene of Leu-hirudin and the signal sequence of the precursor of the alkaline phosphatase from *E. fergusonii*

35

Construction takes place in accordance with the scheme described in Example 1 with the exception that, in place of the primers smompaf1 / f2, there is use of two new primers which, in terms of their specificity for the hirudin gene and their sequence for recognition by the restriction enzyme

EcoRI, have the same characteristics as the smompa primers but code for the required signal sequence of the alkaline phosphatase protein from *E. fergusonii* (Du Bose, R.F. and Hartl, D.L. Mol. Biol. Evol. 7, 547-577 (1990)).

5

This signal sequence differs at five positions from the alkaline phosphatase from *E. coli*.

Primer fergusf1 has the following sequence:

10

5' GCTGAGCTGC CTGATCACCC CGGTGTCCCA GGCGcttacg tatactgact gca 3'
[SEQ ID NO.: 14]

Primer fergusf2 has the sequence:

15

5' tttttgaaat tcATGAAACA GAGCGCGATC GCGCTGGCTC TGCTgAGCTG CCTGATC 3'
[SEQ ID NO.: 15]

The two primers are optimized in terms of the codon choice for *E. coli*, i.e. they do not correspond entirely to the natural sequence of the starting gene. In this case, the primer fergusf1 is used in accordance with Example 1 in PCR1 and primer fergusf2 is used correspondingly in PCR2. The expression is carried out by comparison with Ala-hirudin expression using the *E. coli* strain WCM100 / pCM7053. It is found that the expression yield obtained is only a fraction of the yield obtained in the comparative test. It is a further approximately 50% lower than observed with the construct of signal peptide from *E. coli* alkaline phosphatase and Leu-hirudin.

Example 8: Synthesis of the fusion gene of Leu-hirudin and the signal sequence of the precursor of cyclodextrin glucanotransferase from *Paenibacillus macerans*

Construction takes place in accordance with the scheme described in Example 1 with the exception that, in place of the primers smompaf1 / f2, there is use of two new primers which, in terms of their specificity for the hirudin gene and their sequence for recognition by the restriction enzyme *EcoRI*, have the same characteristics as the smompa primers but code for the required signal sequence of the cyclodextrin glucanotransferase gene from *Paenibacillus macerans* (Takano, T., Fukuda, M., Monma, M.,

Kobayashi, S., Kainuma, K. and Yamane, K. J. Bacteriol. 166, 1118-1122 (1986)).

Primer baccdgf1 has the following sequence:

5

5' CTTTCGCTGA GTATGGCGTT GGGGATTTC A CTGCCCGCAT GGGCActtac gtatactgac tgca 3'
[SEQ ID NO.: 16]

Primer baccdgf2 has the sequence:

10

5' tttttgaat tcATGAAATC GCGGTACAAA CGTTTGACCT CCCTGGCGCT
TTCGCTGAGT ATGGC 3'
[SEQ ID NO.: 17]

- 15 In this case, the primer baccdgf1 is used in accordance with Example 1 in PCR1 and primer baccdgf2 is used correspondingly in PCR2. The expression is carried out by comparison with Ala-hirudin expression using the *E. coli* strain WCM100 / pCM7053. It is found that the expression yield obtained is about one quarter of the yield obtained in the comparative test.
- 20 The synthesized hirudin behaves like Leu-hirudin in the thrombin inhibition assay. This means that the signal peptide has been correctly processed. This does not correspond to the expectation from the theoretical analysis, which indicated an extension of 8 amino acids or, alternatively, a truncation by two amino acids at the N terminus.

25

Example 9: Synthesis of the fusion gene from Leu-hirudin and the signal sequence of the *E. coli* PCFO20 fimbrillin precursor protein (fotA)

- Construction takes place in accordance with the scheme described in
- 30 Example 1 with the exception that, in place of the primers smompaf1 / f2, there is use of two new primers which, in terms of their specificity for the hirudin gene and their sequence for recognition by the restriction enzyme *EcoRI*, have the same characteristics as the smompa primers but code for the required signal sequence of the *E. coli* PCFO20 fimbrillin precursor
- 35 protein (Viboud, G.I., Jonson, G., Dean-Nystrom, E. and Svennerholm, A.M. Infect. Immun. 64 (4), 1233-1239 (1996)).

Primer pcf1-ala has the following sequence:

5' TGGTTTCAGC TTTAGTAAGC GGGGTTGCAT TTGCTCTTAC
GTATACTGAC TGCAC 3'

[SEQ ID NO.: 18]

5 Primer p-pcf2 has the sequence:

5' TTTTGGGAAT TCATGAAAAA GACAATTATG TCTCTGGCTG TGGTTTCAGC
TTTAGTAAGC 3'

[SEQ ID NO.: 19]

10

In this case, the primer pcf1-ala is used in accordance with Example 1 in PCR1 and the primer p-pcf2 is used correspondingly in PCR2. The expression is carried out by comparison with the Ala-hirudin expression using the *E. coli* strain WCM100 / pCM7053. It is found that the expression yield obtained is about 40% of the yield obtained in the comparative test.

15

Example 10: Synthesis of the fusion gene of Leu-hirudin and the signal sequence of *S. typhimurium* outer membrane protein (fimD)

20 Construction takes place in accordance with the scheme described in Example 1 with the exception that, in place of the primers smompaf1 / f2, there is use of two new primers which, in terms of their specificity for the hirudin gene and their sequence for recognition by the restriction enzyme *EcoRI*, have the same characteristics as the smompa primers but code for the required signal sequence of the *S. typhimurium* outer membrane protein (Rioux, C.R., Friedrich, M.J. and Kadner, R.J.; J. Bacteriol. 172 (11), 6217-6222 (1990)).

25

Primer styfimf1 has the following sequence:

30

5' CGGCGCTGAG TCTCGCCTTA TTTTCTCACC TATCTTTTGC Ccttaegtat actgactgca 3'
[SEQ ID NO.: 20]

Primer styfimf2 has the sequence:

35

5' tttttgaat tcaTGTCATT TCATCACCGG GTATTTAAAC TGTCGGCGCT GAGTCTC 3'
[SEQ ID NO.: 21]

In this case, the primer styfimf1 is used in accordance with Example 1 in PCR1 and the primer styfimf2 is used correspondingly in PCR2. The expression is carried out by comparison with the Ala-hirudin expression using the *E. coli* strain WCM100 / pCM7053. It is found that the expression
 5 yield obtained is about 10% of the yield obtained in the comparative test.

Example 11: Expression in *E. coli*

10 The example describes expression of hirudin. For this purpose, 1-5 ml portions of LB medium which contains 25 mg/ml ampicillin and 0.5-2 mM IPTG (isopropyl β -D-thiogalactopyranoside) are inoculated with cells of a transformant and shaken at 220 rpm in an incubating shaker at 28°C for about 20 hours. Subsequently, after determination of the optical density,
 15 the cell suspension is centrifuged and hirudin is determined in the clear supernatant.

Expression of the Ala-hirudin described in European patent 0 448 093 via the plasmid pCM7053 in the secretor mutant WCM100 described in the
 20 patent is carried out in parallel with expression of Refludan. This makes direct comparison of the expression rate possible.

Expression on a larger scale can take place as described in US patent 5,616,476. Refludan can then be purified by the methods described in
 25 Examples 5 and 6 in this patent.

Example 12 : Determination of the hirudin concentration

Determination of the hirudin concentration is carried out by the method of
 30 Griebbach *et al.* (Thrombosis Research 37, 347-350 , 1985). For this purpose, defined amounts of a Refludan standard are included in the series of measurements to construct a calibration plot. It is thus possible to state the yield directly in mg/l.

35 Table 1: Hirudin-encoding DNA sequence with translation into amino acids

1 CTTACGTATACTGACTGCACTGAATCTGGTCAGAACCTGTGCCTGTGCGAAGGATCTAAC 60
 L T Y T D C T E S G Q N L C L C E G S N -

61 GTTTGCGGCCAGGGTAACAAATGCATCCTTGGATCCGACGGTGAAAAGAACCAGTGC GTT 120
V C G Q G N K C I L G S D G E K N Q C V -

5

121 ACTGGCGAAGGTACCCCGAAACCGCAGTCTCATAACGACGGCGACTTCGAAGAGATCCCT 180
T G E G T P K P Q S H N D G D F E E I P -

181 GAGGAATACCTTCAGTAATAGAGCTCGTCGACCTGCAGCCCAAGCTT 227

[SEQ ID NO.:22]

E E Y L Q * * ----- -

10

[SEQ ID NO.: 23]

Table 2:

Example	Signal sequence	Primary structure	Relative yield per ml of culture	SEQ ID NO.:
-	Control: cgtase-Ala-hirudin	MKRNRFNTS AAIAISIALNTFF CSMQTIA	1	24
1	Outer membrane protein / <i>Serratia marcescens</i>	MKKTALALAVALAGFATVAQ A	1.5	25
2	oprF protein / <i>Pseudomonas fluorescens</i>	MKNTLGLAIGSLIAATSFGV LA	1.1	26
3	LambB protein / <i>E. coli</i>	MMITLRKLPL AVAVAAGVMS AQAMA	1	27
4	Fumate reductase / <i>Shewanella putrifaciens</i>	MKKMNLAVCI ATLMGTAGLM GTAVA	1.5	28
5	β - Lactamase / pBR322	MSIQHFRVAL IPFFAAFSLPVFA	0.5	29
8	Alk. phosphatase / <i>E.coli</i>	MKQSTIALAL LPLLFTPVT K A	0.1	30
9	Alk. phosphatase / <i>E. fergusonii</i>	MKQSAIALAL LSCLITPV SQ A	0.05	31
10	Cyclodextrin glucanotransferase / <i>Paenibacillus macerans</i>	MKSRYKRLTS LALSLSMALGI SLPWA	0.25	32
11	Outer membrane protein / <i>S. typhimurium</i>	MSFHHRVFKL SALSLALFSH LSFA	0.11	33

Patent Claims:

1. A hirudin precursor comprising a signal sequence selected from the group comprising the signal sequences of the outer membrane protein of
5 *Serratia marcescens*, the oprF protein of *Pseudomonas fluorescens*, the lamb B protein of *Escherichia coli*, and the fumarate reductase of *Shewanella putrifaciens*, onto which there is C-terminal attachment of the sequence of Leu-hirudin.
- 10 2. A hirudin precursor as claimed in claim 1, where the signal sequence is selected from the group comprising the signal sequence of the outer membrane protein of *Serratia marcescens* and the fumarate reductase of *Shewanella putrifaciens*.
- 15 3. A process for preparing Leu-hirudin, in which a hirudin precursor as claimed in either of claims 1 or 2 occurs as intermediate, in which
 - (a) an expression plasmid comprising a DNA sequence coding for the
20 hirudin precursor is prepared;
 - (b) the expression plasmid from (a) is expressed in a suitable *E. coli* cell;
 - (c) the hirudin precursor is secreted from the *E. coli* and simultaneously
25 processed; and
 - (d) the Leu-hirudin is isolated from the culture medium.
4. The use of a hirudin precursor as claimed in either of claims 1 or 2
30 for preparing Leu-hirudin.
5. The use of a hirudin precursor as claimed in claim 4 in a process as claimed in claim 3.
- 35 6. A process for finding a suitable signal peptide for secretory expression of any desired protein in *E. coli*, wherein
 - (a) hirudin or a hirudin derivative which has an antithrombotic effect and
which has a defined amino acid aa_x at its N terminus which is

connected N-terminally to a signal peptide to be tested is expressed in *E. coli*;

(b) the expression rate is determined by measuring the hirudin activity in the culture supernatant;

5 (c) steps (a) and (b) are repeated with various signal peptides;

(d) a suitable signal peptide is selected by comparing the expression rates represented by the hirudin activities found in step (b).

10 7. The use of hirudin or a hirudin derivative which has an anti-thrombotic effect and which has a defined amino acid aa_x at its N terminus for finding a signal peptide which makes it possible to secrete efficiently a precursor protein consisting of the signal peptide and any other desired protein with the N-terminal amino acid aa_x with simultaneous elimination of the signal peptide from *E. coli*.

15

8. The use as claimed in claim 7, where aa_x is leucine.

9. A process for preparing any desired protein by secretory expression in *E. coli*, wherein

20

(a) a suitable signal peptide is found by the process as claimed in claim 6;

(b) a nucleic acid construct coding for a precursor protein consisting of the suitable signal peptide from (a) and the desired protein is expressed in *E. coli*; and

25

(c) the desired protein is isolated from the culture supernatant.

30 10. The process for preparing any desired protein in *E. coli* as claimed in claim 9, where the N-terminal amino acid of the desired protein is leucine, and the expression takes place via a nucleic acid construct in which the sequence coding for the signal peptide codes for a signal peptide selected from the group comprising the outer membrane protein of *Serratia marcescens*, the oprF protein of *Pseudomonas fluorescens*, the lamb B protein of *Escherichia coli*, and the fumarate reductase of *Shewanella putrifaciens*.

35