Surprisingly, a new type of leader peptide has been found which allows secretion in high yield of a polypeptide in yeast.

In the present context, the expression "leader peptide" is understood to indicate a peptide whose function is to allow the expressed polypeptide to be directed from the endoplasmic reticulum to the Golgi apparatus and further to a secretory vesicle for secretion into the medium, (i.e. exportation of the expressed polypeptide across the cell wall or at least through the cellular membrane into the periplasmic space of the cell).

Claim

1. A DNA expression cassette comprising the following sequence:

\[ 5' - \text{P-SP-LS-PS-*gene*-(T)}_i - 3' \]

wherein

- P is a promoter sequence,
- SP is a DNA sequence encoding a signal peptide,
- LS is a DNA sequence encoding a leader peptide with the general formula I:

\[ .../2 \]
GlnProIle(Asp/Glu)(Asp/Glu)X₁(Glu/Asp)X₂AsnZ(Thr/Ser)X₃ \quad (I)

wherein

X₁ is a peptide bond or a codable amino acid;
X₂ is a peptide bond, a codable amino acid or a sequence of up to 4 codable amino acids which may be the same or different;
Z is a codable amino acid except Pro; and
X₃ is a sequence of from 4 to 30 codable amino acids which may be the same or different;

PS is a DNA sequence encoding a processing site;
*gene* is a DNA sequence encoding a polypeptide;
T is a terminator sequence; and
i is 0 or 1.
SYNTHETIC LEADER PEPTIDE SEQUENCES

The present invention relates to synthetic leader peptide sequences for secreting polypeptides in yeast.
SYNTHETIC LEADER PEPTIDE SEQUENCES

FIELD OF INVENTION

The present invention relates to synthetic leader peptide sequences for secreting polypeptides in yeast.

BACKGROUND OF THE INVENTION

Yeast organisms produce a number of proteins which are synthesized intracellularly, but which have a function outside the cell. Such extracellular proteins are referred to as secreted proteins. These secreted proteins are expressed initially inside the cell in a precursor or a pre-protein form containing a presequence ensuring effective direction of the expressed product across the membrane of the endoplasmic reticulum (ER). The presequence, normally named a signal peptide, is generally cleaved off from the desired product during translocation. Once entered in the secretory pathway, the protein is transported to the Golgi apparatus. From the Golgi the protein can follow different routes that lead to compartments such as the cell vacuole or the cell membrane, or it can be routed out of the cell to be secreted to the external medium (Pfeffer, S.R. and Rothman, J.E. Ann.Rev.Biochem. 56 (1987) 829-852).

Several approaches have been suggested for the expression and secretion in yeast of proteins heterologous to yeast. European published patent application No. 88 632 describes a process by which proteins heterologous to yeast are expressed, processed and secreted by transforming a yeast organism with an expression vehicle harbouring DNA encoding the desired protein and a signal peptide, preparing a culture of the transformed organism, growing the culture and recovering the protein from the culture medium. The signal peptide may be the signal peptide of the desired protein itself, a heterologous signal
peptide or a hybrid of native and heterologous signal peptide.

A problem encountered with the use of signal peptides heterologous to yeast might be that the heterologous signal peptide does not ensure efficient translocation and/or cleavage after the signal peptide.

The Saccharomyces cerevisiae MFA1 (α-factor) is synthesized as a prepro form of 165 amino acids comprising signal- or prepeptide of 19 amino acids followed by a "leader" or prepropeptide of 64 amino acids, encompassing three N-linked glycosylation sites followed by (LysArg(Asp/Glu)Ala)2-3α-factor)4 (Kurjan, J. and Herskowitz, I. Cell 30 (1982) 933-943). The signal-leader part of the preproMFA1 has been widely employed to obtain synthesis and secretion of heterologous proteins in S. cerevisiae.

Use of signal/leader peptides homologous to yeast is known from i.a. US patent specification No. 4,546,082, European published patent applications Nos. 116 201, 123 294, 123 544, 163 529 and 123 289 and DK patent application No. 3614/83.

In EP 123 289 utilization of the S. cerevisiae α-factor pre-20 cursor is described whereas WO 84/01153 indicates utilization of the S. cerevisiae invertase signal peptide and DK 3614/83 utilization of the S. cerevisiae PH05 signal peptide for secretion of foreign proteins.

US patent specification No. 4,546,082, EP 16 201, 123 294, 123 25544 and 163 529 describe processes by which the α-factor signal-leader from S. cerevisiae (MFA1 or MFA2) is utilized in the secretion process of expressed heterologous proteins in yeast. By fusing a DNA sequence encoding the S. cerevisiae MFA1 signal/leader sequence at the 5' end of the gene for the desired protein secretion and processing of the desired protein was demonstrated.
EP 206 783 discloses a system for the secretion of polypeptides from *S. cerevisiae* using an α-factor leader sequence which has been truncated to eliminate the four α-factor units present on the native leader sequence so as to leave the leader peptide itself fused to a heterologous polypeptide via the α-factor processing site LysArgGluAlaGluAla. This construction is indicated to lead to an efficient processing of smaller peptides (less than 50 amino acids). For the secretion and processing of larger polypeptides, the native α-factor leader sequence has been truncated to leave one or two of the α-factor units between the leader peptide and the polypeptide.

A number of secreted proteins are routed so as to be exposed to a proteolytic processing system which can cleave the peptide bond at the carboxy end of two consecutive basic amino acids. This enzymatic activity is in *S. cerevisiae* encoded by the KEX 2 gene (Julius, D.A. et al., *Cell* 37 (1984b) 1075). Processing of the product by the KEX 2 protease is needed for the secretion of active *S. cerevisiae* mating factor α1 (MFα1 or α-factor) whereas KEX 2 is not involved in the secretion of active *S. cerevisiae* mating factor α.

Secretion and correct processing of a polypeptide intended to be secreted is obtained in some cases when culturing a yeast organism which is transformed with a vector constructed as indicated in the references given above. In many cases, however, the level of secretion is very low or there is no secretion, or the proteolytic processing may be incorrect or incomplete. It is therefore the object of the present invention to provide leader peptides which ensure a more efficient expression and/or processing of polypeptides.

**SUMMARY OF THE INVENTION**

Surprisingly, a new type of leader peptide has been found which allows secretion in high yield of a polypeptide in yeast.
Accordingly, the present invention relates to a DNA expression cassette comprising the following sequence:

\[ 5'\text{-P-SP-LS-PS-*gene*-}(T)_{i-3'} \]

wherein

5 P is a promoter sequence,
SP is a DNA sequence encoding a signal peptide,
LS is a DNA sequence encoding a leader peptide with the general formula I:

\[ \text{GlnProIle(Asp/Glu)(Asp/Glu)}X^1\text{(Glu/Asp)}X^2\text{AsnZ(Thr/Ser)}X^3 \]  

10 wherein

\( X^1 \) is a peptide bond or a codable amino acid;
\( X^2 \) is a peptide bond, a codable amino acid or a sequence of up to 4 codable amino acids which may be the same or different;
\( Z \) is a codable amino acid except Pro; and
\( X^3 \) is a sequence of from 4 to 30 codable amino acids which may be the same or different;

PS is a DNA sequence encoding a processing site;
*gene* is a DNA sequence encoding a polypeptide;
T is a terminator sequence; and
\( i \) is 0 or 1.

In the present context, the expression "leader peptide" is understood to indicate a peptide whose function is to allow the expressed polypeptide to be directed from the endoplasmic reticulum to the Golgi apparatus and further to a secretory vesicle for secretion into the medium, (i.e. exportation of the expressed polypeptide across the cell wall or at least through the cellular membrane into the periplasmic space of the cell). The term "synthetic" used in connection with leader peptides is intended to indicate that the leader peptide is one not found in nature.
The term "signal peptide" is understood to mean a presequence which is predominantly hydrophobic in nature and present as an N-terminal sequence of the precursor form of an extracellular protein expressed in yeast. The function of the signal peptide is to allow the expressed protein to be secreted to enter the endoplasmic reticulum. The signal peptide is normally cleaved off in the course of this process. The signal peptide may be heterologous or homologous to the yeast organism producing the protein.

The expression "polypeptide" is intended to indicate a heterologous polypeptide, i.e. a polypeptide which is not produced by the host yeast organism in nature as well as a homologous polypeptide, i.e. a polypeptide which is produced by the host yeast organism in nature and any preform thereof. In a preferred embodiment, the expression cassette of the present invention encodes a heterologous polypeptide.

The expression "a codable amino acid" is intended to indicate an amino acid which can be coded for by a triplet ("codon") of nucleotides.

When, in the amino acid sequences given in the present specification, the three letter codes of two amino acids, separated by a slash, are given in brackets, e.g. (Asp/Glu), this is intended to indicate that the sequence has either the one or the other of these amino acids in the pertinent position.

The present invention also provides a yeast expression vector comprising an expression cassette according to the invention.

The present invention also provides a yeast cell which is capable of expressing a polypeptide and which is transformed with a yeast expression vector according to the invention.

In a further aspect, the present invention relates to a process for producing a polypeptide in yeast, the process comprising culturing a yeast cell, which is capable of expressing a polypeptide and which is transformed with a yeast expression vector as described above including a leader peptide sequence of the invention, in a suitable medium to obtain expression and secretion of the polypeptide, after which the polypeptide is recovered from the medium.
BRIEF DESCRIPTION OF THE DRAWINGS

The present invention is further illustrated with reference to the appended drawings wherein

Fig. 1 schematically shows the plasmid pAK492;

5 Fig. 2 shows part of the DNA sequence encoding the signal peptide/leader/MI3 insulin precursor;

Fig. 3 shows the construction of the plasmid pAK546;

Fig. 4 shows the amino acid sequence of the leader SEQ ID No. 4 and the DNA sequence encoding it;

10 Fig. 5 shows the DNA sequence of S. cerevisiae expression plasmid pAK546 encoding the YAP3 signal peptide, leader SEQ ID No. 4 and the MI3 insulin precursor and the encoded amino acid sequence;

Fig. 6 shows the amino acid sequence of the leader SEQ ID No. 15 6 and the DNA sequence encoding it;

Fig. 7 shows the amino acid sequence of the leader SEQ ID No. 8 and the DNA sequence encoding it;

Fig. 8 shows the amino acid sequence of the leader SEQ ID No. 17 and the DNA sequence encoding it;

20 Fig. 9 shows the amino acid sequence of the leader SEQ ID No. 16 and the DNA sequence encoding it;

Fig. 10 shows the amino acid sequence of the leader SEQ ID No. 19 and the DNA sequence encoding it;

Fig. 11 shows the amino acid sequence of the leader SEQ ID No. 25 20 and the DNA sequence encoding it;
Fig. 12 shows the amino acid sequence of the leader SEQ ID No. 21 and the DNA sequence encoding it;

Fig. 13 shows the DNA fragment of pAK527 used as the direct template in the construction of SEQ ID Nos. 4 and 6;

Fig. 14 shows the DNA fragment of pAK531 used as the direct template in the construction of SEQ ID No. 8;

Fig. 15 shows the DNA fragment of pAK555 used as the direct template in the construction of SEQ ID Nos. 16 and 17;

Fig. 16 shows the DNA fragment of pAK559 used as the direct template in the construction of SEQ ID Nos. 19 and 20; and

Fig. 17 shows the DNA fragment of pAK562 used as the direct template in the construction of SEQ ID No. 21;

Fig. 18 shows the amino acid sequence of the leader SEQ ID No. 27 and the DNA sequence SEQ ID No. 66 encoding it;

Fig. 19 shows the amino acid sequence SEQ ID No. 70 of an N-terminally extended MI3 insulin precursor and the DNA sequence SEQ ID No. 71 encoding it;

Fig. 20 shows the amino acid sequence of the leader SEQ ID No. 67 and the DNA sequence SEQ ID No. 69 encoding it;

Fig. 21 shows the DNA fragment SEQ ID No. 72 of pAK614 used as the direct template in the construction of SEQ ID No. 27; and

Fig. 22 shows the DNA fragment SEQ ID No. 73 of pAK625 used as the direct template in the construction of SEQ ID No. 67.
DETAILED DISCLOSURE OF THE INVENTION

When X\(^1\) in general formula I designates an amino acid, it is preferably Ser, Thr or Ala. When X\(^2\) in general formula I designates one amino acid, it is preferably Ser, Thr or Ala. When X\(^2\) in general formula I designates a sequence of two amino acids, it is preferably SerIle. When X\(^2\) in general formula I designates a sequence of three amino acids, it is preferably SerAlaIle. When X\(^2\) in general formula I designates a sequence of four amino acids it is preferably SerPheAlaThr. In a preferred embodiment, X\(^3\) is an amino acid sequence of the general formula II

\[ x^4 - x^5 - x^6 \]  

(II)

wherein X\(^4\) is a sequence of from 1 to 21 codable amino acids which may be the same or different, X\(^5\) is Pro or one of the amino acid sequences ValAsnLeu or LeuAlaAsnValAlaMetAla, and X\(^6\) is a sequence of from 1 to 8 codable amino acids which may be the same or different.

In general formula II, X\(^4\) is preferably an amino acid sequence which includes one or more of the motifs LeuValAsnLeu, SerValAsnLeu, MetAlaAsp, ThrGluSer, ArgPheAlaThr or ValAlaMetAla; or X\(^4\) is an amino acid sequence which includes the sequence AsnSerThr or AsnThrThr; or X\(^4\) is an amino acid sequence which includes the sequence

\[ (\text{Ser/Leu})\text{ValAsnLeu}, \]
\[ (\text{Ser/Leu})\text{ValAsnLeuMetAlaAsp}, \]
\[ (\text{Ser/Leu})\text{ValAsnLeuMetAlaAspAsp}, \]
\[ (\text{Ser/Leu})\text{ValAsnLeuMetAlaAspAspThrGluSer}, \]
\[ (\text{Ser/Leu})\text{ValAsnLeuMetAlaAspAspThrGluSerIle or} \]
\[ (\text{Ser/Leu})\text{ValAsnLeuMetAlaAspAspThrGluSerArgPheAlaThr}; \]

30 or X\(^4\) is an amino acid sequence which includes the sequence

Asn(Thr/Ser)ThrLeu,
Asn(Thr/Ser)ThrLeuAsnLeu or
Asn(Thr/Ser)ThrLeuValAsnLeu; or any combination
thereof.

In general formula II, X^5 is preferably Pro or an amino acid sequence which includes the sequence ValAsnLeu, LeuAlaAsnValAlaMetAla, LeuAspValValAsnLeuProGly or 5 LeuAspValValAsnLeuIleSerMet.

When X^6, in general formula II, designates one amino acid, it is preferably Ala, Gly, Leu, Thr, Val or Ser. When X^6, in general formula II, designates a sequence of two amino acids, it is preferably GlyAla or SerAla. When X^6, in general formula 10 II, designates a sequence of three amino acids, it is preferably AlaValAla. When X^6, in general formula II, designates a sequence of eight amino acids, it is preferably GlyAlaAspSerLysThrValGlu.

Examples of preferred leader peptides coded for by the DNA sequence LS are:

SEQ ID No. 1 GlnProIleAspGluAspAsnAspThrSerValAsnLeuProAla;
SEQ ID No. 2 GlnProIleAspGluAspThrThrSerValAsnLeuProAla;
SEQ ID No. 3 GlnProIleAspGluSerAsnThrThrSerValAsnLeuProAla;
SEQ ID No. 4 GlnProIleAspGluAsnThrThrSerValAsnLeuProVal;
SEQ ID No. 5 GlnProIleAspGluThrGluAsnThrThrSerValAsnLeuProAla;
SEQ ID No. 6 GlnProIleAspGluThrGluSerAsnThrThrSerValAsnLeuProAla;
SEQ ID No. 7 GlnProIleAspGluAsnThrThrSerValAsnLeuMetAla;
SEQ ID No. 8 GlnProIleAspGluThrGluSerAsnThrThrSerValAsnLeuPro-
GlyAla;
SEQ ID No. 9  GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuMet-Ala;

SEQ ID No. 10  GlnProIleAspAspThrGluSerAsnThrThrSerValAsnValPro-Thr;

SEQ ID No. 11  GlnProIleAspAspThrGluSerAsnThrThrLeuValAsnValPro-Thr;

SEQ ID No. 12  GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuPro-Thr;

SEQ ID No. 13  GlnProIleAspAspThrGluSerAsnThrThrLeuValAsnValPro-GlyAla;

SEQ ID No. 14  GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuMet-AlaProAlaValAla;

SEQ ID No. 15  GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuMet-AspLeuAlaValGlyLeuProGlyAla;

SEQ ID No. 16  GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuMet-AlaAspAspThrGluSerIleAsnThrThrLeuValAsnLeuProGly-Ala;

SEQ ID No. 17  GlnProIleAspAspThrGluSerIleAsnThrThrLeuValAsnLeu-ProGlyAla;

SEQ ID No. 18  GlnProIleAspAspThrGluSerAsnThrThrLeuValAsnLeuPro-GlyAla;

SEQ ID No. 19  GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuMet-AlaAspAspThrGluSerArgPheAlaThrAsnThrThrLeuValAsn-LeuProLeu;

SEQ ID No. 20  GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuMet-AlaAspAspThrGluSerIleAsnThrThrLeuValAsnLeuAlaAsn-
ValAlaMetAla;
SEQ ID No. 21 GlnProIleAspAspThrGluSerAlaIleAsnThrThrLeuValAsn-LeuProGlyAla;
SEQ ID No. 22 GlnProIleAspAspThrGluSerPheAlaThrAsnThrThrLeuVal-
      AsnLeuProGlyAla;
SEQ ID No. 23 GlnProIleAspAspThrGluSerIleAsnThrThrLeuValAsnLeu-
      MetAlaAspAspThrGluSerArgPheAlaThrAsnThrThrLeuValAsnLeuProLeu;
SEQ ID No. 24 GlnProIleAspAspThrGluSerIleAsnThrThrLeuValAsnLeu-
      MetAlaAspAspThrGluSerArgPheAlaThrAsnThrThrLeuAsp-
      ValValAsnLeuProGlyAla;
SEQ ID No. 25 GlnProIleAspAspThrGluSerAlaAlaIleAsnThrThrLeuVal-
      AsnLeuProGlyAla;
SEQ ID No. 26 GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuMet-
      AlaAspAspThrGluSerArgPheAlaThrAsnThrThrLeuValAsn-
      LeuAlaAsnValAlaMetAla;
SEQ ID No. 27 GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuMet-
      AlaAspAspThrGluSerArgPheAlaThrAsnThrThrLeuAspVal-
      ValAsnLeuIleSerMetAla;
SEQ ID No. 28 GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuMet-
      AlaAsnThrThrGluSerArgPheAlaThrAsnThrThrLeuAspVal-
      ValAsnLeuIleSerMetAla; and
SEQ ID No. 67 GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuMet-
      AlaAspAspThrGluSerArgPheAlaThrAsnThrThrLeuAlaLeu-
      AspValValAsnLeuIleSerMetAla.

Particularly preferred leader peptides coded for by the DNA sequence LS are:
SEQ ID No. 15  GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuMet-AspLeuAlaValGlyLeuProGlyAla;

SEQ ID No. 16  GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuMet-AlaAspAspThrGluSerIleAsnThrThrLeuValAsnLeuProGly- Ala;

SEQ ID No. 17  GlnProIleAspAspThrGluSerIleAsnThrThrLeuValAsnLeu-ProGlyAla;

SEQ ID No. 18  GlnProIleAspAspThrGluSerAsnThrThrLeuValAsnLeuPro-GlyAla;

SEQ ID No. 19  GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuMet-AlaAspAspThrGluSerArgPheAlaThrAsnThrThrLeuValAsn-LeuProLeu;

SEQ ID No. 20  GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuMet-AlaAspAspThrGluSerIleAsnThrThrLeuValAsnLeuAlaAsn-ValAlaMetAla;

SEQ ID No. 21  GlnProIleAspAspThrGluSerAlaIleAsnThrThrLeuValAsn-LeuProGlyAla;

SEQ ID No. 22  GlnProIleAspAspThrGluSerPheAlaThrAsnThrThrLeuVal-AsnLeuProGlyAla;

SEQ ID No. 23  GlnProIleAspAspThrGluSerIleAsnThrThrLeuValAsnLeu-MetAlaAspAspThrGluSerArgPheAlaThrAsnThrThrLeuVal-AsnLeuProLeu;

SEQ ID No. 24  GlnProIleAspAspThrGluSerIleAsnThrThrLeuValAsnLeu-MetAlaAspAspThrGluSerArgPheAlaThrAsnThrThrLeuAsp-ValValAsnLeuProGlyAla;

SEQ ID No. 25  GlnProIleAspAspThrGluSerAlaIleAsnThrThrLeuVal-AsnLeuProGlyAla;
The signal sequence (SP) may encode any signal peptide which ensures an effective direction of the expressed polypeptide into the secretory pathway of the cell. The signal peptide may be a naturally occurring signal peptide or functional parts thereof or it may be a synthetic peptide. Suitable signal peptides have been found to be the α-factor signal peptide, the signal peptide of mouse salivary amylase, a modified carboxypeptidase signal peptide, the yeast BARK signal peptide or the Humicola lanuginosa lipase signal peptide or a derivative thereof. The mouse salivary amylase signal sequence is described by Hangenbühle, O. et al., Nature 289 (1981) 643-646. The carboxypeptidase signal sequence is described by Valls, L.A. et al., Cell 48 (1987) 887-897. The BARK signal peptide is disclosed in WO 87/02670. The yeast aspartic protease 3 signal peptide is described in Danish patent application No. 0828/93.

The yeast processing site encoded by the DNA sequence PS may suitably be any paired combination of Lys and Arg, such as LysArg, ArgLys, ArgArg, or LysLys which permits processing of the polypeptide by the KEX2 protease of Saccharomyces cerevisiae or the equivalent protease in other yeast species (Julius, D.A. et al., Cell 27 (1984) 1075). If KEX2 processing is not convenient, e.g. if it would lead to cleavage of the polypeptide product, e.g. due to the presence of two
consecutive basic amino acid internally in the desired product, a processing site for another protease may be selected comprising an amino acid combination which is not found in the polypeptide product, e.g. the processing site for FXa, IleGluGlyArg (cf. Sambrook, J., Fritsch, E.F. and Maniatis, T., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, New York, 1989).

The protein produced by the method of the invention may be any protein which may advantageously be produced in yeast. Examples of such proteins are heterologous proteins such as aprotinin, tissue factor pathway inhibitor or other protease inhibitors, insulin or insulin precursors, human or bovine growth hormone, interleukin, glucagon, GLP-1, IGF-I, IGF-II, tissue plasminogen activator, transforming growth factor α or β, platelet-derived growth factor, enzymes or a functional analogue thereof. In the present context, the term "functional analogue" is meant to indicate a protein with a similar function as the native protein (this is intended to be understood as relating to the nature rather than the level of biological activity of the native protein). The protein may be structurally similar to the native protein and may be derived from the native protein by addition of one or more amino acids to either or both the C- and N-terminal end of the native protein, substitution of one or more amino acids at one or a number of different sites in the native amino acid sequence, deletion of one or more amino acids at either or both ends of the native protein or at one or several sites in the amino acid sequence, or insertion of one or more amino acids at one or more sites in the native amino acid sequence. Such modifications are well known for several of the proteins mentioned above. Also, precursors or intermediates for other proteins may be produced by the method of the invention. An example of such a precursor is the MI3 insulin precursor which comprises the amino acid sequence B(1-29)AlaAlaLysA(1-21) wherein A(1-21) is the A chain of human insulin and B(1-29) is the B chain of human insulin in which Thr(B30) is missing.
Preferred DNA constructs encoding leader sequences are as shown in Figs. 4 - 12 or suitable modifications thereof. Examples of suitable modifications of the DNA sequence are nucleotide substitutions which do not give rise to another amino acid 5 sequence of the protein, but which may correspond to the codon usage of the yeast organism into which the DNA construct is inserted or nucleotide substitutions which do give rise to a different amino acid sequence and therefore, possibly, a different protein structure. Other examples of possible 10 modifications are insertion of one or more codons into the sequence, addition of one or more codons at either end of the sequence and deletion of one or more codons at either end of or within the sequence.

The recombinant expression vector carrying the expression cassette

$$5'\text{-P-SP-LS-PS-*gene*-}(T)_{i-3}'$$

wherein P, SP, LS, *gene*, T and i are as defined above may be any vector which is capable of replicating in yeast organisms. The promoter may be any DNA sequence which shows 20 transcriptional activity in yeast and may be derived from genes encoding proteins either homologous or heterologous to yeast. The promoter is preferably derived from a gene encoding a protein homologous to yeast. Examples of suitable promoters are the *Saccharomyces cerevisiae* MFα1, TPI, ADH or PGK promoters.

The sequences shown above should preferably also be operably connected to a suitable terminator, e.g. the TPI terminator (cf. Alber, T. and Kawasaki, G., *J. Mol. Appl. Genet.* 1 (1982) 419-434).

The recombinant expression vector of the invention further 30 comprises a DNA sequence enabling the vector to replicate in yeast. Examples of such sequences are the yeast plasmid 2μ replication genes REP 1-3 and origin of replication. The vector may also comprise a selectable marker, e.g. the *Schizo*
saccharomyces pombe TPI gene as described by Russell, P.R.,

The methods used to ligate the sequence 5'-P-SP-LS-PS-*gene*-(T)\_i-3' and to insert it into suitable yeast vectors containing
the information necessary for yeast replication, are well known
to persons skilled in the art (cf., for instance, Sambrook, J.,
Fritsch, E.F. and Maniatis, T., op.cit.). It will be understood
that the vector may be constructed either by first preparing a
DNA construct containing the entire sequence 5'-P-SP-LS-PS-
10 *gene*-(T)\_i-3' and subsequently inserting this fragment into a
suitable expression vector, or by sequentially inserting DNA
fragments into a suitable vector containing genetic information
for the individual elements (such as the promoter sequence, the
signal peptide, the leader sequence
15 GlnProIle\_o(Asp/Glu)(Asp/Glu)X\_1(Glu/Asp)X\_2AsnX(Thr/Ser)X\_3, the
processing site, the polypeptide, and, if present, the
terminator sequence) followed by ligation.

The yeast organism used in the method of the invention may be
any suitable yeast organism which, on cultivation, produces
large amounts of the desired polypeptide. Examples of suitable
yeast organisms may be strains of the yeast species
Saccharomyces cerevisiae, Saccharomyces kluveyri,
Schizosaccharomyces pombe or Saccharomyces uvarum. The
transformation of the yeast cells may for instance be effected
by protoplast formation followed by transformation in a manner
known per se. The medium used to cultivate the cells may be any
conventional medium suitable for growing yeast organisms. The
secreted polypeptide, a significant proportion of which will be
present in the medium in correctly processed form, may be
recovered from the medium by conventional procedures including
separating the yeast cells from the medium by centrifugation or
filtration, precipitating the proteinaceous components of the
supernatant or filtrate by means of a salt, e.g. ammonium
sulphate, followed by purification by a variety of
35 chromatographic procedures, e.g. ion exchange chromatography,
affinity chromatography or the like.

The invention is further described in the following examples which are not to be construed as limiting the scope of the invention as claimed.

EXAMPLES

Plasmids and DNA material

All expression plasmids are of the C-POT type. Such plasmids are described in EP patent application No. 171 142 and are characterized in containing the \textit{Schizosaccharomyces pombe} triose phosphate isomerase gene (POT) for the purpose of plasmid selection and stabilization. A plasmid containing the POT-gene is available from a deposited \textit{E. coli} strain (ATCC 39685). The plasmids furthermore contain the \textit{S. cerevisiae} triose phosphate isomerase promoter and terminator (P_{TP1} and T_{TP1}). They are identical to pMT742 (Egel-Mitani, M. et al., \textit{Gene} 72 (1988) 113-120) (see Fig. 1) except for the region defined by the EcoR I-Xba I restriction sites encompassing the coding region for signal/leader/product.

The plasmids pAK527, pAK531, pAK555, pAK559, pAK562, pAK614 and 20 pAK625 were used as DNA templates in the PCR reactions applied in the construction of the leaders described in the examples. The synthetic DNA fragments serving as the direct template are shown in Figs. 13 - 17. With the exception of the shown DNA regions the plasmids are identical to pAK492 shown in Fig. 1.

All other methods and materials used are common state of the art knowledge (see, e.g. Sambrook, J., Fritsch, E.F. and Maniatis, T., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, New York, 1989).

5 EXAMPLE 1

Synthesis of the leader SEQ ID No. 4 for expression of the MI3 insulin precursor in S. cerevisiae (strain yAK546).

The leader SEQ ID No. 4 has the following amino acid sequence:

10 GlnProIleAspAspGluAsnThrThrSerValAsnLeuProVal

The following oligonucleotides were synthesised:

# 94  5'-TAAATCTATAACTACAAAAAACACATA-3'  SEQ ID No. 29
# 333 5'-GACTCTCTTAACTGGCAAGTGACA-3'  SEQ ID No. 30
# 312 5'-AAGTACAAAGCTTCAACCAAGTGAAACCACACCAAGTGTT
15  GGTAAACGAATCTCTTT-3'  SEQ ID No. 31
# 1845 5'-CATACACCAATATAAACGACGG-3'  SEQ ID No. 32

The following polymerase chain reactions (PCR) were performed using the GeneAmp PCR reagent kit (Perkin Elmer, 761 Main AveWalk, CT 06859, USA) according to the manufacturers instructions. During the reaction, the PCR mixtures were overlayed with 100 μl of mineral oil (Sigma Chemical CO, St. Louis MO, USA):

Polymerase chain reaction No. 1

5 μl of oligonucleotide # 94 (50 pmol)
25 5 μl of oligonucleotide # 333 (50 pmol)
10 μl of 10X PCR buffer
16 μl of dNTP mix
0.5 μl of Taq enzyme
0.5 μl of pAK527 plasmid (Fig. 13) as template (0.2 μg of DNA)
63 µl of water

A total of 12 cycles were performed, one cycle was 94°C for 1 min; 37°C for 2 min; 72°C for 3 min. The PCR mixture was then loaded onto a 2% agarose gel and electrophoresis was performed using standard techniques (Sambrook, J., Fritsch, E.F. and Maniatis, T., op.cit.). The resulting DNA fragment was cut out of the agarose gel and isolated using the Gene Clean kit (Bio 101 inc., PO BOX 2284, La Jolla, CA 92038, USA) according to the manufacturers instructions.

Polymerase chain reaction No. 2

5 µl of oligonucleotide # 312 (50 pmol)
5 µl of oligonucleotide # 94 (50 pmol)
10 µl of 10X PCR buffer
16 µl of dNTP mix
15 0.5 µl of Taq enzyme
10 µl of purified DNA fragment from PCR No. 1
53.5 µl of water

A total of 12 cycles were performed, one cycle was 94°C for 1 min; 37°C for 2 min; 72°C for 3 min.

The DNA fragment from polymerase chain reaction No. 2 was isolated and purified using the Gene Clean kit (Bio 101 inc., PO BOX 2284, La Jolla, CA 92038, USA) according to the manufacturers instructions.

The purified PCR DNA fragment was dissolved in 10 µl of water and restriction endonuclease buffer and cut with the restriction endonucleases Asp 718 and Hind III in a total volume of 15 µl according to standard techniques (Sambrook, J., Fritsch, E.F. and Maniatis, T., op.cit.). The 167 bp Asp 718/Hind III DNA fragment was subjected to electrophoresis on agarose gel and purified using The Gene Clean Kit as described. The S. cerevisiae expression plasmid pAK492 (shown in Fig. 1) is a derivative of the previously described plasmid pMT742 in
which the fragment encoding the signal/leader/insulin precursor has been replaced by the EcoR I-Xba I fragment shown in Fig. 2. This fragment has been synthesized on an Applied Biosystems DNA synthesizer in accordance with the manufacturer's instructions.

5 The plasmid pAK492 was cut with the restriction endonucleases Asp 718 and Xba I and the vector fragment of 10986 bp was isolated. The plasmid pAK492 was cut with the restriction endonucleases Hind III and Xba I and the DNA fragment of 140 bp encoding part of the MI3 insulin precursor was isolated. The three DNA fragments were ligated together using T4 DNA ligase under standard conditions (Sambrook, J., Fritsch, E.F. and Maniatis, T., *op.cit.*). The ligation mixture was then transformed into a competent *E. coli* strain (R-, M+) and transformants were identified by ampicillin resistance.

10 Plasmids were isolated from the resulting *E. coli* colonies using standard DNA miniprep technique (Sambrook, J., Fritsch, E.F. and Maniatis, T., *op.cit.*), checked with appropriate restrictions endonucleases i.e. EcoR I, Xba I, Nco I and Hind III. The selected plasmid, pAK546, was shown by DNA sequencing analysis (Sequenase, U.S. Biochemical Corp.) using the primer § 94 to contain a DNA sequence encoding the leader SEQ ID No. 4. For the DNA sequence encoding the leader SEQ ID No. 4, see Fig. 4). The plasmid pAK546 was transformed into *S. cerevisiae* strain MT663 as described in European published patent application No. 214 826 and the resulting strain was named yAK546. The DNA sequence of the protein coding region of the expression plasmid is given in Fig. 5.

**EXAMPLE 2**

Synthesis of the leader SEQ ID No. 6 for expression of the MI3 insulin precursor in *S. cerevisiae* (strain yAK531).

The leader SEQ ID No. 6 has the following amino acid sequence:

GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuProAla
The following oligonucleotide was synthesised:

\[
\text{# 331 5'}-\text{GAATCTCTTAGCTGGCAAGTTGACAGAAAGTAGTGTTAG} \\
\text{TTCAGAGTCGTAATT-3'} \\
\text{SEQ ID No. 33}
\]

The polymerase chain reaction was performed as described in Example 1 with the exception that oligonucleotide # 331 was used instead of oligonucleotide # 333.

The Asp 718/Hind III DNA fragment of 168 bp was subjected to electrophoresis on agarose gel and purified as described in Example 1. The Asp 718/Hind III DNA fragment was subcloned into the S. cerevisiae expression plasmid as described in Example 1. The selected plasmid, pAK531, was shown by DNA sequencing analysis, as described in Example 1, to contain a DNA sequence encoding the leader SEQ ID No. 6. For the DNA sequence encoding the leader SEQ ID No. 6, see Fig. 6. The plasmid pAK531 was transformed into S. cerevisiae strain MT663 as described in European patent application 86306721.1 and the resulting strain was named yAK531. The DNA sequences encoding the signal peptide and the insulin precursor MI3 were the same as those shown in Fig. 5.

20 EXAMPLE 3

Synthesis of the leader SEQ ID No. 8 for expression of the MI3 insulin precursor in S. cerevisiae (strain yAK547).

The leader SEQ ID No. 8 has the following amino acid sequence:

\[
\text{GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuProGlyAla}
\]

The following oligonucleotide was synthesised:

\[
\text{# 345 5'}-\text{AACGAATCTCTTAGCACCTGGCAAGTTGACAGAACT-3'} \text{ SEQ ID No. 34}
\]
The polymerase chain reaction was performed as described in Example 1 with the exception that oligonucleotide # 345 was used instead of oligonucleotide # 333 and plasmid pAK531 (Fig. 14) was used as template.

The Asp 718/Hind III DNA fragment of 171 bp was subjected to electrophoresis on agarose gel and purified as described in Example 1. The Asp 718/Hind III DNA fragment was subcloned into the S. cerevisiae expression plasmid as described in Example 1. The selected plasmid, pAK547, was shown by DNA sequencing analysis, as described in Example 1, to contain a DNA sequence encoding the leader SEQ ID No. 8. For the DNA sequence encoding the leader SEQ ID No. 8, see Fig. 7. The plasmid pAK547 was transformed into S. cerevisiae strain MT663 as described in European patent application No. 86306721.1 and the resulting strain was named yAK547. The DNA sequences encoding the signal peptide and the insulin precursor MI3 were the same as those shown in Fig. 5.

EXAMPLE 4
Synthesis of the leader SEQ ID No. 17 for expression of the MI3 insulin precursor in S. cerevisiae (strain yAK561).

The leader SEQ ID No. 17 has the following amino acid sequence: GlnProIleAspAspThrGluSerIleAsnThrThrLeuValAsnLeuProGlyAla

The following oligonucleotide was synthesised:

25 # 376 5'--AACGAATCTCTTAGCACCCTTGCAAGTTGACCAAAGTAG
TGTGATAGATTGTCGTC-3'  

SEQ ID No. 35

The polymerase chain reaction was performed as described in Example 1 with the exception that oligonucleotide # 376 was used instead of oligonucleotide # 333 and plasmid pAK555 (Fig. 15) was used as template.
The Asp 718/Hind III DNA fragment of 180 bp was subjected to electrophoresis on agarose gel and purified as described in Example 1. The Asp 718/Hind III DNA fragment was subcloned into the S. cerevisiae expression plasmid as described in Example 1. The selected plasmid, pAK561, was shown by DNA sequencing analysis, as described in Example 1, to contain a DNA sequence encoding the leader SEQ ID No. 17. For the DNA sequence encoding the leader SEQ ID No. 17, see Fig. 8. The plasmid pAK561 was transformed into S. cerevisiae strain MT663 as described in European patent application No. 86306721.1 and the resulting strain was named yAK561. The DNA sequences encoding the signal peptide and the insulin precursor MI3 were the same as those shown in Fig. 5.

EXAMPLE 5

15 Synthesis of the leader SEQ ID No. 16 for expression of the MI3 insulin precursor in S. cerevisiae (strain yAK559).

The leader SEQ ID No. 16 has the following amino acid sequence:
GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuMetAlaAspAspThr-
GluSerIleAsnThrThrLeuValAsnLeuProGlyAla

The following oligonucleotide was synthesised:

$375 5'-AACGAATCTCITAGCCTAGCTGGCAAGTTACACAAAGTAGT$
$GTTGATATGATCTGTCGTCAGCCATCAAGTTGAC-3' $ SEQ ID No. 36

The polymerase chain reaction was performed as described in Example 1 with the expection that oligonucleotide $375$ was used instead of oligonucleotide $333$ and plasmid pAK555 (Fig. 15) was used as template.

The Asp 718/Hind III DNA fragment of 222 bp was subjected to electrophoresis on agarose gel and purified as described in Example 1. The Asp 718/Hind III DNA fragment was subcloned into
the *S. cerevisiae* expression plasmid as described in Example 1. The selected plasmid, pAK559, was shown by DNA sequencing analysis, as described in Example 1, to contain a DNA sequence encoding the leader SEQ ID No. 16. For the DNA sequence 5 encoding the leader SEQ ID No. 16, see Fig. 9. The plasmid pAK559 was transformed into *S. cerevisiae* strain MT663 as described in European patent application No. 86306721.1 and the resulting strain was named yAK559. The DNA sequences encoding the signal peptide and the insulin precursor MI3 were the same 10 as those shown in Fig. 5.

**EXAMPLE 6**

Synthesis of the leader SEQ ID No. 19 for expression of the MI3 insulin precursor in *S. cerevisiae* (strain yAK580).

15 The leader SEQ ID No. 19 has the following amino acid sequence:
GlnProIleAspAspGluSerAsnThrThrSerValAsnLeuMetAlaAspThr-GluSerArgPheAlaThrAsnThrThrLeuValAsnLeuProLeu

The following oligonucleotide was synthesised:

# 384 5' - AACGATCTCTTCAATGGGCAATTTAACCAAAGTGTG
TAGTACGAATCTAGATTCTGCGTCGAGCCAT-3'  SEQ ID No. 37

The polymerase chain reaction was performed as described in Example 1 with the exception that oligonucleotide # 384 was used insted of oligonucleotide # 333 and plasmid pAK559 (Fig. 16) was used as template.

25 The Asp 718/Hind III DNA fragment of 228 bp was subjected to electrophoresis on agarose gel and purified as described in Example 1. The Asp 718/Hind III DNA fragment was subcloned into the *S. cerevisiae* expression plasmid as described in Example 1. The selected plasmid, pAK580, was shown by DNA sequencing 30 analysis, as described in Example 1, to contain a DNA sequence
encoding the leader SEQ ID No. 19. For the DNA sequence encoding the leader SEQ ID No. 19, see Fig. 10. The plasmid pAK580 was transformed into S. cerevisiae strain MT663 as described in European patent application No. 86306721.1 and the resulting strain was named yAK580. The DNA sequences encoding the signal peptide and the insulin precursor MI3 were the same as those shown in Fig. 5.

EXAMPLE 7
Synthesis of the leader SEQ ID No. 20 for expression of the MI3 insulin precursor in S. cerevisiae (strain yAK583).

The leader SEQ ID No. 20 has the following amino acid sequence:
GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuMetAlaAspAspThr-GluSerIleAsnThrThrLeuValAsnLeuAlaAsnValAlaMetAla

The following oligonucleotide was synthesised:

\[
\# 3905'-AACGAATCTCTTAGCCATGGCAACGTTAGCCAGTTAA\]

The polymerase chain reaction was performed as described in Example 1 with the exception that oligonucleotide \# 390 was used instead of oligonucleotide \# 333 and plasmid pAK559 (Fig. 16) was used as template.

The Asp 718/Hind III DNA fragment of 231 bp was subjected to electrophoresis on agarose gel and purified as described in Example 1. The Asp 718/Hind III DNA fragment was subcloned into the S. cerevisiae expression plasmid as described in Example 1. The selected plasmid, pAK583, was shown by DNA sequencing analysis, as described in Example 1, to contain a DNA sequence encoding the leader SEQ ID No. 20. For the DNA sequence encoding the leader SEQ ID No. 20, see Fig. 11. The plasmid pAK583 was transformed into S. cerevisiae strain MT663 as
described in European patent application No. 86306721.1 and the resulting strain was named yAK583. The DNA sequences encoding the signal peptide and the insulin precursor MI3 were the same as those shown in Fig. 5.

5 EXAMPLE 8
Synthesis of the leader SEQ ID No. 21 for expression of the MI3 insulin precursor in S. cerevisiae (strain yAK586).

The leader SEQ ID No. 21 has the following amino acid sequence:
10 GlnProIleAspAspThrGluSerAlaIleAsnThrThrLeuValAsnLeuProGlyAla

The following oligonucleotide was synthesised:

# 401 5'-AACGAATCTCTTAGCACCCTGGCAAGTTGACCAGAAATGAG
TGTGATAGCAATTCCAGTGC-3'  SEQ ID No. 39

The polymerase chain reaction was performed as described in Example 1 with the exception that oligonucleotide # 401 was used instead of oligonucleotide # 333 and plasmid pAK562 (Fig. 17) was used as template.

The Asp 718/Hind III DNA fragment of 183 bp was subjected to electrophoresis on agarose gel and purified as described in Example 1. The Asp 718/Hind III DNA fragment was subcloned into the S. cerevisiae expression plasmid as described in Example 1. The selected plasmid, pAK586, was shown by DNA sequencing analysis, as described in Example 1, to contain a DNA sequence encoding the leader SEQ ID No. 21, see Fig. 12. The plasmid pAK586 was transformed into S. cerevisiae strain MT663 as described in European patent application No. 86306721.1 and the resulting strain was named yAK586. The DNA sequences encoding the signal peptide and the insulin precursor MI3 were the same as those shown in Fig. 5.
EXAMPLE 9

Expression of the MI3 insulin precursor using selected leader sequences according to the present invention.

5 yeast strains harbouring plasmids as described above, were grown in YPD medium (Sherman, F. et al., Methods in Yeast Genetics, Cold Spring Harbor Laboratory Press, 1981). For each strain 6 individual 5 ml cultures were shaken at 30°C for 72 hours, with a final OD₆₅₀ of approx. 15. After centrifugation the supernatant was removed for HPLC analysis by which method the concentration of secreted insulin precursor was measured by a method described by Snel, L. et al. Chromatographia 24 (1987) 329-332.

In Table 1 the expression levels of the insulin precursor, MI3, obtained by use of selected leader sequences according to the present invention, are given as a percentage of the level obtained with transformants of pMT742, utilizing the MFα(1) leader of S. cerevisiae.

Table 1

<table>
<thead>
<tr>
<th>Leader</th>
<th>Expression level, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>MT748 α-leader</td>
<td>100</td>
</tr>
<tr>
<td>SEQ ID No. 15</td>
<td>87</td>
</tr>
<tr>
<td>SEQ ID No. 16</td>
<td>215</td>
</tr>
<tr>
<td>SEQ ID No. 17</td>
<td>157</td>
</tr>
<tr>
<td>SEQ ID No. 19</td>
<td>166</td>
</tr>
<tr>
<td>SEQ ID No. 20</td>
<td>86</td>
</tr>
<tr>
<td>SEQ ID No. 21</td>
<td>145</td>
</tr>
<tr>
<td>SEQ ID No. 22</td>
<td>137</td>
</tr>
<tr>
<td>SEQ ID No. 23</td>
<td>121</td>
</tr>
</tbody>
</table>
EXAMPLE 10
Synthesis of the leader SEQ ID No. 27 for expression of the extended MI3 insulin precursor in S. cerevisiae (strain yAK677).

The leader SEQ ID No. 27 has the following amino acid sequence:
GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuMetAlaAspAspThr-
GluSerArgPheAlaThrAsnThrThrLeuAspValValAsnLeuIleSerMetAla

The following oligonucleotides were synthesised:

10 # 440 5'-GGTTAACAACTTTGGAGCTTCAGCTTCAGCTTTCTCTCTTTAGCCAT
    GGAGATCAAGTTAAACCATCCAAGAGTAGTGT-3'    SEQ ID No. 64

and

# 441 5'-CAAGTACAAAAGCTTCAACCAAGTGGGAACCGCACAAGTGTTGGTTAACG
    AACTT-3'       SEQ ID No. 65

15 Polymerase chain reactions were performed as described in Example 1 with the exception that oligonucleotide # 440 was used instead of oligonucleotide # 333 and plasmid pAK614 was used as template. For the second polymerase chain reaction, oligonucleotide # 441 was used instead of oligonucleotide # 20 312.

The purified PCR DNA fragment was isolated and digested with the restriction endonucleases Asp 718 and Hind III as described in Example 1. The Asp 718/Hind III DNA fragment of 268 bp was subjected to electrophoresis on agarose gel and purified as described in Example 1. The Asp 718/Hind III DNA fragment was subcloned into the S. cerevisiae expression plasmid as described in Example 1, with the exception that the 140 bp Hind III/Xba I DNA fragment was derived from pAK602 and encodes AspB28 human insulin. The selected 30 plasmid, pAK616, was shown by DNA sequencing analysis, as
described in Example 1, to contain the DNA sequence encoding the leader SEQ ID No. 27. For the DNA sequence, SEQ ID No. 66, encoding the leader SEQ ID No. 27, see Fig. 18. The Asp 718/Hind III DNA fragment of 268 bp from pAK616 was isolated and ligated with the 10986 bp Asp 718/Xba I DNA fragment from pAK601 and the 140 bp DNA fragment Hind III/Xba I from pAK464 (encoding an extended version of AspB28 human insulin) and named pAK 625. The 180 bp Asp 718/Nco I DNA fragment from pAK625 was isolated and ligated with the 221 bp Nco I/Xba I DNA fragment from pJB146 (encoding and extended version of the insulin precursor) and the 10824 bp Asp 718/Xba I DNA fragment from pAK601 and the resulting plasmid was named pAK677. The plasmid pAK677 was transformed into \textit{S. cerevisiae} strain MT663 as described in European patent application 86306721.1 and the resulting strain was named yAK677. With the exception of the DNA sequence encoding the leader, the DNA sequence encoding the signal peptide is as described in Fig. 5. The DNA sequence coding for the extended MI3 insulin precursor is as described in Fig. 19.

**EXAMPLE 11**

Synthesis of the leader SEQ ID No. 67 for expression of the extended MI3 insulin precursor in \textit{S. cerevisiae} (yAK680)

The leader SEQ ID No. 67 has the following amino acid sequence:

\text{GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuMetAlaAspAspThrGluSerArgPheAlaThrAsnThrThrLeuAlaLeuAspValValAsnLeuIleSerMet Ala}

The following oligonucleotide was synthesised:

\text{30 # 577 5'-TCTTTAGCCATGAGATCAAGTTAACAACATCCAAAG CCAAAGTACTGTT-3'}
The PCR was performed as described in Example 1 with the exception that oligonucleotide # 577 was used instead of oligonucleotide # 333 and plasmid pAK625 was used as template and the second PCR was not performed. The PCR fragment was 5 digested with the restriction endonucleases Asp 718 and Nco I as described in Example 1.

The Asp 718/Nco I DNA fragment of 190 bp was subjected to electrophoresis on agarose gel and purified as described in Example 1 expect that the 10824 bp Asp 718/Xba I vector DNA fragment was isolated from and from pAK601. The 190 bp Asp 718/Nco I DNA fragment was subcloned into the S. cerevisiae expression plasmid as described in Example 1, expect that the 221 bp DNA fragment Nco I/Xba I (encoding an extended version of the MI3 insulin precursor) was isolated from pAK677 and used instead of the Hind III/Xba I DNA fragment. The selected plasmid was shown by DNA sequencing analysis as described in Example 1 to contain the DNA sequence encoding the leader SEQ ID No. 67 and named pAK680. For the DNA sequence, SEQ ID No. 69, encoding the leader SEQ ID No. 67, see Fig. 20. The plasmid pAK680 was transformed into S. cerevisiae strain MT663 as described in European patent application 86306721.1 and the resulting strain was named yAK680. With the exception of the DNA sequence encoding the leader, the DNA sequence encoding the signal peptide is as described in Fig. 5 and the extended insulin precursor MI3 DNA sequence is as described in Fig. 19.

**EXAMPLE 12**

Expression of N-terminally extended MI3 insulin precursors using the leader sequences SEQ ID No. 27 and SEQ ID No. 67 according to the present invention.

Yeast strains harbouring plasmids as described above, were grown in YPD medium (Sherman, F. et al., *Methods in Yeast*
Genetics, Cold Spring Harbor Laboratory Press, 1981). For each strain 6 individual 5 ml cultures were shaken at 30°C for 72 hours, with a final OD₆₀₀ of approximately 15. After centrifugation the supernatant was removed for HPLC analysis by which method the concentration of secreted insulin precursor was measured by a method described by Snel, L. et al. *Chromatographia* 24 (1987) 329-332.

In Table 2 the expression levels of some N-terminally extended MI3 insulin precursors, obtained by use of the 10 leader sequences SEQ ID No. 27 and SEQ ID No. 67 according to the present invention, are given as a percentage of the level obtained with transformants of pMT742, utilizing the Mα(1) leader of *S. cerevisiae*.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Signal peptide</th>
<th>Leader</th>
<th>Extension</th>
<th>Relative to MT748</th>
</tr>
</thead>
<tbody>
<tr>
<td>MT748</td>
<td>α</td>
<td>α</td>
<td></td>
<td></td>
</tr>
<tr>
<td>yAK675</td>
<td>YAP3</td>
<td>SEQ ID No. 27</td>
<td>EEAEEAEAP K</td>
<td>251%</td>
</tr>
<tr>
<td>yAK677</td>
<td>YAP3</td>
<td>SEQ ID No. 27</td>
<td>EEAEEAEK PK</td>
<td>224%</td>
</tr>
<tr>
<td>yAK681</td>
<td>YAP3</td>
<td>SEQ ID No. 67</td>
<td>EEAEEAEAP K</td>
<td>248%</td>
</tr>
<tr>
<td>yAK680</td>
<td>YAP3</td>
<td>SEQ ID No. 67</td>
<td>EEAEEAEK PK</td>
<td>362%</td>
</tr>
</tbody>
</table>
SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:
   (A) NAME: Novo Nordisk A/S
   (B) STREET: Novo Allé
   (C) CITY: DK-2880 Bagsvaerd
   (E) COUNTRY: Denmark
   (G) TELEPHONE: +45 44448888
   (H) TELEFAX: +45 44490555
   (I) TELEX: 37173

(ii) TITLE OF INVENTION: SYNTHETIC LEADERS PEPTIDE SEQUENCES

(iii) NUMBER OF SEQUENCES: 73

(iv) CORRESPONDENCE ADDRESS:
   (A) ADDRESSEE: Novo Nordisk A/S
      Corporate Patents
   (B) STREET: Novo Allé
   (C) CITY: DK-2880 Bagsvaerd
   (E) COUNTRY: Denmark

(v) COMPUTER READABLE FORM:
   (A) MEDIUM TYPE: Floppy disk
   (B) COMPUTER: IBM PC compatible
   (C) OPERATING SYSTEM: PC-DOS/MS-DOS
   (D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:
   (A) APPLICATION NUMBER:
   (B) FILING DATE:
   (C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:
   (A) APPLICATION NUMBER: DK 0705/94 and US 08/282,852
   (B) FILING DATE: 16-JUN-1994 and 29-JUL-1994

(viii) ATTORNEY/AGENT INFORMATION:
   (A) NAME: Jørgensen, Dan et al.
   (C) REFERENCE/DOCKET NUMBER: 4085-W0, DJ

(ix) TELECOMMUNICATION INFORMATION:
   (A) TELEPHONE: +45 44448888
   (B) TELEFAX: +45 44493256

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 15 amino acids
   (B) TYPE: amino acid
   (D) TOPOLOGY: linear
(11) MOLECULE TYPE: peptide

(xv) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Gln Pro Ile Asp Glu Asp Asn Asp Thr Ser Val Asn Leu Pro Ala
1  5  10  15

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xv) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Gln Pro Ile Asp Glu Asn Thr Ser Val Thr Ser Val Asn Leu Pro Ala
1  5  10  15

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xv) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Gln Pro Ile Asp Glu Ser Asn Thr Ser Val Asn Leu Pro Ala
1  5  10  15

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xv) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Gln Pro Ile Asp Glu Asn Thr Ser Val Asn Leu Pro Val
1  5  10  15

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Gln Pro Ile Asp Asp Thr Glu Asn Thr Thr Ser Val Asn Leu Pro Ala
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Gln Pro Ile Asp Asp Thr Glu Ser Asn Thr Thr Ser Val Asn Leu Pro
1 5 10 15
Ala

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Gln Pro Ile Asp Asp Glu Asn Thr-Thr Ser Val Asn Leu Met Ala
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:
<table>
<thead>
<tr>
<th>1</th>
<th>5</th>
<th>10</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gln Pro Ile Asp Asp Thr Glu Ser Asn Thr Ser Val Asn Thr Leu Pro</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Gly Ala

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:
Gln Pro Ile Asp Asp Thr Glu Ser Asn Thr Ser Val Asn Leu Met
1 5 10 15
Ala

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:
Gln Pro Ile Asp Asp Thr Glu Ser Asn Thr Ser Val Asn Val Pro
1 5 10 15
Thr

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:
Gln Pro Ile Asp Asp Thr Glu Ser Asn Thr Thr Leu Val Asn Val Pro
1      5     10     15
Th

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 17 amino acids
    (B) TYPE: amino acid
    (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Gln Pro Ile Asp Asp Thr Glu Ser Asn Thr Thr Ser Val Asn Leu Pro
1      5     10     15
Thr

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 18 amino acids
    (B) TYPE: amino acid
    (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Gln Pro Ile Asp Asp Thr Glu Ser Asn Thr Thr Leu Val Asn Val Pro
1      5     10     15
Gly Ala

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 21 amino acids
    (B) TYPE: amino acid
    (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:
Gln Pro Ile Asp Asp Thr Glu Ser Asn Thr Thr Ser Val Asn Leu Met
1 5 10 15

Ala Pro Ala Val Ala
20

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 25 amino acids
   (B) TYPE: amino acid
   (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(x) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Gln Pro Ile Asp Asp Thr Glu Ser Asn Thr Thr Ser Val Asn Leu Met
1 5 10 15

Asp Leu Ala Val Gly Leu Pro Gly Ala
20 25

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 33 amino acids
   (B) TYPE: amino acid
   (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(x) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Gln Pro Ile Asp Asp Thr Glu Ser Asn Thr Thr Ser Val Asn Leu Met
1 5 10 15

Ala Asp Thr Glu Ser Ile Asn Thr Leu Val Asn Leu Pro Gly 20 25 30

Ala

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 19 amino acids
   (B) TYPE: amino acid
   (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(x) SEQUENCE DESCRIPTION: SEQ ID NO:17:
Gln Pro Ile Asp Asp Thr Glu Ser Ile Asn Thr Thr Leu Val Asn Leu
1   5   10   15
Pro Gly Ala

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 18 amino acids
   (B) TYPE: amino acid
   (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:18:
   Gln Pro Ile Asp Asp Thr Glu Ser Asn Thr Thr Leu Val Asn Leu Pro
   1   5   10   15
   Gly Ala

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 35 amino acids
   (B) TYPE: amino acid
   (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:19:
   Gln Pro Ile Asp Asp Thr Glu Ser Asn Thr Ser Val Asn Leu Met
   1   5   10    
   Ala Asp Asp Thr Glu Ser Arg Phe Ala Thr Asn Thr Thr Leu Val Asn
   20  25    30
   Leu Pro Leu
   35

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 36 amino acids
   (B) TYPE: amino acid
   (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:20:
Gln Pro Ile Asp Asp Thr Glu Ser Asn Thr Thr Ser Val Asn Leu Met
1 5 10 15

Ala Asp Asp Thr Glu Ser Ile Asn Thr Thr Leu Val Asn Leu Ala Asn
20 25 30

Val Ala Met Ala
35

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 20 amino acids
    (B) TYPE: amino acid
    (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Gln Pro Ile Asp Asp Thr Glu Ser Ala Ile Asn Thr Thr Leu Val Asn
1 5 10 15

Leu Pro Gln Ala
20

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 21 amino acids
    (B) TYPE: amino acid
    (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:22:

Gln Pro Ile Asp Asp Thr Glu Ser Phe Ala Thr Asn Thr Thr Leu Val
1 5 10 15

Asn Leu Pro Gly Ala
20

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 36 amino acids
    (B) TYPE: amino acid
    (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

Gln Pro Ile Asp Asp Thr Glu Ser Ile Asn Thr Thr Leu Val Asn Leu
1 5 10 15

Met Ala Asp Thr Glu Ser Arg Phe Ala Thr Asn Thr Thr Leu Val
20 25 30

Asn Leu Pro Leu
35

(2) INFORMATION FOR SEQ ID NO:24:

(1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 39 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:24:

Gln Pro Ile Asp Asp Thr Glu Ser Ile Asn Thr Thr Leu Val Asn Leu
1 5 10 15

Met Ala Asp Thr Glu Ser Arg Phe Ala Thr Asn Thr Thr Leu Asp
20 25 30

Val Val Asn Leu Pro Gly Ala
35

(2) INFORMATION FOR SEQ ID NO:25:

(1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ixii) SEQUENCE DESCRIPTION: SEQ ID NO:25:

Gln Pro Ile Asp Asp Thr Glu Ser Ala Ala Ile Asn Thr Thr Leu Val
1 5 10 15

Asn Leu Pro Gly Ala
20

(2) INFORMATION FOR SEQ ID NO:26:

(1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 39 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(11) MOLECULE TYPE: peptide

(x1) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

Gln Pro Ile Asp Asp Thr Glu Ser Asn Thr Thr Ser Val Asn Leu Met
  1    5   10   15

Ala Asp Asp Thr Glu Ser Arg Phe Ala Thr Asn Thr Thr Leu Val Asn
  20   25   30

Leu Ala Asn Val Ala Met Ala
  35

(2) INFORMATION FOR SEQ ID NO: 27:

(1) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 39 amino acids
   (B) TYPE: amino acid
   (D) TOPOLOGY: linear

(11) MOLECULE TYPE: peptide

(x1) SEQUENCE DESCRIPTION: SEQ ID NO: 27:

Gln Pro Ile Asp Asp Thr Glu Ser Asn Thr Thr Ser Val Asn Leu Met
  1    5   10   15

Ala Asp Asp Thr Glu Ser Arg Phe Ala Thr Asn Thr Thr Leu Asp Val
  20   25   30

Val Asn Leu Ile Ser Met Ala
  35

(2) INFORMATION FOR SEQ ID NO: 28:

(1) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 39 amino acids
   (B) TYPE: amino acid
   (D) TOPOLOGY: linear

(11) MOLECULE TYPE: peptide

(x1) SEQUENCE DESCRIPTION: SEQ ID NO: 28:

Gln Pro Ile Asp Asp Thr Glu Ser Asn Thr Thr Ser Val Asn Leu Met
  1    5   10   15

Ala Asn Thr Thr Glu Ser Arg Phe Ala Thr Asn Thr Thr Leu Asp Val
  20   25   30
Val Asn Leu Ile Ser Met Ala
35

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(x) SEQUENCE DESCRIPTION: SEQ ID NO:29:

TAAATCTATA ACTACAAAAA ACACATA

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(x) SEQUENCE DESCRIPTION: SEQ ID NO:30:

GACTCTCTTA ACTGGCAAGT TGACA

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 56 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(x) SEQUENCE DESCRIPTION: SEQ ID NO:31:

AAGTACAAAG CTTCAACCA GTGAGAACCA CACAAAGTT GGTTAAGAA TCTCTT

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:
CATACACAAT ATAAACGACG G

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 55 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:
GAATCTCTTA GCTGGCAAAGT TGACAGAAGT AGTGTTAGTT TCAGAGTCGT CAATT

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:
AACGAATCTC TTAGCACCCTG GCAAGTTGAC AGAAGT

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 60 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:
AACGAATCTC TTAGCACCCTG GCAAGTTGAC CAAAGTASTG TTGATAGATT CAGTGTGCTC

(2) INFORMATION FOR SEQ ID NO:36:
(1) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 75 base pairs
   (B) TYPE: nucleic acid
   (C) STRANDEDNESS: single
   (D) TOPOLOGY: linear

(11) MOLECULE TYPE: cDNA

(11) SEQUENCE DESCRIPTION: SEQ ID NO:36:

AACGAATCTC TTAGCACCTG GCAAGTTAAC CAAAATAGTG TTGATAGATT CAGTGTCGTC
AGCCATCAAG TTGAC

(2) INFORMATION FOR SEQ ID NO:37:

(1) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 72 base pairs
   (B) TYPE: nucleic acid
   (C) STRANDEDNESS: single
   (D) TOPOLOGY: linear

(11) MOLECULE TYPE: cDNA

(11) SEQUENCE DESCRIPTION: SEQ ID NO:37:

AACGAATCTC TTCAATGGCA AGTTAACCAA AGTAGTGTTA GTAGCGAATC TAGATTCA GT
GTCGCAGC AT

(2) INFORMATION FOR SEQ ID NO:38:

(1) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 45 base pairs
   (B) TYPE: nucleic acid
   (C) STRANDEDNESS: single
   (D) TOPOLOGY: linear

(11) MOLECULE TYPE: cDNA

(11) SEQUENCE DESCRIPTION: SEQ ID NO:38:

AACGAATCTC TTAGCCATG GCAAGTTGC CAAAGTAAACC AAAGT

(2) INFORMATION FOR SEQ ID NO:39:

(1) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 61 base pairs
   (B) TYPE: nucleic acid
   (C) STRANDEDNESS: single
   (D) TOPOLOGY: linear

(11) MOLECULE TYPE: cDNA
(x1) SEQUENCE DESCRIPTION: SEQ ID NO:39:

AACGAATCTC TTAGCACCTG GCAAAGTTGAC CAAAAGTAGTG TTGATAGCAG ATTCAGTGTC

G

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 372 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 82..351

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:40:

GAATTACATTC AAGAATGTT CAAACAAGAA GATTACAAAC TATCAAATTC ATACAACAATA

TAAACGACGG GTACCAAAAT A ATG AAA CTG AAA ACT GTA AGA TCT GCC GTC

Met Lys Leu Lys Thr Val Arg Ser Ala Val

1 5 10

CTT TCG TCA CTC TTT GCA TCT CAG GTC CTT GGC CAA ACA ATA GAC GAA

Leu Ser Ser Leu Phe Ala Ser Gln Val Leu Gly Gln Pro Ile Asp Glu

15 20 25

GAC AAC GAC ACT TCT TCC ATG GCT AAG AGA TTC GTT AAC CAA CAC TTG

Asp Asn Asp Thr Ser Ser Met Ala Lys Arg Phe Val Asn Gln His Leu

30 35 40

TGC GGT TCC CAC TTG GAA GCT TG TGC GTT GGC GGT GAA AGA

Cys Gly Ser His Leu Val Glu Ala Ala Tyr Leu Val Cys Gly Glu Arg

45 50 55

GCT TTC TCC ACT ACT CCT AAG GCT GCT AAG GGT ATT GTC GAG CAA TGC

Gly Phe Phe Tyr Thr Pro Lys Ala Ala Lys Ile Val Glu Gln Cys

60 65 70

TGT ACC TCC ATC TCC TTG TAC CAA TTG GAA AAC TAC TGC AAC TAGACGCAGC

Cys Thr Ser Ile Cys Ser Leu Tyr Glu Leu Glu Asn Tyr Cys Asn

75 80 85 90

CCGCAGGCTC TAGA

(2) INFORMATION FOR SEQ ID NO:41:

(1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 89 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(i1) MOLECULE TYPE: peptide

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:41:

Met Lys Leu Lys Thr Val Arg Ser Ala Val Leu Ser Ser Leu Phe Ala
1 5 10
Ser Glu Val Leu Gly Gln Pro Ile Asp Glu Asp Asn Thr Ser Ser
20 25 30
Met Ala Lys Arg Phe Val Asn Gln His Leu Cys Gly Ser His Leu Val
35 40 45
Glu Ala Leu Tyr Leu Val Cys Gly Glu Arg Gly Phe Phe Tyr Thr Pro
50 55 60
Lys Ala Ala Lys Gly Ile Val Glu Gln Cys Cys Thr Ser Ile Cys Ser
65 70 75 80
Leu Tyr Gln Leu Glu Asn Tyr Cys Asn
85

(2) INFORMATION FOR SEQ ID NO:42:

(1) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 45 base pairs
   (B) TYPE: nucleic acid
   (C) STRANDEDNESS: single
   (D) TOPOLOGY: linear

(i1) MOLECULE TYPE: cDNA

(ix) FEATURE:
   (A) NAME/KEY: CDS
   (B) LOCATION: 1..45

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:42:

CAA CCA ATT GAC GAC GAA AAC ACT ACT TCT GTC AAC TTG CCA GTT
Gln Pro Ile Asp Asp Glu Asn Thr Thr Ser Val Asn Leu Pro Val
1 5 10

(2) INFORMATION FOR SEQ ID NO:43:

(1) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 15 amino acids
   (B) TYPE: amino acid
   (D) TOPOLOGY: linear
(11) MOLECULE TYPE: peptide

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:43:

Gln Pro Ile Asp Asp Glu Asn Thr Thr Ser Val Asn Leu Pro Val
1  5  10  15

(2) INFORMATION FOR SEQ ID NO:44:

(1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 297 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..276

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:44:

ATG AAA CTG AAA ACT GTA AGA TCT GCG GTC CTT TCG TCA CTC TTT GCA 48
Met Lys Leu Lys Thr Val Arg Ser Ala Val Leu Ser Ser Leu Phe Ala
1  5  10  15

TCT CAG GTC CTT GGC CAA CCA ATT GAC GAC GAA AAC ACT ACT TCT GTC 96
Ser Gin Val Leu Gly Gin Pro Ile Asp Asp Glu Asn Thr Thr Ser Val
20  25  30

AAC TTG CCA GTT AAG AGA TTC GTT AAC CAC TTG GTG TGT TTC CAC 144
Asn Leu Pro Val Lys Arg Ala TTC GAA GGT TTC TAC
35  40  45

TTG GAA GCT TTG TAC TTG GTC GTG GAT GAA AGA GGT TTC TAC 192
Leu Val Gin Ala Leu Tyr Leu Val Cys Gin Glu Gly Phe Phe Tyr
50  55  60

ACT CCT AAG GCT AAG GGT ATT GTC GAA CAA TGC TGT ACC TCC ATC 240
Thr Pro Lys Ala Ala Lys Gly Ile Val Glu Gin Cys Thr Ser Ile
65  70  75  80

TGC TCC TTG TAC CAA TTG GAA AAC TAC TGC AAC TAGACGCA GCGCAGGCTC 293
Cys Ser Leu Tyr Gin Leu Glu Asn Tyr Cys Gin
85  90

TAGA

(2) INFORMATION FOR SEQ ID NO:45:

(1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 91 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

```
Met Lys Leu Lys Thr Val Arg Ser Ala Val Leu Ser Ser Leu Phe Ala
1    5    10    15
Ser Glu Val Leu Gly Gln Pro Ile Asp Asp Glu Asn Thr Thr Ser Val
20   25    30
Asn Leu Pro Val Lys Arg Phe Val Asn Gln His Leu Cys Gly Ser His
30   40    45
Leu Val Glu Ala Leu Tyr Leu Val Cys Gly Glu Arg Gly Phe Phe Tyr
50   55    60
Thr Pro Lys Ala Ala Lys Gly Ile Val Glu Gln Cys Cys Thr Ser Ile
65   70    75    80
Cys Ser Leu Tyr Glu Leu Glu Asn Tyr Cys Asn
85   90
```

(2) INFORMATION FOR SEQ ID NO:46:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 51 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

(A) NAME/KEY: CDS

(1) LOCATION: 1..51

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

```
CAA CCA ATT GAC GAC ACT GAA TCT AAC ACT TCT GTC AAC TTG CCA
Gl n Pro Ile Asp Asp Thr Gl u Ser Asn Thr Thr Ser Val Asn Leu Pro
1    5    10    15

GCT
Al a
```

(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 17 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

Gln Pro Ile Asp Asp Thr Glu Ser Asn Thr Thr Ser Val Asn Leu Pro
1  5  10  15
Ala

(2) INFORMATION FOR SEQ ID NO:48:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 54 base pairs
   (B) TYPE: nucleic acid
   (C) STRANDEDNESS: single
   (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
   (A) NAME/KEY: CDS
   (B) LOCATION: 1..54

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

CAA CCA ATT GAC GAC ACT GAA TCT AAT GTC AAC TTG CCA
Gln Pro Ile Asp Asp Thr Glu Ser Asn Thr Thr Ser Val Asn Leu Pro
1  5  10  15

GGT GCT
Gly Ala

(2) INFORMATION FOR SEQ ID NO:49:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 57 base pairs
   (B) TYPE: nucleic acid
   (C) STRANDEDNESS: single
   (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
   (A) NAME/KEY: CDS
   (B) LOCATION: 1..57

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

CAA CCA ATT GAC GAC ACT GAA TCT ATC AAC ACT TTG GTC AAC TTG
Gln Pro Ile Asp Asp Thr Glu Ser Ile Asn Thr Thr Leu Val Asn Leu
1  5  10  15
CCA  GGT  GCT
Pro  Gly  Ala

(2) INFORMATION FOR SEQ ID NO:50:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 99 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..99

(x) SEQUENCE DESCRIPTION: SEQ ID NO:50:

CAA  CCA  ATT  GAC  GAC  ACT  GAA  TCT  AAC  ACT  ACT  TCT  GTC  AAC  TTG  ATG
Gln  Pro  Ile  Asp  Asp  Thr  Glu  Ser  Asn  Thr  Thr  Ser  Val  Asn  Leu  Met
1    5      10   15

GCT  GAC  GAC  ACT  GAA  TCT  ATC  AAC  ACT  ACT  TCT  GCT  ATT  AAC  TTG  CCA  GGT
Ala  Asp  Asp  Thr  Glu  Ser  Ile  Asn  Thr  Thr  Leu  Val  Asn  Leu  Pro  Gly
20   25     30

GCT
Ala

(2) INFORMATION FOR SEQ ID NO:51:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 105 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..105

(x) SEQUENCE DESCRIPTION: SEQ ID NO:51:

CAA  CCA  ATT  GAC  GAC  ACT  GAA  TCT  AAC  ACT  ACT  TCT  GTC  AAC  TTG  ATG
Gln  Pro  Ile  Asp  Asp  Thr  Glu  Ser  Asn  Thr  Thr  Ser  Val  Asn  Leu  Met
1    5      10   15

GCT  GAC  GAC  ACT  GAA  TCT  ATC  AAC  ACT  ACT  TCT  GCT  ATT  AAC  TTG  CCA  GGT
Ala  Asp  Asp  Thr  Glu  Ser  Ile  Asn  Thr  Thr  Leu  Val  Asn  Leu  Pro  Gly
20   25     30

GCT
Ala
GAC GAC ACT GAA TCT AGA TTC GCT ACT AAC ACT TTG GTT AAC
Ala Asp Asp Thr Glu Ser Arg Phe Ala Thr Asn Thr Thr Leu Val Asn
20 25 30
TTG CCA TTG
Leu Pro Leu
35

(2) INFORMATION FOR SEQ ID NO:52:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 108 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..108

(x) SEQUENCE DESCRIPTION: SEQ ID NO:52:

CAA CCA ATT GAC GAC ACT GAA TCT AAC ACT ACT TCT GTC AAC TTG ATG
Gln Pro Ile Asp Asp Thr Glu Ser Asn Thr Thr Ser Val Asn Leu Met
1 5 10 15

GCT GAC GAC ACT GAA TCT ATC AAC ACT ACT TTG TTG AAC TTG GCC AAC
Ala Asp Asp Thr Glu Ser Ile Asn Thr Thr Leu Val Asn Leu Ala Asn
20 25 30

GTT GCC ATG GCT
Val Ala Met Ala
35

(2) INFORMATION FOR SEQ ID NO:53:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 60 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..60

(x) SEQUENCE DESCRIPTION: SEQ ID NO:53:
CAACA ATT GAC GAC ACT GAA TCT GCT ATC AAC ACT ACT TTG GTC AAC 48
Gln Pro Ile Asp Asp Thr Glu Ser Ala Ile Asn Thr Thr Leu Val Asn
1 5 10 15

TTG CCA GGT GCT 60
Leu Pro Gly Ala
20

(2) INFORMATION FOR SEQ ID NO:54:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 276 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 113..274

(x) SEQUENCE DESCRIPTION: SEQ ID NO:54:

TTAAATCTAT AACTACAAAA AACACATACA GGAATTCATT CAAGAATGTT TCAAACAAGA 60
AGATTACAAAA CTATCAATT CATAACACAAT ATAAACGACG GGTAACAAAA TAT AG
Net
1

AAA CTG AAA ACT GTA AGA TCT GCG GTC CTT TCG TCA CTC TTT GCA TCT 163
Lys Leu Lys Thr Val Arg Ser Ala Val Leu Ser Ser Leu Phe Ala Ser
5 10 15

CAG GTC CTT GTT GCA CCA ATT GAC GAC GAA AAC ACT ACT TCT GTT AAC 211
Gln Val Leu Gly Gln Pro Ile Asp Asp Glu Asn Thr Thr Ser Val Asn
20 25 30

TTG CCA GCT AAG AGA TTC GTT GAC AAC CAA CAC TTG TGC GGT TCC CAC TTG 259
Leu Pro Ala Lys Arg Phe Val Asn Gln Leu Cys Gly Ser His Leu
35 40 45

GTT GAA GCT TTG TAC TT 276
Val Glu Ala Leu Tyr
50

(2) INFORMATION FOR SEQ ID NO:55:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 54 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(x1) SEQUENCE DESCRIPTION: SEQ ID NO:55:

Met Lys Leu Lys Thr Val Arg Ser Ala Val Leu Ser Ser Leu Phe Ala 1 5 10 15
Ser Glu Val Leu Gly Gln Pro Ile Asp Asp Glu Asn Thr Thr Ser Val 20 25 30
Asn Leu Pro Ala Lys Arg Phe Val Asn Glu His Leu Cys Gly Ser His 35 40 45
Leu Val Glu Ala Leu Tyr 50

(2) INFORMATION FOR SEQ ID NO:56:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 282 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 113..280

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:56:

TTAAATCTAT AACTACAAAA AACACATACA GGAATCATT CAAGAATAGT TCAAACAAGA

AGATTACAAAA CTATCAATTT CATACACAAT ATAAACGACG GGTACCAAAAA TA ATG

Met 1

AAA CTG AAA ACT GTA AGA TCT GGC GTC CTT TCG TCA CTC TTT GCA TCT
Lys Leu Lys Thr Val Arg Ser Ala Val Leu Ser Ser Leu Phe Ala Ser 5 10 15

CAG GTC CTT GGC CAA CCA ATT GAC GAC ACT GAA TCT AAC ACT ACT TCT
Gln Val Leu Gly Gln Pro Ile Asp Asp Thr Glu Ser Asn Thr Thr Ser 20 25 30

GTC AAC TTG CCA GCT AAG AGA TTC GTT AAC CAA CACC TTG TGC GGT TCC
Val Asn Leu Pro Ala Lys Arg Phe Val Asn Glu His Leu Cys Gly Ser 35 40 45

CAC TTG GTT GAA GCT TTG TAC TT
His Leu Val Glu Ala Leu Tyr 50 55

(2) INFORMATION FOR SEQ ID NO:57:
(1) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 56 amino acids
   (B) TYPE: amino acid
   (D) TOPOLOGY: linear

(i) MOLECULE TYPE: peptide

(x) SEQUENCE DESCRIPTION: SEQ ID NO:57:

Met Lys Leu Lys Thr Val Arg Ser Ala Val Leu Ser Ser Leu Phe Ala
    1 5 10

Ser Glu Val Leu Gly Gln Pro Ile Asp Thr Glu Ser Asn Thr Thr
    20 25 30

Ser Val Asn Leu Pro Ala Lys Arg Phe Val Asn Gln His Leu Cys Gly
    35 40 45

Ser His Leu Val Glu Ala Leu Tyr
    50 55

(2) INFORMATION FOR SEQ ID NO:58:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 282 base pairs
   (B) TYPE: nucleic acid
   (C) STRANDEDNESS: single
   (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
   (A) NAME/KEY: CDS
   (B) LOCATION: 113..280

(x) SEQUENCE DESCRIPTION: SEQ ID NO:58:

TTAAATCTAT AACTACAAA AACACATACA GGAATTCATT CAAGAATAGT TCAAAACAAGA
AGATTACAAA CTATCAATTT CATACACAAT ATAAAACGACG GGTACCAAAA TA ATG
    115

Met

AAA CTG AAA ACT GTA AGA TCT GCG GTC TTT TCA CTC TTG CTA CTC TTT GCA TCT
Lys Leu Lys Thr Val Arg Ser Ala Val Leu Ser Ser Leu Phe Ala Ser
    5 10 15

CAG GTC CTT GGC CAA CCA ATT GAC GAC ACT GAA GCT ACT ACT ACT TCT
Gln Val Leu Gly Gln Pro Ile Asp Asp Thr Glu Ser Asn Thr Thr Ser
    20 25 30

GTC AAC TTG ATG GCT AAG AGA TTC GGT AAC CAA CAC TTG TGC GGT TCC
Val Asn Leu Met Ala Lys Arg Phe Val Asn Gln His Leu Cys Gly Ser
    35 40 45
CAC TTG GTT GAA GCT TTG TAC TT
His Leu Val Glu Ala Leu Tyr
50 55

(2) INFORMATION FOR SEQ ID NO:59:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 56 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

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

(2) INFORMATION FOR SEQ ID NO:60:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 330 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 113..328

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

TTAAATCTAT AACTACAAAA AACACATACA GGAATTCAT CAAGAATAGT TCAAAACAAGA
1
AGATTACAAA CTATCAATT CATAACAAAT ATAAACGACG GTACAAAAA TA ATG
115
 Met

AAA CTG AAA ACT GTA AGA TCT GCG GTC CTT TCG TCA CTC TTT GCA TCT
Lys Leu Lys Thr Val Arg Ser Ala Val Leu Ser Ser Leu Phe Ala Ser
5 10 15
163
### (2) INFORMATION FOR SEQ ID NO:61:

#### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 72 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

#### (ii) MOLECULE TYPE: peptide

#### (vi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

```
  Met Lys Leu Lys Thr Val Arg Ser Ala Val Leu Ser Ser Leu Phe Ala
  1  5  10  15
 Ser Gln Val Leu Gly Gln Pro Ile Asp Thr Glu Ser Asn Thr Thr
 20  25  30
 Ser Val Asn Leu Met Ala Asp Thr Glu Ser Ile Asn Thr Thr Leu
 35  40  45
Val Asn Leu Pro Gly Ala Lys Arg Phe Val Asn Gln His Leu Cys Gly
 50  55  60
 Ser His Leu Val Glu Ala Leu Tyr
 65  70
```
(B) LOCATION: 113..286

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:62:

```
TAAATCTAT AACTAAAAA AACACATACA GGAATTCTT CAAATAGT TCAAAACAAGA
  60
AGATTACAAA CTATCAATT TATACACAAAT ATAAACGAGC GGTACCAAAA TA ATG
Met
  115
AAA CTG AAA ACT GTA AGA TCT GCG GTC CTT TCG TCA CTC TTT GCA TCT
Lys Leu Lys Thr Val Arg Ser Ala Val Leu Ser Ser Leu Phe Ala Ser
  163
5     10    15
CAG GTC CTT GGC CAA CCA ATT GAC GAC ACT GAA TCT ATC AAC ACT ACT
Gln Val Leu Gly Gln Pro Ile Asp Asp Thr Glu Ser Ile Asn Thr Thr
  211
20    25    30
TTG GTC AAC TTG CCA GGT GCT AAG AGA TTC GTT AAC CAA CAC TTG TGC
Leu Val Asn Leu Pro Gly Ala Tyr Arg Phe Val Asn Gln His Leu Cys
  259
35    40    45
GGT TCC CAC TTG GAA GCT TTG TAC TT
Gly Ser His Leu Val Glu Ala Leu Tyr
  288
50    55
```

(2) INFORMATION FOR SEQ ID NO:63:

(1) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 58 amino acids
   (B) TYPE: amino acid
   (D) TOPOLOGY: linear

(11) MOLECULE TYPE: peptide

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:63:

```
Met Lys Leu Lys Thr Val Ser Ala Val Leu Ser Ser Leu Phe Ala
  1     5    10    15
Ser Gln Val Leu Gly Gln Pro Ile Asp Asp Thr Glu Ser Ile Asn Thr
  20    25    30
Thr Leu Val Asn Leu Pro Gly Ala Tyr Arg Phe Val Asn Gln His Leu
  35    40    45
Cys Gly Ser His Leu Val Glu Ala Leu Tyr
  50    55
```

(2) INFORMATION FOR SEQ ID NO:64:

(1) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 82 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i) MOLECULE TYPE: cDNA

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:64:
GTTAACGAA CTTGGAAGCT TCAGGCTTCAG CTTTTCTCT CGTAGCCATG GAGATCAAGT
TAACAAACATC CAAAGTAGTG TT

(2) INFORMATION FOR SEQ ID NO:65:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 54 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i) MOLECULE TYPE: cDNA

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:65:
CAAGTACAAA GCTTCAACCA AGTGGGAACC GCACAAGTG TGGTTAACGA ACTT

(2) INFORMATION FOR SEQ ID NO:66:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 117 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i) MOLECULE TYPE: cDNA

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:66:
CAACCAATTG ACGACACTGA ATCTAACACT ACTTTCGTCA ACTTGATGGC TGACGACACT
GAATCTAGAT TGCTACTAA CACTACTTTG GATGTTGTA ACTTGATCTC CATGGCT

(2) INFORMATION FOR SEQ ID NO:67:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 41 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(i) MOLECULE TYPE: peptide

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:67:
Gln Pro Ile Asp Asp Thr Glu Ser Asn Thr Thr Ser Val Asn Leu Met
1      5      10     15

Ala Asp Asp Thr Glu Ser Arg Phe Ala Thr Asn Thr Thr Leu Ala Leu
20     25     30

Asp Val Val Asn Leu Ile Ser Met Ala
35          40

(2) INFORMATION FOR SEQ ID NO:68:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 51 base pairs
   (B) TYPE: nucleic acid
   (C) STRANDEDNESS: single
   (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:68:

TCTCTTAGCC ATGGAGATCA AGTAAACAAC ATCCAAAGCC AAAGTAGGTG T

(2) INFORMATION FOR SEQ ID NO:69:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 123 base pairs
   (B) TYPE: nucleic acid
   (C) STRANDEDNESS: single
   (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:69:

CAACCAATGG ACGACACTGA ATCTAAACACT ACTTCTGTCA ACTTGAGGCC TGACGACACT
60
GAATCTAGAT TC6GTACTAA CACTACTTTG CATTTGGATG TTGTTAACTT GATCTCCATG
120
GCT
123

(2) INFORMATION FOR SEQ ID NO:70:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 65 amino acids
   (B) TYPE: amino acid
   (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:70:
Lys Arg Glu Glu Ala Glu Ala Glu Ala Glu Pro Lys Phe Val Asn Gln  
1  5  10  15  
His Leu Cys Gly Ser His Leu Val Glu Ala Leu Tyr Leu Val Cys Gly  
20  25  30  
Glu Arg Gly Phe Phe Tyr Thr Pro Lys Ala Ala Lys Gly Ile Val Glu  
35  40  45  
Gln Cys Cys Thr Ser Ile Cys Ser Leu Tyr Gln Leu Glu Asn Tyr Cys  
50  55  60  
Asn  
65  

(2) INFORMATION FOR SEQ ID NO:71:  

(1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 219 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  

(11) MOLECULE TYPE: cDNA  

(v) FRAGMENT TYPE: internal  

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:71:  
AAGAGAGAGG AAGCTGAGG TGAAGCTGAA CCAAAAGTGG TTAACCAACA CTTGTGTGGT  
TCTCACTGG TTGAAGCTTT GTACTTGGTT TCCGTTGAAA GAGGTTCCTT CTACACTCCT  
AAGGCTGCTA AGGTTATGT C8AACAATGC TGTACTCCCA TCTGCTCCTT GTACCAATTG  
GAAAAACTACT GCAACTAGAC GCAGCCCGCA GGCTCTAGA  

(2) INFORMATION FOR SEQ ID NO:72:  

(1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 348 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  

(11) MOLECULE TYPE: cDNA  

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:72:  
TTAAATCTAT AACTACAAA AACCACATAA GGAATCCCAT TCAAGAATAG TTCAAACAAG  
AAGATTACAA ACTATCAATT TCATAACCAA TATAAACGAC GGTACCAAATA TAATGAAACT  
GAAAACTGTA AGATCTCGGG TCTTTTGGTC ACTCTTTGCA TCTCAGGTCCT TGGCCAACCC
AATTGACGAC ACTGAATCTA ACACACTTCC TGTCAACTTG ATGGCTGACG ACACCTGATC 240
TAGATTCCAGT ACTAAACACTA CTTGGAATCGT ATGGCCACCC AACACCTGTG 300
TGTTTCTCAC TTGGTTGAGG CTTTGTACCT ATGGCTAAGA GATTCGTT 348

(2) INFORMATION FOR SEQ ID NO:73:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 379 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(x) SEQUENCE DESCRIPTION: SEQ ID NO:73:

TTAAATCTAT AACTACAABA AACACATACA GGAATTCCAT TCAAGAATAG TTCAACAAG 60
AAGATTCAAA ACTATCAATT TCATAACAA TATAACGAC GGTACCAAAA TAATGAAACT 120
GAAAACCTCTGTA AGATCTGCGG TCCTTTGCTG ACTCTTTGCA TCTCAGGTCC TTGGCCAACC 180
AATTGACGAC ACTGAATCTA ACACACTTCC TGTCAACTTG ATGGCTGACG ACACCTGATC 240
TAGATTCCAGT ACTAAACACTA CTTGGAATCGT ATGGCTAAGA GATTCGTT 300
AGAAGGTGAA GCTGAAGCTG AACCAAGATC CGTAAACCAAA CACTTGCTTG GTTTCACCTT 360
GGTTGAGCT TTTGACTTTG 379
CLAIMS

1. A DNA expression cassette comprising the following sequence:

\[ 5'-P-SP-LS-PS-*gene*-(T)_i-3' \]

wherein
P is a promoter sequence,
SP is a DNA sequence encoding a signal peptide,
LS is a DNA sequence encoding a leader peptide with the general formula I:

\[ 10 \text{ GlnProIle(Asp/Glu)(Asp/Glu)X}^1\text{(Glu/Asp)X}^2\text{AsnZ(Thr/Ser)}X^3 \] (I)

wherein
\( X^1 \) is a peptide bond or a codable amino acid;
\( X^2 \) is a peptide bond, a codable amino acid or a sequence of up to 4 codable amino acids which may be the same or different;
\( Z \) is a codable amino acid except Pro; and
\( X^3 \) is a sequence of from 4 to 30 codable amino acids which may be the same or different;

PS is a DNA sequence encoding a processing site;

2. An expression cassette according to claim 1, wherein \( X^1 \), in general formula I, is Ser, Thr or Ala.

25 3. An expression cassette according to claim 1, wherein \( X^2 \), in general formula I, is Ser, Thr or Ala.

4. An expression cassette according to claim 1, wherein \( X^2 \), in general formula I, is SerIle.
5. An expression cassette according to claim 1, wherein $X^2$, in general formula I, is SerAlaIle.

6. An expression cassette according to claim 1, wherein $X^2$, in general formula I, is SerPheAlaThr.

7. An expression cassette according to claim 1, wherein $X^3$, in general formula I, is an amino acid sequence of the general formula II

$$X^4-X^5-X^6$$

wherein

10 $X^4$ is a sequence of from 1 to 21 codable amino acids; $X^5$ is Pro or an amino acid sequence including the amino acid sequence ValAsnLeu, LeuAlaAsnValAlaMetAla, LeuAspValValAsnLeuProGly, or LeuAspValValAsnLeuIleSerMet; and $X^6$ is a sequence of from 1 to 8 codable amino acids.

15 8. An expression cassette according to claim 7, wherein $X^4$, in general formula II, is an amino acid sequence including one or more of the motifs LeuValAsnLeu, SerValAsnLeu, MetAlaAsp, ThrGluSer, ArgPheAlaThr and ValAlaMetAla.

9. An expression cassette according to claim 7, wherein $X^4$, 20 in general formula II, is an amino acid sequence including the sequence AsnSerThr or AsnThrThr.

10. An expression cassette according to claim 7, wherein $X^4$, in general formula II, is an amino acid sequence including the sequence

25 (Ser/Leu)ValAsnLeu,
   (Ser/Leu)ValAsnLeuMetAlaAsp,
   (Ser/Leu)ValAsnLeuMetAlaAspAsp,
   (Ser/Leu)ValAsnLeuMetAlaAspAspThrGluSer,
   (Ser/Leu)ValAsnLeuMetAlaAspAspThrGluSerIle, or

30 (Ser/Leu)ValAsnLeuMetAlaAspAspThrGluSerArgPheAlaThr.
11. An expression cassette according to claim 7, wherein \( X_5 \)
in general formula II, is an amino acid sequence including
the sequence
\[
\text{Asn(Thr/Ser)ThrLeu,}
\]
\[
\text{Asn(Thr/Ser)ThrLeuAsnLeu, or}
\]
\[
\text{Asn(Thr/Ser)ThrLeuValAsnLeu.}
\]

12. An expression cassette according to claim 7, wherein \( X_5 \)
in general formula II, is Pro.

13. An expression cassette according to claim 7, wherein \( X_5 \)
in general formula II, is the amino acid sequence ValAsnLeu.

14. An expression cassette according to claim 7, wherein \( X_5 \)
in general formula II, is the amino acid sequence
LeuAlaAsnValAlaMetAla.

15. An expression cassette according to claim 7, wherein \( X_5 \)
in general formula II, is the amino acid sequence
LeuAspValValAsnLeuProGly.

16. An expression cassette according to claim 7, wherein \( X_5 \)
in general formula II, is the amino acid sequence
LeuAspValValAsnLeuIleSerMet.

17. An expression cassette according to claim 7, wherein \( X_6 \)
in general formula II, is Ala, Gly, Leu, Thr, Val or Ser.

18. An expression cassette according to claim 7, wherein \( X_6 \)
in general formula II, is GlyAla or SerAla.

19. An expression cassette according to claim 7, wherein \( X_6 \)
in general formula II, is AlaValAla.

20. An expression cassette according to claim 7, wherein \( X_6 \)
in general formula II, is GlyAlaAspSerLysThrValGlu.
21. An expression cassette according to claim 1, wherein the leader peptide coded for by the DNA sequence LS is selected from the group comprising:

SEQ ID No. 1  GlnProIleAspGluAspAsnAspThrSerValAsnLeuProAla;

SEQ ID No. 2  GlnProIleAspGluAsnThrThrSerValAsnLeuProAla;

SEQ ID No. 3  GlnProIleAspGluSerAsnThrThrSerValAsnLeuProAla;

SEQ ID No. 4  GlnProIleAspGluAsnThrThrSerValAsnLeuProVal;

SEQ ID No. 5  GlnProIleAspGluAsnThrThrSerValAsnLeuProAla;

SEQ ID No. 6  GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuProAla;

SEQ ID No. 7  GlnProIleAspGluAsnThrThrSerValAsnLeuMetAla;

SEQ ID No. 8  GlnProIleAspGluSerAsnThrThrSerValAsnLeuProGlyAla;

SEQ ID No. 9  GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuMetAla;

SEQ ID No. 10 GlnProIleAspGluSerAsnThrThrSerValAsnVal-ProThr;

SEQ ID No. 11 GlnProIleAspAspThrGluSerAsnThrLeuValAsnVal-ProThr;

SEQ ID No. 12 GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeu-ProThr;
SEQ ID No. 13  GlnProIleAspAspThrGluSerAsnThrThrLeuValAsnVal-ProGlyAla;

SEQ ID No. 14  GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeu-MetAlaProAlaValAla;

SEQ ID No. 15  GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeu-MetAlaAspAspThrGluSerIleAsnThrThrLeuValAsnLeu-ProGlyAla;

SEQ ID No. 16  GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeu-MetAlaAspAspThrGluSerIleAsnThrThrLeuValAsnLeu-ProGlyAla;

SEQ ID No. 17  GlnProIleAspAspThrGluSerIleAsnThrThrLeuValAsnLeu-ProGlyAla;

SEQ ID No. 18  GlnProIleAspAspThrGluSerAsnThrThrLeuValAsnLeu-ProGlyAla;

SEQ ID No. 19  GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeu-MetAlaAspAspThrGluSerArgPheAlaThrAsnThrThrLeuValAsnLeu-ProLeu;

SEQ ID No. 20  GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeu-MetAlaAspAspThrGluSerIleAsnThrThrLeuValAsnLeu-AlaAsnValAlaMetAla;

SEQ ID No. 21  GlnProIleAspAspThrGluSerAlaIleAsnThrThrLeuVal-AsnLeu-ProGlyAla;

SEQ ID No. 22  GlnProIleAspAspThrGluSerPheAlaThrAsnThrThr-LeuValAsnLeu-ProGlyAla;

SEQ ID No. 23  GlnProIleAspAspThrGluSerIleAsnThrThrLeuValAsn-LeuMetAlaAspAspThrGluSerArgPheAlaThrAsnThrThr-LeuValAsnLeu-ProLeu;
SEQ ID No. 24  GlnProIleAspThrGluSerIleAsnThrThrLeuVal- 
AsnLeuMetAlaAspThrGluSerArgPheAlaThrAsnThr- 
ThrLeuAspValValAsnLeuProGlyAla;

SEQ ID No. 25  GlnProIleAspThrGluSerAlaAlaIleAsnThrThrLeu- 
ValAsnLeuProGlyAla;

SEQ ID No. 26  GlnProIleAspThrGluSerAsnThrThrSerValAsnLeu- 
MetAlaAspThrGluSerArgPheAlaThrAsnThrThrLeu- 
ValAsnLeuAlaAsnValAlaMetAla;

SEQ ID No. 27  GlnProIleAspThrGluSerAsnThrThrSerValAsnLeu- 
MetAlaAspThrGluSerArgPheAlaThrAsnThrThrLeu- 
AspVal-ValAsnLeuIleSerMetAla;

SEQ ID No. 28  GlnProIleAspThrGluSerAsnThrThrSerValAsnLeu- 
MetAlaAsnThrThrGluSerArgPheAlaThrAsnThrThrLeu- 
AspValValAsnLeuIleSerMetAla; and

SEQ ID No. 67  GlnProIleAspThrGluSerAsnThrThrSerValAsnLeu- 
MetAlaAspThrGluSerArgPheAlaThrAsnThrThrLeu- 
AlaLeuAspValValAsnLeuIleSerMetAla.

Particularly preferred leader peptides coded for by the DNA sequence LS are:

SEQ ID No. 15  GlnProIleAspThrGluSerAsnThrThrSerValAsnLeu 
MetAspLeuAlaValGlyLeuProGlyAla;

SEQ ID No. 16  GlnProIleAspThrGluSerAsnThrThrSerValAsnLeu- 
MetAlaAspThrGluSerIleAsnThrThrLeuValAsnLeu- 
ProGlyAla;

SEQ ID No. 17  GlnProIleAspThrGluSerIleAsnThrThrLeuValAsn- 
LeuProGlyAla;

SEQ ID No. 18  GlnProIleAspThrGluSerAsnThrThrLeuValAsnLeu
Pro Gly Ala
SEQ ID No. 19 Gln Pro Ile Asp Asp Thr Glu Ser Asn Thr Thr Ser Val Asn Leu Met Ala Asp Asp Thr Glu Ser Arg Phe Ala Thr Asn Thr Thr Leu Val Asn Leu Pro Leu;
5 SEQ ID No. 20 Gln Pro Ile Asp Asp Thr Glu Ser Asn Thr Thr Ser Val Asn Leu Met Ala Asp Asp Thr Glu Ser Ile Asn Thr Thr Leu Val Asn Leu Ala Asn Val Ala Met Ala;
SEQ ID No. 21 Gln Pro Ile Asp Asp Thr Glu Ser Ala Ile Asn Thr Thr Leu Val Asn Leu Pro Gly Ala;
10 SEQ ID No. 22 Gln Pro Ile Asp Asp Thr Glu Ser Phe Ala Thr Asn Thr Thr Leu Val Asn Leu Pro Gly Ala;
SEQ ID No. 23 Gln Pro Ile Asp Asp Thr Glu Ser Ile Asn Thr Thr Leu Val Asn Leu Met Ala Asp Asp Thr Glu Ser Arg Phe Ala Thr Asn Thr Thr Leu Val Asn Leu Pro Leu;
15 SEQ ID No. 24 Gln Pro Ile Asp Asp Thr Glu Ser Ile Asn Thr Thr Leu Val Asn Leu Met Ala Asp Asp Thr Glu Ser Arg Phe Ala Thr Asn Thr Thr Leu Asp Val Val Asn Leu Pro Gly Ala;
SEQ ID No. 25 Gln Pro Ile Asp Asp Thr Glu Ser Ala Ala Ile Asn Thr Thr Leu Val Asn Leu Pro Gly Ala;
20 SEQ ID No. 26 Gln Pro Ile Asp Asp Thr Glu Ser Asn Thr Thr Ser Val Asn Leu Met Ala Asp Asp Thr Glu Ser Arg Phe Ala Thr Asn Thr Thr Leu Val Asn Leu Ala Asn Val Ala Met Ala;
SEQ ID No. 28 Gln Pro Ile Asp Asp Thr Glu Ser Asn Thr Thr Ser Val Asn Leu Met Ala Asp Asp Thr Glu Ser Arg Phe Ala Thr Asn Thr Thr Leu Asp Val Val Asn Leu Ile Ser Met Ala and
25 SEQ ID No. 67 Gln Pro Ile Asp Asp Thr Glu Ser Asn Thr Thr Ser Val Asn Leu Met Ala Asp Asp Thr Glu Ser Arg Phe Ala Thr Asn Thr Thr Leu Ala Leu Asp Val Val Asn Leu Ile Ser Met Ala.

22. An expression cassette according to claim 1, wherein SP is a DNA sequence encoding the \( \alpha \)-factor signal peptide, the signal peptide of mouse salivary amylase, the carboxypeptidase signal peptide, the yeast aspartic Protease 3 signal peptide or the yeast BAR1 signal peptide.

23. An expression cassette according to claim 1, wherein PS is a DNA sequence encoding Lys Arg, Arg Lys, Arg Arg, Lys Lys or Ile Glu Gly Arg.

24. An expression cassette according to claim 1, wherein the polypeptide is selected from the group consisting of aprotinin, tissue factor pathway inhibitor, or other protease inhibitors, insulin or insulin precursors, insulin-like growth factor I, insulin-like growth factor II, human or bovine growth hormone, interleukin, glucagon, glucagon-like peptide 1, tissue plasminogen activator, transforming
growth factor α or β, platelet-derived growth factor, enzymes, or a functional analogue thereof.

25. A yeast expression vector comprising an expression cassette according to any of the preceding claims.

26. A yeast cell which is capable of expressing a polypeptide and which is transformed with a yeast expression vector according to claim 25.

27. A process for producing a polypeptide in yeast, the Process comprising culturing a yeast cell, which is capable of expressing the desired polypeptide and which is transformed with a yeast expression vector according to claim 25, in a suitable medium to obtain expression and secretion of the polypeptide, after which the polypeptide is recovered from the medium.

28. A polypeptide produced by the process of claim 27.

29. A DNA expression cassette, substantially as hereinbefore described with reference to any one of the examples.

30. A yeast expression vector, substantially as hereinbefore described with reference to any one of the examples.

31. A yeast cell which is capable of expressing a polypeptide, substantially as hereinbefore described with reference to any one of the examples.

32. A process for producing a polypeptide in yeast, substantially as hereinbefore described with reference to any one of the examples.

Dated 14 January, 1997
Novo Nordisk A/S

Patent Attorneys for the Applicant/Nominated Person
SPRUSON & FERGUSON
pAK492
11283 bp
Fig. 2

**EcoR I**

GAATTCATTCAAGAATAGT
CTTAAGTAAGTTCTATTCA

TCAAACAAGAAGATTCAAAACTATCAATTTCAATACACAATATAAACGACGGGTACCCAAA
AGTTTTGTTCTTCTAATGTTTGTAGTTAAGTATGTGTTATATTTGCTGCCATGGTTTT

TAATGAAACTGAAGACTGTAAGATCTGACGCCTCCATTGCACTCTTGGTACCTACGTCTATTCT
ATTACCTTTGACTTTTTGACATTCTAGAGCGCAAGGAAAGCAGTGAGAAACGTTAGAGTCCAGGAA

MetLysLeuLysThrValArgSerAlaValLeuSerSerLeuPheAlaSerGlnValLeu

GGCCAAACCAATAGACGAAGACAAACGACACTTCTTCCATGGCTAAGGATTTCAACAA
CCGGTTGTTATCTGCGCTTCTGCTGCTGTAAGAGGTACCAGATTCTCTAAAGCAAMTTGTT
GlyGlnProIleAspGluAspAsnAspThrSerSerMetAlaLysArgPheValAsnGln

CACTTGTGCCGTCCACTTGGTGAAGCTTTGTACTTGTGTTTGGTGTAAGGAGTTTC
GTGAACACGCAAAGGTGAACCTTGCACCAAATGAAACAAACCCACTTTTCTCCAAAG
HisLeuCysGlySerHisLeuValGluAlaLeuLeuLeuValCysGlyGluArgGlyPhe

TTCTACACTCCTAAGGCTGTAAGGATTTGTCAGCAATGCTGTACCTCATCTGCTCC
AAGATGTGAGATCCAGACTCCATAACAGCTGTACAGATGAGACGAGG

PheTyrThrProLysAlaAlaAlaLysIleValGluGlnCysCysThrSerIleCysSer

**Xba I**

TTGTAACCAATTGAAAAACTACTGGAACGCGACGCCGCCAGCTCCTAGA **SEQ ID No.40**
AACCAGTTAAACCCTTTGATGACGTTGATCTGCCTGGCGGCTCCAGATCT

LeuTyrGlnLeuGluAsnTyrCysAsn* **SEQ ID No.41**
Fig. 3

Asp 718-Hind III leader PCR DNA fragment

Hind III-Xba I insulin DNA fragment
Fig. 4.
CAACCAATTGACGACGAAACACTACTTCTGTCAACTTGCCAGTT
GTTGGTTAACCTGCTGCTTTTTGGATGAAGAAGCAGTTGAACGGTCAA
GlnProIleAspAspGluAsnThrThrSerValAsnLeuProVal

SEQ ID No. 42

Fig. 5.
ATGAACATGAAACTGTAAGATCTGCGTGCTTTTCTGACTACTTTTGCA
TACTTCTGACCTTTTGCATTTCTAGACGACGAAACTACGTACGAGAAGCGT
MetLysLeuLysThrValArgSerAlaValLeuSerSerLeuPheAla
TCTCAGGTCTTCGAGCAACAATTTGACGACGAAACACTACTTCTGTC
AGAGACCAAGCCGGTTGTTAACTGCTGTCTTTTTGATGAAGACAG
SerGlnValLeuGlyGlnProIleAspAspGluAsnThrThrSerVal

SEQ ID No. 43

AATCTGCCAGTTAAAGAGATCTGTTAAACCAACACTTTGTGGTTTCAC
TTGAACGTTCAATTCTCTAAAGCAATTTGTTGTAACACACCAAGAGTG
AsnLeuProValLysArgPheValAsnGlnHisLeuCysGlySerHis

SEQ ID No. 44

TTGATTGAAGCGTTGTTACTTGGTTGCGGTGAAAGAGTTTCTCTTAC
AACCACCTCCGAAACTGAAACCACAGCCACTTTCTCCAAAAGAGATG
LeuValGluAlaLeuTyrLeuValCysGlyGluArgGlyPhePheTyr

SEQ ID No. 45

ACTCCTAAAGGCTGCTAAGGATTGTTGTCGAAATATGCTGTATACCTCCATC
TGAGGATTCCGACGATTTCCATAACAGCTTTGTACGATGGAGTAG
ThrProLysAlaAlaLysGlyIleValGluGlnCysThrSerIle

SEQ ID No. 46

TGCTCATTGTTAACCTGGAAAACCTACTGCAACTAGACGACGCGCCGCA
ACGAGAAGACATGGTTAACCTTTTGATGAGTTGATCTGCCTCGGGCGT
CysSerLeuTyrGlnLeuGluAsnTyrCysAsn*

SEQ ID No. 47

GGCTCTAGA
CCGAGATCT
Fig. 6.
CAACCATGACGACACTGAATCTAACAACACTTCTGTCACCTTGCCAGCT SEQ ID No. 46
GTTGGTTAACTGCTGTGACTTAGATTGTGATGAAAGACAGTTGACCGTCA
GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuProAla SEQ ID No. 47

Fig. 7.
CAACCATGACGACACTGAATCTAACAACACTTCTGTCACCTTGCCA
GTTGGTTAACTGCTGTGACTTAGATTGTGATGAAAGACAGTTGACCGTCA
GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuPro

GGTGCT                          SEQ ID No. 48
CCACGA                          SEQ ID No. 5
GlyAla                          SEQ ID No. 8

Fig. 8.
CAACCATGACGACACTGAATCTATCAAACACTACTTTTGTCAACCTTG
GTTGGTTAACTGCTGTGACTTAGATTGTGATGAAACCAGTTGACCGTCA
GlnProIleAspAspThrGluSerIleAsnThrThrLeuValAsnLeu

CCAGGTGCT                  SEQ ID No. 49
GGTCCCACGA                SEQ ID No. 17
ProGlyAla                  SEQ ID No. 17

Fig. 9.
CAACCATGACGACACTGAATCTAACAACACTACTTCTGTCACCTTGATG
GTTGGTTAACTGCTGTGACTTAGATTGTGATGAAAGACAGTTGACCGTCA
GlnProIleAspGluSerAsnThrThrSerValAsnLeuMet

GCTGACGACACTGAATCTAACAACACTACTTTTGTGTAACCTTGCCAGGTGCT SEQ ID No. 50
CGACTGCTGTGACTTAGATTGTGATGAAACCAGTTGACCGTCCACGA
AlaAspAspThrGluSerIleAsnThrThrLeuValAsnLeuProGlyAla SEQ ID No. 18
**Fig. 10.**

```
CAACCAATTGACGACACTGAATCTAACACTACTTCTGTCAACTTTGGCTGACGACACT
GTTGTTAACTGCTGTGACCTTAGATGGATGATGAAAGACCTACCGACTGCTGTGA
GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuMetAlaAspAspThr
```

**SEQ ID No. 51**

```
GAATCTAGATTCCCTACTAACACTACTTTGGTTAACTTGGCCATGTG
CTTAGATCTAACGCCATGTGATGAAACCAATTGAAACGGTAAAC
GluSerArgPheAlaThrAsnThrThrLeuValAsnLeuProLeu
```

**SEQ ID No. 19**

**Fig. 11**

```
CAACCAATTGACGACACTGAATCTAACACTACTTCTGTCAACTTTGGCTGACGACACT
GTTGTTAACTGCTGTGACCTTAGATGGATGATGAAAGACCTACCGACTGCTGTGA
GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuMetAlaAspAspThr
```

**SEQ ID No. 52**

```
GAATCTATCAACACTACTTTGGTTAACTTGGCTAAGCTTTGCCATGTGCT
CTTAGATAGTTCTGATGAAACCAATTGAAACCCGATTCAACCG
GluSerIleAsnThrThrLeuValAsnLeuAlaAsnValAlaMetAla
```

**SEQ ID No. 20**

**Fig. 12.**

```
CAACCAATTGACGACACTGAATCTGCTATCAAACACTACTTTGGGTCAC
GTTGTTAACTGCTGTGACCTTAGATGGATGATGAAACCAATTG
GlnProIleAspAspThrGluSerAlaIleAsnThrThrLeuValAsn
```

**SEQ ID No. 53**

```
TTGCCAGGGTCT
AACGGTCCACGA
LeuProGlyAla
```

**SEQ ID No. 21**

**Fig. 13**

```
pAKS27:
TTAAAATCTATAACTAAAAACACATACAGGAATTTCATCAAGATATAGTTCAAACAA
AATTTAGATATTTGTGTTTTGTGTATGCTCTAAGTAA GTTCTTATCAAGTTGTT
```

**SEQ ID No. 22**
GAAGATTACAACATCTATCATAATTCATACATAATATAAAGCACCGGTACCAATAATAATGAAA
CTTCTAATGTTTGTAGTTAAGATATGTGTTATATTGCTGGCCATGGTTTTATATTCTTT
MetLys

CTGAAAACTGTAAGATCTGCGGTCTTTTCTGTCACCTCTTTGCATCTCAGGTCTTTGGCCAAGACCTTGACATTCTAGCACCCAGGAAAGCAGTAGAGTAAGAGTCCAGGAAACCGGTT
LeuLysThrValArgSerAlaValLeuSerSerLeuPheAlaSerGlnValLeuGlyGln

CCAATGGACGAGAAAAACACTACTTCTCCTGTTAACTTGCAGCTAAGAGATTTCTTTGAAACCAGGGTACGCTTTTGATGAGACATAATGAAACTCTGATGCACTACATGCAACATGAA
GTTAAGCTGCTGTTTTGTGTGATGAGACATAATGAAACTCTGATGCACTACATGCAACATGAA
ProIleAspAspGluAsnThrThrSerValAsnLeuProAlaLysArgPheValAsnGln

CACTTGTGCGGTCCACTTGGTGGAAGCTTTGTACTT
SEQ ID No. 54

GTGAAACACGCGCCAAGGTGAACCAACTTCCGAAACATGAA
HisLeuCysGlySerHisLeuValGluAlaLeuTyr
SEQ ID No. 55

**Fig. 14**

pAK531:

TTTAATTACTATAACTAACAAAACACATACAGGAATTCTATCTCAAGAATAGTTCAAAACAA
AATTTAGATATTGATTTTTTTGTATGCTCTTAAGTAAGTTCTTTATCAAGTGTGTT

GAAGATTACAACATCTATCATAATTCATACATAATATAAAGCACCGGTACCAATAATAATGAAA
CTTCTAATGTTTGTAGTTAAGATATGTGTTATATTGCTGGCCATGGTTTTATATTCTTT
MetLys

CTGAAAACTGTAAGATCTGCGGTCTTTTCTGTCACCTCTTTGCATCTCAGGTCTTTGGCCAAGACCTTGACATTCTAGCACCCAGGAAAGCAGTAGAGTAAGAGTCCAGGAAACCGGTT
LeuLysThrValArgSerAlaValLeuSerSerLeuPheAlaSerGlnValLeuGlyGln

CCAATGGACGACACTGAATCTAACAACACTACTTCTCCTGTAACCTTTGCCAGCTAAGAGATTTCGTT
GGTTAAGCTGCTGCTGTTTTGTGTGATGAGACATAATGAAACTCTGATGCACTACATGCAACATGAA
ProIleAspAspThrGluSerAsnThrThrSerValAsnLeuProAlaLysArgPheVal
Fig. 15

pAK555

TTAAATCTATAACTACAACAAAAACACATACAGGAATTCTACCATGCAAGATAGTTCAAAAACAA
AATTTAGATATTAGTGGTTTTGTATGTTCTCTTAAAGTTAGTCTTATCAAGTTTGGTT

GAAGATTACAACACTATATTTCTATACACATAAAGCAAGACGGTGACCACAAAATAATGAAA
CTTCTAAAGTTGGAGTTAAGTGTGTTATATTTGTGCTCCATGGTTTTATTACTTT

MetLys

CTGAAAACCTGTAAGATCTGCCGTTCTTTCTGCTACTCTTTCATCTCAGGTCTCTGGCCAA
GACTTTGACATCTCAGGCGGAGAAACGTGAGATCCAGAAGACGCACCGGTTC
LeuLysThrValArgSerAlaValLeuSerSerLeuPheAlaSerGlnValLeuGlyGln

CCATTGACGACACTGAATCTAAACACTACTTCGATCGGCTGATGCGATAGATTCTGGTT
GGTTAACGTGTTGACTTAGATTGTGATCAGACAGACACTACGTTACTGATCCTCTAAGCA
ProIleAspAspThrGluSerAsnThrThrSerValAsnLeuMetAlaLysArgPheVal

Fig. 16

pAK559:

TTAAATCTATAACTACAACAAAAACACATACAGGAATTCTACCATGCAAGATAGTTCAAAAACAA
AATTTAGATATTAGTGGTTTTGTATGTTCTCTTAAAGTTAGTCTTATCAAGTTTGGTT

GAAGATTACAACACTATATTTCTATACACATAAAGCAAGACGGTGACCACAAAATAATGAAA
CTTCTAAAGTTGGAGTTAAGTGTGTTATATTTGTGCTCCATGGTTTTATTACTTT

MetLys
CTGAAAACTGTAAGATCTGGGTCCCTTTCTGCTACTCTTTGCTACCTCAGGTTCTCTTGCCAA
GACCTTTCTAGATTTTAGGACTGGAGGCTAGCTGAGAGAGTGGCAGAGTGCTGACTGAACCTAT
CCAATGACGACACTGAATCTAACACTACTTTCTGTCAACTTTGATGGCTGACGACACTGAA
GGTTACTGCTGTGACTTAGATTGATGAAGACAGGTAACGTGCCTGGAGCTCTTT
ProIleAspAspThrGluSerAsnThrSerValAsnLeuMetAlaAspAspThrGlu
TCTATCAACACACTACTTGGTTAACTCTGGCAGTTGCTATAGAGATTCTGTTAACAAACACTTG
AGATAGTTGATGAACCAATTGAAGGTCACGGTCCAGATTCTCTAAAGGATTGTTGTAAC
SerIleAsnThrThrLeuValAsnLeuProGlyAlaLysArgPheValAsnGlnHisLeu

TGCCTCCACCTTGGTGGAAGCTTTTGTACTT
ACGCCAAGGTTGAAACCAACCTTCGAAACATGAA
CysGlySerHisLeuValGluAlaLeuTyr

Fig. 17
pAK562:
TTAAACTATAACTACAAAAACACATAACAGGAATTCATCAAGAATAGTTCAAACAA
AATTTTAGATATGATGTTTTTGTGTATGCCTTAAGTAAGTTTCTTTATCAAGTTTGGT
GAAGATGACAAACTATCAATTTCTACACAATATAAAACGACGGGTACAAAAATAATGAAA
CTTCTAAATGTTGGATGTTAAGATGATGTTATTTTTGCTGCCCATGTTTTATTACTTT
MetLys

CTGAAAACTGTAAGATCTGGGTCCCTTTCTGCTACTCTTTGCTACCTCAGGTTCTCTTGCCAA
GACCTTTCTAGATTTTAGGACTGGAGGCTAGCTGAGAGAGTGGCAGAGTGCTGACTGAACCTAT
CCAATGACGACACTGAATCTAACACTACTTTCTGTCAACTTTGATGGCTGACGACACTGAA
GGTTACTGCTGTGACTTAGATTGATGAAGACAGGTAACGTGCCTGGAGCTCTTT
ProIleAspAspThrGluSerIleAsnThrThrLeuValAsnLeuProGlyAlaLysArg
**Fig. 18**

CAACCAATTTGACGACACTGAATCTAAACACTACTTCTTGTAACCTTGAT
GTTGGTTAATGTGACTTGACTTAGATTGTGAGAGACAGTTGAAACTAC
GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuMet

GCTGACGACACTGAATCTAGATTTGCTACTAAACACTACTTTGGATGT
CGACTGCTGTGACTTAGATCTAAAGGATGTTGATGAAACCTACAA
AlaAspAspThrGluSerArgPheAlaThrAsnThrThrLeuAspVal

**Fig. 19**

AAGAGAGAAAGACGTGAAGCTGAACTGAAACCAAGTTCCGTTAACC
TTCTCTTCTTCTTGACTTGCTTGTTTCTAGGATTGAGTTGGTG
LysArgGluAlaGluAlaGluAlaGluProLysPheValAsnGln

CAGCTGCTGCTTCTACATTTGCTGATTTTTGACTTTGCTTTGG
GTGAAACACCAAGATGTGAGATTCCACAGGATGCCCATCACAACGGA
HisLeuCysGlySerHisLeuValGluAlaLeuTyrLeuValCysGly

GAAAGAGGTTTCTTCTACACTCTGTAAGGCTGATTGATG
CTTTTCTCAGAAGAGATGTGAGATTCCACAGGATGCCCATCACAACGGA
GluArgGlyPhePheTyrThrProLysAlaAlaLysGlyIleValGlu

CAATGCTGTAATCTCCATCTGCTCTTGTTACAAATTGAAAAACTACTG
GTACGACATGAGGATGACGAGAAACATGTTACCTTTGGATGACG
GlnCysCysThrSerIleCysSerLeuTyrGlnLeuGluAsnTyrCys
**Xba I**

AACTAGACGCACCCGCAGGCTCTAGA
TTGATCTGGCGGGGCTCGAGATCT
Asn***

**Fig. 20**

CAACCAATTGACGTGAAATCTAACAACACTACTTTCTGCTCAACTGGATG
GTTGTTAATGCTGTGACTTAGATTGTGATGAAGACAGTGGACTAC
GlnProIleAspAspThrGluSerAsnThrThrSerValAsnLeuMet

GCTGAGCAGACACTGAATCTAGATTGCTACTAACAACACTACTTTGCGCTTG
CGACTGCTGTGACTTAGATCTAAAGGGATGATTGTGATGAAGACAGAACC
AlaAspAspThrGluSerArgPheAlaThrAsnThrThrLeuAlaLeu

GATGTTGTTAACTTGATCTCCATGGCT
CTACAACAAATGGAACTAGAGGTACCGA
AspValValAsnLeuIleSerMetAla

**SEQ ID No. 71**

**SEQ ID No. 70**

**Fig. 21**

pAK514

TTAAATCTATAACTACAAAAAACACATACAGGAATCCATCCATCAAGA
AATTAGATATTGAGTGTGTGTGTGTATGTCCTAAGGTAAGTCT

ATATGTCAATGACGATCAGGAAACTGTCAGATTGTGATGAAGATGAGTATGTTAT
ATACAGCTTTCTCTTCTATGTGGATGTTAAAGGTATGTGTTAT

AACGACGGTACCAAAATAATGAAAACTGATAAGATCTGGCTGC
TTGGATCGGCTGTGTCTATTACCCTTGCATATCTGGACAGCGGAG
MetLysLeuLysThrValArgSerAlaVal

CTTTCGCTACTCTTTGCAATCTAGGTCCTTGCCACACACATGGACGAC
GAAAGCGTGAAGACGTAGGAGTCCAGGAACCGTGGTTAACTGCTG
LeuSerSerLeuPheAsaSerGlnValLeuGlyGlnProIleAspAsp

**SEQ ID No. 69**

**SEQ ID No. 67**
ACTGAACTCTACACTACTTCTCTGCACTTGGTGGCTGACGACACTGAA
TGACTTGAAGATGGTGAAGACACGGTACCTCAGCTGGACTTT
ThrGluSerAsnThrThrSerValAsnLeuMetAlaAspAspThrGlu
TCTAGATTCGCTACTACACTACTTTGGTTAAGCTTGCTACGTTGCC
AGATCTAACGATGGATTGTGATGAAACCAGTGAAACCGATTGCAACGG
SerArgPheAlaThrAsnThrThrLeuValAsnLeuAlaAsnValAla
AACCAACACTTTGGTTGGTTCTCACCTTTGGTTGGACACTTTTTGTTACTTATGG
TACCCATCTCTCAAGCAATGGTTGATTGAACACACCAAGATGTAACCAA
MetAlaLysArgPheValAsnGlnHisLeuCysGlySerHisLeuVal
CTAAGAGATTCGTT
CTTCGAAACATGAA
GluAlaLeuTyr

SEQ ID No. 72

Fig. 22
pAX525
TTAATCTTAAACTACAACACACCATACAGGAATTCGATTCAAG
AATTTAGATATGGATTTTTTTTGTTATGCTCTTAAGGTAGTTCT
ATAGTTCAACTACGAAGAAGATTACACAACTATCAATTTGATACACAATATA
TATCAAGTGTGTTCTCTTAATGTTGATAGTTAAAGTATGTTATAT
AAGCGACGTACCAAAATAATGAAACTGGAATGTCTGAGATCTCGCGTC
TTGCTGCCATGTTTTATTACTTTTGAGCTTTTCAGATACACGAGCAG
MetLysLeuLysThrValArgSerAlaVal
CTTTGCACTCTTTGGCATCTCACGGTGCTCTGCTGGGCAACCAAATGACGAC
GAAAGCAGTGAAGAAGGATAAGTTTGGTCCAGGAAACCGTTTGTTAATCTGCTG
LeuSerSerLeuPheAlaSerGlnValLeuGlyGlnProIleAspAsp
ACTGAATCTAAACACTACTTTCTGTCAAACTTGTGGATGCTGACGACTGAA
TGACTTATTTTGTGATAGACAGAGTATCATCACGACTTGCTTGACTT
ThrGluSerAsnThrThrSerValAsnLeuMetAlaAspAspThrGlu

TCTAGATTCTGCTACTAAACACTACTTTTGATGTTGTATCTCC
AGATCTAAGCGATGATTGTGATGAAACCTACAAACTGAACTAGAGG
SerArgPheAlaThrAsnThrThrLeuAspValValAsnLeuIleSer

ATGGCTAAAGAGAGAAAGCTGAAGCTGAAGCTGAACCAAGTTTGGTT
TACCGATTCTCTCTCTCTCGACCTCGACTTGTTTCAAGCAA
MetAlaLysArgGluAlaGluAlaGluAlaGluProLysPheVal

AACCAACACTTTGTGTTCTCTACTTTGTTGAAGCTTTGACTTG
TTGTTGATGAACAACACAGAGTGAAACCAACTTGAACAACTGAC
AsnGlnHisLeuCysGlySerHisLeuValGluAlaLeuTyr

SEQ ID No. 73
**INTERNATIONAL SEARCH REPORT**

**A. CLASSIFICATION OF SUBJECT MATTER**

<table>
<thead>
<tr>
<th>IPC:</th>
<th>C12N 15/81, C12N 15/79 // C12N 15/17</th>
</tr>
</thead>
<tbody>
<tr>
<td>According to International Patent Classification (IPC) or to both national classification and IPC</td>
<td></td>
</tr>
</tbody>
</table>

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)

<table>
<thead>
<tr>
<th>IPC:</th>
<th>C12N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched</td>
<td></td>
</tr>
<tr>
<td>SE, DK, FI, NO classes as above</td>
<td></td>
</tr>
</tbody>
</table>

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

**STN, EDOC, WPI, CLAIMS, STRAND**

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

<table>
<thead>
<tr>
<th>Category</th>
<th>Citation of document, with indication, where appropriate, of the relevant passages</th>
<th>Relevant to claim No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>US 5037743 A (SUSAN K. WELCH ET AL), 6 August 1991 (06.08.91), the whole document, especially column 7, line 13-25</td>
<td>1-27</td>
</tr>
</tbody>
</table>

* Special categories of cited documents:
  * "A" document defining the general state of the art which is not considered to be of particular relevance
  * "B" earlier document but published on or after the international filing date
  * "L" later document published on or after the international filing date or priority date which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
  * "O" document referring to an oral disclosure, use, exhibition or other means
  * "P" document published prior to the international filing date but later than the priority date claimed

* Later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

* "X" document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve inventive step when the document is taken alone

* "Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

* "Z" document member of the same patent family

**Date of the actual completion of the international search**

25 Sept. 1995

**Name and mailing address of the IHA/ Swedish Patent Office**

Box 1089, S-108 88 STOCKHOLM

Fax number: 46 8 668 02 88

**Date of mailing of the international search report**

02-10-1995

**Authorized officer**

Patrick Andersson

Telephone No. 46 8 789 35 00
<table>
<thead>
<tr>
<th>Patent document cited in search report</th>
<th>Publication date</th>
<th>Patent family member(s)</th>
<th>Publication date</th>
</tr>
</thead>
<tbody>
<tr>
<td>US-A- 5037743</td>
<td>06/08/91</td>
<td>AU-B- 660172</td>
<td>15/06/95</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AU-A- 1498792</td>
<td>21/01/93</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AU-A- 2334088</td>
<td>18/05/89</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CA-A- 1314830</td>
<td>23/03/93</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DE-D.T- 3887425</td>
<td>23/06/94</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EP-A,A,A 0310137</td>
<td>05/04/89</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SE-T3 0310137</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ES-T- 2061585</td>
<td>16/12/94</td>
</tr>
<tr>
<td></td>
<td></td>
<td>JP-A- 2016984</td>
<td>19/01/90</td>
</tr>
<tr>
<td>WO-A1- 9211378</td>
<td>09/07/92</td>
<td>AU-B- 660161</td>
<td>15/06/95</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AU-A- 9134891</td>
<td>22/07/92</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CZ-A- 9301192</td>
<td>16/02/94</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EP-A- 0563175</td>
<td>06/10/93</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FI-D- 932831</td>
<td>00/00/00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HU-A- 68751</td>
<td>28/07/95</td>
</tr>
<tr>
<td></td>
<td></td>
<td>JP-T- 6503957</td>
<td>12/05/94</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AU-A- 2765088</td>
<td>10/03/94</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DE-D.T- 3885728</td>
<td>16/11/94</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ES-T- 2059547</td>
<td>16/12/94</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HK-A- 139994</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>IE-B- 62087</td>
<td>14/12/94</td>
</tr>
<tr>
<td></td>
<td></td>
<td>JP-A- 2002339</td>
<td>08/01/90</td>
</tr>
</tbody>
</table>