(54) A yeast cell of the genus schwanniomyces
Hefezellen der Gattung-Schwanniomyces
Cellules de levure du genre schwanniomyces

(84) Designated Contracting States:
AT BE CH DE ES FR GB GR IT LI LU NL SE

(43) Date of publication of application:
31.10.1990 Bulletin 1990/44

(73) Proprietor: RHEIN BIOTECH GESELLSCHAFT
FUR NEUE BIOTECHNOLOGISCHE PROZESSE
UND PRODUKTE MBH
40595 Düsseldorf (DE)

(72) Inventors:
• Hollenberg, Cornelius P., Prof. Dr.
  D-4000 Düsseldorf (DE)
• Strasser, Alexander, Dr.
  D-4000 Düsseldorf (DE)

(74) Representative: Grüneker, Kinkeldey,
Stockmair & Schwahnhäuser Anwaltssozietät
Maximilianstrasse 58
80538 München (DE)

(56) References cited:
EP-A- 0 256 421
EP-A- 0 257 115
EP-A- 0 260 404

• SCIENCE, vol. 229, 20th September 1985, pages
  1219-1224, Washington, DC, US; R. A. SMITH et
  al.: "Heterologous protein secretion from yeast"
• CRC - CRITICAL REVIEWS IN
  BIOTECHNOLOGY, vol. 5, no. 2, 1987, pages
  159-176, Boca Raton, Florida, US; W.M.
  INGLEDEW.
Description

The present invention relates to yeast cells of the genus Schwanniomyces and a process for the production of polypeptides.

For expression of eukaryotic proteins of biotechnological relevance there exists a variety of host systems, which are selected in compliance with the specific requirements. The most frequently used host organism is Escherichia coli, being the best known representative of prokaryotic host organisms. Saccharomyces cerevisiae and animal tissue cultures are used as eukaryotic host cells. For certain purposes, e.g. where expression of a glycoprotein is desired, expression in E. coli or other prokaryotic organisms is not appropriate, since they lack any glycosylation capability.

On the other hand, the mass production of protein in tissue cultures still is cumbersome and expensive; therefore, the organism of choice in many instances is yeast of the well characterized genus Saccharomyces. S. cerevisiae is able to secrete efficiently small proteins, e.g. polypeptides like MG1 (Kurjan and Herskowitz, 1982), β-endorphin or α-interferon (Bitter et al., 1984), however proteins of higher molecular weight are captured in the periplasmic space (Emr et al., 1981) or the cytoplasm, thus causing decreased growth rates or even death of the cell.

Therefore it is an object of the present invention to provide an improved eukaryotic expression system, which is optionally capable of secreting proteins encoded by a foreign DNA into the culture medium.

This object has been solved by a yeast cell of the genus Schwanniomyces, wherein said yeast cell contains at least one expression cassette comprising:

a) a first DNA sequence serving as a regulon,
b) optionally a second DNA sequence coding for a signal peptide 
c) a third DNA sequence coding for a foreign protein, and 
d) optionally a fourth DNA sequence serving as a terminator, wherein the first, second and fourth DNA sequence are derived from yeast.

The invention, which comprises further subjects, is now described in a more detailed manner by the following description, examples and figures.

BRIEF DESCRIPTION OF THE FIGURES

Fig. 1

A: Restriction map of the AMY1 gene encoding α-amylase and its 5’ and 3’ noncoding region

B: Nucleotide sequence of the AMY1 gene and its flanking region. In the 5’ region the transcription initiation site is indicated by a dashed line above the sequence. Within the structural gene the presumptive signal peptidase cleavage sites are indicated by star-symbols. The two potential glycosylation sites, the cystein residues presumably forming four disulfide bonds (concluded by homology to the A. oryzae enzyme) are underlined. Within the 3’ end region the transcription termination signal TATG...TATGT...TTT is indicated by a solid line.

Fig. 2

A: Restriction map of the GAM1 gene encoding glucoamylase and its 5’ and 3’ noncoding region

B: Nucleotide sequence of the GAM1 gene and its flanking region. In the 5’ region the transcription initiation site is indicated by a dashed line. Within the structural gene the presumptive signal peptides cleavage sites are indicated by star-symbols.

Fig. 3

M13 XhoI. The EcoRI-ScaI fragment, containing a 1.8 kb promoter fragment and approx. 0.4 kb structural gene fragment of the α-amylase gene, was subcloned into M13mp8 previously digested with EcoRI and Smal. By site directed mutagenesis a XhoI-site was inserted just in front of the ATG codon:
Positions where bases have been exchanged by site directed mutagenesis, are indicated by asterisks.

Fig. 4
pMaGAM-B/S. The BglII fragment containing the glucoamylase gene and its 5' and 3' flanking region was subcloned into the BamHI site of vector pMACG5-S (Kramer et al., 1984). By site directed mutagenesis using the gapped duplex DNA method (Kramer et al., 1984) a BamHI site and a SalI site were inserted just in front of the translation initiation codon ATG.

Positions affected by the site specific mutagenesis are indicated by asterisks.

Fig. 5
YRp7GAM. The GAM gene was isolated from pMaGAM-B/S as a SalI fragment and ligated into the XhoI site of M13xXho (Fig. 3). The BglII fragment harbouring the α-amylase promoter GAM gene fusion was ligated into the single BamHI site of the S. cerevisiae vector YRp7 resulting in plasmid YRp7GAM.

Fig. 6
Construction of S. cerevisiae vectors for expression of foreign genes using the promoter and the signal sequence of the glucoamylase gene.

a. Insertion of the 4.0 kb EcoRI-Pvull fragment from pBRSwARS GAM, containing the complete GAM promoter and the first 208 bp of the GAM encoding sequence into pRR322, cleaved with EcoRI/Pvull, resulting in vector pBRGAM.

b. Insertion of the 3.4 kb Pvull fragment from plasmid pCT603, containing the structural gene coding for cellulase from base pair position +112, into the Pvull site of plasmid pBRGAM, resulting in vector pBRGC1.

c. Insertion of the 5.5 kb BglII-PstI fragment from plasmid pBRGC1 into BamHI/PstI cut pCJD5-1. The resulting plasmid, pMPGC1-1 contains 321 bp of the GAM 5' non coding region and the first 208 bp of the GAM coding region fused to position +112 of the celD gene.

d. Insertion of the 6.5 kb BamHI-PstI fragment from plasmid pBRGC1 into BamHI/PstI cut pCJD5-1. The resulting plasmid, pMPGC1-2 contains 1.3 kb of the GAM 5' non coding region and the first 208 bp of the GAM coding region fused to position +112 of the celD.

Fig. 7
Promoter studies using the qualitative endoglucanase D activity assay (cono red staining) with different transformants: NGA58: pMPAGC, NGA58: pMPPGC, NGA58: pMPGC1-1, NGA58: pMPGC1-2

a. Growth on 2% maltose
b. Growth on 4% glucose

Fig. 8
Construction of a S. cerevisiae vector for expression of foreign genes under control of the ADH1 promoter.

a. Insertion of the 1.45 kb BamHI-Sall ADH1 promoter fragment of plasmid pFM2-1 into pCJD5-1, previously cut with BamHI/Sall resulting in plasmid pMPADH1.

b. Insertion of the 3.2 kb Sall fragment of pJDcg-15 into the Sall site of pMPADH1, resulting in pMPAG.
c: Exchange of the GAM promoter in plasmid pMPGC1-2 (Fig 6d) for the ADH1 promoter from plasmid pMPAG by \textit{in vivo} recombination, resulting in plasmid pMPAGC.

Fig. 9

Construction of a \textit{S. occidentalis} vector for the expression of foreign genes under control of the PDC promoter.

a: Insertion of the 6.0 kb Sphl-Aval fragment from plasmid pJDcg-15, containing the PDC promoter fused to the GAM gene into YRpJD2 previously cleaved with Sphl-Aval, resulting in plasmid pMPPG.

b: Exchange of the GAM promoter in plasmid pMPGC1-2 (Fig. 6d) for the PDC promoter from plasmid pMPPG by \textit{in vivo} recombination, resulting in plasmid pMPPGC.

Fig. 10

Construction of an autonomous replicating vector (containing SwARS2) for transformation of \textit{S. occidentalis}. The 3.2 kb TRP5 BamHI fragment from plasmid YRpJD2 was ligated with the 5.5 kb BamHI-BglII fragment of plasmid pBRSwARSAGAM (Fig 6a), resulting in pMPTS2-A and pMPTS2-B.

ABBREVIATIONS USED THROUGHOUT THE APPLICATION

Restriction endonucleases :

\begin{align*}
\text{Aa} & : \text{AatII} \\
\text{A} & : \text{Asp718} \\
\text{B} & : \text{BamHI} \\
\text{Bg} & : \text{BglII} \\
\text{C} & : \text{ClaI} \\
\text{E} & : \text{EcoRI} \\
\text{EV} & : \text{EcoRV} \\
\text{H} & : \text{HindIII} \\
\text{P} & : \text{PvuII} \\
\text{Ps} & : \text{PstI} \\
\text{S} & : \text{SalI} \\
\text{Sc} & : \text{ScaI} \\
\text{Sm} & : \text{SmaI} \\
\text{Sp} & : \text{SphI} \\
\text{X} & : \text{XhoI}
\end{align*}

Throughout this application various publications are referenced by mention of the first author and the year of publication within parentheses. Full citation of the references may be found at the end of the Specification as an annex, listed according to their alphabetical order immediately following the specification. The disclosures of these publications in their entireties are hereby incorporated by reference into this application in order to more fully describe the state of the art as known to those skilled therein as of the date of the invention described and claimed herein.

The yeast cell according to the present invention is characterized by containing an expression cassette which in turn is composed of three to four different components, being arranged in operable linkage in order to ensure efficient expression of a desired polypeptide, optionally followed by secretion of same. These components are defined to be the following:

a) A first DNA sequence serving as a regulon. The term "a regulon" as used throughout the application comprises any cis-acting DNA sequence involved in the regulation of expression of a given structural gene. The term thus embraces sequences preceding a given coding sequence, for example, a promoter and sequences which are recognised by transcription factors or other proteins involved in transcription regulation. The "first DNA sequence" is thus not limited to sequences corresponding to known promoters. Also included are DNA sequences corresponding to natural occurring regulons, which have undergone insertion, deletion or substitution events, but still retain their activity as regulon. Further the term "regulon" includes DNA sequences involved in regulation of transcription, which have been synthesised chemically or obtained by methods of gene technology.

b) An optional second DNA sequence coding for a signal peptide. The meaning of this second DNA sequence
includes various DNA sequences, directing the secretion of any polypeptide containing the thus encoded peptide at its N-terminal end.

c) A third DNA sequence coding for a foreign protein. This coding sequence coding for the signal peptide in order to retain a proper reading frame for the third DNA sequence. The fusion thus requires application of methods or recombination DNA technology known to every person skilled in the art. For methods enabling to design proper fusions it is in general referred to Maniatis et al. (1982), wherein appropriate methods, e.g. various ways to treat restriction fragment ends, cloning of linker molecules and similar procedures are described in great detail. The foreign protein coded for by the third DNA sequences may be any protein which to express there exists some need. Since the secretion system is capable of secretion not only of small molecules, but also of large proteins, expression of any desired protein should be possible. For examples it is referred to the discussion below.

d) Optionally a fourth DNA sequence serving as a terminator. As the terminator any sequence may be used, which efficiently terminates transcription in the respective host organism. Terminators may be for example sequences prone to formation of hairpin structures and/or polyadenylation sites. In a preferred embodiment the terminator is derived from the same gene as the second DNA sequence, i.e. the reguler, is derived from.

The four components are combined with each other using known techniques. Optionally the proper fusion will be verified for example by sequencing the boundaries.

The yeast cell according to the present invention contains at least one expression cassette comprising some or all of the above mentioned DNA sequences. It provides a convenient tool for the industrial expression of any desired protein.

In a preferred embodiment said first and/or second and/or fourth DNA sequence are derived from a gene coding for an amylolytic enzyme, for example α-amylase or glucoamylase, which enzymes in most organisms are induced upon contact with starch. The DNA sequences may be derived from a yeast of a member of one of the genera Debaryomyces, Saccharomyces, Lipomyces, Pichia, or Saccharomyces, but yeast of the genus Schwanniomyces are preferred.

Yeast species of the genus Schwanniomyces have an efficient expression and secretion apparatus, enabling them to grow on starch as a sole carbon source. Starch is hydrolysed to glucose by two amylolytic enzymes, secreted into the supernatant, namely α-amylase (α- 1.4-glucan 4-glucanohydrolase, E.C.3.2.1.1) and glucoamylase (syn. amyloglucosidase) (α-1.4-glucan-glucohydrolase, E.C.3.2.1.3; with debranching activity E.C.3.2.1.9.). The amylolytic enzymes of Schwanniomyces occidentalis, formerly also called Schwanniomyces castelli and Schwanniomyces aluivus (Price et al., 1976), are well documented (Oteig-Gyang et al., 1981; Silla et al., 1982, 1984a, 1984b; Wilson et al., 1982).

The genes for α-amylase and glucoamylase from a yeast of the genus Schwanniomyces have been disclosed by the present inventors in EP-A-0260404. However, at that time the DNA or peptide sequences directing secretion have not been identified. Furthermore their use for the establishment of an efficient expression and secretion system in yeast was not known then.

It is preferred, to use as a first DNA sequence part or all of a 1.8 kb BglII-Xhol fragment which precedes the structural gene coding for the Schwanniomyces α-amylase, part or all of a 1.3 kb BamHI-Pvull fragment preceding the structural gene coding for the Schwanniomyces occidentalis glucoamylase. Further if only smaller fragments are to be used, it is preferred, to use a DNA sequence comprising part or all of the DNA sequence corresponding to bases -1 to -540 of the region preceding the structural gene coding for α-amylase or part or all of the DNA sequence corresponding to bases -1 to -320 of the region preceding the structural gene coding for glucoamylase. The respective DNA fragments are fully active as promoters, inducible by starch, dextrin, maltose.

It is obvious, that replacement of bases or base pairs, not involved in the transcriptional regulation, still are within the spirits of the present invention. Thus natural mutations or synthetic equivalents also are embraced by the present invention.

In yeast cells of the genus Schwanniomyces it is also possible to use promoters derived from yeast genes obtained from many other yeast genera. Preferred are regulons involved in the expression of ADH1, ADH2, PDC1, GAL1/10, PGK and GAPDH, or LAC4 which are obtained from Saccharomyces cerevisiae, Kluyveromyces Lactis or Schwanniomyces occidentalis.

Furthermore, it is possible to use viral regulons and/or terminators, for example regulons and terminators obtained from E. coli T-phages. The phages possess extraordinarily strong functional sequences for performance of their lytic functions. In a preferred embodiment sequences of phage T7 are used, which are regulated by expression of T7 RNA polymerase.

In order to increase the over all expression level and furthermore, in order to facilitate purification of produced polypeptides, it is most desirable, that the polypeptides are secreted into the culture medium. The present invention provides signal sequences having the invaluable advantage of being derived from genes of amylolytic and thus induc-
ible enzymes, which upon induction have to be secreted very efficiently. A further advantage is, that induction of amy-
loytic enzymes does not require complicated media, lacking for example a specific amino acid or phosphate, but only
require the culture to be transferred to a starch containing medium. Signal sequences, which are preferably used for
secretion of large proteins include part or all of at least one of the following peptides:

\[ \begin{align*}
\text{a)} & \quad \text{MetArgPheSerThrGluGlyPheThrSerLysValValAlaAlaIle} \\
& \quad \text{LeuAlaPheSerArgLeuValSerAla;} \\
\text{b)} & \quad \text{MetArgPheSerThrGluGlyPheThrSerLysValValAlaAlaIle} \\
& \quad \text{LeuAlaPheSerArgLeuValSerAlaGlnProIlePheAspMetArg;} \\
\text{c)} & \quad \text{MetIlePheLeuLysLeuIleLysSerIleValIleGlyLeuGlyLeu} \\
& \quad \text{ValSerAla;} \\
\text{d)} & \quad \text{MetIlePheLeuLysLeuIleLysSerIleValIleGlyLeuGlyLeu} \\
& \quad \text{ValSerAlaIleGlnAlaAlaProAla;} \\
\text{e)} & \quad \text{MetIlePheLeuLysLeuIleLysSerIleValIleGlyLeuGlyLeu} \\
& \quad \text{ValSerAlaIleGlnAlaAlaProAlaSerIleGlySerSerAlaSerAla}
\end{align*} \]

In a further embodiment the above mentioned peptides are coded for by DNA sequences, which correspond to all
or part of the following DNA sequences:

\[ \begin{align*}
\text{a)} & \quad \text{ATGAGATTTTCACTGAAGGATTTTCAAAGTTAAGTGTTGCAAGCAATTTT} \\
& \quad \text{AGCATTTCACAGATTGATCTGCT;} \\
\text{b)} & \quad \text{ATGAGATTTTCACTGAAGGATTTTCAAAGTTAAGTGTTGCAAGCAATTTT} \\
& \quad \text{TAGCATTTCACAGATTGATCTGCTCAACCGATTACACACATGAGA;} \\
\text{c)} & \quad \text{ATGATTTTTGCTGAAGCTGATTAAAAGTAATAGTAACTGCTTGGGAAT} \\
& \quad \text{GTTAGTGCT;} \\
\end{align*} \]
d) ATGATTTTCTGAAGCTGATTAAAAGTATAGTAAATTGGTTTGGATTA
   GTTAGCTATCCAAACGCCCCTGCC;

5 e) ATGATTTTCTGAAGCTGATTAAAAGTATAGTAAATTGGTTTGGATTA
   GTTAGCTATCCAAACGCCCCTGCTTTGGATATTGCTCAG
   GCA.

It will be understood, that the replacement of amino acids by residues having comparable properties and use of alternative codons in compliance with the possibilities provided by the genetic code will not have any impact on the yeast cell according to the present invention.

The polypeptide to be expressed in a yeast cell according to the present invention is coded for by a third DNA sequence. Said third DNA sequence may comprise natural or synthetic DNA, even DNA containing intervening sequences. Examples for proteins being expressed by the yeast cell according to the present invention comprise proteins of enormous importance for scientific or medical purposes, for example cellulase, interleukine, interferon, insulin like growth factor, lymphokine, human growth factor, nerve growth factor, aprotin, insulin, hirudin, hormones, blood clotting factors, hepatitis B surface or core antigens, viral or bacterial vaccines or human granulocyte/macrophage colony stimulating factor.

The yeast cell according to the present invention may optionally comprise a terminator as component of said expression cassette. Preferred terminators are derived from genes of one of the following genera: Saccharomyces, Pichia, Hansenula, but preferably Schwanniomyces. Examples of preferred terminators are those derived from Schwanniomyces α-amylase gene or Schwanniomyces glucoamylase gene, part or all of which may be used. Such terminators are comprised by nucleotides 1537 to 1740 of the Schwanniomyces α-amylase gene or nucleotides 2875 to 3320 of the Schwanniomyces glucoamylase gene.

It is obvious, that in case of any of the DNA sequences mentioned modifications, which do not impair the biological function of the respective DNA fragment still are covered by the present application. There are numerous possibilities to change the DNA sequences mentioned above slightly without impairing the biological function of said fragment.

The yeast cell according to the present invention, carrying an expression cassette composed of the above discussed components, may be carried by a circular or linear vector, which optionally comprises one or more selective marker genes. The present inventors demonstrated, that a lot of marker genes so far used for transformation of Saccharomyces, for example TRP5, LEU2, ADE1, ADE2, HIS3, HIS4, URA3, LYS2, which may be obtained from Saccharomyces, Hansenula, Pichia or Schwanniomyces, are functional in Schwanniomyces occidentalis. Furthermore it is also possible to use the amylolytic properties of α-amylase or glucoamylase in order to select transformed organisms. The corresponding genes have been isolated and characterized. AMY1 and GAM1 are preferably obtained from Schwanniomyces.

The yeast cell according to the present invention may contain the expression cassette as an insert into one of the endogenous yeast chromosomes. Furthermore, there is however the possibility, that a DNA sequence capable of controlling autonomous replication and even also stable maintenance of the vector in the host organism, is introduced into the yeast cell. Said DNA sequences are the so called ARS sequences, capable of controlling autonomous replication and stable maintenance in the respective host organism. Said DNA sequences are preferably derived from Saccharomyces, Pichia Hansenula or Kluyveromyces, most preferably from Schwanniomyces.

In order to ensure effective replication in the respective host organism, it is preferred to use an ARS, derived from the respective host organism.

The present invention provides one new SwARS sequence. The DNA sequence SwARS2 is capable of controlling autonomous replication and stable maintenance of the vector in a Schwanniomyces yeast cell. The detailed sequence of SwARS has not yet been determined. SwARS2 is localised in the Schwanniomyces genome in the 3' region of the glucoamylase gene, as shown for example in Fig. 10. SwARS2 is easy to identify and efficiently allows autonomous replication of any plasmid, exhibiting this sequence. The so far known Schwanniomyces autonomously replicating sequences, SwARS1 and SwARS2 are comprised on a CiaI-BamHI fragment (SwARS1) adjacent to the EcoRI fragment shown in Fig 2a (see Fig. 10) and/or an EcoRI-BglII fragment contained in the EcoRI fragment shown in Fig 2a (SwARS2).

For the SwARS sequences it is also not necessary, in any case to use the sequence as it occurs naturally. DNA sequences, that have been obtained by any kind of modifying treatment, for example by deletion, insertion or substitution of one or more nucleotides or base pairs, are still within the scope of the present invention, provided the modi-
fications retain or improve the biological function of the yeast cell according to the present invention.

In a preferred embodiment, the yeast cell according to the present invention comprises a DNA sequence, which is homologous to genomic Schwanniomyces DNA. Provision of structures of homologous DNA enable the DNA fragment flanked by such homologous sequences to be inserted by homologous recombination. It is thus preferred to have homologous sequences, which do not represent any essential gene of the Schwanniomyces genome DNA, but rather any gene, whose expression product may be supplemented by additives to the respective medium or were lack of the respective expression product does not have any serious impact on the transformed yeast cell.

If no SwARS sequence or any equivalent sequence, for example originating from any other yeast species, is available, the expression cassette can be used without an autonomously replicating agent. In this case the expression cassette tends to integrate into the genome, for example via homologous recombination. Preferably there are several copies inserted into one or more chromosomes.

In this case it is preferred, to have multiple copies of the expression cassette inserted, in order to ensure simultaneous transcription and translation of several copies.

When the yeast cell according to the present invention is derived from a yeast of the genus Schwanniomyces the polypeptide products observed following expression of the respective foreign gene in said yeast cells exhibit glyco-sylation patterns, which closely resemble or even match those of the desired eucaryotic protein. Yeast cells of the genus Schwanniomyces are thus preferred for expression of eucaryotic proteins.

The above discussed yeast cell according to the present invention can be used in a process for the production of a polypeptide, wherein a yeast host organism is cultivated under suitable conditions and the polypeptide is recovered in a manner known per se. The conditions of cultivation depend on the respective host organism used. Examples for cultivation conditions for yeast of the genus Schwanniomyces are provided in the examples below. The polypeptide produced by the yeast according to the present invention, is recovered by conventional purification methods, which depend on the properties of the expressed protein, the availability of specific absorbing agents and so on.

In a preferred embodiment yeast cells of the species Schwanniomyces occidentalis are used as a host organism and are cultivated on starch dextrin, maltose and/or plant biomass as the respective carbon source. As commonly known, these carbon sources are available at low costs.

Furthermore, there is a possibility to use cells of the genus Schwanniomyces for the production of single cell protein. The single cell protein may then be subjected to any further use as it is required.

In a preferred embodiment the polypeptide produced in a process according to the present invention is an amylolytic enzyme, for example an α-amylase or glucoamylase. These enzymes are commercially used to debranch and degrade starch, which thereupon may be further processed, for example in brewery or baking processes. There is a high need for the provision of starch degrading enzymes.

In a further embodiment of the process according to the present invention production of a polypeptide may be achieved by providing expression cassettes as discussed above containing a yeast α-amylase or glucoamylase signal sequence and inserting the respective expression cassette into a yeast of one of the genera Saccharomyces, Kluyveromyces, Hansenula, Pichia or Schizosaccharomyces. Production and secretion of foreign proteins in the respective yeast host organism is obtained by applying the appropriate inductive measure, whereupon the produced protein is secreted due to the presence of the yeast α-amylase or glucoamylase signal sequence. Since it is known from experiments performed by the present inventor, that these signal sequences are recognised in virtually any yeast genus, the α-amylase or glucoamylase signal sequences may be applied in many different yeast with success. Also in this case, it is preferred, to use the α-amylase or glucoamylase signal sequences derived from the respective genes of Schwanniomyces occidentalis.

In summary the yeast cells according to the present invention are well suited for transformation and expression of any desired foreign gene. As already outlined above, expression of procaryotic as well as eucaryotic genes is possible. However, for expression of eucaryotic proteins yeast cells of the genus Schwanniomyces containing some or all of the DNA sequences mentioned above are the most preferred yeast cells.

CHARACTERISATION OF THE α-AMYLASE AND GLUCOAMYLASE GENE EXPRESSION IN DIFFERENT YEAST GENERA:

The N-termini of the deduced amino acid sequences of the α-amylase gene (Fig. 1B) and glucoamylase gene (Fig. 2B) exhibit the features of typical leader sequences of secretory proteins (Von Heijne 1983, Perlman and Halvorson 1983). The N-terminal amino acids of mature α-amylase secreted by S. occidentalis are determined to be AspValSer. This indicates processing between Arg33 and Asp31 which is unusual and not according to the Von Heijnes model of site specificity for peptide signal cleavage. This putative cleavage site was confirmed also for α-amylase expressed in S. cerevisiae, Pichia stipitis, K. lactis, S. pombe and H. polymorpha. However, in these yeast species processing is also found to occur between Ala 25 and Gln 26, resulting in the N-terminal sequence GinProlle.

α-amylase isolated from a culture of S. occidentalis migrates as a single band of 55 kDa on SDS polyacrylamide
gels: following Endoglucosidase H treatment and subsequent gel electrophoreses a single band of 54 kDa is observed, indicating that only one of the two potential N-glycosylation sites of the protein is used for glycosylation. It is noteworthy that high mannose glycosylation, as very often found by proteins secreted in S. cerevisiae is not accompanying secretion in S. occidentalis.

IDENTIFICATION OF THE TRANSCRIPTION INITIATION SITES AND PROMOTER FUNCTION OF THE α-AMYLASE AND GLUCOAMYLASE GENE.

By S1 mapping of polyA-mRNA (Sharp et al., 1980) the transcription initiation site of the α-amylase gene could be localised to be in the region of -42 to -30 (indicated by an interrupted line above the relevant sequence in Fig.1B) and the transcription initiation site of the glucoamylase gene was determined to be in the region -7 to -10 (underlined in Fig.2B).

CONSTRUCTION OF AN α-AMYLASE - GLUCOAMYLASE HYBRID GENE

In order to facilitate use of the α-amylase promoter and glucoamylase structural gene in further constructions additional restriction sites were inserted in front of the translation initiation codon of both genes using oligonucleotide directed mutagenesis as described in the AMERSHAM protocol. Correct mutagenesis was verified by sequencing according to Sanger (1977). The boundaries between the original Schwanniomyces genes and the inserted polylinker are shown below:

\[ \alpha\text{-amylose} : \]

\[ \text{**} * \text{**} \]

\[ 5' \ldots T A A T A A A G C T C G A G A T G A G A T T T \ldots 3' \]

\[ \_XhoI \_ \]

\[ \text{Glucosamylase :} \]

\[ \text{*} \text{*****} \text{*} \]

\[ 5' \ldots C T C A T G A C T G T G C G A C G G A T C C A A G A T G A T T T T \ldots 3' \]

\[ \_SaiI__BamHI \_ \]

The manipulations resulted in plasmids M13xXho (Fig.3) and pMaGAM-B/S (Fig.4).

Thus, both the α-amylase promoter and the structural glucoamylase gene can be isolated from plasmid M13xXho and pMaGAM-B/S as suitable DNA fragments. The GAM gene was isolated from pMaGAM B/S as a 3.2 kb Sall fragment and ligated into the Xhoi site of M13xho. The BglII fragment harbouring the α-amylase promoter/GAM gene fusion was ligated into the single BamHI site of the S. cerevisiae vector YRp7 (Struhl et al., 1979) resulting in plasmid YRp7γGAM (Fig 5). S. cerevisiae strain YNN27 was transformed with this plasmid selecting for TRP prototrophy. Transformants secreted active glucoamylase into the culture supernatant (50 mU/ml) indicating that the α-amylase promoter fragment can direct expression and secretion of foreign proteins.

SECRETION OF α-AMYLASE AND GLUCOAMYLASE IN DIFFERENT YEASTS USING SUITABLE PROMOTERS.

Expression of the α-amylase and glucoamylase gene and secretion of the gene products can be detected in K. lactis under control of the promoter of the α-galactosidase (LAC4) gene (Breunig et al., 1984). Expression of the α-amylase and glucoamylase gene and secretion of the both proteins can also be detected in S. pombe under control of the ADH1 promoter (Russel et al., 1983) or of the GAL1/10 promoters (Johnston and Davis 1984). After fusion with the GAL1/10 promoter (Johnston and Davis 1984), DHAS promoter (Janowicz et al., 1985), MOX promoter (Ledeboer et al., 1985) or FMDH promoter (EP 87 110 417.0) expression of the α-amylase and glucoamylase
gene and secretion of both proteins can be detected in \textit{trans} in \textit{S. cerevisiae}.

\textbf{EXPRESSION AND SECRETION STUDIES IN S. OCCIDENTALIS USING THE CELLULASE ENDOGLUCANASE D (EGD) OF CLOSTRIDIUM THERMOCELLUM AS A MODEL FOR FOREIGN GENE EXPRESSION.}

As a first example for the expression of a foreign gene in \textit{S. occidentalis}, the celD gene from \textit{C. thermocellum} (Millet et al., 1985; Jolliet et al., 1986b), which codes for a thermostable cellulase, namely an endoglucanase (EGD) was used. Advantages in using this system are:

- specific reaction of EGD, which can easily be monitored.
- a rapid qualitative assay for colonies with EGD activity (see material and methods).
- availability of antibodies against EGD for specific detection of translation products.

\textbf{CHARACTERISATION OF THE GLUCOAMYLASE PROMOTER AFTER FUSION WITH THE celD GENE}

Plasmids were constructed containing a replicon from \textit{S. cerevisiae} (SwARS1) (EP 87110370.1) the TRP5 gene from \textit{S. cerevisiae} (Zalkin and Yanovski, 1982) as a selective marker and, in addition, a GAM/celD gene fusion under the control of different promoters. In the first step the 4.0 kb BcorI-PvuII fragment from plasmid pBRSwARS3GAM (Fig. 6a) was isolated and inserted into the 2296 bp pBR322 EcoRI-PvuII fragment, containing the amp gene and the bacterial origin, resulting in plasmid pBRPGAM (Fig. 6a). In addition to the pBR322 sequence this plasmid carries 3.6 kb from the 5' non coding region of the GAM gene and the first 206 bp coding for the N-terminal part (including the signal sequence) of the glucoamylase. A 3.4 kb PvuII fragment from plasmid pCT6C3 (Jolliet et al., 1986) containing the coding region of the celD gene missing the 5' 111 bp was inserted into the PvuI single site of pBRPGAM resulting in pBRGC1 (Fig. 6b). E. coli transformed with this plasmid results in weak Carboxymethylcellulase (CMCase) activity, analysed by Congo red staining (see material and methods). The glucoamylase promoter is slightly expressed in E. coli. For the construction of a \textit{S. occidentalis} expression vector the plasmid pCD5-1 (EP 87110370.1) was cleaved with BamHI/PstI and ligated with the 5.5 kb BglII-PstI fragment or 6.5 kb BamHI-PstI fragment from pBRGC1, respectively. Both plasmids were transformed into frame fusion of GAM/celD but differ in the length of the 5' upstream region of the GAM gene. \textit{S. occidentalis} strain NGA56 (trp5 ade) isolated after UV mutagenesis of strain NGA23 was transformed with these plasmids and transformants selected for tryptophan prototrophy and analyzed by means of Congo red assay for cellulase activity. Transformants produced active EGD when grown under inducing conditions in 2% maltose (Fig. 7a). By analysis with the Congo red assay no activity could be detected under repressed conditions in 4% glucose (Fig. 7b). NGA56: pMPGC1-1 and NGA56: pMPGC1-2 transformants secrete active EGD into the culture supernatant when grown in maltose under induced conditions. The intact or complete glucoamylase promoter resides on plasmid pMPGC1-2 as shown by 20 times higher activity compared to the shorter promoter fragment in gene fusion of pMPGC1-1 (table 1). However the 0.3 kb promoter fragment still harbours regulatory units which lead to catabolite repression by glucose (Fig. 7b). Comparison of the N-terminal GAM sequence to the signal peptidase cleavage sites of several other secretory proteins (Van Heijne 1983) reveals three potential signal peptidase cleavage sites after the amino acid residues Ala 19, Ala 25 and Ala 34. For determination of the N-terminal amino acid of the mature glucoamylase the protein was isolated for N-terminal amino acid sequencing. However the N-terminals is blocked and therefore could not be sequenced. To confirm the correct cleavage site of the GAM signal sequence in \textit{S. occidentalis}, the regions between the potential signal peptidase cleavage sites and the PvuII gene fusion site are deleted in the GAM/celD fusion by in vitro mutagenesis to study the influence on secretion of active EGD.

\textbf{STUDIES OF THE USE OF SUITABLE S. CEREVISIAE PROMOTERS IN THE EXPRESSION OF FOREIGN GENES IN S. OCCIDENTALIS}

The ADH1 promoter (Hitzeban et al., 1981) was isolated as a 1.45 kb BamHI-Sall fragment from plasmid pFM2-1 (Mueller et al., 1987) and ligated into plasmid pCD5-1 (EP 87 110370.1) after cleavage with BamHI and Sall resulting in pMPADH1 (Fig. 5a). The GAM gene was excised as a Sall fragment from pJDC9-15 (see below) and inserted into the single Sall site of pMPADH1 (Fig. 5b) resulting in plasmid pMPAG (Fig. 5b). By means of in vivo recombination (Ma et al., 1987) the GAM promoter from pMPGC1-2 (Fig. 6a) was replaced by the ADH1 promoter of pMPAG leading to plasmid pMPAGC (Fig. 5c). Plasmid pMPGC1-2 was linearized with BamHI; the unique BamHI site of pMPGC1-2 is localized between the TRP5 coding sequence and the glucoamylase promoter. The 2.75 kb CiaI fragment of pMPAG containing the ADH1 promoter shares homologies to the TRP5 gene as well as to the GAM coding region up to position +206. For in vivo recombination NGA56 was transformed with BamHI linearized pMPGC1-2 together with approx. 10 fold excess of CiaI fragment using the transformation protocol of Ito et al., 1983.
Recombinants were screened for their ability to hydrolyze CMC on 4% glucose. Under these conditions celD expression is repressed in transformants harbouring recircularized pMPPG1-2. From 146 transformants 7 colonies expressing cellulase in the presence of 4% glucose were isolated. NGA58 containing the *in vivo* constructed plasmid pMPPG1 expresses active EGD constitutively (see Fig. 7). This result demonstrates that the newly constructed plasmid pMPADH1 serves as a vector for a foreign gene expression.

In another experiment a recombinant plasmid was constructed by *in vivo* recombination to study the expression of the GAM/celD gene fusion under control of the PDC promoter (Das and Hollenberg, 1982). Therefore the 6.0 kb Sphl-Aval fragment from plasmid pJDcg-15 (EP 87 110370.1) carrying the glucoamylase gene fused to the PDC promoter was subcloned into the respective restriction sites of YRp7D2 (EP 87 110370.1) resulting in plasmid pMPPG (Fig. 9a).

For *in vivo* recombination pMPPG was cleaved with ClaI and used together with BamHI linearized pMPPG1-2 for transformation of NGA58. Homologous recombination results in plasmid pMPPG (Fig. 9b) harbouring a GAM-celD fusion under control of the PDC promoter. This construction leads to an expression in Schwanniomyces occidentalis NGA58 inducible by glucose and partly repressible by maltose (Table 1). Transformants were analysed for expression of active EGD using the qualitative congo red assay (Fig. 7). The results indicate that the 0.3 kb GAM promoter fragment leads to regulated expression, however less efficient that the 1.3 kb fragment (for comparison see table 1 and Fig. 2A). Expression under the ADH1 promoter leads to efficient constitutive expression. The expression level is comparable with that of induced expression directed by the 1.3 kb GAM promoter.

The fact that plasmids can be constructed by *in vivo* recombination is a proof for homologous recombination and hence offers the possibility for stable site directed integration into the genome of *S. occidentalis*, a prerequisite for the industrial production of foreign proteins.

The expression of the α-amylase gene and glucoamylase gene and secretion of the both gene products in *S. occidentalis* could also be achieved under control of the phosphoglycerate kinase (PGK) promoter (Dobson et al., 1982), glyceraldehyde-3-phosphate dehydrogenase (GAPDH) promoter (Holland and Holland 1979), copperchelatin (CUP1) promoter (Karlin et al., 1984, Butt et al., 1984) and alcohoholdehydrogenase 2 (ADH2 or ADH3) promoter (Russell et al., 1983, Beier et al., 1985) or galactokinase (GAL1/10) promoter (Johnston and Davis, 1984).

Construction of pJDcg-15:

The GAM gene was isolated as a 3.2 kb BamHI-Sall DNA fragment from plasmid pMaGAM B/S (Fig. 4) and inserted into the respective restriction sites of plasmid pBM272 (Johnston & Davis, 1984). The PDC promoter was isolated from plasmid pCP202 (Kellermann & Hollenberg, 1986) as a Sall-Sphl fragment and ligated into the respective single sites of vector YRp7 (Struhl et al., 1979). To be able to isolate the PDC1 promoter as a BamHI fragment the 3.2 kb Sall fragment of pMaGAM B/S containing the GAM gene was fused to the PDC promoter via the single Sall site. The resulting BamHI fragment containing the PDC-promoter was ligated into the single BamHI site of pBM272 containing the GAM gene. The final plasmid is called pJDcg-15 (Fig. 8b).

EXPRESSION OF HUMAN GRANULOCYTE/MACROPHAGE COLONY STIMULATING FACTOR (hGM-CSF)

The gene coding for hGM-CSF (Cantrell et al., 1985) was obtained from British Biotechnology Ltd. as a HindIII-EcoRI fragment. Position +52 of hGM-CSF gene was fused to position +102 of the GAM signal sequence under control of the ADH1 or GAM promoter. This fusion results in a hybrid protein in which the endopeptidase cleavage site of the GAM signal sequence is linked with a Met included in the synthetic gene offered by British Biotechnology Ltd., followed by the first amino acid of mature hGM-CSF (Ala 18). This fusion was inserted into a *S. occidentalis* expression vector and transformed into NGA58, Western blot analysis revealed secretion of hGM-CSF protein.

EXPRESSION OF INTERLEUKINE-2

The gene coding for IL-2 (Taniguchi et al., 1983) was obtained from British Biotechnology Ltd., as a EcoRI-HindIII fragment. Position +61 of IL-2 gene was fused to position +102 of the GAM signal sequence under control of the ADH1 or GAM promoter. This fusion results in a hybrid protein in which the endopeptidase cleavage site of the GAM signal sequence is linked with a Met included in the synthetic gene offered by British Biotechnology Ltd., followed by the first amino acid of mature IL-2 (Ala 21). This fusion was inserted into a *S. occidentalis* expression vector and transformed into NGA58. Western blot analysis revealed secretion of IL-2 protein.

ISOLATION OF AN ADDITIONAL SwARS ELEMENT AND A NEW SELECTABLE MARKER

A new SwARS sequence (SwARS2), localized downstream of the coding region of the GAM gene, could be iden-
tified. To analyze the function of SwARS2 a vector was constructed containing the Trps gene and SwARS2 in pBR322, resulting in pMPTS2-A and pMPTS2-B, respectively (Fig. 10). For the construction of pMPTS2-A and pMPTS2-B the 5.5 kb BamHI-BglII fragment of pBRSwARS-GAM, constructed by insertion of the EcoRI fragment shown in Fig. 2a into the single EcoRI site of plasmid pBRSwAR51, (disclosed in EP 87 110370.1), was isolated and ligated with 3.2 kb BamHI TRP5 fragment obtained from YRPJD2. High frequency transformation of S. occidentalis is obtained using this plasmid.

As an additional selectable marker the LEU2 gene of S. occidentalis was isolated. The LEU2 gene, coding for isopropyl malate dehydrogenase, was cloned by functional complementation of the mutant AH22 (his4leu2) (Hinnen et al., 1976) of S. cerevisiae by transformation with a cosmid genomic library from S. occidentalis.

The LEU2 gene was found to reside on a 9 kb EcoRI fragment of Schwanniomyces genomic DNA; its function could be abolished by insertion of the TRP5 gene of S. cerevisiae. The disrupted LEU2 gene was integrated into the genome of NGA58 by homologous recombination resulting in a new strain NGA581 (TRP5, LEU2, and ADE). An analogous experiment with the cloned HIS4 gene of S. occidentalis (EP 87 110 370.1) led to a strain NGA58h (TRP5, HIS4, ADE).

IMPROVED EXPRESSION OF HETEROLOGOUS GENES USING THE T7 RNA POLYMERASE AND T7 PROMOTER AND TERMINATOR FOR THE EXPRESSION OF FOREIGN GENES IN S. OCCIDENTALIS

T7 RNA polymerase (Fuerst et al., 1986) can be expressed in S. occidentalis after fusion of the structural gene (Dunn and Studier, 1984) using the BamHI restriction site near the translation initiation codon with suitable promoters asu-amyrase, GAM (0.3 kb fragment) or PDC and homologous integration into the genome via any homologous cloning gene, e.g. HIS4, LEU2, AMY1 and GAM. A suitable fragment, containing the fusion flanked by homologous cloned sequences can be isolated and used together with pCJD5-1 (Fig. 6c) for co-transformation of NGA58, selecting on YNB containing 0.5% casaminoacids. Colonies containing the integrated T7 RNA polymerase gene can be identified by Southern analysis after isolation of transformants mutated in the target gene.

For the expression of foreign genes a S. occidentalis expression vector was constructed containing SwARS2, the ColE1 replication origin of E. coli, a suitable selective marker for S. occidentalis (e.g. TRPS, HIS4, LEU2) and a cassette containing the T7 polymerase promoter and terminator sequences isolated from plasmid pARD529 (Fuerst et al., 1986), separated by the polylinker sequence from M13mp19. To study the efficiency of the T7 system we expressed a GAM/ celD fusion (GAM signal sequence and structural celD gene obtained as a BamHI-Asp 718 fragment from pMPGC1-2) under control of the T7 promoter by insertion into the BamHI-Asp718 site of the polylinker sequence. Using this system active cellulase is secreted into the culture medium.

MATERIALS AND METHODS

1. The microorganisms used are as follows:

Schwanniomyces occidentalis NG23 (DSM 3792)
Schwanniomyces occidentalis NG8
Kluyveromyces lactis SD11 (DSM 3795)
Schizosaccharomyces pombe LEU1-32, HIS3-303 (DSM 3796)
Saccharomyces cerevisiae YNN27 (URA3-52, TRP1-298, GAL2)

The E. coli strain used is as follows E. coli HB101 F-, hsd S20 (rB-, mB-), recA13, ara-14, proA2, leu1, galK2, rspL20 (Smr), xyl1, mtl1, supF44

2. Culture media:

Schwanniomyces occidentalis strain NG28 (isolated by UV mutagenesis of strain NG23) was grown at 30°C in YNB (0.67% yeast nitrogen base without amino acids) supplemented with appropriate amino acids and with 2% of soluble starch when indicated.

S. cerevisiae strain YNN27 (URA3-52, TRP1-298, GAL2) (Stinchcomb et al 1980); K. lactis strain SD11 TRP1 (Das and Hollenberg, 1982); S. pombe strains LEU1-32, HIS3-303 (obtained from W.D. Heyer). These strains were grown under conditions as described above for Schwanniomyces occidentalis with the replacement of 2% soluble starch by
2% glucose, where indicated.
The culture media for \( S. \) cerevisiae and \( H. \) polymorphus were buffered with 0.2 M NaPO₄ buffer pH 6.2; those for \( S. \) parapumila and \( K. \) lactis were buffered with 0.1 M citrate buffer pH 6 and those for \( S. \) pombe with 0.05 M acetate buffer pH 6.
The pH optimum for the cellulase (Joliff et al., 1968a), \( \alpha \)-amylase and glucoamylase (Wilson and Ingledew, 1982) is approx. pH 6.

\( E. coli \) HB101 (Bolivar et al., 1977) was grown in LB medium (Maniatis et al., 1982) with penicillin G (150 \( \mu \)g/ml).

3. Miscellaneous methods

Plasmid DNA was purified either by CsCl gradient centrifugation (Maniatis et al., 1982) or by rapid alkaline extraction of plasmid DNA described by Birnboim and Doly (1979).

Yeast crude extracts were prepared by the method of Schatz (1979). Yeast miniylisates were prepared by the method of Sherman et al., (1983).

DNA fragments were isolated by the "low melting agarose" procedure as described by Gaten et al., (1983). Southern hybridization was carried out as described by Southern (1975).

Yeast transformation was performed according to Klebe et al. (1983) or Ito et al. (1983). The transformants or integrants were tested for expression and secretion of cellulase using either the qualitative congo red assay (Teather and Wood, 1982: halo formation on carboxymethylcellulose agar plates) or a quantitative cellulase assay according to Joliff et al. (1968a). One unit (U) is defined as the amount of enzyme that released 1 umol para-nitrophenol per min at 60°C. Western analysis was performed as described by Towbin et al. (1979).

The transformants or integrants were tested for expression and secretion of \( \alpha \)-amylase using either a standardized \( \alpha \)-amylase enzyme test from Merck, W. Germany or a qualitative starch degradation test (halo formation on starch agar plates after staining with iodine), as follows: in this test the rate of formation of 2-chloro-4-nitrophenol is determined by photometry at 405 nm in 0.1 M potassium phosphate at 37°C. The enzyme unit (U) is defined as the amount of enzyme catalysing the formation of 1 umol 2-chloro-4-nitrophenol per min at 37°C.

Glucoamylase activity was measured by a stop assay method:

After incubation of the sample in 10% soluble starch in 0.05M \( \text{KH}_2\text{PO}_4 \) NaOH (pH 5.0) at 50°C the amount of glucose produced is determined by the glucose dehydrogenase method (system glucose Merck). The quantity of NAD was formed is proportional to the glucose concentration. The enzyme unit (U) is defined as the amount of enzyme catalysing the formation of 1 umol glucose/min at 50°C.

<table>
<thead>
<tr>
<th>Transformant</th>
<th>OD₄₀₀ nm</th>
<th>EDG-activity mU/ml</th>
</tr>
</thead>
<tbody>
<tr>
<td>NGA58:</td>
<td>2.3</td>
<td>0.12</td>
</tr>
<tr>
<td>pMPPG1-1 (0.3 kb promoter)</td>
<td>5.3</td>
<td>0.73</td>
</tr>
<tr>
<td>NGA58:</td>
<td>2.0</td>
<td>0.53</td>
</tr>
<tr>
<td>pMPPG1-2 (1.3 kb GAM promoter)</td>
<td>7.9</td>
<td>1.83</td>
</tr>
<tr>
<td>NGA58:</td>
<td>3.0</td>
<td>0.52</td>
</tr>
<tr>
<td>pMPAGC ADH1 promoter</td>
<td>5.9</td>
<td>1.07</td>
</tr>
<tr>
<td>NGA58:</td>
<td>2.2*</td>
<td>0.40*</td>
</tr>
<tr>
<td>pMPPGC</td>
<td>4.5*</td>
<td>0.64*</td>
</tr>
<tr>
<td>PDC1-promoter</td>
<td>2.9</td>
<td>0.19</td>
</tr>
<tr>
<td></td>
<td>7.2</td>
<td>0.31</td>
</tr>
</tbody>
</table>

* values for transformants grown in 4% glucose

Transformants were grown in YNB, 0.5% Casaminoacids, 2% maltose, 20 mg/l adenine and 0.2 M NaP buffer pH 6.2

LIST OF REFERENCES


Bolivar et al. (1977), Gene 2, 75-93.
Breunig (1984), NUCl. Acids Res. 12, 2327-2341.
Dobson et al. (1982), Nucleic Acids Research 10, 2625-2637.
Holland & Holland (1979), J. Biol. Chem. 254, 5466-5474.
Janowicz et al. (1985), Nucleic Acids Res. 9, 3043-3062.
Joliff et al. (1986a), Bio/Technology 4, 696-900.
Joliff et al. (1986b), NUCl. Acids Res. 14, 8605-8613.
Klebe et al. (1983), Gene 25, 333-341.
Kramer et al. (1984), Nucl. Acids Res. 12, 9441-9456.
Kurjan & Herskowitz (1982), Cell 30, 933-943.
Ledeboer et al. (1985), Nucl. Acids Res. 9, 3063-3082.
Ma et al. (1987), Gene 59, 201-216.
Russell et al. (1983), J. Biol. Chem. 258, 2674-2682.
Schatz (1979), Meth. Enzymol. 175, 40-50.
Southern (1975), J. Mol. Biol. 98, 503-517.
Taniguchi et al. (1983), Nature 302, 305-309.

Claims

Claims for the following Contracting States : AT, BE, CH, LI, DE, FR, GB, GR, IT, LU, NL, SE

1. A yeast cell of the genus Schwanniomyces, characterized in that said yeast cell contains at least one expression cassette comprising:
   a) a first DNA sequence serving as a regulon,
   b) optionally a second DNA sequence coding for a signal peptide
   c) a third DNA sequence coding for a foreign protein, and
   d) optionally a fourth DNA sequence serving as a terminator,

wherein the first, second and fourth DNA sequence are derived from yeast.
2. The yeast cell according to claim 1, characterized in that said first and/or second and/or fourth DNA sequence are derived from a gene coding for an amylolytic enzyme.

3. The yeast cell according to claim 2, characterized in that said amylolytic enzyme is an α-amylase or glucoamylase.

4. The yeast cell according to any of claims 1 to 3, characterized in that said first and/or second and/or fourth DNA sequences are derived from a yeast of a member of the genera Debaromyces, Saccharomycopsis, Lipomyces, Pichia, or Saccharomyces, preferably Schwanniomyces.

5. The yeast cell according to claim 4, characterized in that said yeast is Schwanniomyces occidentalis.

6. The yeast cell according to any of claims 1 to 5, characterized in that said first DNA sequence is comprised by a 1.8 kb BglII-XhoI-fragment, if derived from Schwanniomyces occidentalis α-amylase gene, or a 1.3 kb BamHII-PvuII fragment if derived from Schwanniomyces occidentalis glucoamylase gene.

7. The yeast cell according to claim 5, characterized in that said first DNA sequence is comprised by part or all of the DNA sequence corresponding to bases -1 to -540 of the region preceding the structural gene coding for α-amylase or part or all of the DNA sequence corresponding to bases -1 to -320 of the region preceding the structural gene coding for glucoamylase.

8. The yeast cell according to any of claims 1 to 5, characterized in that said first DNA sequence corresponds to one of the regulons regulating expression of:

   a) ADH1, ADH2, PDC1, GAL1/10, PGK, GAPDH, wherein the regulon is preferably obtained from Saccharomyces cerevisiae,

   b) LAC4, wherein the regulon is preferably obtained from Kluyveromyces lactis, or

   c) corresponding regulons, preferably obtained from Schwanniomyces.

9. The yeast cell according to any of claims 1 to 8, characterized in that said second DNA sequence is coding for all or part of at least one of the following peptides:

   a) MetArgPheSerThrGluGlyPheThrSerLysValValAlaAlaIle
   LeuAlaPheSerArgLeuValSerAla;

   b) MetArgPheSerThrGluGlyPheThrSerLysValValAlaAlaIle
   LeuAlaPheSerArgLeuValSerAlaGlnProIleIlePheAspMetArg;
c) MetIlePheLeuLysLeuIleLysSerIleValIleGlyLeuGlyLeuValSerAla;

d) MetIlePheLeuLysLeuIleLysSerIleValIleGlyLeuGlyLeuValSerAlaIleGlnAlaAlaProAla;

e) MetIlePheLeuLysLeuIleLysSerIleValIleGlyLeuGlyLeuValSerAlaIleGlnAlaAlaProAlaSerIleGlySerSerAlaSerAla

10. A yeast cell according to claim 9, characterized in that the DNA sequence coding for said peptide corresponds to part or all of one of the following DNA sequences:

a) ATGAGATTTTCACTGAAGGATTTCACAAGTAAAGTTGTTGCAGCAATT
TTAGCATTCTCAAGATTGTATCTGCT,

b) ATGAGATTTTCACTGAAGGATTTCACAAGTAAAGTTGTTGCAGCAATT
TTAGCATTCTCAAGATTGTATCTGCTCAACCGATTATTTTTGACATGAGA;

c) ATGATTTTTTCTGAAGCTGATTTAAAGTATAGTAATTGTTTGGGATTA
GTAGTGTCT;

d) ATGATTTTTTCTGAAGCTGATTTAAAGTATAGTAATTGTTTGGGATTA
GTAGTGTCTATCCAAGCAGCCCTGCCC;

e) ATGATTTTTTCTGAAGCTGATTTAAAGTATAGTAATTGTTTGGGATTA
GTAGTGTCTATCCAAGCAGCCCTGCCCCTGCTCGATTGAGATCTAGTGCCTCA
GCA.

11. The yeast cell according to any of claims 1 to 10, characterized in that said third DNA sequence coding for a foreign protein is a natural or synthetic DNA sequence encoding a cellulase, interleukin, insulin like growth factor, interferon, lymphokine, human growth factor, nerve growth factor, aprotinin, insulin, hirudin, hormones, blood clotting factors, hepatitis B surface or core antigens, viral or bacterial vaccines or human granulocyte/macrophage colony stimulating factor.

12. The yeast cell according to any of claims 1 to 11, characterized in that
said fourth DNA sequence originates from a yeast of a member of the genera *Saccharomyces*, *Pichia*, *Hansenula*, preferably *Schwanniomyces*.

13. The yeast cell according to claim 12,
characterized in that
said fourth DNA sequence corresponds to part or all of the terminator comprised by nucleotides 1537 to 1740 of the *Schwanniomyces* α-amylose gene or part or all of the terminator comprised by nucleotides 2875 to 3220 of the *Schwanniomyces* glucoamylase gene.

14. The yeast cell according to any of claims 1 to 13,
characterized in that
said first and/or second and/or third and/or fourth DNA sequence has been modified by insertion or deletion or substitution of one or more nucleotides while the biological activity of the DNA sequence is retained or improved.

15. The yeast cell according to any of claims 1 to 14,
characterized in that
said expression cassette is carried by a circular or linear vector, which optionally comprises one or more selective marker genes.

16. The yeast cell according to claim 15,
characterized in that
said vector comprises any of the following selective marker genes: TRP5, LEU2, ADE1, ADE2, HIS3, HIS4, URA3, LYS2, preferably obtained from *Saccharomyces*, *Hansenula Pichia* or *Schwanniomyces*, AMY1 or GAM1, preferably obtained from *Schwanniomyces*.

17. The yeast cell according to any of claims 15 and 16,
characterized in that
said vector further comprises a DNA sequence capable of controlling autonomous replication and stable maintenance of the vector in the host organism.

18. The yeast cell according to claim 17
characterized in that
said DNA sequence is derived from the genome of a member of the genera *Saccharomyces*, *Pichia*, *Hansenula* or *Kluveromyces*, preferably *Schwanniomyces*.

19. A yeast cell according to any of claims 17 or 18,
characterized in that
said DNA sequence is selected from the autonomously replicating sequences ARS, preferably an ARS derived from the respective host organism.

20. The yeast cell according to claim 19,
characterized in that
said DNA sequence is SwARS1 or SwARS2, capable of controlling autonomous replication and stable maintenance of the vector in *Schwanniomyces*, wherein SwARS1 is comprised on a Clal-BamHI fragment adjacent to the EcoRI fragment shown in Fig. 2a, and SwARS2 is comprised on an EcoRI-BglII fragment contained in the EcoRI fragment shown in Fig. 2a.

21. The yeast cell according to any of claims 17 to 20,
characterized in that
said DNA sequence has been modified by deletion, insertion or substitution of one or more nucleotides, while retaining or improving its biological function.

22. The yeast cell according to any of claims 1 to 21,
characterized in that
said expression cassette is integrated into the genome, optionally in multiple copies, preferably via homologous recombination.

23. A process for the production of polypeptide, wherein a yeast host organism is cultivated under suitable conditions
and the polypeptide is recovered in a manner known per se,
characterized in that
said yeast host organism is a yeast cell according to any of claims 1 to 22.

24. The process according to claim 23
characterized in that
yeast cells of the species Schwanniomyces occidentalis are used as a host organism and are preferably cultivated
on starch dextrin, maltose and/or plant biomass.

25. The process according to any of claims 23 or 24,
characterized in that
yeast cells of the genus Schwanniomyces are used for the production of single cell protein.

26. The process according to any of claims 23 to 25,
characterized in that
said polypeptide is amylase or glucoamylase.

27. Use of a yeast cell according to any of claim 1 to 22 for transformation and expression of foreign genes.

Claims for the following Contracting State: ES

1. A process for preparing a yeast cell capable of expressing foreign protein,
characterized in that
at least one expression cassette comprising:

   a) a first DNA sequence serving as a regulon,

   b) optionally a second DNA sequence coding for a signal peptide

   c) a third DNA sequence coding for a foreign protein, and

   d) optionally a fourth DNA sequence serving as a terminator,

   is inserted into a yeast cell of the genus Schwanniomyces, wherein the first, second and fourth DNA sequence
   are derived from yeast.

2. The process according to claim 1,
characterized in that
said first and/or second and/or fourth DNA sequence are derived from a gene coding for a amylolytic enzyme.

3. The process according to claim 2,
characterized in that
said amylolytic enzyme is an α-amylase or glucoamylase.

4. The process according to any of claims 1 to 3,
characterized in that
said first and/or second and/or fourth DNA sequences are derived from a yeast of a member of the genera Debar-
omyces, Saccharomyces, Lipomyces, Pichia, or Saccharomyces, preferably Schwanniomyces.

5. The process according to claim 4,
characterized in that
said yeast is Schwanniomyces occidentalis.

6. The process according to any of claims 1 to 5,
characterized in that
said first DNA sequence is comprised by a 1.8 kb BglII-Xhol-fragment, if derived from Schwanniomyces occidentalis
α-amylase gene, or a 1.3 kb BamHI-PvuII fragment if derived from Schwanniomyces occidentalis glucoamylase.
7. The process according to claim 5, characterized in that said first DNA sequence is comprised by part or all of the DNA sequence corresponding to bases -1 to -540 of the region preceding the structural gene coding for α-amylose or part or all of the DNA sequence corresponding to bases -1 to -320 of the region preceding the structural gene coding for glucoamylose.

8. The process according to any of claims 1 to 5, characterized in that said first DNA sequence corresponds to one of the regulons regulating expression of:

   a) ADH1, ADH2, PDC1, GAL1/10, PGK, GAPDH, wherein the regulon is preferably obtained from *Saccharomyces cerevisiae*,

   b) LAC4, wherein the regulon is preferably obtained from *Kluyveromyces lactis*, or

   c) corresponding regulons, preferably obtained from *Schwanniomyces*.

9. The process according to any of claims 1 to 8, characterized in that said second DNA sequence is coding for all or part of at least one of the following peptides:

   a) MetArgPheSerThrGluGlyPheThrSerLysValValAlaAlaIleLeuAlaPheSerArgLeuValSerAla;

   b) MetArgPheSerThrGluGlyPheThrSerLysValValAlaAlaIleLeuAlaPheSerArgLeuValSerAlaGlnProIlePheAspMetArg;

   c) MetIlePheLeuLysLeuIleLysSerIleValIleGlyLeuGlyLeuValSerAla;

   d) MetIlePheLeuLysLeuIleLysSerIleValIleGlyLeuGlyLeuValSerAlaIleGlnAlaAlaProAla;

   e) MetIlePheLeuLysLeuIleLysSerIleValIleGlyLeuGlyLeuValSerAlaIleGlnAlaAlaProAlaSerIleGlySerSerAlaSerAla

10. A process according to claim 9, characterized in that the DNA sequence coding for said peptide corresponds to part or all of one of the following DNA sequences:

   a) ATGAGATTTTGCAAGGATTATGCTAAGTTAGTTGTGCAGCAATT
       TTAGCATTTCAGATTGGTATCTGCT,
b) ATGAGATTTTCAACTGAAGGATTTCACAAGTAAAGTTGTGGCAGCAATTT
   TAGCATTCTCAAAGATTTGATCTGCTCAACCGATTATTTTGTACATGAGA;

c) ATGATTTTTCTGAAGCTGATTAAAAGATTAGTAATTGGTTTGGGATTA
   GTTACTGCT;

d) ATGATTTTTCTGAAGCTGATTAAAAGATTAGTAATTGGTTTGGGATTA
   GTTACTGCTATCCAAGCAGCCCTGCC;

e) ATGATTTTTCTGAAGCTGATTAAAAGATTAGTAATTGGTTTGGGATTA
   GTTACTGCTATCCAAGCAGCCCCTGCTCTTCCGATTTGGATCTAGTGCCTCA
   GCA.

11. The process according to any of claims 1 to 10,
    characterized in that
    said third DNA sequence coding for a foreign protein is a natural or synthetic DNA sequence encoding a cellulase,
    interleukin, insulin like growth factor, interferon, lymphokine, human growth factor, nerve growth factor, aprotinin,
    insulin, hirudin, hormones, blood clotting factors, hepatitis B surface or core antigens, viral or bacterial vaccines
    or human granulocyte/macrophage colony stimulating factor.

12. The process according to any of claims 1 to 11,
    characterized in that
    said fourth DNA sequence originates from a yeast of a member of the genera Saccharomyces, Pichia, Hansenula,
    preferably Schwanniomyces.

13. The process according to claim 12,
    characterized in that
    said fourth DNA sequence corresponds to part or all of the terminator comprised by nucleotides 1537 to 1740 of
    the Schwanniomyces α-amylase gene or part or all of the terminator comprised by nucleotides 2875 to 3320 of
    the Schwanniomyces glucoamylase gene.

14. The process according to any of claims 1 to 13,
    characterized in that
    said first and/or second and/or third and/or fourth DNA sequence has been modified by insertion or deletion or
    substitution of one or more nucleotides while the biological activity of the DNA sequence is retained or improved.

15. The process according to any of claims 1 to 14,
    characterized in that
    said expression cassette is carried by a circular or linear vector, which optionally comprises one or more selective
    marker genes.

16. The process according to claim 15,
    characterized in that
    said vector comprises any of the following selective marker genes: TRP5, LEU2, ADE1, ADE2, HIS3, HIS4, URA3,
    LYS2, preferably obtained from Saccharomyces, Hansenula Pichia or Schwanniomyces, AMY1 or GAM1, preferably
    obtained from Schwanniomyces.

17. The process according to any of claims 15 and 16,
characterized in that
said vector further comprises a DNA sequence capable of controlling autonomous replication and stable maintenance of the vector in the host organism.

18. The process according to claim 17
characterized in that
said DNA sequence is derived from the genome of a member of the genera *Saccharomyces*, *Pichia*, *Hansenula* or *Kluveromyces*, preferably *Schwanniomyces*.

19. A process according to any of claims 17 or 18,
characterized in that
said DNA sequence is selected from the autonomously replicating sequences ARS, preferably an ARS derived from the respective host organism.

20. The process according to claim 19,
characterized in that
said DNA sequence is SwARS1 or SwARS2, capable of controlling autonomous replication and stable maintenance of the vector in *Schwanniomyces*, wherein SwARS1 is comprised on a ClaI-BamHI fragment adjacent to the EcoRI fragment shown in Fig. 2a, and SwARS2 is comprised on an EcoRI-BgII fragment contained in the EcoRI fragment shown in Fig 2a.

21. The process according to any of claims 17 to 20,
characterized in that
said DNA sequence has been modified by deletion, insertion or substitution of one or more nucleotides, while retaining or improving its biological function.

22. The process according to any of claims 1 to 21,
characterized in that
said expression cassette is integrated into the genome, optionally in multiple copies, preferably via homologous recombination.

23. A process for the production of polypeptide, wherein a yeast host organism is cultivated under suitable conditions and the polypeptide is recovered in a manner known per se,
characterized in that
said yeast host organism is a yeast cell according to any of claims 1 to 22.

24. The process according to claim 23
characterized in that
yeast cells of the species *Schwanniomyces occidentalis* are used as a host organism and are preferably cultivated on starch dextrin, maltose and/or plant biomass.

25. The process according to any of claims 23 or 24,
characterized in that
yeast cells of the genus *Schwanniomyces* are used for the production of single cell protein.

26. The process according to any of claims 23 to 25,
characterized in that
said polypeptide is amylase or glucoamylase.

27. Use of a yeast cell according to any of claim 1 to 22 for transformation and expression of foreign genes.
EP 0 394 538 B1

onkassette enthält, umfassend:

a) eine erste DNA-Sequenz, die als Regulon dient,

b) gegebenenfalls eine zweite DNA-Sequenz, die für ein Signalpeptid kodiert,

c) eine dritte DNA-Sequenz, die für ein Fremdprotein kodiert, und

d) gegebenenfalls eine vierte DNA-Sequenz, die als ein Terminator dient,

wobei die erste, zweite und vierte DNA-Sequenz von Hefe abgeleitet sind.

2. Hefezelle nach Anspruch 1, dadurch gekennzeichnet, daß die erste und/oder zweite und/oder vierte DNA-Sequenz von einem für ein amyloytisches Enzym kodierenden Gen abgeleitet sind.

3. Hefezelle nach Anspruch 2, dadurch gekennzeichnet, daß das amyloytische Enzym eine α-Amylase oder Glucoamyrase ist.


5. Hefezelle nach Anspruch 4, dadurch gekennzeichnet, daß die Hefe Schwanniomyces occidentalis ist.

6. Hefezelle nach einem der Ansprüche 1 bis 5, dadurch gekennzeichnet, daß die erste DNA-Sequenz von einem 1,8 kb BglII-XhoI-Fragment umfaßt ist, wenn sie vom Schwanniomyces occidentalis-α-Amylasegen abgeleitet ist, oder von einem 1,3 kb BamHI-PvuII-Fragment, wenn sie von einem Schwanniomyces occidentalis-Glucoamyrasegen abgeleitet ist.


8. Hefezelle nach einem der Ansprüche 1 bis 5, dadurch gekennzeichnet, daß die erste DNA-Sequenz einem der Regulons entspricht, die die Expression regulieren von

   a) ADH1, ADH2, PDC1, GAL1/10, PGK, GAPDH, wobei das Regulon bevorzugt von Saccharomyces cerevisiae erhalten ist,

   b) LAC4, wobei das Regulon bevorzugt von Kluyveromyces lactis erhalten ist, oder

   c) entsprechenden Regulons, bevorzugt aus Schwanniomyces erhalten.

9. Hefezelle nach einem der Ansprüche 1 bis 8, dadurch gekennzeichnet, daß die zweite DNA-Sequenz für die Gesamtheit oder einen Teil mindestens eines der folgenden Peptide kodiert:

   a) MetArgPheSerThrGluGlyPheThrSerLysValValAlaAlaIleLeuAlaPheSerArgLeuValSerAla;

   b) MetArgPheSerThrGluGlyPheThrSerLysValValAlaAlaIleLeuAlaPheSerArgLeuValSerAlaGlnProIleIlePheAspMetArg;
c) MetIlePheLeuLysLeuIleLysSerIleValIleGlyLeuGlyLeu ValSerAla;

d) MetIlePheLeuLysLeuIleLysSerIleValIleGlyLeuGlyLeu ValSerAlaIleGlnAlaAlaProAla;

e) MetIlePheLeuLysLeuIleLysSerIleValIleGlyLeuGlyLeu ValSerAlaIleGlnAlaAlaProAlaSerSerIleGlySerSerAlaSerAla

10. Hefe zelle nach Anspruch 9, dadurch gekennzeichnet, daß die für das Peptid kodierende DNA-Sequenz einem Teil oder der Gesamthheit einer der folgenden DNA-Sequenzen entspricht:

a) ATGAGATTTTCAACTGAAGGATTTTACAAGTAAAGTTGTTTGACACATT TTACGATTCTCAAGATTGGAATTGCTT;

b) ATGAGATTTTCAACTGAAGGATTTTACAAGTAAAGTTGTTTGACACATT TTACGATTCTCAAGATTGGAATTGCTT;

c) ATGATTTTTCTGAAGCTGATTTAAAGATAGTAAATTGTTGGGATTA GTTAGTGC;

d) ATGATTTTTCTGAAGCTGATTTAAAGATAGTAAATTGTTGGGATTA GTTAGTGC ATCAGCGACCCCTGCC;

e) ATGATTTTTCTGAAGCTGATTTAAAGATAGTAAATTGTTGGGATTA GTTAGTGC ATCAGCGACCCCTGCC

11. Hefe zelle nach einem der Ansprüche 1 bis 10, dadurch gekennzeichnet, daß die dritte DNA-Sequenz, die für ein Fremdprotein kodiert, eine natürliche oder synthetische DNA-Sequenz ist, die für eine Zellulase, Interleukin, insulinähnlichen Wachstumsfaktor, Interferon, Lymphokin, menschlichen Wachstumsfaktor, Nervenwachstumsfaktor, Aprotinin, Insulin, Hirudin, Hormone, Blutgerinnungsfaktoren, Hepatitis-B-Oberflächen- oder -Kernantigene, virale oder bakterielle Impfstoffe oder humanen Granulocyten/Makrophagen-Kolonie-stimulierenden Faktor kodiert.


13. Hefe zelle nach Anspruch 12, dadurch gekennzeichnet, daß die vierte DNA-Sequenz einem Teil oder der Gesamt-

14. Hefezelle nach einem der Ansprüche 1 bis 13, dadurch gekennzeichnet, daß die erste und/oder zweite und/oder dritte und/oder vierte DNA-Sequenz durch Insertion oder Deletion oder Substitution eines oder mehrerer Nukleotide modifiziert worden ist, während die biologische Aktivität der DNA-Sequenz beibehalten oder verbessert wird.

15. Hefezelle nach einem der Ansprüche 1 bis 14, dadurch gekennzeichnet, daß die Expressionskassette von einem zirkulären oder linearen Vektor getragen wird, der gegebenenfalls ein oder mehrere selektive Marker gene umfaßt.

16. Hefezelle nach Anspruch 15, dadurch gekennzeichnet, daß der Vektor eines der folgenden selektiven Markergene umfaßt:
   TRP5, LEU2, ADE1, ADE2, HIS3, HIS4, URA3, LYS2, bevorzugt erhalten von Saccharomyces, Hansenula, Pichia oder Schwanniomyces, Amy1 oder GAl1, bevorzugt erhalten von Schwanniomyces.


19. Hefezelle nach einem der Ansprüche 17 oder 18, dadurch gekennzeichnet, daß die DNA-Sequenz aus den autonomen replizierenden Sequenzen ARS, bevorzugt einer ARS, die von dem jeweiligen Wirtsorganismus abgeleitet ist, ausgewählt ist.

20. Hefezelle nach Anspruch 19, dadurch gekennzeichnet, daß die DNA-Sequenz SwARS1 oder SwARS2 ist, die die autonome Replikation und stabile Beibehaltung des Vektors in Schwanniomyces steuern können, wobei SwARS1 von einem ClaI-BamHI-Fragment umfaßt ist, das dem in Figur 2a gezeigten EcoRI-Fragment benachbart ist, und SwARS2 von einem EcoRI-BglII-Fragment umfaßt ist, das in dem in Figur 2a gezeigten EcoRI-Fragment enthalten ist.

21. Hefezelle nach einem der Ansprüche 17 bis 20, dadurch gekennzeichnet, daß die DNA-Sequenz durch Deletion, Insertion oder Substitution eines oder mehrerer Nukleotide modifiziert worden ist, während ihre biologische Funktion beibehalten oder verbessert ist.

22. Hefezelle nach einem der Ansprüche 1 bis 21, dadurch gekennzeichnet, daß die Expressionskassette in das Genom integriert wird, gegebenenfalls in mehreren Kopien, bevorzugt über homologe Rekombination.

23. Verfahren zum Erzeugen von Polypeptid, wobei ein Hefewirtsorganismus unter geeigneten Bedingungen kultiviert und das Polypeptid in an sich bekannter Weise gewonnen wird, dadurch gekennzeichnet, daß der Hefewirtsorganismus eine Hefezelle gemäß einem der Ansprüche 1 bis 22 ist.


27. Verwendung einer Hefezelle nach einem der Ansprüche 1 bis 22 zur Transformation und Expression von Fremdgenen.
Patentansprüche für folgenden Vertragsstaat: ES

1. Verfahren zum Herstellen einer Hefezelle, die Fremdprotein exprimieren kann, dadurch gekennzeichnet, daß mindestens eine Expressionskassette in einer Hefezelle der Gattung Schwanniomyces inseriert wird, wobei die Expressionskassette umfaßt:
   a) eine erste DNA-Sequenz, die als Regulon dient,
   b) gegebenenfalls eine zweite DNA-Sequenz, die für ein Signalpeptid kodiert,
   c) eine dritte DNA-Sequenz, die für ein Fremdprotein kodiert, und
   d) gegebenenfalls eine vierte DNA-Sequenz, die als Terminator dient,

wobei die erste, zweite und vierte DNA-Sequenz von Hefe abgeleitet sind.

2. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß die erste und/oder zweite und/oder vierte DNA-Sequenz von einem für ein amylolysches Enzym kodierenden Gen abgeleitet sind.

3. Verfahren nach Anspruch 2, dadurch gekennzeichnet, daß das amylolysche Enzym eine α-Amylase oder Glucoamylase ist.


5. Verfahren nach Anspruch 4, dadurch gekennzeichnet, daß die Hefe Schwanniomyces occidentalis ist.

6. Verfahren nach einem der Ansprüche 1 bis 5, dadurch gekennzeichnet, daß die erste DNA-Sequenz von einem 1.8 kb Bgill-Xhol-Fragment umfaßt ist, wenn sie vom Schwanniomyces occidentalis-α-Amylasegen abgeleitet ist, oder von einem 1.3 kb BamHI-PvuII-Fragment umfaßt ist, wenn sie vom Schwanniomyces occidentalis-Glucoamylasegen abgeleitet ist.


8. Verfahren nach einem der Ansprüche 1 bis 5, dadurch gekennzeichnet, daß die erste DNA-Sequenz einem der Regulons entspricht, die die Expression regulieren von
   a) ADH1, ADH2, PDC1, GAL1/10, PGK, GAPDH, wobei das Regulon bevorzugt von Saccharomyces cerevisiae erhalten ist,
   b) LAC4, wobei das Regulon bevorzugt von Kluyveromyces lactis erhalten ist, oder
   c) korrespondierenden Regulons, bevorzugt von Schwanniomyces erhalten.

9. Verfahren nach einem der Ansprüche 1 bis 8, dadurch gekennzeichnet, daß die zweite DNA-Sequenz für die Gesamtheit oder einen Teil mindestens eines der folgenden Peptide kodiert:

   a) MetArgPheSerThrGluGlyPheThrSerLysValValAlaAlaIle
      LeuAlaPheSerArgLeuValSerAla;
b) MetArgPheSerThrGluGlyPheThrSerLysValValAlaAlaIleLeuAlaPheSerArgLeuValSerAlaGlnProIleIlePheAspMetArg;

c) MetIlePheLeuLysLeuIleLysSerIleValIleGlyLeuGlyLeuValSerAla;

d) MetIlePheLeuLysLeuIleLysSerIleValIleGlyLeuGlyLeuValSerAlaIleGlnAlaAlaProAla;

e) MetIlePheLeuLysLeuIleLysSerIleValIleGlyLeuGlyLeuValSerAlaIleGlnAlaAlaProAlaSerIleGlySerSerAlaSerAla

10. Verfahren nach Anspruch 9, dadurch gekennzeichnet, daß die für das Peptid kodierende DNA-Sequenz einem Teil oder der Gesamtheit einer der folgenden DNA-Sequenzen entspricht:

a) ATGAGATTTTCAACTGAAGGGATTTTACAAAGTAAAAGTTGTTGCGAGCAATT
TTAGCATTCTCAAGATTTGTATCTGCT,

b) ATGAGATTTTCAACTGAAGGAGGATTTTACAAAGTAAAAGTTGTTGCGAGGAATTT
TAGCATTCTCAAGATTTGTATCTGCTCAACCGATTATTTTGACATGAGA;

c) ATGATTTTTTCTGAAGCTGATTTAAAAAGTATAGTAATTTGTTTTGGGATTA
GGTAGTGCT;

d) ATGATTTTTCTGAAGCTGATTTAAAAAGTATAGTAATTTGTTTTGGGATTA
GGTAGTCTATCCAAGCAGCCCTGCC;

e) ATGATTTTTCTGAAGCTGATTTAAAAAGTATAGTAATTTGTTTTGGGATTA
GGTAGTCTATCCAAGCAGCCCTGCCCTCTTGATTTGCTATGCTGCTGCA.

11. Verfahren nach einem der Ansprüche 1 bis 10, dadurch gekennzeichnet, daß die dritte DNA-Sequenz, die für ein Fremdprotein kodiert, eine natürliche oder synthetische DNA-Sequenz ist, die für eine Zellulase, Interleukin, insulinnählichen Wachstumsfaktor, Interferon, Lymphokin, menschlichen Wachstumsfaktor, Nervenwachstumsfaktor,
Aprotinin, Insulin, Hirudin, Hormone, Blutgerinnungsfaktoren, Hepatitis-B-Oberflächen- oder -Kernantigene, virale oder bakterielle Impfstoffe oder humanen Granulozyten/Makrophagen-Kolonie-stimulierenden Faktor kodiert.


14. Verfahren nach einem der Ansprüche 1 bis 13, dadurch gekennzeichnet, daß die erste und/oder zweite und/oder dritte und/oder vierte DNA-Sequenz durch Insertion oder Deletion oder Substitution eines oder mehrerer Nukleotideme modifiziert worden ist, während die biologische Aktivität der DNA-Sequenz beibehalten oder verbessert wird.

15. Verfahren nach einem der Ansprüche 1 bis 14, dadurch gekennzeichnet, daß die Expressionskassette von einem zirkulären oder linearen Vektor getragen wird, der gegebenenfalls ein oder mehrere selektive Markergene umfaßt.

16. Verfahren nach Anspruch 15, dadurch gekennzeichnet, daß der Vektor eines der folgenden selektiven Markergene umfaßt:

TRP1, LEU2, ADE1, ADE2, HIS3, HIS4, URA3, LYS2, bevorzugt erhalten von Saccharomyces, Hansenula, Pichia oder Schwanniomyces, AMY1 oder GAM1, bevorzugt erhalten von Schwanniomyces.


19. Verfahren nach einem der Ansprüche 17 oder 18, dadurch gekennzeichnet, daß die DNA-Sequenz aus den autonomen replizierenden Sequenzen ARS, bevorzugt einer ARS, die von dem jeweiligen Wirtsorganismus abgeleitet ist, ausgewählt wird.

20. Verfahren nach Anspruch 19, dadurch gekennzeichnet, daß die DNA-Sequenz SwARS1 oder SwARS2 ist, die die autonome Replikation und stabile Beibehaltung des Vektors in Schwanniomyces steuern können, wobei SwARS1 von einem Clal-BamHI-Fragment umfaßt ist, das dem in Figur 2a gezeigten EcoRI-Fragment benachbart ist, und SwARS2 von einem EcoRI-BglII-Fragment umfaßt ist, das dem in Figur 2a gezeigten EcoRI-Fragment enthalten ist.

21. Verfahren nach einem der Ansprüche 17 bis 20, dadurch gekennzeichnet, daß die DNA-Sequenz durch Deletion, Insertion oder Substitution eines oder mehrerer Nukleotide modifiziert worden ist, während ihre biologische Funktion beibehalten oder verbessert wird.

22. Verfahren nach einem der Ansprüche 1 bis 21, dadurch gekennzeichnet, daß die Expressionskassette in das Genom integriert ist, gegebenenfalls in mehrfachen Kopien, bevorzugt über homologe Rekombination.

23. Verfahren zum Erzeugen eines Polypeptides, wobei ein Hefewirtsorganismus unter geeigneten Bedingungen kul- tiviert und das Polypeptid in an sich bekannter Weise gewonnen wird, dadurch gekennzeichnet, daß der Hefe- wirtsorganismus eine Hefazelle gemäß einem der Ansprüche 1 bis 22 ist.


27. Verwendung einer Hefezelle gemäß einem der Ansprüche 1 bis 22 zur Transformation und Expression von Fremdgenen.

Revendications

Revendications pour les États contractants suivants : AT, BE, CH, LI, DE, FR, GB, GR, IT, LU, NL, SE

1. Cellule de levure du genre Schwanniomyces, caractérisée en ce que ladite cellule de levure contient au moins une cassette d'expression comprenant :
   a) une première séquence d'ADN servant de régulon,
   b) facultativement, une deuxième séquence d'ADN codant pour un peptide signal
   c) une troisième séquence d'ADN codant pour une protéine étrangère, et
   d) facultativement, une quatrième séquence d'ADN servant de terminator,

les première, deuxième et quatrième séquences d'ADN étant obtenues à partir d'une levure.

2. Cellule de levure selon la revendication 1, caractérisée en ce que lesdites première et/ou deuxième et/ou quatrième séquences d'ADN sont obtenues à partir d'un gène codant pour une enzyme amylolytique.

3. Cellule de levure selon la revendication 2, caractérisée en ce que ladite enzyme amylolytique est une α-amylase ou la glucoamylase.

4. Cellule de levure selon l'une quelconque des revendications 1 à 3, caractérisée en ce que lesdites première et/ou deuxième et/ou quatrième séquences d'ADN sont obtenues à partir d'une levure d'un élément des genres Debamomyces, Saccharomycopsis, Lipomyces, Pichia ou Saccharomyces, de préférence Schwanniomyces.

5. Cellule de levure selon la revendication 4, caractérisée en ce que ladite levure est Schwanniomyces occidentalis.

6. Cellule de levure selon l'une quelconque des revendications 1 à 5, caractérisée en ce que ladite première séquence d'ADN est constituée d'un fragment BglII-XhoI de 1,8 kb, si elle est obtenue à partir du gène de la α-amylase de Schwanniomyces occidentalis, ou d'un fragment BamHI-PvuII de 1,3 kb si elle est obtenue à partir du gène de la glucoamylase de Schwanniomyces occidentalis.

7. Cellule de levure selon la revendication 5, caractérisée en ce que ladite première séquence d'ADN est constituée en partie ou en totalité de la séquence d'ADN correspondant aux bases -1 à -540 de la région précédant le gène structural codant pour la α-amylase, ou en partie ou en totalité de la séquence d'ADN correspondant aux bases -1 à -320 de la région précédant le gène structural codant pour la glucoamylase.

8. Cellule de levure selon l'une quelconque des revendications 1 à 5, caractérisée en ce que ladite première séquence d'ADN correspond à l'un des régulons régulant l'expression :
   a) de ADH1, ADH2, PDC1, GAL110, PGK, GAPDH, le régulon étant de préférence obtenu à partir de Saccharomyces cerevisiae,
   b) de LAC4, le régulon étant de préférence obtenu à partir de Kluyveromyces lactis, ou
   c) de régulons correspondants, obtenus de préférence à partir de Schwanniomyces.

9. Cellule de levure selon l'une quelconque des revendications 1 à 8, caractérisée en ce que ladite deuxième séquence d'ADN code pour tout ou partie d'au moins l'un des peptides suivants :
a) MetArgPheSerThrGluGlyPheThrSerLysValValAlaAlaIleLeuAlaPheSerArgLeuValValSerAla;

b) MetArgPheSerThrGluGlyPheThrSerLysValValAlaAlaIleLeuAlaPheSerArgLeuValValSerAlaGlnProIleIlePheAspMetArg;

c) MetIlePheLeuLysLeuIleLysSerIleValIleGlyLeuGlyLeuValValSerAla;

d) MetIlePheLeuLysLeuIleLysSerIleValIleGlyLeuGlyLeuValValSerAlaIleGlnAlaAlaProAla;

e) MetIlePheLeuLysLeuIleLysSerIleValIleGlyLeuGlyLeuValValSerAlaIleGlnAlaAlaProAlaSerSerIleGlySerSerAlaSerAla

10. Cellule de levure selon la revendication 9, caractérisée en ce que la séquence d'ADN codant pour l'édit peptide correspond à une partie ou à la totalité des séquences d'ADN suivantes :

a) ATGAGATTTTTCAACTGAAGGATTTTCAAGTAAAGTTGTTGAGCAATT
    TTAGCATTCTCAAGATTTGATCTGCT,

b) ATGAGATTTTTCAACTGAAGGATTTTCAAGTAAAGTTGTTGAGCAATT
    TTAGCATTCTCAAGATTTGATCTGCTCAACCAGATTATTTTGACATGAGA;

c) ATGATTTTTCTGAAGCAGTTAAAAGTATAAGTAATTGGTGGGATTA
    GTTAGTGCT;

d) ATGATTTTTTCTGAAGCTGATTAAAAGTATAAGTAATTGGTGGGATTA
    GTTAGTGCTATCCAAGCACGGCCTGCC;

e) ATGATTTTTTCTGAAGCTGATTAAAAGTATAAGTAATTGGTGGGATTA
    GTTAGTGCTATCCAAGCACGGCCTGCC

11. Cellule de levure selon l'une quelconque des revendications 1 à 10, caractérisée en ce que l'adite troisième séquence d'ADN codant pour une protéine étrangère est une séquence d'ADN naturelle ou synthétique codant
pour une celluase, une interleukine, un facteur de croissance semblable à l'insuline, un interféron, la lymphokine, un facteur de croissance humain, un facteur de croissance des nerfs, l'aprotinine, l'insuline, l'hirudine, les hormones, les facteurs de coagulation du sang, les antigènes de surface ou de coeur de l'hépatite B, les vaccins viraux ou bactériens, ou le facteur stimulant une colonie de granulocytes/macrophages humain.

12. Cellule de levure selon l'une quelconque des revendications 1 à 11, caractérisée en ce que ladite quatrième séquence d'ADN provient d'une levure d'un élément des genres Saccharomyces, Pichia, Hansenula, de préférence Schwanniomyces.

13. Cellule de levure selon la revendication 12, caractérisée en ce que ladite quatrième séquence d'ADN correspond à une partie ou à la totalité du terminateur constitué des nucléotides 1537 à 1740 du gène de la α-amylose de Schwanniomyces, ou à une partie ou à la totalité du terminateur constitué des nucléotides 2875 à 3320 du gène de la glucoamylase de Schwanniomyces.

14. Cellule de levure selon l'une quelconque des revendications 1 à 13, caractérisée en ce que lesdites première et/ou deuxième et/ou troisième et/ou quatrième séquences d'ADN ont été modifiées en insérant ou en délétant ou en substituant un ou plusieurs nucléotides tout en retenant ou en améliorant l'activité biologique de la séquence d'ADN.

15. Cellule de levure selon l'une quelconque des revendications 1 à 14, caractérisée en ce que ladite cassette d'expression est portée par un vecteur circulaire ou linéaire, qui comprend facultativement un ou plusieurs gènes de marqueur de sélection.

16. Cellule de levure selon la revendication 15, caractérisée en ce que ledit vecteur comprend l'un quelconque des gènes de marqueur de sélection suivants : TRP5, LEU2, ADE1, ADE2, HIS3, HIS4, URA3, LYS2, obtenus de préférence à partir de Saccharomyces, Hansenula, Pichia ou Schwanniomyces, AMY1 ou GAM1, obtenus de préférence à partir de Schwanniomyces.

17. Cellule de levure selon l'une quelconque des revendications 15 à 16, caractérisée en ce que ledit vecteur comprend de plus une séquence d'ADN capable de contrôler la réplication autonome et le maintien stable du vecteur dans l'organisme hôte.

18. Cellule de levure selon la revendication 17, caractérisée en ce que ladite séquence d'ADN est obtenue à partir du génome d'un élément des genres Saccharomyces, Pichia, Hansenula ou Kluyveromyces, de préférence Schwanniomyces.

19. Cellule de levure selon l'une quelconque des revendications 17 et 18, caractérisée en ce que ladite séquence d'ADN est choisie parmi les séquences se répliquant de façon autonome ARS, de préférence une ARS obtenue à partir de l'organisme hôte respectif.

20. Cellule de levure selon la revendication 19, caractérisée en ce que ladite séquence d'ADN est SwARS1 ou SwARS2, capable de contrôler la réplication autonome et le maintien stable du vecteur dans Schwanniomyces, dans laquelle SwARS1 est constituée d'un fragment ClaI/BamHI adjacent au fragment EcoRI présenté à la figure 2a, et SwARS2 est constituée d'un fragment EcoRI/BglII contenu dans le fragment EcoRI présenté à la figure 2a.

21. Cellule de levure selon l'une quelconque des revendications 17 à 20, caractérisée en ce que ladite séquence d'ADN a été modifiée en délétant, insérant ou substituant un ou plusieurs nucléotides, tout en retenant ou en améliorant sa fonction biologique.

22. Cellule de levure selon l'une quelconque des revendications 1 à 21, caractérisée en ce que ladite cassette d'expression est intégrée dans le génome, facultativement en plusieurs copies, de préférence par recombinaison homologue.

23. Procédé pour la préparation d'un polypeptide, dans lequel un organisme hôte de levure est cultivé dans des conditions appropriées et le polypeptide est récupéré d'une manière connue d'elle-même, caractérisé en ce que ledit organisme hôte de levure est une cellule de levure selon l'une quelconque des revendications 1 à 22.

24. Procédé selon la revendication 23, caractérisé en ce que des cellules de levure de l'espèce Schwanniomyces.
occidentalis sont utilisées en tant qu'organisme hôte et sont de préférence cultivées sur de l'amidon, de la dextrine, du maltose et/ou de la biomasse végétale.

25. Procédé selon l'une quelconque des revendications 23 et 24, *caractérisé en ce que* des cellules de levure du genre Schwanniomyces sont utilisées pour la production d'une seule protéine de cellule.

26. Procédé selon l'une quelconque des revendications 23 à 25, *caractérisé en ce que* ledit polypeptide est l'amylase ou la glucoamylase.

27. Utilisation d'une cellule de levure selon l'une quelconque des revendications 1 à 22 pour la transformation et l'expression de gènes étrangers.

Recommandations pour l'État contractant suivant: ES

1. Procédé pour préparer une cellule de levure capable d'exprimer une protéine étrangère, *caractérisé en ce qu'* au moins une cassette d'expression comprenant :
   a) une première séquence d'ADN servant de régulon,
   b) facultativement, une deuxième séquence d'ADN codant pour un peptide signal
   c) une troisième séquence d'ADN codant pour une protéine étrangère, et
   d) facultativement, une quatrième séquence d'ADN servant de terminator,

   est insérée dans une cellule de levure du genre Schwanniomyces, les première, deuxième et quatrième séquences d'ADN étant obtenues à partir d'une levure.

2. Procédé selon la revendication 1, *caractérisé en ce que* lesdites première et/ou deuxième et/ou quatrième séquences d'ADN sont obtenues à partir d'un gène codant pour une enzyme amylolytique.

3. Procédé selon la revendication 2, *caractérisé en ce que* ladite enzyme amylolytique est une α-amylase ou la glucoamylase.

4. Procédé selon l'une quelconque des revendications 1 à 3, *caractérisé en ce que* lesdites première et/ou deuxième et/ou quatrième séquences d'ADN sont obtenues à partir d'une levure d'un élément des genres Debaryomyces, Saccharomyces, Lipomyces, Pichia ou Saccharomyces, de préférence Schwanniomyces.

5. Procédé selon la revendication 4, *caractérisé en ce que* ladite levure est Schwanniomyces occidentalis.

6. Procédé selon l'une quelconque des revendications 1 à 5, *caractérisé en ce que* ladite première séquence d'ADN est constituée d'un fragment BglI-Hoxcl de 1.8 kb, si elle est obtenue à partir du gène de la α-amylase de Schwanniomyces occidentalis, ou d'un fragment BamHI-PvuI de 1.3 kb si elle est obtenue à partir du gène de la glucoamylose de Schwanniomyces occidentalis.

7. Procédé selon la revendication 5, *caractérisé en ce que* ladite première séquence d'ADN est constituée en partie ou en totalité de la séquence d'ADN correspondant aux bases -1 à -540 de la région précédant le gène structural codant pour la α-amylase, ou en partie ou en totalité de la séquence d'ADN correspondant aux bases -1 à -320 de la région précédant le gène structural codant pour la glucoamylose.

8. Procédé selon l'une quelconque des revendications 1 à 5, *caractérisé en ce que* ladite première séquence d'ADN correspond à l'un des régulons régulant l'expression :
   a) de ADH1, ADH2, PDC1, GAL1/10, PGK, GAPDH, le régulon étant de préférence obtenu à partir de Saccharomyces cerevisiae,
   b) de LAC4, le régulon étant de préférence obtenu à partir de Kluyveromyces lactis, ou
   c) de régulons correspondants, obtenus de préférence à partir de Schwanniomyces.

9. Procédé selon l'une quelconque des revendications 1 à 8, *caractérisé en ce que* ladite deuxième séquence d'ADN code pour tout ou partie d'au moins l'un des peptides suivants ;
a) MetArgPheSerThrGluGlyPheThrSerLysValValAlaAlaIleLeuAlaPheSerArgLeuValSerAla;

b) MetArgPheSerThrGluGlyPheThrSerLysValValAlaAlaIleLeuAlaPheSerArgLeuValSerAlaGlnProIleIlePheAspMetArg;

c) MetIlePheLeuLysLeuIleLysSerIleValIleGlyLeuGlyLeuValSerAla;

d) MetIlePheLeuLysLeuIleLysSerIleValIleGlyLeuGlyLeuValSerAlaIleGlnAlaAlaAlaProAla;

e) MetIlePheLeuLysLeuIleLysSerIleValIleGlyLeuGlyLeuValSerAlaIleGlnAlaAlaAlaProAlaSerIleGlySerSerAlaSerAla

10. Procédé selon la revendication 9, caractérisé en ce que la séquence d’ADN codant pour ledit peptide correspond à une partie ou à la totalité d’une des séquences d’ADN suivantes :

a) ATGAGATTTTCGACTGAGGATTATTTACAAAGTAAAGTTGGCTGAGCAATT TTAGCATTCTCAAGATTTGCTATCTGCT,

b) ATGAGATTTTCGACTGAGGATTATTTACAAAGTAAAGTTGGCTGAGCAATT TTAGCATTCTCAAGATTTGCTATCTGCTCAACGATTATTTTGTACAGA;

c) ATGATTTTTTCGAAGCTGATTTAAAAGTAGTAAATGGTGGGGGATTA GTTAGTGCT;

d) ATGATTTTTTCGAAGCTGATTTAAAAGTAGTAAATGGTGGGGGATTA GTTAGTGCTATCCAAGCAGCCCTGCC;

e) ATGATTTTTTCGAAGCTGATTTAAAAGTAGTAAATGGTGGGGGATTA GTTAGTGCTATCCAAGCAGCCCTGCC.
11. Procédé selon l'une quelconque des revendications 1 à 10, caractérisé en ce que ladite troisième séquence d'ADN codant pour une protéine étrangère est une séquences d'ADN naturelle ou synthétique codant pour une cellulose, une interleukine, un facteur de croissance semblable à l'insuline, un interféron, la lymphokine, un facteur de croissance humain, un facteur de croissance des nerfs, l'aprotinine, l'insuline, l'hirudine, les hormones, les facteurs de coagulation du sang, les antigènes de surface ou de coeur de l'hépatite B, les vaccins vitaux ou bactériens, ou le facteur stimulant une colonie de granulocytes/macrophages humain.

12. Procédé selon l'une quelconque des revendications 1 à 11, caractérisé en ce que ladite quatrième séquence d'ADN provient d'une levure d'un élément des genres Saccharomyces, Pichia, Hansenula, de préférence Schwanniomyces.

13. Procédé selon la revendication 12, caractérisé en ce que ladite quatrième séquence d'ADN correspond à une partie ou à la totalité du terminateur constitué des nucléotides 1537 à 1740 du gène de la α-amylase de Schwanniomyces, ou à une partie ou à la totalité du terminateur constitué des nucléotides 2875 à 3320 du gène de la glucoamylase de Schwanniomyces.

14. Procédé selon l'une quelconque des revendications 1 à 13, caractérisé en ce que les dites première et/ou deuxième et/ou troisième et/ou quatrième séquences d'ADN ont été modifiées en insérant ou en délimitant ou en substituant un ou plusieurs nucléotides tout en retenant ou en améliorant l'activité biologique de la séquence d'ADN.

15. Procédé selon l'une quelconque des revendications 1 à 14, caractérisé en ce que ladite cassette d'expression est portée par un vecteur circulaire ou linéaire, qui comprend facultativement un ou plusieurs gènes de marqueur de sélection.

16. Procédé selon la revendication 15, caractérisé en ce que le dit vecteur comprend l'un quelconque des gènes de marqueur de sélection suivants : TRP5, LEU2, Ade1, ADE2, HIS3, HIS4, URA3, LYS2, obtenus de préférence à partir de Saccharomyces, Hansenula, Pichia ou Schwanniomyces, AMY1 ou GAM1, obtenus de préférence à partir de Schwanniomyces.

17. Procédé selon l'une quelconque des revendications 15 à 16, caractérisé en ce que le dit vecteur comprend de plus une séquence d'ADN capable de contrôler la réplication autonome et le maintien stable du vecteur dans l'organisme hôte.

18. Procédé selon la revendication 17, caractérisé en ce que ladite séquence d'ADN est obtenue à partir du génome d'un élément des genres Saccharomyces, Pichia, Hansenula ou Kluyveromyces, de préférence Schwanniomyces.

19. Procédé selon l'une quelconque des revendications 17 et 18, caractérisé en ce que ladite séquence d'ADN est choisie parmi les séquences se répliquant de façon autonome ARS, de préférence une ARS obtenue à partir de l'organisme hôte respectif.

20. Procédé selon la revendication 19, caractérisé en ce que ladite séquence d'ADN est SwARS1 ou SwARS2, capable de contrôler la réplication autonome et le maintien stable du vecteur dans Schwanniomyces, dans laquelle SwARS1 est constituée d'un fragment Clal-BamHI adjacent au fragment EcoRI présenté à la figure 2a, et SwARS2 est constituée d'un fragment EcoRI-BglII contenu dans le fragment EcoRI présenté à la figure 2a.

21. Procédé selon l'une quelconque des revendications 17 à 20, caractérisé en ce que ladite séquence d'ADN a été modifiée en délimitant, insérant ou substituant un ou plusieurs nucléotides, tout en retenant en améliorant sa fonction biologique.

22. Procédé selon l'une quelconque des revendications 1 à 21, caractérisé en ce que ladite cassette d'expression est intégrée dans le génome, facultativement en plusieurs copies, de préférence par recombinaison homologue.

23. Procédé pour la préparation d'un polypeptide, dans lequel un organisme hôte de levure est cultivé dans des conditions appropriées et le polypeptide est récupéré d'une manière connue d'elle-même, caractérisé en ce que ledit organisme hôte de levure est une cellule de levure selon l'une quelconque des revendications 1 à 22.

24. Procédé selon la revendication 23, caractérisé en ce que des cellules de levure de l'espèce Schwanniomyces occidentalis sont utilisées en tant qu'organisme hôte et sont de préférence cultivées sur de l'amidon, de la dextrine,
du maltose et/ou de la biomasse végétale.

25. Procédé selon l'une quelconque des revendications 23 et 24, caractérisé en ce que des cellules de levure du genre Schwanniomyces sont utilisées pour la production d'une seule protéine de cellule.

26. Procédé selon l'une quelconque des revendications 23 à 25, caractérisé en ce que l'édit polypeptide est l'amylase ou la glucoamylase.

27. Utilisation d'une cellule de levure selon l'une quelconque des revendications 1 à 22 pour la transformation et l'expression de gènes étrangers.
-540
GCTAAATTTA GAAACCGGCTA TAGATCCGGCT TGTCTAAAGA AGAGATAATG AAGAAAAACAA
-480
TTAACCGAGCC ACTCTTATTA AGTTTTTTTC TATTATTTT TGCTCCTACT TCAATAATT
-420
ATCTAAATTB TATTGTGCAT TAGATCAGAA TGTACTGATA ACAGAGAATA TTATCATACA
-360
CTTGTGGATT TCAAAAGGCG BAATCAAAAQ CATACGTAGT CAAACCCCTTG GTTATTGGAT
-300
GCAATTAAGG TTGTAOTCGT TCTTACCGAT CCAATCAATT ACACCACACG GTTTCATGGT
-240
ATGTAGGTGT TTCAATAAGTB AAATACAATB AATTGTTTGB TAATGCTGTA TGTTGATCAG
-180
TAATTATGTT AAACAAATTAA GTCTGAAAAT TTATTAATA TTATACCTACA AATTAAGCCG
-120
AAATCCAATC GAAGGTGCCG CCCAGCTGGT GTATAAAATTA CTCTTGAAT TCAAAGTTGAA
-60
CGTTGATCTC TCTAAAGCAA AGCTGTATT TTACAATACT AAATAAATA AAAGCAAGAC
-1
FIG. 1B, 1
ATGAGATTTTCAACTGAGGATTTTACAGAAGGTGTGGTGCAGCAATTTTACAGCTCA
MetArgPheSerThrGluGlyPheThrSerLysValValAlaAlaIleLeuAlaPheSer
10  20  30  40  50  60
AGATTGGTATCTGCTCAACCCGATTATTTTTGACATAGAGATGTAGCTCGTCAGCTGAT
ArgLeuValSerAlaGlnProIleIlePheAspMetArgAspValSerSerSerAlaAsp
70  80  90  100  110  120
AAATGGAAAGACCAATCGATTATCAATCGTTACTGATAGGTTTGGCCAGATCTGATGGC
LysTrpLysAspGlnSerIleTyrGlnIleValThrAspArgPheAlaArgSerAspGly
130  140  150  160  170  180
TCGACCACAGCTGACTGTATATAGTGATGTCAGCAAGTACTGTTGGTGGATCTTTATAAAGGG
SerThrThrAlaAspCysLeuValSerAspArgLysTyrCysGlyGlySerTyrLysGly
190  200  210  220  230  240
ATTATCGACAAAGTGATATATATCAAGGATGATGGTTTACTGCACTGAGATCTCCCCCA
IleIleAspLysLeuAspTyrIleGlnGlyMetGlyPheThrAlaIleTrpIleSerPro
250  260  270  280  290  300
310 320 330 340 350 360
GTTGTGACAAAATTCCTGACAAATCCTGCTTTGTTATGCTTACCATGGTTATTGGATG
ValValGluGlnIleProAspAsnThrAlaTyrGlyTyrAlaTyrHisGlyTyrTrpMet

370 380 390 400 410 420
AAAAATATGTGAATGGACACTAATTATTTTGTTACCGCTGATGAATTGAAACAATTAGCT
LysAsnIleAspGluLeuAsnThrAsnPheGlyThrAlaAspGluLeuLysGlnLeuAla

430 440 450 460 470 480
ACGCAATTGCATTCCAGAACGATGTTATGTGTGCTGACGTGTTACCAAACTTATGCT
SerGluLeuHisSerArgSerMetLeuLeuMetValAspValValTyrAsnHisTyrAla

490 500 510 520 530 540
TGGAACGGAGATGTTCAAGCGTAGATTATTCTTAGTTCACTCCATCAAACATCT
TrpAsnGlyAspGlySerSerValAspTyrSerSerPheThrProPheAsnGlnGlnSer

550 560 570 580 590 600
TACTTCCCCAGATTATTGGATTACCAAAATTATAATGATAAAACCAATGGTTGAAGATGTT
TyrPheHisAspTyrCysLeuIleThrAsnTyrAsnAspGlnThrAsnValGluAspCys

610 620 630 640 650 660
TGGGAAGGTGAACTGACGTTCCCTCCAGATTATTGAAGTACCAGGATAATGAAGTTATA
TrpGluGlyAspThrGluValSerLeuProAspLeuSerThrGluAspAsnGluValIle
670 680 690 700 710 720
GGAGTATTTCAAACTTGGGTGTCAGATTGGTCTCAAAACTTTATATTTAATCAATCGATGGTTAAGA
GlyValPheGlnThrTrpValSerAspPheValGlnAsnTyrSerIleAspGlyLeuArg

730 740 750 760 770 780
ATTGATAGTGCAAAGCGAGGTGATCCACCGTCTTCAATTTACGGCAATAGTTGGAGGCAGCTTCTGGTA
IleAspSerAlaLysHisValAspThrAlaSerLeuThrLysPheGluAspAlaSerGly

790 800 810 820 830 840
TTTATAACTTTAGGGTGTATTATCAAGAGATCCCAACTTTATATTTAGTTGCTCCATATCAAGAT
ValTyrAsnLeuGlyGluValTyrGlnGlyAspProThrTyrThrCysProTyrGlnAsn

850 860 870 880 890 900
TATATGAAAGGAGTTACCAAACTATATCCATTATATACCAGATATTATAGTTGCTCCATATGAT
TyrMetLysGlyValThrAsnTyrProLeuTyrTyrProValTyrArgPhePheSerAsp

910 920 930 940 950 960
ACTTCGGCGAATTCGTAAGTTAATTTCAATGATCCACCTGAGATTGCCATATTGCTCG
ThrSerAlaThrSerSerGluLeuThrSerMetIleSerThrLeuGlnSerSerCysSer

970 980 990 1000 1010 1020
GACGTTCTTTTGGGAGACTTTATGAAAACCATGACCAAGTTGGATTTTCCATACGTT
\spValSerLeuLeuGlyAsnPheIleGluAsnHisAspGlnValArgPheProSerVal

1030 1040 1050 1060 1070 1080
ACCTCAGACACATCTTTGATTTTGAGATGTGTTTATTTATTTGGGTTGATGTTACT
ThrSerAspThrSerLeuIleLysAsnAspMetAlaPheIleIleLeuGlyAspGlyIle
1510 1520 1530 1540 1550 1560
TCCTCTGTTTTTCGGGATCTGGCATCTGCAATCAATAGATTGATCCAGCGCTAACCCCT
SerSerValLeuSerGlySerGlyIleCysAsnGln

1570 1580 1590 1600 1610 1620
TTTTTAGCAACGACAAGTTTATTTTAGAAAAAGTTTTTCTAAGAATGGTCAAAAACAAGTTC

1630 1640 1650 1660 1670 1680
TTATTACTTCTATGTCCTGGATATCTGTGTTCAATGTTCTGACTCCACATTCCCTCAT

1690 1700 1710 1720 1730 1740
GTTTAGTTCTCTATTTTTGTGATCTTTCTAAGTTTTTTTTATTTCTTAATTTAATCCAAA
FIG. 2B

-310
AGATCTACATTTAAACCCCA

-250
TCTACTCCAGATATTGGAGTTATATAACCCCAATTCTTACCCTTATATCCATGACCCCGCATCBA

-200
AATTTTCAAAGGBATTTCGABGAAATTCTTTTCCTAAAAATACGAAGTGTATTGTGATTCA

-150
ATTACTACGGAACCTACTCATATGGTATAGTGATGTGATGACTGCAATTGTATTTT

-100
GCAGAAGTTATAGTAATAGTTTGCAAAACTGGAGAAATTTTTTCAATTATTGGGAATAATATAAA

-50
TAAAGGCAAGTATCCATTGAAATTAAAAATGAACACTGACTGTATTATACAAACAGCAAG

-1

CCACTGGTGGAAGGTGATGCACAATCTTAATACCTCGACAATTCCGATTTGATTTCGAA
ProLeuValGluGlyAspAlaGlnSerTyrAsnPheAspAsnSerAspLeuValPheGlu

TACTCTAATACGTACTCTCCTTTGAAGTAGTTAGATCATCTACTAAAGAAGGTTTTTTAAA
TyrSerAsnThrAspPheSerPheGluValIleArgSerSerThrLysGluValLeuPhe

TCTACTAAAGGTAAATCCATTGGTTTTTTCAATCAAATTCAATTCAATTCGTCATTG
SerThrLysGlyAsnProLeuValPheSerAsnGlnPheIleGlnPheAsnSerSerLeu

CCAAAGAACCATTATATTACCTGGCTTTGTTGGAATCTATTTACGGATTGTAAACGACCA
ProLysAsnHisValIleThrGlyLeuGlyGluSerIleHisGlyLeuValAsnGluPro

GGTAGGCTTTAAAACATTATTTGCTAATGATTTGTTGATCAACTCGATGTTATAATTAT
GlySerValLysThrLeuPheAlaAsnAspValGlyAspProIleAspGlyAsnIleTyr

GGTGTCATCCAGTATTTATCTTGATCAAGAGATATGACACTGAAACTACCCATGCTTTAT
GlyValHisProValTyrLeuAspGlnArgTyrAspThrGluThrHisAlaValTyr

TGGAGAACCTCTGCTATTCAAGAGATTATTAATCTCGGGATGAAATCTAACTTTGAGAGCT
TrpArgThrSerAlaIleGlnGluValLeuIleGlyGluGluSerIleThrTrpArgAla
<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cys</td>
<td>1</td>
</tr>
<tr>
<td>Tyr</td>
<td>2</td>
</tr>
<tr>
<td>Thr</td>
<td>3</td>
</tr>
<tr>
<td>Tyr</td>
<td>4</td>
</tr>
<tr>
<td>Tyr</td>
<td>5</td>
</tr>
<tr>
<td>Tyr</td>
<td>6</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

2b,4
1270 1280 1290 1300 1310 1320
TTCTTAAGAATCCAGATGGTCATTATATATTGGCTGGTTTGCGGACTACTGTT
PheLeuLysAsnProAspGlySerLeuTyrIleGlyAlaValTrpGlnValThrLeuPhe

1330 1340 1350 1360 1370 1380
TCCAGATTCTTCTGCAAGAACATTCCAGATATGGATAAAAGTCATTAAAGATTTGGATAGAA
SerArgPheLeuSerArgLysHisSerAspMetAspLysValIleLysAspTrpTyrGlu

1390 1400 1410 1420 1430 1440
TTAACTCCTTTTGATGGGATTTGGGCTGATATGAATGAATGATCCTCATCATCTCTGTTG
LeuThrProPheAspGlyIleTrpAlaAspMetAsnGluValSerSerPheCysValGly

1450 1460 1470 1480 1490 1500
TCTTGTTGCTACTGGTAAATATCTTCGAAAACCCAGCATACTCCATTACTTACTGGGAAGT
SerCysGlyThrGlyLysTyrPheGluAsnProAlaTyrProProPheThrValGlySer

1510 1520 1530 1540 1550 1560
AAAGCTACCTTTATCCGGTTGTTCGATTGGTTTTCTAAGGCATCTGAATGGAAATCTATT
LysAlaThrSerTyrProValGlyPheAspValSerAsnAlaSerGluTrpLysSerIle

1570 1580 1590 1600 1610 1620
CAAAAGCTCATTCTGCTACTGCTAAAGACTCTTTCAACTTCTTCCGTATCGTCTGCCT
GlnSerSerIleSerAlaThrAlaLysThrSerSerThrSerValSerSerSerSer

1630 1640 1650 1660 1670 1680
TCCACAATCGATTATATAGACACTTTAGCTCCAGGTAAAGGTAAATATAATTATCTCCACCA
SerThrIleAspTyrMetAsnThrLeuAlaProGlyLysGlyAsnIleAsnTyrProPro
2110  2120  2130  2140  2150  2160
BCTATTTGACAGAACATATTGCTGGGATCATCAGGCTGAGCTACTAGAATTCTTAGT
AlaIleAspGlnGluProTyrValTrpGluSerValAlaGluAlaThrArgThrSerMet

2170  2180  2190  2200  2210  2220
GCCATTAGATACCTTTATATTACCATATTAACACTTTATATTACATGAAATCTCATACT
AlaIleArgTyrLeuLeuLeuLeuProTyrTyrTyrThrLeuLeuHisGluSerHisThrThr

2230  2240  2250  2260  2270  2280
GGTTCACAAATCTTAAGAGCTTTTCTCGTGCAATCTCCCTAAACGATCGTTCCCTAAAGTG
GlyLeuProIleLeuArgAlaPheSerTrpGlnPheProAsnAspArgSerLeuSerGly

2290  2300  2310  2320  2330  2340
GTCTGATAACCATTTTTGTTCGTATCGTTTATACTCTGTCTTTAGAACCTGGT
ValAspAsnGlnPheProValGlyAspGlyLeuValValThrProValLeuGluProGly

2350  2360  2370  2380  2390  2400
GGTTCATAGTTAAAGGTGTTTCCGAGAGCTGTGAAAGAAGGAGTTTACTACGACTGG
ValAspPheValLeuGlyValPheProGlyAlaGlyLysGluGluValTyrTyrAspTrp

2410  2420  2430  2440  2450  2460
TACACCAAAAGAGAGTTCACTTTAAAGAGATGAAATGAAACTTATTAGATGCACTTA
TyrThrGlnArgGluValHisPheLysAspGlyLysAsnGluThrLeuAspIleProLeu

2470  2480  2490  2500  2510  2520
GGTTCATATTCCATTACACATAGAGGTGGAAGCTGCTTCCGCAACTCAAGAGCAGGTTAT
GlyHisIleProLeuHisIleArgGlyGlyAsnValLeuProThrGlnGluProGlyTyr
M13

$\alpha$Amylase
0.2 kb of the structural gene

$\alpha$Amylase - Promoter

E  Eco RI
P$_S$  Pst I
Bg  Bgl II
EV  Eco RV

Aa  Aat II
A  Asp 718
B  Bam HI
Sm  Sma I
S  Sal I
H  Hind III
Sc  Scal
X  Xho I

FIG. 3
FIG. 4
EP 0 394 538 B1

YRp7\alpha GAM
Bam HI
11.2 kb

TRP/ARS
\alpha Amylase-Promoter
Glucoamylase structural gene
YRp7 Struhl et al.1979
\alpha Amylase 0.4kb of the structural gene
E  Eco RI
B  Bam HI
Bg Bgl II
EV Eco RV
X  Xho I
S  Sal I

FIG. 5

55
C. thermocellum

pBR 332

Schwanniomyces

pUC8

DNA

Fig6b, 2

DNA
FIG. 9.4.2

pBR 332

Saccharomyces

E. coli

EcoRI

BglII

BamHI

PstI

SmaI

Trp5

Cam

14 kb

PmpG

DNA

DNA

Schwannomycetes

EM 0 354 538 B1