PROCESS FOR EXPRESSION AND
SECRETION OF PROTEINS BY THE
NON-CONVENTIONAL YEAST
ZYGOSACCHAROMYCES BAILII

Inventors: Paola Branduardi, Mailand (IT);
Danilo Porro, Erba (IT); Minoska
Valli, Calcinate (IT); Lilia Alberghina,
Mailand (IT)

Correspondence Address:
HAMILTON, BROOK, SMITH & REYNOLDS,
P.C.
530 VIRGINIA ROAD
P.O. BOX 9133
CONCORD, MA 01742-9133 (US)

Appl. No.: 10/534,171
PCT Filed: Nov. 6, 2003
PCT No.: PCT/EP03/12377
§ 371(c)(1), (2), (4) Date: May 6, 2005

Foreign Application Priority Data
Nov. 7, 2002 (DE).............................. 102 52 245.6

Publication Classification

Int. Cl.
C12P 21/06 (2006.01)
C07H 21/04 (2006.01)
C07K 14/47 (2006.01)
C12N 15/74 (2006.01)
C12N 1/18 (2006.01)

U.S. Cl. ................. 435/69.1; 435/483; 435/254.2;
530/350; 536/23.7

ABSTRACT

Herein is disclosed a method for the production of proteins.
The protein is expressed by a yeast belonging to the species
Zygosaccharomyces bailii. The yeast secretes the protein
produced into the culture medium from where it is isolated,
thereby simplifying the isolation process. Preferably the
yeast is cultivated in chemically defined medium, thereby
further simplifying the isolation process significantly.
Figure 1

S. cerevisiae TPI promoter

\( pZ_3 \)  
(7950 bp)

\( pZ_{3kl} \)  
(7998 bp)

\( pZ_{3pp\alpha} \)  
(8215 bp)
Figure 2

S. cerevisiae TPI promoter

pZ3kIL-1β
(8470 bp)

KanR

Amp

col Elori

S. cerevisiae TPI promoter

pZ3ppαIL-1β
(8685 bp)

KanR

Amp

col Elori

pre-pro-αF

pre-pro-αL

poly A

S. cerevisiae TPI promoter

pZ3ppαGFP
(8980 bp)

KanR

Amp

col Elori

pre-pro-αF

pre-pro-αL

poly A

GFP
**Figure 3**

**a**

S. cerevisiae TPI promoter

pZ₃GAA

(10720 bp)

poly A

ARS/CEN

Kan⁸

Amp

col Elori

**b**

S. cerevisiae TPI promoter

pZ₃LacZ

(11800 bp)

poly A

ARS/CEN

Kan⁸

Amp

col Elori
Figure 4

**a**

Z. bailii TPI prom.

pZ₃bT

(7450 bp)

**b**

Z. bailii TPI prom.

pZ₃bTLacZ

(11300 bp)
Figure 6

(a) YPD Glc 2%
- Z. bailii
- S. cerevisiae
- 16 24 24 hours

ii) 16 24 24 hours

ii) 16 24 24 hours

iii) 15 22 37 63 15 22 37 63 hours

iv) 15 22 37 63 15 22 37 63 hours

(b) YNB Glc 2%
- Z. bailii
- pZ3ppαL-1β
- pZ3ppαGFP
- 47.5 kDa
- 32.5 kDa
- cells
- sup
- cells
- sup

S. cerevisiae
- pZ3ppαL-1β
- S. cerevisiae
- pZ3ppαL-1β
- Z. bailii
- pZ3ppαL-1β
Figure 7

Batch Z. bailii pZ3kIII-L-1β 266g/l glc

Glucose (g/l) vs. time (h)
- OD • biomass g/l
- glc g/l △ EthOH g/l

Graph a: Glucose levels over time for Batch Z. bailii pZ3kIII-L-1β with biomass, glucose, and ethanol concentrations.

Graph b: IL-1β expression at different time points (T0, 16, 25, 44 hours).

Figure 8

Glucoamylase activity

a

b
Figure 10

a

b

Plasmid stability (%)

- pZ3 LacZ: 70
- p195IaZ: 7
- pEZ-JAIaZ: 29
- pEZ-JAFLacZ: not determined
- pEZ2 LacZ: 44
- pEZ2-IBLacZ: not determined
Figure 11

(a) YPD Glc 2%, YNB Glc 5%

1. Z. bailii pZ3kbIL-1β
2. S. cerevisiae pZ3kbIL-1β
3. Z. bailii pZ3kbIL-1β
4. S. cerevisiae pZ3kbIL-1β

(hours: 15, 23, 37, 63)

(b) pZ3kbIL-1β, pZ3klIL-1β, pZ3ppαIL-1β

(hours: 23, 37, 63)
Figure 12

Graph showing the comparison of different strains:
- Z. baillii [pZ3-STA2]
- S. cer [pZ3-STA2]
- Z. baillii [pZ3-kSTA2]
- S. cer [pZ3-kSTA2]
- Z. baillii [pZ3]
- S. cer [pZ3]
PROCESS FOR EXPRESSION AND SECRETION OF PROTEINS BY THE NON-CONVENTIONAL YEAST ZYGOSACCHAROMYCES BAILII

[0001] High level production of proteins from engineered organisms (recombinant, mutagenised, . . .) provides an alternative to the extraction of the proteins from natural sources. Natural sources of proteins are often limited, and furthermore the concentration of the desired product is generally low so extraction is regularly very cost-intensive or even impossible. Besides, extraction might bear the danger of toxic or infectious contamination depending on the natural origin of the protein.

[0002] With the advent of molecular cloning in the mid-70s, it became possible to produce foreign proteins in new hosts. Recombinant DNA (rDNA) technologies (genetic, protein and metabolic engineering) allow the production of a wide range of peptides and proteins from naturally-non-producing cells. In fact the first biotech-products on the world market made by means of rDNA were pharmaceutical products (for example insulin, interferons, erythropoietin, vaccine against hepatitis B) and industrial enzymes (for example used for the treatments of food, feed, detergents, paper-pulp and health care). World-sales of the top-20 recombinant pharmaceutical products in 2000 was about 13 billions Euro, while the world-wide market for the industrial enzymes was about 2.0 and it is projected to reach about 8 billions Euro in 2008.

[0003] Microorganisms as well as cultured cells from higher organisms (such as mammals, insects or plants) represent the mainly conceivable hosts for the production of heterologous as well as homologous proteins.

[0004] Several processes using mammalian cell culture for the production of proteins have been developed and many in such a manner produced proteins are on the market. Among them, several vaccines, monoclonal antibodies, interferon, blood factors, urokinase and tPA, hormones and growth factors.

[0005] The main advantage of a mammalian cell based expression system is the ability of mammalian cells to process the proteins in a proper way (correct folding, appropriate post-translational modification, correct glycosylation, specific proteolytic activities, etc.). A cloned protein expressed from recombinant DNA of mammalian origin (human) is usually correctly processed and folded and commonly secreted into the medium, allowing a fast recovery and purification. On the other hand the costs of production are generally quite high due to a usually low level of expression, costs of the mammalian medium components, very slow growth rates and demanding culture conditions. Furthermore, production in mammalian cells bears the danger of toxic or infectious contamination of the product.

[0006] Microorganisms (prokaryotic as well as eukaryotic) are advantageous hosts for the production of proteins because of high growth rates and commonly ease of genetic manipulation. But, in particular, bacterial hosts lack the ability of a correct protein processing and in a lot of cases heterologously produced proteins build up inclusion bodies inside of the bacterial cells, whereupon the proteins are lost, because their enzymatic activity can in most instances not be reconstituted. Due to their incorrect structure any use of such proteins for the treatment of humans is also excluded.

[0007] Yeast hosts can combine the advantages of unicellular organisms (i.e., ease of genetic manipulation and growth) with the capability of a protein processing typical for eukaryotic organisms (i.e. protein folding, assembly and post-translational modifications), together with the absence of endotoxins as well as oncogenic or viral DNA. Starting from the early 80s, the majority of recombinant proteins produced in yeast have been expressed using Saccharomyces cerevisiae (Hitzeman, R. A. et al., 1981, Nature 293, 717-22). The choice was determined by the familiarity of molecular biologists to this yeast together with the accumulated knowledge about its genetics and physiology. Furthermore, S. cerevisiae is an organism generally regarded as safe (GRAS). However, this yeast is not an optimal host for the large-scale production of foreign proteins, especially due to its characteristics regarding fermentation needs. In particular, growth of S. cerevisiae shows a very pronounced Crabtree effect, therefore fed-batch fermentation is required to attain high-cell densities (see for example Porro, D., et al., 1991, Res. Microbiol. 142, 555-9). Furthermore, this yeast is comparatively sensitive regarding the culture conditions, for example regarding the pH value and the temperature. Therefore, its cultivation is complicated and requires a highly sophisticated equipment. In addition, the proteins produced by S. cerevisiae are often hyper-glycosylated and retention of the products within the periplasmic space is frequently observed (Reiser, J. et al., 1990, Adv. Biochem. Eng./Biophys. 43, 75-102 and Romanos, M. A. et al., 1992, Yeast 8, 423-88). Furthermore, due to the partial retention of the protein in S. cerevisiae, a fraction of the protein is commonly degraded. These respective degradation products are generally very difficult to remove from the desired product. Disadvantages such as these have promoted a search for alternative hosts, trying to exploit the great biodiversity existing among the yeasts, and starting the development of expression systems in the so-called “non conventional” yeasts. Prominent examples are Hansenula polymorpha (Buckholz, R. G. et al., 1991, Bio/Technology 9, 1067-72); Pichia pastoris (Fleer, R., 1992, Curr. Opin. Biotechnol. 3, 486-96); Kluyveromyces lactis (Gellissen, G. et al., 1997, Gene 190, 87-97); Yarrowia lipolytica (Muller, S. et al., 1998, Yeast 14, 1267-83) among others. Another yeast genus under investigation is the genus Zygomonasccharomyces. Eleven species, which appear to be evolutionary quite close to S. cerevisiae and not so far from K. lactis have been classified so far (James, S. A. et al., 1994, Yeast 10, 871-81, Steels, H., et al., 1999, Int. J. Syst. Bacteriol. 49, 319-27 and Kurtzman, C. P. et al., 2001, FEMS Yeast research 1, 133-8). An exceptional resistance to several stresses renders some of the Zygomonasccharomyces species potentially interesting for industrial purposes. For example Z. rouxii is known to be salt tolerant (osmophilic) and Z. bailii is known to tolerate high sugar concentrations and acidic environments as well as relatively high temperatures of growth (Makdesi, A. K. et al. 1996, Int. J. Food Microbiol. 33, 169-81 and Sousa, M. J. et al., 1996, Appl. Environm. Microbiol. 62, 3152-7). However, the data available related to the molecular biology of these yeasts are very poor. While expression and secretion of a heterologous protein could be achieved in Z. rouxii (Ogawa, Y. et al. 1990, Agric. Biol. Chem. 54, 2521-9), for Z. bailii just the first molecular tools to successfully transform this yeast and to express heterologous proteins intracellularly have been developed (WO 00/41477). Since purification of intracellular proteins is very elaborate, the use of
this host for industrial production processes remains limited. Furthermore, while a lot of such non-conventional yeasts show specific advantages regarding their cultivation requirements, a lot of times these advantages are foiled by unexpected negative characteristics or unsolvable problems in their handling. In a lot of instances the tools for transformation of the organisms or expression of heterologous genes are not developed or the development fails due to unfavourable natural properties of the organism in question. The secretory capabilities often impose further problems for the production of proteins in industrial scale. If the organism does not allow the efficient secretion of the desired protein, the isolation of the product is significantly complicated. In addition, some very interesting products, such as Interleukin 1-β, turned out to be toxic for the cells as long as they are intracellularly located (Fleer, R et al, 1991, Gene 107, 285-95). Production of such proteins is therefore only possible if the host comprises a highly potent secretory system that can be exploited. Another problem comes from a potentially different codon usage or codon frequency that can hamper the expression of heterologous genes in such organisms decisively.

[0008] In consideration of the state of the art, the problem to be solved by the present invention was to provide a new, easy and economical method for the production of proteins. Apart of being cost effective that method should be easy to perform and allow the production of highly pure proteins in a high yield.

[0009] This problem as well as all further not explicitly mentioned problems, that are easily deduced from the introductory explicated contents, are solved by the objects outlined in the claims of the instant invention.

[0010] An advantageous process for the production of a protein is provided by a method as outlined in claim one. This method comprises culturing a Zygospaccharomyces bailii strain expressing and secreting the protein and isolating the protein. This process is particularly advantageous in that Z. bailii can be cultured yieldingly in a chemically defined medium without the addition of complex ingredients that have to be separated tediously from the protein produced. Surprisingly, the secretory capacity of this yeast in chemically defined medium is significantly superior to the secretory capacity of S. cerevisiae under identical conditions. A further important advantage is the surprising fact that the protein produced by Z. bailii is not only readily secreted but also near to completion, what is not the case for S. cerevisiae under identical conditions. Through efficient secretion of the desired protein by Z. bailii also no degradation of the protein takes place. Subsequently, the purification of the product is significantly simplified.

[0011] Further major advantages of Z. bailii as host organism for protein production, and in particular for production of heterologous proteins are a naturally favourable codon usage as deduced from the examples presented herein and the comparatively low demands on the culture conditions. This is in particular due to a high acid and temperature tolerance as well as a weak Crabtree effect allowing the cultivation with a high sugar concentration from the beginning (i.e. batch instead of fed-batch cultivation) and the omission of extremely sophisticated regulations of the culture conditions such as temperature or pH. Accordingly, this method allows a cost effective production of proteins in an easy way even in industrial scale yielding proteins of high purity.

[0012] The term “expression” of a protein by a host cell is well known to the skilled artisan. Usually expression of a protein comprises transcription of a DNA sequence into a mRNA sequence followed by translation of the mRNA sequence into the protein. A more detailed description of the process can be found for example in Knippers, R. et al, 1990, Molekulare Genetik, Chapter 3, Georg Thieme Verlag, Stuttgart.

[0013] The term “secretion” of a protein as known in the art means translocation of the protein produced, from inside of the cell to outside of the cell, thereby accumulating the protein in the culture medium. A more detailed description of the process can be found for example in Stryer, L. 1991, Biochemie, Chapter 31, Spektrum Akad. Verlag, Heidelberg, Berlin, New York.

[0014] The protein produced might be any protein known in the art for which an industrial production is desirable. For example the protein might be useful in the pharmaceutical field, such as medication or vaccine or in pre-clinical or clinical trials among others (examples are growth hormones, tissue plasminogen activator, hepatitis B vaccine, interferones, erythropoietin). The protein produced might also be useful in industry for example in the area of food production (e.g. β-galactosidase, chymosin, amylases, glucoamylase, amylo-β-glucosidase, invertase) or textile and paper production (proteases, amylases, cellulases, lipases, catalases, etc.). Enzymes are useful among others as detergents (proteases, lipases and surfactants) and their characteristics of stereospecificity are furthermore exploitable in a wide number of biocatalyses, yielding a desired chiral compound. Another promising application of recombinant enzymes that can be produced by the method of the instant invention is the development of biosensors.

[0015] The proteins secreted can vary greatly in size (molecular weight). The herein described method functions well for very small proteins (e.g. IL-1β, 17 kDa, see FIG. 5), but also for quite large proteins (e.g. GAA, 67.5 kDa, see FIG. 8a). The secreted proteins may or may not comprise consensus sites for glycosylation. Such consensus sites might occur naturally or might be introduced by genetic engineering. Depending on the intended use of the protein produced it might also be advantageous to remove naturally occurring consensus sites for glycosylation by genetic engineering, thereby preventing for example hyper-glycosylation of the protein. Remarkably, the herein described method leads to proteins that conserve their desired catalytic characteristics after the secretion (e.g. GAA, see FIG. 8a).

[0016] In one embodiment of the present invention the Z. bailii strain is transformed with a vector comprising a DNA sequence coding for the protein, functionally linked to a signal sequence leading to the secretion of the protein and further functionally linked to a promoter leading to the expression of the protein.

[0017] The term “vector” refers to any agent as such a plasmid, cosmids, virus, phage, or linear or circular single-stranded or double-stranded DNA or RNA molecule, derived from any source that carries nucleic acid sequences into a host cell. Preferably a vector is capable of genomic integra-
tion or autonomous replication. Such a vector is capable of introducing a 5' regulatory sequence or promoter region and a DNA sequence for a selected gene product into a cell in such a manner that the DNA sequence is transcribed into a functional mRNA, which may or may not be translated and therefore expressed. Preferably the vector is an extra-chromosomal plasmid. Such a plasmid comprises preferably an autonomously replicating sequence (ARS) and advantageously a centromeric sequence (CEN) in addition. More preferably the plasmid is a 2μ-like episomal multicopy plasmid. Even more preferably the plasmid is derived from an endogenous episomal plasmid from a Z. bailii strain such as pSB2 (Utatsu, I. et al., 1987, J. Bacteriol. 169, 5537-45) and more preferably from pZB1 or pZB2 (see FIG. 9).

[0018] The plasmid pZB1 was extracted from NCYC 1427 and partially sequenced. Accordingly, the plasmid comprises preferably at least 35, more preferably at least 55 and even more preferably at least 75 and even more preferably at least 100 bases from at least one of the sequences selected from the list of SEQ ID No.: 63, SEQ ID No.: 64, SEQ ID No.: 65, SEQ ID No.: 66, SEQ ID No.: 67, SEQ ID No.: 68, SEQ ID No.: 69, SEQ ID No.: 70 or SEQ ID No.: 71.

[0019] Yeast multicopy plasmids (also referred to as 2μ or 2 μ-like plasmids) isolated from different yeast genus or species usually show a well conserved structural homology while having a low sequence homology. Some regulatory elements were identified as necessary and sufficient to build a functional multicopy plasmid. These are:

[0020] the recombinase promoting amplification of these plasmids, encoded by the FLP gene. (Blanc H., et al., 1979, Mol. Gen. Genet. 176, 335-42 and Broach J. R. et al., 1980, Cell 21, 501-8);

[0021] two inverted repeats (IR-sequences);


[0024] Within the scope of the instant invention these key elements of the 2μ plasmid are preferably derived from Z. bailii, even more preferably from Z. bailii NCYC1427 or ATCC6947. Particularly preferred these sequences correspond to SEQ ID No.: 71 (IR-ARS), SEQ ID No.: 72 (FLP), SEQ ID No.: 74 (TFB) and SEQ ID No.: 76 (TFC), respectively. The expressed recombinase and the expressed regulatory proteins exhibit preferably the amino acid sequence shown in SEQ ID No.: 73 (FLP), SEQ ID No.: 75 (TFB) and SEQ ID No.: 77 (TFC), respectively. Preferably the plasmid additionally comprises the homologue upstream regions of the FLP and the TFB/TFC genes, in order to obtain an optimal control of the transcription level.

[0025] Generally speaking the plasmid preferably comprises sequences for (autonomous) replication, stabilization and/or plasmid copy number control, obtainable from a Z. bailii strain.

[0026] Preferably the plasmid is pEZ1, (see FIG. 9c).

[0027] Particularly preferred is the plasmid pEZ2, (see FIG. 9d). One preferred way to construct pEZ2 is to amplify the IR/ARS region and the TFC/FLP genes including their homologous promoters by PCR with the oligos

[0028] 5'-AGAATCAATCATTTAGTGGCGAGGAG3' (SEQ ID NO.: 90) and

[0029] 5'-TAAAGCTGCCCCCTATTCGCTTC3' (SEQ ID NO.: 91, IRAARS),

[0030] 5'-AGAATGACTCAGATCTCTCCTTG3' (SEQ ID NO.: 86) and

[0031] 5'-CTATGGCAGGTTCAGCCTG3' (SEQ ID NO.: 85, FLP/TFC).

[0032] and to substitute the ARS/CEN cassette from pZ4 with these amplified products. Another way is to substitute the 2μ ori sequence from the plasmid p195 with the aforementioned PCR-products.

[0033] Advantageously, the vector comprises a selectable marker. The term selectable marker refers to a nucleic acid sequence whose expression confers a phenotype facilitating identification of cells containing the nucleic acid sequence. Selectable markers include those which confer resistance to toxic chemicals (= dominant marker, e.g. G418, hygromycin, formamidenehyde, phleomycin or fluoracetasate like reviewed in Van den Berg, M. et al. 1997, Yeast 13, 551-9) or complement an auxotrophy (= auxotrophic marker, e.g. uracil, histidine, leucine, tryptophane). Auxotrophic selection markers can be used for naturally auxotrophic Z. bailii strains or strains that have been rendered auxotrophic by genetic manipulation, in particular by (partial) deletion or mutagenisation of an essential gene, e.g. HIS3 (Branduardi, P., 2002, Yeast 19, 1165-70). As complementing marker sequence the homologous gene from Z. bailii or a heterologous gene might be employed. Auxotrophic markers are preferred since no component has to be added to the medium to keep the selective pressure during the cultivation.

[0034] The term "promoter" or "promoter region" refers to a DNA sequence, usually found upstream (5') to a coding sequence, that controls expression of the coding sequence by controlling production of messenger RNA (mRNA) by providing the recognition site for RNA polymerase and/or other factors necessary for start of transcription at the correct site. The promoter can be derived from any organism. Preferably the promoter is derived from a yeast, even more preferably from Saccharomyces, Kluyveromyces or Zygoseaachromyces and most preferably from Z. rouxii or Z. bailii. The promoter can be constitutive, inducible or repressible. Inducible promoters can be induced by the addition to the medium of an appropriate inducer molecule or by an appropriate change of the chemical or physical growth environment (such as the temperature or pH value, which will be determined by the identity of the promoter. Repressible promoters can be repressed by the addition to the medium of an appropriate repressor molecule or by an appropriate change of the chemical or physical growth environment (such as the temperature or pH value), which will be determined by the identity of the promoter. Constitutive promoters are preferred, as the use of an appropriate repressor or inducer molecule or an appropriate change of the chemical or physical growth environment is not required.
Preferably the promoter is selected from the list of: triose-phosphate isomerase (TPI), glyceraldehyde dehydrogenase (GAPDH), alcohol dehydrogenase 1 (ADH1), phosphoglycerate kinase (PGK), glyceraldehyde-3-phosphate dehydrogenase (GAP), GAL1, GAL10, acid phosphatase (PHOS), cytochrome C-1 (CYC1), copper-binding metallothionen (CUP1) or a-fact maturing pheromone pre-cursor (Msh1) promoter or the hybrid promoters GAL1/CYC1, such as GAL1-10/CYC1, GAL1/GAP1/GAL, GAP1/ADH1, GAP1/PHOS or CYC1/GRE either from S. cerevisiae, Z. rouxii or Z. bailii, but preferred from Z. bailii.

Especially preferred promoters are the TPI promoters either from S. cerevisiae corresponding to SEQ ID NO.: 78 or Z. bailii corresponding to SEQ ID NO.: 79, but particularly preferred is the TPI promoter from Z. bailii (SEQ ID NO.: 79). Further particularly preferred promoters are the GAPDH promoters from Z. rouxii (SEQ ID NO.: 92) or Z. bailii.

[0035] Furthermore the vector comprises preferably a transcriptional terminator sequence following the coding sequence for the desired protein for efficient mRNA 3' end formation. Such a terminator sequence is preferably derived from a yeast, more preferably from Saccharomyces or Zygosaccharomyces, even more preferably from S. cerevisiae or Z. bailii and most preferably from Z. bailii. A preferred example for a terminator sequence comprises the following tripartite consensus sequence: TAG . . . (T-rich) . . . TATGT . . . (AT-rich) . . . TTT. Another preferred example comprises the sequence motif TTTTATA.

[0036] Further the vector comprises a signalling sequence (=leader sequence; upon expression translated into signal peptide or leader peptide). Such sequences lead to the direction of expressed proteins from the cytosol into the culture medium. In other words signal sequences cause the secretion of the proteins and their accumulation in the medium. Signal sequences generally code for a continuous stretch of amino acids, typically 15 to 60 residues long (up to 150), which characteristically include one or more positively charged amino acid(s) followed by a stretch of about 5 to 10 hydrophobic amino acids, which may or may not be interrupted by non-hydrophobic residues. Preferably the signal peptide comprises 15-45 amino acids, even more preferably 15 to 30 amino acids. Even though their amino acid sequences can vary greatly, the signal peptides of all proteins having the same destination in one organism are functionally interchangeable: physical properties, such as hydrophobicity or the pattern of charged amino acids, often appear to be more important in the signal-recognition process than the exact amino acid sequence.


 Particularly preferred the DNA sequence coding for the signal peptide is selected from the list of SEQ ID NO.: 1, SEQ ID NO.: 3, SEQ ID NO.: 21 or SEQ ID NO.: 35 corresponding the amino acid sequence of the signal peptide is preferably selected from the list of SEQ ID NO.: 2, SEQ ID NO.: 4, SEQ ID NO.: 22 or SEQ ID NO.: 36.

[0038] The signal peptide is preferably removed from the finished protein. This can occur through activity of a specialised signal peptidase. The signal peptidase can be of homologous or heterologous origin. Therefore, the signal peptide comprises preferably a processing site or a cleavage site that allows for recognition by a specific endopeptidase.

[0039] In a preferred embodiment of the present invention the Z. bailii strain is transformed with a vector comprising the DNA sequence coding for the protein, functionally linked to the signalling pre-sequence (16 aa) of the alphahemolysin of the K. lactis (Stark M.J. et al., 1986, EMBO J. 5,1995-2002, SEQ ID NO.: 35 (DNA) and SEQ ID NO.: 36 (peptide)) and further functionally linked to the TPI promoter from S. cerevisiae. More preferably the vector is pHZM4 (FIG. 16). Even more preferably the Z. bailii strain is transformed with a vector comprising the DNA sequence coding for the protein, functionally linked to the signal sequence of the K. lactis and further functionally linked to the GAPDH promoter from Z. rouxii. Even more preferably the Z. bailii strain is transformed with a vector comprising the DNA sequence coding for the protein, functionally linked to the signal sequence of the K. lactis and further functionally linked to the TPI promoter from Z. bailii. Preferably said vector is derived from pHZM7 (FIG. 4a).

[0040] In another preferred embodiment of the present invention the Z. bailii strain is transformed with a vector comprising the DNA sequence coding for the protein, functionally linked to the signal sequence of the pre-pro α-factor of S. cerevisiae and further functionally linked to the TPI promoter from S. cerevisiae. Preferably the vector is pHZppox (FIG. 1c). Even more preferably the Z. bailii strain is transformed with a vector comprising the DNA sequence coding for the protein, functionally linked to the signal sequence of the pre-pro α-factor of S. cerevisiae and further functionally linked to the GAPDH promoter from Z. rouxii. Even more preferably the Z. bailii strain is transformed with a vector comprising the DNA sequence coding for the protein, functionally linked to the signal sequence of the pre-pro α-factor of S. cerevisiae and further functionally linked to the TPI promoter from Z. bailii. Particularly preferred said vector is derived from pHZM7 (FIG. 4a).

[0041] In yet another preferred embodiment of the present invention the Z. bailii strain is transformed with a vector comprising the DNA sequence coding for the protein, functionally linked to the zygocin killer toxin pre-sequence of Z. bailii (SEQ ID NO.: 59) and further functionally linked to a promoter functional in Z. bailii. Preferably said promoter is
the TPI promoter from *S. cerevisiae*. Even more preferably said promoter is the TPI promoter from *Z. bailii*. Most preferred is the GAPDH promoter from *Z. rouxii*.

[0042] The DNA sequence coding for the protein can be derived from animal, bacterial, fungal, plant or viral sources, more preferably from metazoan, mammalian or fungal sources. The expressed protein might therefore be homologous or heterologous to *Z. bailii*.

[0043] Any yeast belonging to the species *Z. bailii* can be used for the production of proteins in the scope of the present invention. In a preferred embodiment of the invention the *Z. bailii* strain is transformed. “Transformation” refers to a process of introducing an exogenous nucleic acid sequence (of homologous and/or heterologous origin, recombinant or not) into a cell in which the exogenous nucleic acid is incorporated into a chromosome or is capable of autonomous replication. A cell that has undergone transformation, or a descendant of such a cell, is “transformed” or “recombinant”. If the exogenous nucleic acid comprises a coding region encoding a protein and the protein is produced in the transformed yeast such a transformed yeast is functionally transformed. Preferred methods to transform *Z. bailii* are electroporation, as described in [WO00/14177](http://tto.trends.com), or the chemical LiAc/PEG/ssDNA method as described by Agatep, R. et al., 1998, Technical Tips Online (http://ttto.trends.com).

[0044] Preferably the *Z. bailii* strain that is being transformed is selected from the list of ATCC 36947, ATCC 60483, ATCC 8766, FRRI 1292, ISA 1307, NCYC 128, NCYC 563, NCYC 1416, NCYC 1427, NCYC 1766, NRRL Y-2227, NRRL Y-2228, NRRL Y-7239, NRRL Y-7254, NRRL Y-7255, NRRL Y-7256, NRRL Y-7257, NRRL Y-7258, NRRL Y-7259, NRRL Y-7260, NRRL Y-7261, NRRL Y-7264; particularly preferred are ATCC 36947, ATCC 60483, ATCC 8766 and NCYC 1427.

[0045] (ATCC: American Type Culture Collection, Manassas VA, USA; FRRI: Korean Culture Collection, North Ryde NSW, Australia; ISA: Culture Collection of the Instituto Superior de Agronomía, Lisbon, Portugal; NCYC: National Collection of Yeast Cultures, Norwich, UK; NRRL: Agricultural Research Service Culture Collection, Peoria Ill., USA.)

[0046] Within the scope of the present invention the *Z. bailii* strain can be subjected to a selection process for improved secretion. Screening for and isolation of such a “super-secreting” phenotype can occur before or after transformation of the respective *Z. bailii* strain.

[0047] In a preferred embodiment of the present invention the *Z. bailii* gene's homologous to GAS1 from *S. cerevisiae* are identified and disrupted. GAS1 is one example for the few cases wherein the key molecules involved in the intriguingly complex secretory pathway have been identified. It was possible to influence the whole secretory mechanism modifying the GAS1 expression level in *S. cerevisiae* (Vai M., et al., 2000, Appl. Environ. Microbiol. 66, 5477-9) due to a resultant modification of the organisation of the cell wall structure, namely it was demonstrated that gas1 mutants show a “super-secreting” phenotype (Popolo L., et al., 1997, J. Bacteriol. 180, 163-6; Ram A. F. J., et al., 1998, J. Bacteriol. 180, 1418-24).

[0048] In another preferred embodiment of the present invention the *Z. bailii* strain has undergone one or more mutagenisation/selection cycle(s) to obtain super secreting mutants, comprising chemical or physical mutagenesis. Preferably the mutagenisation is caused by orthovanadate. Orthovonadate is a molecule known to affect the glycosylation process and the cell wall construction in *S. cerevisiae* (Kanik-Fanullat, C. et al., 1990, Mol. Cell. Biol. 10, 898-909). Methods involving orthovanadate mutagenisation to obtain cells with changed cell wall construction/secretory properties that are useful in the scope of the present invention are disclosed in more detail for example for *S. cerevisiae* (Willsky, G. R., et al., 1985, J. Bacteriol. 164, 611-7) and *K. lactis* (Uccellletti, D., et al., 1999, Res. Microbiol. 150, 5-12; Uccellletti, D., et al., 2000, Yeast 16, 1161-71).

[0049] Culturing techniques and media suitable for yeast are well known in the art. Typically, but it is not limited to, culturing is performed by aseptic fermentation in an appropriate vessel. Examples for a typical vessel for yeast fermentation comprise a shake flask or a bioreactor.

[0050] The culture is typically performed at a temperature between 20°C and 40°C, preferably between 25°C and 35°C and even more preferably between 28°C and 32°C.

[0051] The medium in which the *Z. bailii* strain is cultured can be any medium known in the art to be suitable for this purpose. The medium might contain complex ingredients or might be chemically defined. Chemically defined media are preferred. The medium comprises any component required for the growth of the yeast. In particular the medium comprises a carbon source, such as fructose, glucose or other carbohydrates (such as sucrose, lactose, D-galactose, or hydrolysates of vegetable matter, among others). Typically, the medium also comprises further a nitrogen source, either organic or inorganic, and optionally the medium may also comprise macro nutrients and/or micro nutrients such as amino acids; purines; pyrimidines; corn steep liquor; yeast extract; protein hydrolysates, such as peptone; vitamins (water-soluble and/or water-insoluble), such as B complex vitamins; or inorganic salts such as chlorides, hydrochlorides, phosphates, or sulphates of Ca, Mg, Na, K, Fe, Ni, Co, Cu, Mn, Mo, or Zn, among others. Antifoam might be added, if necessary. Further components known to one of ordinary skill in the art to be useful in yeast culturing or fermentation can also be included. The medium may or may be not buffered. A preferred medium comprises yeast extract, peptone and glucose (=YPD). A more preferred medium comprises yeast extract, peptone and fructose (=YPF). An even more preferred medium comprises glucose and Yeast Nitrogen Base (YNB, Difco Laboratories, Detroit, Mich. #919-15). Another even more preferred medium comprises fructose and YNB.

[0052] Particularly preferred is a medium comprising high fructose corn syrup as carbon source (for example Isosweet® 100 42% High Fructose (80% solids) or Isosweet® 5500 55% Fructose from Tate & Lyle PLC or IsoClear® 42% High Fructose Corn Syrup or IsoClear® 55% High Fructose Corn Syrup from Cargill, Inc.).

[0053] The compositions of preferred media for batch/fed batch cultivation of *Z. bailii* according to the instant invention are as follows: the batch phase medium comprises 4% w/V Glucose, 0.5% w/V (NH₄)₂SO₄, 0.05% w/V MgSO₄·7H₂O, 0.3% w/V KH₂PO₄, vitamins according to Verduny, C., et al., 1992, Yeast 8, 501-17, wherein the final concentration of vitamins will be 3 times in respect to the indicated concen-
trations and trace elements according to Verduyn, C., et al., 1992, Yeast 8, 501-17, wherein the final concentration will also be 3 times in respect to the indicated concentrations. The pH control (value: pH 5) is performed by the addition of 2M KOH. The fed-batch medium comprises 50% w/V Glucose, 15.708 g/l KH$_2$PO$_4$, 5 g/l KCl, 5.831 g/l MgSO$_4$, 1.2 g/l CaCl$_2$, 1 g/l Yeast Extract, 0.4447 g/l NaCl, 1 g/l Glutamate, 0.05 g/l ZnSO$_4$, 0.04 g/l CuSO$_4$, 0.05 g/l MnCl$_2$, 0.001 g/l CoCl$_2$, 0.5 g/l molybdenum, 0.1 g/l thiamine hydrochloride, 0.02 g/l pyridoxol hydrochloride, 0.04 g/l Ca-D(+)-D-glucuronate, 0.004 g/l D-biotin, 0.09 g/l nicotinic acid. The pH control (value: pH 5) is performed by the addition of 2M NH$_4$OH.

[0054] In case of selection for the dominant G418 marker 200 mg/l G418 is added to the respective medium.

[0055] The use of a defined medium, of which the components are adjusted to the needs of the organism is preferred. The purification of the protein is thereby significantly simplified.

[0056] Preferably, the pH of the culture medium ranges between 2 and 9, more preferably between 3 and 8 and even more preferably between 4 and 7. The pH can be regulated or partially regulated or not be regulated during the course of fermentation; accordingly the pH can be kept constant at a preferred value or can change during fermentation. A significant advantage of Z. bailii is its surprising capacity to grow as well as express and secrete proteins at low pH. Therefore, the demand of this organisms for a strictly controlled pH is not very pronounced.

[0057] The cultivation can take place in batch, fed-batch or continuous mode as is known to the ordinary skilled artisan.

[0058] During the course of the fermentation, the desired protein is expressed, properly processed (i.e. folded, modified, cut, etc.) and secreted (accumulated in the medium). While the protein produced may be partially retained within the yeast cells it is preferred that a substantial amount of the protein is secreted. Even more preferred is that the protein is entirely secreted.

[0059] After culturing has progressed for a sufficient length of time to produce a desired concentration of the protein in the yeast and/or the culture medium, the protein is isolated. "Isolated," as used herein to refer to the protein, means being brought to a state of greater purity by separation of the protein from at least one other component of the yeast or the medium. Preferably, the isolated protein is at least about 80% pure as based on the weight, more preferably at least about 90% pure as based on the weight and even more preferably at least about 95% pure as based on the weight. Evidence of purity can be obtained by SDS-PAGE, 2D electrophoresis, IF, CIEF, mass spectrometry, capillary electrophoresis or other methods well known in the art.

[0060] "Purity" refers to the absence of contaminants in the final purified protein. Typical contaminants to be separated from the desired product are proteins, pyrogens, nucleic acids and more.

[0061] The protein is isolated from the culture medium, preferably without lysing the cells. Such an isolation comprises purifying the protein from the medium. Purification can be achieved by techniques well-known in the art, such as filtration (e.g. microfiltration, ultrafiltration, nanofiltration), crystallisation or precipitation, centrifugation, extraction, chromatography (e.g. ion exchange, affinity, hydrophobic exchange), among others.

[0062] Upon removal of the cells, the culture broth might also directly serve as the product (e.g. enzyme solution), without further purification. The medium components can be adjusted appropriately prior to the cultivation.

[0063] If the protein is not completely secreted, the protein can also be isolated from both the yeast cells and the medium. Methods for lysing the yeast cells are known in the art and comprise chemical or enzymatic treatment, treatment with glass beads, sonication, freeze/thaw cycling, or other known techniques. The protein can be purified from the various fractions of the yeast lysate by appropriate techniques, such as filtration (e.g. microfiltration, ultrafiltration, nanofiltration), crystallisation or precipitation, centrifugation, extraction, chromatography (e.g. ion exchange, affinity, hydrophobic exchange), among others.

[0064] Another embodiment of the present invention relates to a Z. bailii strain, expressing and secreting a heterologous protein.

[0065] The Z. bailii strain might be transformed with a vector comprising a DNA sequence coding for the heterologous protein, functionally linked to a signal sequence leading to the secretion of the protein and further functionally linked to a promoter.

DESCRIPTION OF THE FIGURES

[0066] FIG. 1 : Expression and Secretion Vectors


[0068] a) pZ$_2$: the backbone of the plasmid is the pYX022 S. cerevisiae expression plasmid (R&D Systems, Inc., Wiesbaden, D); the expression cassette is based on the constitutive S. cerevisiae TPI promoter and the corresponding polyA signal, as indicated in the Figure). The ARS/CEN fragment, from YCP333 (Gietz, R. D., et al., 1988, Gene 74, 527-34) ensures replication and stability of the plasmid, while the Kan cassette, derived from pFA6-KanMX4 (Wach, et al., 1994, Yeast 10, 1793-808) allows a G418-based selection of the transformants.

[0069] b) pZ$_2$kl: a pZ$_2$ expression vector comprising the signal sequence of the K. lactis K1 killer toxin (kl) for leading the secretion of the protein of interest.

[0070] c) pZ$_2$ppa: a pZ$_2$ expression vector comprising the pro-pro leader sequence of the S. cerevisiae phenomone α-factor (pre-pro-αf) for leading the secretion of the protein of interest.

[0071] (Amp=ampicillin resistance cassette; MCS=multi cloning site; coli 1 ori, E. coli replication origin)

[0072] FIG. 2: Expression and Secretion Vectors

[0074] a) pZ3klL-lβ: a pZ3kl vector where the sequence encoding for the human IL-1β was sub-cloned into the MCS.

[0075] b) pZ3ppaL-lβ: a pZ3ppa vector where the sequence encoding for the human IL-1β was sub-cloned into the MCS.

[0076] c) pZ3ppaGFP: a pZ3ppa vector where the sequence encoding for the GFP was sub-cloned into the MCS.

[0077] FIG. 3: Expression and Secretion Vectors


[0079] a) pZ3GAA: a pZ3 vector where the sequence encoding for the glucoamylase (GAA) was sub-cloned into the MCS.

[0080] b) pZ3lacZ: a pZ3 vector where the sequence encoding for the β-galactosidase was sub-cloned into the MCS.

[0081] FIG. 4: Expression Vectors

[0082] Schematic maps of the plasmids constructed for the expression of proteins in *Z. bailii* based on the *Z. bailii* TPI promoter.

[0083] a) pZ3bT: a pZ3 vector where the *S. cerevisiae* TPI promoter was substituted by the *Z. bailii* TPI promoter.

[0084] b) pZ3bTlacZ: a pZ3bT expression vector where the sequence encoding for the β-galactosidase was sub-cloned into the MCS.

[0085] FIG. 5: IL-1β secretion

[0086] a) Growth kinetics in minimal (YNB) and rich (YPD) medium, with glucose 5% (w/v) as a carbon source: the cellular growth was measured as optical density (OD 660 nm, circles) and the residual glucose (g/l, squares) was evaluated. Comparison between *S. cerevisiae* (open symbols) and *Z. bailii* (full symbols).

[0087] b) Western Blot analyses performed on cellular extracts of *S. cerevisiae* and *Z. bailii* cells transformed with the plasmid pZ3klL-lβ (expressing IL-1β preceded by the leader sequence from the *K. lactis* killer toxin) and with the corresponding empty plasmid (pZ3α), as a negative control. In the first lane a positive control (IL-1β, human recombinant (*E. coli*), Roche cat n° 1 457 756) was loaded. Samples were collected at the indicated times and from the indicated media, corresponding to the kinetics showed in (a). The loaded volumes were rectified for a corresponding OD value of 0.08. The blotted membranes were probed with an α-IL-1β polyclonal antibody.

[0088] c) as above, were the loaded samples represent the corresponding supernatant.

[0089] d) as above, were the samples were loaded with an equal volume of medium (30 μl).

[0090] FIG. 6: Leading of the pre-pro-α-factor signal sequence to the secretion of IL-1β and of GFP in *Z. bailii*

[0091] a) Western Blot analyses performed on cellular extracts (i) and on supernatants (ii) of *Z. bailii* and *S. cerevisiae* cells transformed with the plasmid pZ3ppaL-lβ (and with the corresponding empty plasmid pZ3α) and growing on YPD medium (glucose 2% w/v). Samples were taken at the indicated times. First lane: positive control (IL-1β, human recombinant (*E. coli*), Roche cat n° 1 457 756). The blotted membranes were probed with an α-IL-1β polyclonal antibody. Western Blot analyses performed on cellular extracts (iii) and on supernatants (iv) of *Z. bailii* and *S. cerevisiae* cells transformed with the plasmid pZ3ppaL-lβ (and with the corresponding empty plasmid pZ3α) and growing on YNB medium (glucose 5% w/v). Samples were taken at indicated times. First lane: positive control (IL-1β, human recombinant (*E. coli*), Roche cat n° 1 457 756). The blotted membranes were probed with an α-IL-1β polyclonal antibody.

[0092] b) Western Blot analyses performed on cellular extracts (cells) and on supernatants (sup) of *Z. bailii* cells growing on YNB medium (glucose 2% w/v) transformed with the control plasmid pZ3 (1st and 2nd lanes) and with the plasmid pZ3ppaGFP (3rd) and (4th) lanes. The blotted membrane was probed with an α-GFP polyclonal antibody. An arrow indicates the expected positive signal.

[0093] FIG. 7: Batch cultivations of *Z. bailii* cells comprising the pZ3klL-lβ expression plasmid on chemically defined medium in high sugar concentration

[0094] a) Culture OD (full circles), dry weight (open circles), glucose consumption (full squares) and ethanol production (open triangle).

[0095] b) Western Blot analyses performed on the growth medium (lane 2 to 5) and on the cell extracts (lanes 6 to 9) of *Z. bailii* cells. Samples were collected at the indicated times of the kinetics, and an equal volume (30 μl) for the supernatants and 15 μl for the cell extracts, respectively was loaded. The blotted membranes were probed with an α-IL-1β polyclonal antibody.

[0096] First lane: positive control (IL-1β, human recombinant (*E. coli*), Roche cat n° 1 457 756).

[0097] FIG. 8: Enzymatic activity of heterologous enzymes expressed in *Z. bailii* cells

[0098] a) Determination of the *A. adeninivorans* glucoamylase activity (μU/OD) present in the growth medium (YNB, glucose 2% w/v) of *Z. bailii* cells transformed with the plasmid pZ3GAA (and the respective empty plasmid pZ3α, as a control). Three independent clones were analysed (C1.1, C1.3 and C1.5).

[0099] b) Determination of the β-galactosidase activity (Miller U/OD) in cell extracts of *Z. bailii* cells transformed with the plasmid pZ3lacZ (two independent clones) and with the plasmid pZ3bTlacZ (three independent clones), and the respective empty plasmid pZ3α as a control. Cells were grown in YPD medium (glucose 2% w/v), and samples were collected at indicated times.

[0100] On the left panel the *Z. bailii* strain ATCC 36047, on the right panel the strain *Z. bailii* ATCC 60483 were tested, respectively.
FIG. 9: Construction of a Z. bailii multicopy plasmid

Schematic maps of the endogenous plasmids isolated from Z. bailii ATCC 36947, named pZB1 (a) and from Z. bailii NCYC 1427, named pZB3 (b).

Z. bailii multicopy expression vector comprising the genes and the sequences necessary and sufficient for a stable and autonomous high copy number replication. The expression cassette is based on the Z. bailii constitutive TPI promoter and the polyA, as indicated in the Figure. The marker for selection is the KanR cassette.

Z. bailii multicopy expression vector. The expression cassette is based on the Z. bailii constitutive TPI promoter and the polyA, as indicated in the Figure. Furthermore, the vector comprises the IR/ARS region and the TFC/FLP genes including their homologous promoters as indicated.

FIG. 10: Influence of the promoter or the plasmid constituents, respectively, on β-galactosidase activity.

Shown is the relative β-galactosidase activity in cell extracts of Z. bailii ATCC 36947 cells transformed with the indicated plasmids. The β-galactosidase activity of cells transformed with pZβLacZ was set to 1 and the other activities were related to that value. Cells were grown in YPD medium (glucose 2% w/v), and samples were collected as the cultures reached an OD600 value between 1 and 2.

a) Different promoters in the same plasmid. pZβ: ScTPI, pZβB: ZbTPI, pZβG: ZbGAPDH.

b) Different plasmid constituents. pZ3: Sc ARS/CEN, p195: Sc 2 µm ori sequence, pEZ-LA: Zb 2 µm ori sequence (IR-A), pEZ-LA: Zb 2 µm ori sequence (IR-A)+FLP, pEZ-2: Zb 2 µm ori sequence (IR-A)+FLP+TFC, pEZ-2: Zb 2 µm ori sequence (IR-A)+FLP+TFC+IR-B. The table indicates the determined plasmid stability of the respective constructs.

FIG. 11: Leading of the zygocin pre-sequence to the secretion of IL-1β and comparison of different leader sequences

a) Western Blot analyses performed on cellular extracts (i) and on supernatants (ii) of Z. bailii and S. cerevisiae cells transformed with the plasmid pZβ1IL-1β and with the corresponding empty plasmid pZβ and growing on YPD medium (glucose 2% w/v). Samples were taken at the indicated times. First lane: positive control (IL-1β), human recombinant (E. coli, Roche cat n° 1 457 756). The blotted membranes were probed with an α-IL-1β polyclonal antibody.

Western Blot analyses performed on cellular extracts (iii) and on supernatants (iv) of Z. bailii and S. cerevisiae cells transformed with the plasmid pZβ1IL-1β and with the corresponding empty plasmid pZβ and growing on YNB medium (glucose 5% w/v). Samples were taken at the indicated times. The blotted membranes were probed with an α-IL-1β polyclonal antibody.

b) Western Blot analyses performed on supernatants of Z. bailii cells growing on YNB medium (glucose 2% w/v) transformed with the indicated plasmids. The blotted membranes were probed with an α-IL-1β polyclonal antibody.

FIG. 12: Glucoamylase Sta2 activity in transformed Z. bailii or S. cerevisiae cells, respectively

Determination of the S. cerevisiae var. diastaticus glucoamylase Sta2 activity (U/OD) in the growth medium (YNB, fructose 2% w/v) of Z. bailii and S. cerevisiae cells transformed with the plasmids pZ, STA2 and pZ, KIST2 and the respective empty plasmid pZΔ, as a control (as indicated). In the first plasmid the protein is lead to secretion from its own leader sequence, in the second from the K. lactis killer toxin pre-leader sequence. Measurements were repeated more times and on independent clones, and variation levels are indicated with error bars.

EXAMPLES

The following examples are included to demonstrate preferred embodiments of the invention. It should be appreciated by those of skill in the art that the techniques disclosed in the examples which follow represent techniques discovered by the inventors to function well in the practice of the instant invention, and thus can be considered to constitute preferred modes for its practice. However, those of skill in the art should, in light of the present disclosure, appreciate that many changes can be made in the specific embodiments which are disclosed and still obtain a like or similar result without departing from the spirit and scope of the invention.

Example 1

Construction of Z. bailii Expression Plasmids

The Backbone of the new vector pZ3 (FIG. 1a) is the basic S. cerevisiae expression plasmid YX022 (R&D Systems, Inc., Wiesbaden, D).

The ARS1-CEN4 fragment was taken from Yeplac33 (ATCC 87623, Genbank accession no. X75456 L26352), it was cut Clal-blunt/Spel and cloned into pX022 opened DraI-blunt/Spel (this way the plasmid lost completely the HIS gene).

The plasmid obtained was opened KpnI-blunt, and here the Kan cassette, derived from pFA6-KanMX4 (Wach et al., 1994 Yeast 10, 1793-1808) was inserted. The respective fragment was taken out cutting with SphI/SacI-blunt. This kanMX module contains the known kan’ open reading-frame of the E. coli transposon Tn903 fused to transcriptional and translational control sequences of the TEF gene of the filamentous fungus Ashbya gossypii (e.g. NRRL Y-1056). The described hybrid module permits efficient selection of transformants resistant against geneticin (G418).

The expression cassette based on the constitutive S. cerevisiae TPI promoter and the respective polyA, interspaced by the multi cloning site (MCS), as indicated in the Figure derives from the original pYX022 plasmid (see supplier’s information). All the other plasmids indicated in the FIGS. 1 to 4 derive from pZ3.

For the construction of the plasmid pZKl (FIG. 1b), the signalling pre-sequence (16 aa) of the alpha-subunit
of the K1 killer toxin of K. lactis (Stark M. J. et al., 1986, EMBO J. 5, 1995-2002) was functionally linked to the TPI promoter of the pZ3 plasmid, in order to lead the secretion of the protein of interest.

For the construction of the plasmid pZ3pppa (FIG. 1a), the pre-pro-α-factor signal sequence was similarly utilised and functionally inserted. The sequence was taken from the plasmid pPICZαA (Invitrogen BV, The Netherlands).

For the construction of the plasmid pZ3kIII-1β (FIG. 2a), the coding sequence for the protein already fused with the killer toxin of K. lactis signal sequence was taken cutting Xhol/EcoRI-blunted from the plasmid pCXJ-kan (Fleer R. et al., 1991, Gene 107, 285-95) and sub-cloned into the plasmid pZ3 EcoRI blunted and de-phosphorylated.

For the construction of the plasmid pZ3pppeGFP (FIG. 2c), the fragment containing the α-factor pre-pro leader sequence in frame with the GFP coding sequence was cut HindIII blunted/BamHI from the plasmid pPICAGFP and sub-cloned in the plasmid pZ3 opened EcoRI blunted/BamHI and de-phosphorylated. The plasmid pPICAGFP was constructed according to Passolunghi, S., et al. by introduction of a PCR amplified GFP sequence in frame into the plasmid pPICZαA (Invitrogen BV, The Netherlands). The PCR technique is known in the art. Exemplary reference is made to Gelfand, D. H., et al., PCR Protocols: A Guide to Methods and Applications, 1990, Academic Press and Dianne, C.


For the construction of the plasmid pZ3pppekIII-1β (FIG. 2b), the IL-1β was PCR amplified from the plasmid pZ3kIII-1β.

The oligos for the amplification are the following:

**Primer: DrdI-IL (SEQ ID NO.: 80)**

5' AACAGATCCCAACGCGGCCAGCTG 3' * Tm: 63°C.

**Primer: IL-C-TERM (SEQ ID NO.: 81)**

5' AACAGATTAGAAGGACAAAATCTCACGTGG 3' * Tm: 61°C.

The following program was used for the amplification:

<table>
<thead>
<tr>
<th>Temperature (°C)</th>
<th>Time (s)</th>
<th>Cycle</th>
</tr>
</thead>
<tbody>
<tr>
<td>94</td>
<td>5</td>
<td>10 cycles</td>
</tr>
<tr>
<td>94</td>
<td>45</td>
<td></td>
</tr>
<tr>
<td>72</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>94</td>
<td>45</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>45</td>
<td>20 cycles</td>
</tr>
<tr>
<td>72</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>72</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>30</td>
<td></td>
</tr>
</tbody>
</table>

In this way a DrdI cutting site for sub-cloning the coding sequence of the IL-1β protein in frame with the α-factor pre-pro leader sequence was introduced. The plasmid pZ3pppeGFP was opened EcoRI blunted/BamHI.

The PCR fragment was cut DrdI blunted/BamHI. Combination resulted in the plasmid pZ3pppeGFP.

In the plasmid pZ3kIII-1β, the coding sequence of the interleukin was functionally linked to the deduced pre leader sequence of the Z. bailii killer toxin zygocin (Genebank accession no.: AF515592; Weiler F. et al., 2002, Mol Microbiol. 46, 1095-105.). Essentially oligonucleotides were synthesized corresponding to the deduced leader sequence of the Z. bailii killer toxin zygocin (SEQ ID NO.: 59) and cloned into the plasmid pZ3. Subsequently, the IL-1β was PCR amplified as explicated before and cloned in-frame to the zygocin pre-sequence.

In the plasmid pZ3GAA (FIG. 3a), the coding sequence of the A. adeninivorans α-glucosamylase was cut BamHI blunted from the plasmid pTA32x-β (Bui D. M., et al., 1996, Appl Microbiol. Biotechnol. 45, 102-6) and inserted in the plasmid pZ3 opened EcoRI blunted and de-phosphorylated. For the construction of the plasmid pZ3STA2, the coding sequence of the S. cerevisiae var. diastaticus amylase (comprising its own leader sequence) was cut Xhol/AseI-blunt from the plasmid pMV35 (Vanoni M. et al., 1989, Biochim Biophys Acta 1008, 168-76) and inserted in the plasmid pZ3 opened EcoRI-blunt. For the construction of the plasmid pZ3kISTA2, the coding sequence of the same amylase but functionally linked to the K. lactis killer toxin leader sequence was cut Xhol/AseI-blunt from the plasmid pMV57 (Venturini M. et al., 1997, Mol Microbiol. 23, 997-1007) and inserted in the plasmid pZ3 opened EcoRI-blunt.

For the construction of the plasmid pZ3LacZ (FIG. 3b), the coding sequence of the bacterial β-galactosidase was cut HindIII blunted/BamHI from the plasmid pSUβ-galactosidase (Promega, Inc.) and inserted into the plasmid pZ3 opened EcoRI blunted/BamHI and de-phosphorylated.

In the plasmid pZ3lgt (FIG. 4a), the TPI promoter of S. cerevisiae was substituted with the endogenous TPI promoter from Z. bailii. The sequence was PCR amplified from the genomic DNA of the Z. bailii strain ISA 1307, and the primers were designed according to the literature (Merico A., et al., 2001, Yeast 18, 775-80). Extraction of genomic DNA was performed according to the protocol published by Hoffman, C. S., et al., 1987, Gene 57, 267-72).

The oligos for the amplification are the following:

**TPIproβ5 (SEQ ID NO: 82)**

5' ACCCTACACCTACTACACACTACATG 3' * Tm: 59.6°C.

**TPIproβ2 (SEQ ID NO: 83)**

5' TTGGATACGACGCTGACGAGGATGTC 3' * Tm: 59.6°C.

The following program was used for the amplification:

<table>
<thead>
<tr>
<th>Temperature (°C)</th>
<th>Time (s)</th>
<th>Cycle</th>
</tr>
</thead>
<tbody>
<tr>
<td>94</td>
<td>5</td>
<td>25 cycles</td>
</tr>
<tr>
<td>94</td>
<td>45</td>
<td></td>
</tr>
<tr>
<td>57</td>
<td>45</td>
<td></td>
</tr>
<tr>
<td>72</td>
<td>1 min 30 s</td>
<td></td>
</tr>
<tr>
<td>72</td>
<td>7 min</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>30</td>
<td></td>
</tr>
</tbody>
</table>
[0135] The PCR fragment was sub-cloned into the vector pST-Blue1 (Novagen, Perfect Blunt cloning Kit cat. no. 70191-4), according to the included protocol. Therefrom, the promoter was cut SnaBI/SacI and sub-cloned into the pZ2 opened AatII blunted/SacI (so to remove the S. cerevisiae TPI promoter), obtaining the desired plasmid.

[0136] For the construction of the plasmid pZ2bTLacZ (FIG. 4b), the coding sequence of the bacterial β-galactosidase was cut HindIII/BamHI blunted from the plasmid pSV-β-galactosidase (Promega, Inc.; Genbank accession no.: X65335) and inserted into the plasmid pZ2bT opened Nhel blunted and de-phosphorylated.

[0137] In the plasmid pZ2rG, the TPI promoter of S. cerevisiae was substituted with the GAPDH promoter from Z. rouxii. The sequence was PCR amplified from genomic DNA of the Z. rouxii strain LST 1, and the primers were designed according to the literature (Ogawa Y. et al., 1990, Agric Biol Chem. 54, 2521-9). Extraction of genomic DNA was performed according to the protocol previously mentioned. (Another possible strain is Z. rouxii NRRL Y-229.)

[0138] The oligos for the amplification are the following:

pErzAPDH fwd (SEQ ID NO.: 93) 5' TGCAAAAGGCCCCTAGGCT 3' Tm: 60.3° C.
pErzGAPDH rev (SEQ ID NO.: 94) 5' TACACTAGTATTTTATTATTAGNG 3' Tm: 59.2° C.

[0139] The following program was used for the amplification:

| 94° C. | 5 min |
| 94° C. | 15 s  |
| 57° C. | 30 s  |
| 72° C. | 45 s  |
| 72° C. | 7 min |
| 4° C.  | 10 min |

25 cycles

[0140] The obtained PCR fragment (708 bp) was sub-cloned into the vector pST-Blue1 (Novagen, Perfect Blunt cloning Kit cat. no. 70191-4), according to the included protocol. Therefrom, the promoter was cut SnaBI/SacI and sub-cloned into the pZ2 opened AatII blunted/SacI (so to remove the S. cerevisiae TPI promoter), obtaining the desired plasmid.

[0142] For the construction of the plasmid pZ2aLacZ (FIG. 4b), the coding sequence of the bacterial β-galactosidase was cut HindIII/BamHI blunted from the plasmid pSV-β-galactosidase (Promega, Inc.; Genbank accession no.: X65335) and inserted into the plasmid pZ2aG opened XhoI blunted and de-phosphorylated.

[0143] DNA manipulation, transformation and cultivation of E. coli (DH5α), were performed following standard protocols (Sambrook J., et al., Molecular Cloning: A Laboratory Manual, 2nd edn., Cold Spring Harbor Laboratory, New York, 1989). Also other basic molecular biology protocols were performed following this manual if not otherwise stated. All the restriction and modification enzymes utilised are from NEB (New England Biolabs, UK) or from Roche Diagnostics.

Example 2

Transformation of Z. bailii

[0144] Transformations of all the Z. bailii and the S. cerevisiae NRRL Y-30230 strains were performed basically according to the LiAc/PEG/ss-DNA protocol (Agapek, R., et al., 1998, Transformation of Saccharomyces cerevisiae by the lithium acetate/single-stranded carrier DNA/polycytidylylne glycol (LiAc/ss-DNA/PEG) protocol. Technical Tips Online (http://ttt.trends.com)). After the transformation, Z. bailii cells were recovered with an incubation of 16 hours in YP medium, comprising 2% w/v of fructose as carbon source (YPF), and 1 M sorbitol, at 30° C. The cell suspension was then plated on selective YPF plates with 200 mg/l G418 (Gibco BRL, cat. 11811-031). Single clones appeared after 2-3 days at 30° C. From then on the transformants were grown either in rich or in minimal medium having glucose as carbon source and 200 mg/l G418 for maintenance of the selection. For S. cerevisiae cells, the procedure was the same, except for the carbon source, that remained glucose in all the steps, and for the G418 concentration, optimised for our strain to 500 mg/l.

Example 3

Expression and Secretion of Interleukin 1-β in Z. bailii

[0145] In order to check the secretory capability of the yeast Z. bailii and to compare it with the well known host S. cerevisiae, both yeasts were transformed (according to Example 2) with the plasmid pZ2klll-1β (FIG. 2a). Independent transformants were shake flask cultured in minimal medium (YNB, 1.34% w/v YNB from Difco Laboratories, Detroit, Mich. #919-15, 5% w/v Glucose, complemented with Histidine, Uracil and Leucine, FIG. 5a, left panel) or in rich medium (YPD, 5% w/v Glucose, 2% w/v Peptone, 1% w/v Yeast extract, FIG. 5a, right panel). FIG. 5a shows the cell density (OD 660 nm) and the glucose consumption during the kinetics of growth. The glucose consumption was determined using a commercially available enzymatic kit from Boehringer Mannheim GmbH, Germany (Cat #716251), according to the manufacturer’s instructions. During the kinetics, samples were collected at the indicated times (see “hours” of FIG. 5b, c, d). Cells were harvested (a culture volume corresponding to 10⁶ cells) by centrifugation (10 min 10,000 rpm). 1 volume 2x Laemmlı Buffer (Laemmli, U.K., 1970, Nature 227, 680-5) was added to the supernatants of said samples, they were boiled 3-5 minutes and stored at -20° C. until loading or loaded directly on a polyacrylamide gel.

[0146] The cell pellets of said samples were resuspended in 5 ml 20% TCA, centrifuged (10 min at 3000 rpm) and the resulting pellets were resuspended in 150 µl 5% TCA. Samples were subsequently centrifuged for 10 min at 3000 rpm, and the pellet was resuspended in Laemmlı Buffer (100 µl). In order to neutralise the samples, 1 M Tris base was added (50 µl). After 3-5 min at 99° C. the samples are ready to be loaded on a polyacrylamide gel (alternatively, they can be stored at -20° C.).

[0147] Samples were loaded on standard polyacrylamide gels (SDS-PAGE, final concentration of the separating gel: 15%); after protein separation, gels were blotted (1 h, 250
mA) to nitrocellulose membranes (protran BA 85, Schleicher & Schuell). Immunodetection: after 1 h (RT) of saturation in TBS 1× (1.2 g/l Tris base; 9 g/l NaCl)+5% NFM (non fat milk). 0.2% Tween-20, the membranes were incubated overnight at 4°C with the primary antibody against interleukin (mabbi polyclonal antibody IL-1p(H-153) from Santa Cruz Biotechnology, Inc. cat. no. sc-7884) diluted 1:200 in TBS 1× (1.2 g/l Tris base; 9 g/l NaCl)+5% NFM. After intensive and repeated washes in TBS+0.2% Tween-20, the secondary antibody (antirabbit IgG horseradish peroxidase-conjugated, Amersham Biosciences, UK cat no. NA934) was added (1:100000 in TBS 1×+5% NFM) and left in incubation for 1 h (RT). The proteins were visualised using ECL Western Blotting System (Amersham Biosciences, UK) according to the manufacturer’s instructions.

[0148] The data obtained by Western Blot performed on the supernatant highlight the surprisingly good secretory capability of Z. bailii cells (see FIG. 5c), both in minimal and in rich medium. Remarkably, the signal corresponding to the secreted protein is significantly more intense compared to the signal obtained from S. cerevisiae cells, in agreement with the lower signal revealed in Z. bailii crude cell extracts (FIG. 5b). Moreover, the difference in the secreted levels of proteins is even more pronounced in minimal medium respect than in rich medium (for a comparison: FIG. 5c, left and right panel). These conclusions can be done either considering samples loaded rectifying the OD (FIG. 5c) or either considering equal volumes of loaded samples (FIG. 5f).

[0149] Similarly, Z. bailii and S. cerevisiae cells were transformed with the plasmid pZpHIS-L-β. In this case the same protein (interleukin) is functionally fused with the leader sequence of the S. cerevisiae α-factor pheromone. As previously described, cells were shake flask cultured in rich YPD or in minimal YNB medium, samples were collected and prepared for protein SDS-PAGE separation.

[0150] The Western Blot (FIG. 6a) once more revealed the surprisingly better secretion occurring in Z. bailii if compared to S. cerevisiae: the signals obtained from the crude extracts (i for YPD, iii for YNB medium) are more intense in the latter strain, suggesting that the product is shorter retained and therefore more efficiently secreted from Z. bailii cells. This observation is consistent with the fact that the signals corresponding to the product secreted into the medium are more intense in Z. bailii samples than in S. cerevisiae ones (ii for YPD, iv for YNB medium; in this case a positive signal is present only in Z. bailii samples).

[0151] Importantly, the process of expression, secretion and accumulation of heterologous proteins in the culture medium can be obtained not only by changing the leader sequence, but also by utilising the same leader sequence but changing the heterologous protein expressed. Z. bailii cells were transformed with the plasmid pZpHIS-GFP, shake flask cultured in minimal YNB medium, samples were collected and prepared for protein SDS-PAGE separation. The Western Blot analyses performed as previously described, except for the primary antibody utilised (anti-GFP, Clontech, Inc.) and its concentration (1:500), show a band of the expected dimension that is present only in the supernatant of the Z. bailii cells expressing the GFP heterologous protein (FIG. 6b) and not in the control strain, transformed with the empty plasmid.

[0152] The data obtained underline the possibility to utilise Z. bailii as a host for the process to express different heterologous proteins and to secrete them, leading the secretion with heterologous leader sequences. Remarkably, the level of secreted proteins is higher compared with the levels obtained in S. cerevisiae, and the difference is even more pronounced, in chemically defined culture medium.

Example 4

Expression and Secretion of Interleukin 1-β in a Z. bailii Bioreactor Batch Cultivation with High Sugar Concentration.

[0153] Z. bailii cells transformed (according to Example 2) with the plasmid pZpHIS-L-β (FIG. 2a) and previously analysed for interleukin 1-β production in shake flask culture (see Example 3), were batch cultivated in a 2.1 laboratory bioreactor (fermentor, Biolafitte & Moritz, Mod. Predex—France) in a chemically defined medium with high glucose content (27% w/v Glucose, 4% w/v (NH₄)₂SO₄, 0.4% w/v MgSO₄, 2.4% w/v KH₂PO₄, vitamins according to Verduyn, C., et al., 1992, Yeast 8, 501-17, wherein the final concentration of vitamins was set to be 24 times in respect to the indicated concentrations and trace elements according to Verduyn, C., et al., 1992, Yeast 8, 501-17, wherein the final concentration of trace elements was also set to be 24 times in respect to the indicated concentrations. (Depending on the salt tolerance of the production strain it might be useful in this context to add only a partial quantity of the salts with the glucose to the initial medium and to add the rest of the salts after the bioreaction (fermentation) has proceeded a sufficient amount of time.) The pH control (value: pH 5) is performed by the addition of 2M KOH, G418 was added to a concentration of 200 mg/l G418, antifoam was added as necessary). The inoculum was prepared by pre-growing the yeast in shake flask (with a headspace-to-culture volume ratio of 4) in YPD rich medium (see above), with the addition of 200 mg/l G418. Cells were harvested, washed with deionised water and inoculated in the final medium at OD 1.68 in the bioreactor. Cell culture was flushed with 90 l/h of air and the dissolved oxygen concentration was maintained at 40% of air saturation, varying the stirrer speed. FIG. 7a shows the growth kinetics (cell density, OD 600 nm), together with the glucose consumption, the ethanol production and the biomass produced (dry weight g/l). The glucose consumption and the ethanol production were determined by using commercial enzymatic kits (Boehringer Mannheim GmbH, Germany, Kits Cat #116251 and 0176290, respectively), according to the manufacturer’s instructions. The determination of the cellular dry weight (biomass) was performed as described before (Rodrigues, F. et al., 2001, Appl. Environ. Microbiol. 67, 2123-8). Samples were collected at the indicated times and prepared for protein SDS-PAGE separation. The Western Blot analysis (performed as described in Example 3) shows a very strong and clean signal accumulating during time corresponding to the secreted product (lanes 2 to 5), and confirms the minimal retention of heterologous protein produced within the cells (lanes 6 to 9, FIG. 7b). This example shows the surprising and advantageous characteristic of Z. bailii cells to be able to grow as well as express and secrete a heterologous protein even at very high sugar concentrations. Reportedly S. cerevisiae does not grow any more or can
Example 5

**Procedures and Secretion of Glucoamylase in Z. bailii**

**[0154]** Z. bailii cells were transformed (according to Example 2) with the plasmid pZ2GAA (Fig. 3), and with the empty plasmid pZ2, as a control. Independent transformants were shake flask cultured in minimal NYB medium with 2% w/V Glucose as a carbon source (0.67% w/V YNB and an, according to the manufacturer’s protocol) till mid-exp phase (also referred to as mid-log). The β-glucosamylase activity was determined as follows: after cell density determination, the cells were harvested in order to rescue the culture supernatant. 15 μl/ml 3M NaAc, pH 5.2 and 20 μl/ml 1% w/V Starch (Fluka 85642—high solubility—) were added. Subsequently, the samples were mixed well and incubated at the desired temperature (this experiment: 30°C). At time zero and every following 20 min, 1 ml of the incubated medium is taken, ice-cooled for 2 min, 50 μl of Lugol solution (Fluka 62659) were added, shaken quickly and read at the spectrophotometer at λ580 nm. The slope of the resulting values corresponds to the glucoamylase activity. Fig. 8 shows the glucoamylase activity of three independent clones expressing the GAA and one negative control. The enzymatic activity is expressed in mU/Od, and it is calculated considering that 1U corresponds to the variation of 1 OD in 1 min. The values reported in the graphic were subtracted from the basic activity level of Z. bailii, as measured in the control sample.

**[0155]** Z. bailii and S. cerevisiae cells were transformed (according to Example 2) with the plasmids pZ2,STA2 and pZ2,kiSTA2, and with the empty plasmid pZ2, as a control. Independent transformants were shake flask cultured in minimal NYB medium with 2% w/V fructose as a carbon source (0.67% w/V YNB and an, according to the manufacturer’s protocol) till mid-exp phase (also referred to as mid-log). The α-glucosamylase activity was determined according to the literature (Modena et al., 1986, Arch of Biochem. And Biophys. 248, 138-50) as follows: after cell density determination, the cells were harvested in order to rescue the culture supernatant, and an aliquot of said supernatant is used for preparing the following reaction mix:

<table>
<thead>
<tr>
<th>Component</th>
<th>Volume (μl)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Supernatant</td>
<td>100 μl</td>
</tr>
<tr>
<td>Maltotriose 400 mM</td>
<td>6.3 μl</td>
</tr>
<tr>
<td>NaAc 200 mM pH 4.6</td>
<td>125 μl</td>
</tr>
<tr>
<td>H2O</td>
<td>18.7 μl</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>250 μl</strong></td>
</tr>
</tbody>
</table>

The mix is incubated for 1 hour at 37°C under slow agitation, and after that time an aliquot of said mixture is used to evaluate the reaction. The product of maltotriose degradation is glucose, and its concentration can be determined using a commercially available enzymatic kit from Boehringer Mannheim GmbH, Germany (Cat #716251).IU of glucoamylase specific activity is the quantity of enzyme necessary to release 1 μmol min⁻¹ of glucose in said condition.

Example 6

**Expression of β-galactosidase (β-gal) in Z. bailii**

**[0157]** Z. bailii cells were transformed (according to Example 2) with the plasmid pZ1,1acZ (Fig. 3b), with the plasmid pZ2,TLacZ (Fig. 4a), with the plasmid pZ3GlacZ, and with the empty plasmid pZ2, as a control. Independent transformants were shake flask cultured in YPD medium (see description above) with 2% w/V Glucose as a carbon source till mid-exp phase. β-galactosidase activity determination: after cell density determination, 1 ml culture is harvested into an eppendorf tube, spun for 5 minutes (to get a hard pellet), aspirated with a pipet, (not using the vacuum line!), washed in 1 ml Z buffer [w/o BME—betamercaptoethanol—; Z buffer: 16.1 g/L NaHPO4·2H2O, 5.5 g/L NaH2PO4·H2O, 0.75 g/L KCl, 0.246 g/L MgSO4·7H2O], repelleted, suspended in 150 μl Z buffer (with BME, 27 μl/10 ml), 50 μl chloroform are added, 20 μl 0.1% SDS and vortexed vigorously for 15 sec. 700 μl of pre-warmed ONPG (o-nitrophenyl β-D-galactopyranoside, Sigma N-1127, 1 mg/ml in Z+5BME) are added, and the reaction is started at 30°C (20 to 3 hr), checking the time. When the suspension turns yellow the reaction is stopped by addition of 0.5 ml of 1 M NaCO3; after centrifugation for 10 min at maximum speed the sample is read at the spectrophotometer at λ420.

**[0158]** Fig. 8b shows the β-gal activity of three independent clones expressing the β-gal under control of the Z. bailii TPI promoter, two independent clones expressing the β-gal under control of the S. cerevisiae TPI promoter and one negative control (see the legend of the figure for indications of the respective clones). The enzymatic activity is expressed as Miller Unit/OD and it is calculated according to the following formula:

$$\text{Miller Units} = \frac{A_{420} \times 1000}{A_{660} \times \text{time (min)} \times \text{Vol(mL)}}$$

**[0159]** As it is readily visible, the expression from the endogenous TPI promoter is much stronger (4-5 times) than from the respective promoter from S. cerevisiae.

**[0160]** A similar series of experiments was performed in order to evaluate the efficiency of the plasmids based on the sequences of the endogenous Z. bailii plasmid in improving the expression levels of heterologous proteins. Z. bailii cells were transformed (according to Example 2) with the following plasmids: pZ2, LacZ (Fig. 3b), p195LacZ, pEZ, LALacZ, pEZ-LAFLacZ, pEZ,LacZ and pEZ,IL3LacZ. Independent transformants were grown till mid-log phase and β-galactosidase activity measured, as previously described. The corresponding data are reported in Fig. 10b.

Example 7

**Isolation of an Endogenous Z. bailii Plasmid**

**[0161]** Z. bailii strains ATCC 36947 and NCYC 1427 were cultivated and their endogenous plasmid was extracted, resulting in the plasmids pZ2, and pZ3 (see Fig. 8a and b). The protocol used was a modification of a protocol by Lorincz, A., 1985, BRI. Focus 6, 11, and uses glass beads to
break the cells. After the DNA extraction, samples were loaded on an agarose gel and the band corresponding to the plasmid was eluted (Qiagen, QIAquick Gel Extraction Kit cat no 28704). The plasmid extracted from NCYC 1427 was cut with EcoRI and some of the fragments were sequenced. These sequences correspond to SEQ ID No.: 63, SEQ ID No.: 64, SEQ ID No.: 65, SEQ ID No.: 66, SEQ ID No.: 67, SEQ ID No.: 68, SEQ ID No.: 69 or SEQ ID No.: 70, respectively.

Example 8

Sequence Amplification of the Open Reading Frames and of Structural Sequences of the Endogenous Z. bailii Plasmids

[0162] The genomic DNA extracted from the Z. bailii strains ATCC 36947 and NCYC 1427 were used as a template for the amplification of the open reading frames and of structural sequences of the endogenous Z. bailii plasmids.

[0163] The oligos for the amplification are the following:

5FLP (SEQ ID NO. : 84) 5'-TAGCTATCCTCTCCGAGGCTCAGTACCAGATGTC-3' Tm: 63.4
3FLP (SEQ ID NO. : 85) 5'-CCATATCCGGCAAGGCTTCGAGGCTCCAG-3' Tm: 64.6
5TPC (SEQ ID NO. : 86) 5'-AAGATCAACTTACAGGACTGCTTCGAGGCTCAGTACCAGATGTC-3' Tm: 59.7
3TPC (SEQ ID NO. : 87) 5'-ATTCTGATGAGATATTACGCCGAC-3' Tm: 58.4
5TFB (SEQ ID NO. : 88) 5'-GTTTTTTTTTGACGTCAGCGCTATTTTTACG-3' Tm: 58.6
3TFB (SEQ ID NO. : 89) 5'-ATTCTGATGAGATATTACGCCGAC-3' Tm: 61.6
5IRAAG (SEQ ID NO. : 90) 5'-AGAACTTACAGGACTGCTTCGAGGCTCAGTACCAGATGTC-3' Tm: 61.9
3IRAAG (SEQ ID NO. : 91) 5'-TAAAAACTGCGCCGCAATTTACGCTC-3' Tm: 61.3

[0164] The following program was used for the amplification:

94° C. 5 min
94° C. 15 s
58° C. 30 s
72° C. 2 min
72° C. 7 min
4° C. ∞

}\{ 25 cycles

[0165] The amplified fragments, sub-cloned into the vector pST-Blue1 (Novagen, Perfect Blunt cloning Kit cat. no. 70191-4), were sequenced and correspond to SEQ ID No.: 71 (IR-ARS), SEQ ID No.: 72 (FLP), SEQ ID No.: 74 (TFB) and SEQ ID No.: 76 (TFC), respectively.

[0166] These coding sequences are used for the construction of the expression plasmid pEZL, according to FIG. 9b.

Example 9

Construction of Expression Plasmids Based on Replication and Stability Sequences from the Z. bailii pSB2 Plasmid

[0167] The backbone of the new vectors is the basic S. cerevisiae multicopy plasmid YEp2195 (Gietz and Sugino, 1988, Gene 74, 527-34) modified to the expression plasmid pBR195, as described in Branduardi (2002, Yeast 19, 1165-70).

[0168] For the construction of the plasmid p195, the plasmid pBR195 was cut at AatII/Apal-blunt in order to excise the URA marker and the KanR cassette, excised SphI/SalI-blunt from pFA6-KanMX4 (Wach et al., 1994 Yeast 10, 1793-1808) was here inserted. From this plasmid derives the plasmid p195LacZ: the LacZ gene was sub-cloned from the plasmid pZLacZ cut SphI/Nhel into the new plasmid p195, opened with the same enzymes.

[0169] For the construction of the plasmids pEZ-1A and pEZ-IA/LacZ, the plasmids p195 and p195LacZ were opened NsiI/Hind-blunt, in order to remove the S. cerevisiae 2 μm ori. The PCR fragment corresponding to the IR-A and ARS sequence from the pSB2 (see previous example for amplification detail) was excised EcoRI-blunt from the pST-Blue1 plasmid and sub-cloned into the opened vectors just described.

[0170] For the construction of the plasmid pEZ-IA/LacZ, the plasmid pEZ-IA/LacZ was Smal opened, and there the fragment corresponding to the FLP and the sequence containing its promoter, derived from the pST-Blue1 plasmid opened AccI-blunt/SmaI, was there sub-cloned. Said sequence was PCR amplified from the genomic DNA extracted from the Z. bailii strains ATCC 36947.

[0171] The oligos for the amplification are the following:

3FLP (SEQ ID NO. : 85) 5'-CCATATCCGGCAAGGCTTCGAGGCTCCAG-3' Tm: 64.6

[0173] The following program was used for the amplification:

94° C. 5 min
94° C. 15 s
58° C. 30 s
72° C. 1 min 30 s
72° C. 7 min
4° C. ∞

}\{ 29 cycles

[0174] For the construction of the plasmids pEZ2 and pEZ2/LacZ, the plasmids pEZ-1A and pEZ-IA/LacZ were opened Smal and the PCR fragment corresponding to the sequences of FLP and TFC and the respective promoters was excised SnaBII/SalI-blunt from the pST-Blue1 plasmid and sub-cloned into the opened vectors just described.
The oligos for the amplification are the following:

5’-TAGGGCGCTCTGGGGTGGGCTAGGAG-3’ Tm: 61.4
5’-ATGGATCGGATGCTCGGGGCTAGGAG-3’ Tm: 58.4

The following program was used for the amplification:

\[
\begin{array}{ccc}
94^\circ C. & 5 \text{ min} \\
94^\circ C. & 15 \text{ s} \\
58^\circ C. & 30 \text{ s} \\
72^\circ C. & 1 \text{ min 30 s} \\
72^\circ C. & 7 \text{ min} \\
4^\circ C. & \infty \\
\end{array}
\]

28 cycles

For resulting in the plasmid pEZ2 an additional cloning step was required, in order to re-insert the polyA; the polyA was excised NaeI/NheI-blunt from the plasmid pYX022 and was sub-cloned in the transitory plasmid BamHI-blunt and de-phosphorylated.

For the construction of the plasmid pEZ2-IBLacZ, the plasmid pEZ2LacZ was opened SalI-blunt and de-phosphorylated, and the fragment IR-B was therein subcloned. That fragment was EcoRI-blunt extracted from pST-Blue1 (see previous example).

Example 10

Plasmid Stability Determination

The stability of the plasmids described in the previous example was determined as follows: independent Z. bailii transformants bearing the different plasmids were inoculated at a cellular density of 5x10⁵ cells/ml in rich media (YPD) and in rich selective media (YPD+G418), respectively. At T₀ of the inoculum and then after 10 and 20 generations, 500 cells from any culture were plated 3 times on selective and non-selective agar plates, and subsequently incubated at 30°C. till the colonies became visible. The ratio between the mean of the colony number grown on selective medium and the mean of the colony number grown on non-selective medium gives the percentage of mitotic stability.

**SEQUENCE LISTING**

```plaintext
<160> NUMBER OF SEQ ID NOS: 95
<210> SEQ ID NO 1
<211> LENGTH: 57
<212> TYPE: DNA
<213> ORGANISM: Saccharomyces cerevisiae

<400> SEQUENCE: 1
cgcacgttaa tttgccgaa tttaataagct tcacttggag aacagtggag cacgtgga ccagtggac 57
```

```plaintext
<210> SEQ ID NO 2
<211> LENGTH: 19
<212> TYPE: PRT
<213> ORGANISM: Saccharomyces cerevisiae

<400> SEQUENCE: 2
Metr Arg Phe Pro Ser Ile Phe Thr Ala Val Leu Phe Ala Ala Ser Ser
1 5 10 15
```

```plaintext
<210> SEQ ID NO 3
<211> LENGTH: 255
<212> TYPE: DNA
<213> ORGANISM: Saccharomyces cerevisiae

<400> SEQUENCE: 3
cgcacgttaa tttgccgaa tttaataagct tcacttggag aacagtggag cacgtgga ccagtggac 57
```

```plaintext
gtgcactotc attatatca acataaatata caagtggac ctacttggag aacagtggag cacgtgga ccagtggac 57
```

```plaintext
120
```

```plaintext
gtcgagaaaa accaanagag accanagtct cgtggtgtaag tacataata aagggccttt
```

```plaintext
180
```

```plaintext
tcgcccaaat ctcataatttt atcgtgcttt cgtgtctttt acgtgcaaat aatagtggag
```

```plaintext
240
```

```plaintext
acaacagatc tacaaccgctt
```

```plaintext
255
```
Met Arg Phe Pro Ser Ile Phe Thr Ala Val Leu Phe Ala Ala Ser Ser
  1      5      10      15
Ala Leu Ala Ala Pro Val Asn Thr Thr Glu Asp Glu Thr Ala Gln
  20     25      30
Ile Pro Ala Glu Ala Val Ile Gly Tyr Leu Asp Leu Glu Gly Asp Phe
  35     40      45
Asp Val Ala Val Leu Pro Phe Ser Asn Asn Ser Thr Asn Asn Gly Leu Leu
  50     55      60
Phe Ile Asn Thr Thr Ile Ala Ser Ile Ala Ala Lys Glu Glu Gly Val
  65     70      75      80
Ser Leu Asp Lys Arg
  85

Met Met Ala Trp Trp Ser Leu Phe Leu Tyr Gly Leu Gln Val Ala
  1      5      10      15
Ala Pro Ala Leu Ala
  20

Met Gln Asn Thr Ala Lys Asn Ser Ile Trp Gln Arg Val Arg His Ser
  1      5      10      15
Ala Ile Ala Leu Ser Ala Leu Ser Leu Ser Phe Gly Leu Gln Ala
  20     25     30
<210> SEQ ID NO 9
<211> LENGTH: 63
<212> TYPE: DNA
<213> ORGANISM: Aspergillus oryzae

<400> SEQUENCE: 9
atgcagtca tcaggtgtac ctggtcttc acagggagtca tcotcctcgg ctggtcgggt 60
gcc

<210> SEQ ID NO 10
<211> LENGTH: 21
<212> TYPE: PRT
<213> ORGANISM: Aspergillus oryzae

<400> SEQUENCE: 10
Met Gln Ser Ile Lys Arg Thr Leu Leu Leu Gly Ala Ile Leu Pro 1 5 10 15
Ala Val Leu Gly Ala 20

<210> SEQ ID NO 11
<211> LENGTH: 75
<212> TYPE: DNA
<213> ORGANISM: Bacillus amyloliquefaciens

<400> SEQUENCE: 11
atgaaacag tgggtgtaat ctggctcacc ggaattttta tgaattttgt ttgggactc 60
tttggtttt cggct

<210> SEQ ID NO 12
<211> LENGTH: 25
<212> TYPE: PRT
<213> ORGANISM: Bacillus amyloliquefaciens

<400> SEQUENCE: 12
Met Lys Arg Val Leu Leu Ile Leu Val Thr Gly Leu Phe Met Ser Leu 1 5 10 15
Cys Gly Ile Thr Ser Ser Val Ser Ala 20 25

<210> SEQ ID NO 13
<211> LENGTH: 51
<212> TYPE: DNA
<213> ORGANISM: Saccharomyces fibuligera

<400> SEQUENCE: 13
atggtatga tggcgagct ggttgctttg ctgggtcgg t gacggtgcc tttgtgtggtct t 51

<210> SEQ ID NO 14
<211> LENGTH: 17
<212> TYPE: PRT
<213> ORGANISM: Saccharomyces fibuligera

<400> SEQUENCE: 14
Met Leu Met Ile Val Gln Leu Leu Val Phe Ala Leu Gly Leu Ala Val 1 5 10 15
Ala
<210> SEQ ID NO 15
<211> LENGTH: 51
<212> TYPE: DNA
<213> ORGANISM: Saccharomyces fibuligera

<400> SEQUENCE: 15
atgttgtga ttttgacact tttagtactt attataggcc ttagagtgc t 51

<210> SEQ ID NO 16
<211> LENGTH: 17
<212> TYPE: PRT
<213> ORGANISM: Saccharomyces fibuligera

<400> SEQUENCE: 16
Met Leu Leu Ile Leu Glu Leu Leu Val Leu Ile Ile Gly Leu Gly Val 1 5 10 15

Ala

<210> SEQ ID NO 17
<211> LENGTH: 66
<212> TYPE: DNA
<213> ORGANISM: Hypocrea pecorina

<400> SEQUENCE: 17
atgcgcccct cagttacact gccgttgacc agggcctacct cggccattgc cggccctgctc 60
gccgcc 66

<210> SEQ ID NO 18
<211> LENGTH: 22
<212> TYPE: PRT
<213> ORGANISM: Hypocrea pecorina

<400> SEQUENCE: 18
Met Ala Pro Ser Val Thr Leu Pro Leu Thr Thr Ala Ile Leu Ala Ile 1 5 10 15

Ala Arg Leu Val Ala Ala 20

<210> SEQ ID NO 19
<211> LENGTH: 63
<212> TYPE: DNA
<213> ORGANISM: Hypocrea pecorina

<400> SEQUENCE: 19
atgacaaagt cgcgtgtcc cttgctgctt gcagcgttca tactatagc cggccctgctc 60
gca 63

<210> SEQ ID NO 20
<211> LENGTH: 21
<212> TYPE: PRT
<213> ORGANISM: Hypocrea pecorina

<400> SEQUENCE: 20
Met Aan Lys Ser Val Ala Pro Leu Leu Ala Ala Ser Ile Leu Tyr 1 5 10 15

Gly Gly Ala Val Ala 20
---continued

<210> SEQ ID NO 21
<211> LENGTH: 48
<212> TYPE: DNA
<213> ORGANISM: Arxula adenivorans

<400> SEQUENCE: 21
atgcgtcagt tgctagcact tgctgtgct tgtccctag cggtgga  48

<210> SEQ ID NO 22
<211> LENGTH: 18
<212> TYPE: PRT
<213> ORGANISM: Arxula adenivorans

<400> SEQUENCE: 22
Met Arg Glu Phe Leu Ala Leu Ala Ala Ala Ala Ala Ser Ile Ala Val Ala
1  5  10  15
Asp Ser

<210> SEQ ID NO 23
<211> LENGTH: 57
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 23
atgcagcgac tgtgtgtgta tggtgtgata tttgcactgg tgtggcgc gcttctctc  57

<210> SEQ ID NO 24
<211> LENGTH: 19
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 24
Met Glu Arg Leu Cys Val Tyr Val Leu Ile Phe Ala Leu Ala Leu Ala
1  5  10  15
Ala Phe Ser

<210> SEQ ID NO 25
<211> LENGTH: 75
<212> TYPE: DNA
<213> ORGANISM: Rhizopus oryzae

<400> SEQUENCE: 25
atgcgaactgt tcaatgtgac attgasagtt tcaaccttgc tgcctccttc tcaacctttct  60
tgctgcttt ctgct  75

<210> SEQ ID NO 26
<211> LENGTH: 25
<212> TYPE: PRT
<213> ORGANISM: Rhizopus oryzae

<400> SEQUENCE: 26
Met Glu Leu Phe Asn Leu Pro Leu Lys Val Ser Phe Phe Leu Val Leu
1  5  10  15
Ser Tyr Phe Ser Leu Leu Leu Val Ser Ala
20  25

<210> SEQ ID NO 27
<211> LENGTH: 40
<212> TYPE: DNA
<213> ORGANISM: Aspergillus niger
atgcagactc tccctgtgac ctcgctttgg gtccctctcg ctgcggcc

Met Gln Thr Leu Leu Val Ser Ser Leu Val Val Ser Leu Ala Ala Ala
1 5 10 15

atgagtgagg taaccttttt tttccttttt tttttttttta gctcgctttta ttcc

Met Lys Trp Val Thr Phe Ile Ser Leu Leu Phe Leu Phe Ser Ser Ser Ala
1 5 10 15

Tyr Ser

tatgaagaac caacccaaagtt attagattag tcgctagta tattatattttt ctaaacatta
tcgtatcctcg tgcgggcg

Met Thr Lys Pro Thr Gln Val Leu Val Arg Ser Ser Val Ile Leu Phe
1 5 10 15

Phe Ile Thr Leu Leu His Leu Val Val Ala
20 25

tatggcact taggctcttt ttcagttttag atgctgtagta tgsatatagc tagcgttga
goac
<210> SEQ ID NO 34
<211> LENGTH: 21
<212> TYPE: PRT
<213> ORGANISM: Saccharomyces cerevisiae

<400> SEQUENCE: 34

Met Gly His Leu Ala Ile Leu Phe Ser Ile Ile Ala Val Leu Asn Ile
1   5   10  15

Ala Thr Ala Val Ala
20

<210> SEQ ID NO 35
<211> LENGTH: 48
<212> TYPE: DNA
<213> ORGANISM: Kluyveromyces lactis

<400> SEQUENCE: 35

ataaatgaat tatattttac atattttttgt ttttctgttc atctgttc
48

<210> SEQ ID NO 36
<211> LENGTH: 17
<212> TYPE: PRT
<213> ORGANISM: Kluyveromyces lactis

<400> SEQUENCE: 36

Met Asn Ile Phe Tyr Ile Phe Leu Phe Leu Leu Ser Phe Val Gln Gly
1   5   10  15

Leu

<210> SEQ ID NO 37
<211> LENGTH: 69
<212> TYPE: DNA
<213> ORGANISM: Kluyveromyces lactis

<400> SEQUENCE: 37

ataaatgaat tatattttac atattttttgt ttttctgttc atctgttcagatgtgggc
60

atactcatc
69

<210> SEQ ID NO 38
<211> LENGTH: 23
<212> TYPE: PRT
<213> ORGANISM: Kluyveromyces lactis

<400> SEQUENCE: 38

Met Lys Ile Tyr His Ile Phe Ser Val Cys Tyr Leu Ile Thr Leu Cys
1   5   10  15

Ala Ala Ala Ala Thr Thr Ala
20

<210> SEQ ID NO 39
<211> LENGTH: 54
<212> TYPE: DNA
<213> ORGANISM: Saccharomyces cerevisiae

<400> SEQUENCE: 39

aggttgct tagcttccat cagctgtttgca atcaaggtttg aggggttttct ggg
54

<210> SEQ ID NO 40
<211> LENGTH: 18
<212> TYPE: PRT
<213> ORGANISM: Saccharomyces cerevisiae

<400> SEQUENCE: 40
Met Phe Ala Phe Tyr Leu Thr Ala Cys Ile Ser Leu Lys Gly Val
1   5   10   15
Phe Gly

<210> SEQ ID NO 41
<211> LENGTH: 54
<212> TYPE: DNA
<213> ORGANISM: Saccharomyces cerevisiae

<400> SEQUENCE: 41
atgtttaagttggttta ggcgttgcctt tagttataagcagttggttta
54

<210> SEQ ID NO 42
<211> LENGTH: 18
<212> TYPE: PRT
<213> ORGANISM: Saccharomyces cerevisiae

<400> SEQUENCE: 42
Met Phe Lys Ser Val Val Tyr Ser Val Leu Ala Ala Ala Leu Val Asn
1   5   10   15
Ala Gly

<210> SEQ ID NO 43
<211> LENGTH: 51
<212> TYPE: DNA
<213> ORGANISM: Saccharomyces cerevisiae

<400> SEQUENCE: 43
atgtttaaactggttta ttcasatitgga ggcgttccctg tggcactgc a
51

<210> SEQ ID NO 44
<211> LENGTH: 17
<212> TYPE: PRT
<213> ORGANISM: Saccharomyces cerevisiae

<400> SEQUENCE: 44
Met Phe Lys Ser Val Val Tyr Ser Ile Leu Ala Ala Ala Ser Leu Ala Asn
1   5   10   15
Ala

<210> SEQ ID NO 45
<211> LENGTH: 48
<212> TYPE: DNA
<213> ORGANISM: Kluyveromyces lactis

<400> SEQUENCE: 45
atgtatatactttcagttgag cttatatcttttactagga ccocatcgcc
48

<210> SEQ ID NO 46
<211> LENGTH: 16
<212> TYPE: PRT
<213> ORGANISM: Kluyveromyces lactis

<400> SEQUENCE: 46
Met Leu Ser Ile Leu Leu Ser Leu Leu Ser Leu Ser Gly Thr His Ala
<210> SEQ ID NO 47
<211> LENGTH: 48
<212> TYPE: DNA
<213> ORGANISM: Kluyveromyces lactis

<400> SEQUENCE: 47  
atgcatatct ttcgtgtggg tttattatca ctatcagga cccatgcgc 48

<210> SEQ ID NO 48
<211> LENGTH: 16
<212> TYPE: PRT
<213> ORGANISM: Kluyveromyces lactis

<400> SEQUENCE: 48  
Met Leu Ser Ile Leu Leu Gly Leu Leu Ser Leu Ser Gly Thr His Ala  
1  5  10  15

<210> SEQ ID NO 49
<211> LENGTH: 54
<212> TYPE: DNA
<213> ORGANISM: Aspergillus niger

<400> SEQUENCE: 49  
atgpgcgtct ctgctgtttct acttcccttg tttctctgt cttggagctcc ctcc 54

<210> SEQ ID NO 50
<211> LENGTH: 18
<212> TYPE: PRT
<213> ORGANISM: Aspergillus niger

<400> SEQUENCE: 50  
Met Gly Val Ser Ala Val Leu Leu Pro Leu Tyr Leu Leu Ser Gly Val  
1  5  10  15

Thr Ser

<210> SEQ ID NO 51
<211> LENGTH: 57
<212> TYPE: DNA
<213> ORGANISM: Saccharomyces cerevisiae

<400> SEQUENCE: 51  
atgttctttgc aagcttcttct tttctttttg gctggttttg cagccaaat atctgca 57

<210> SEQ ID NO 52
<211> LENGTH: 20
<212> TYPE: PRT
<213> ORGANISM: Saccharomyces cerevisiae

<400> SEQUENCE: 52  
Pro Met Leu Leu Gln Ala Phe Leu Phe Leu Leu Ala Gly Phe Ala Ala  
1  5  10  15

Lys Ile Ser Ala

<210> SEQ ID NO 53
<211> LENGTH: 63
<212> TYPE: DNA
<213> ORGANISM: Saccharomyces cerevisiae
<400> SEQUENCE: 53
atgcaagac cattctact cgcttatattg gctcttggc tcttatattaa ctcggctttg
60
ggt
63

<210> SEQ ID NO: 54
<211> LENGTH: 21
<212> TYPE: PRT
<213> ORGANISM: Saccharomyces cerevisiae

<400> SEQUENCE: 54
Met Gln Arg Pro Phe Leu Leu Ala Tyr Leu Val Leu Ser Leu Leu Phe
1 5 10 15
Aas Ser Ala Leu Gly
20

<210> SEQ ID NO: 55
<211> LENGTH: 63
<212> TYPE: DNA
<213> ORGANISM: Saccharomyces cerevisiae

<400> SEQUENCE: 55
atgcaagac cattctact cgcttatattg gctcttggc tcttatattaa ctcggctttg
60
ggt
63

<210> SEQ ID NO: 56
<211> LENGTH: 21
<212> TYPE: PRT
<213> ORGANISM: Saccharomyces cerevisiae

<400> SEQUENCE: 56
Met Gln Arg Pro Phe Leu Leu Ala Tyr Leu Val Leu Ser Leu Leu Phe
1 5 10 15
Aas Ser Ala Leu Gly
20

<210> SEQ ID NO: 57
<211> LENGTH: 96
<212> TYPE: DNA
<213> ORGANISM: Saccharomyces cerevisiae

<400> SEQUENCE: 57
atgtaggagc tcataaatgc attatacgac atagtgcac gacatctttc tctgctttat
60
ttgagctcttt gcgtctatt tacetctgcct ttggt
96

<210> SEQ ID NO: 58
<211> LENGTH: 32
<212> TYPE: PRT
<213> ORGANISM: Saccharomyces cerevisiae

<400> SEQUENCE: 58
Met Val Gly Leu Lys Asn Pro Tyr Thr His Thr Met Gln Arg Pro Phe
1 5 10 15
Leu Leu Ala Tyr Leu Val Leu Ser Leu Leu Phe Aas Ser Ala Leu Gly
20 25 30

<210> SEQ ID NO: 59
<211> LENGTH: 63
<212> TYPE: DNA
---continued---

<213> ORGANISM: Zygosaccharomyces bailii

<400> SEQUENCE: 59

atgaaagac cccaaatatt aacagcaagt atagtaagct tattgccaa atatacagtt 60
gct 63

<210> SEQ ID NO: 60
<211> LENGTH: 21
<212> TYPE: PRT
<213> ORGANISM: Zygosaccharomyces bailii

<400> SEQUENCE: 60

Met Lys Ala Ala Gln Ile Leu Thr Ala Ser Ile Val Ser Leu Pro
1 5 10 15
Ile Tyr Thr Ser Ala 20

<210> SEQ ID NO: 61
<211> LENGTH: 415
<212> TYPE: DNA
<213> ORGANISM: Zygosaccharomyces bailii

<400> SEQUENCE: 61

atgaaagac cccaaatatt aacagcaagt atagtaagct tattgccaa atatacagtt 60
gctagaaaca ttatacaag agataacaa gcaaaagaa taaatatctgc ttattggagat 120
gagagacat ttaaagttt ggcgtactac attcaagttta acgtcagttt gcaattttgc 180
ttcatctt ataatcattt cattttttgct gtaacgcgtc tagaasacoa agatattat 240
agataaat agccatatt ttaacaacag acaaaatata cagagattgag gcctattcgag 300
tactttaaa attcagttt gacgatttgac aagataatcc taactcagta taatttattc 360
gtgaacacta gtactgtat aagcataactc acagacagct atgactccg acgct 415

<210> SEQ ID NO: 62
<211> LENGTH: 139
<212> TYPE: PRT
<213> ORGANISM: Zygosaccharomyces bailii

<400> SEQUENCE: 62

Met Lys Ala Ala Gln Ile Leu Thr Ala Ser Ile Val Ser Leu Pro
1 5 10 15
Ile Tyr Thr Ser Ala Arg Asn Ile Leu Asp Arg Glu Tyr Thr Ala Asn 20 25 30
Glu Leu Lys Thr Ala Phe Gly Asp Glu Glu Ile Phe Thr Asp Leu Thr 35 40 45
Tyr His Ile His Val Asn Val Ser Gly Ile Asp Ser Tyr Tyr His 50 55 60
Asn Leu Val Asn Phe Val Asn Asl Ala Leu Ala Asn Lys Asp Ile Asn 65 70 75 80
Arg Tyr Ile Tyr Ala Ile Phe Thr Gln Gin Thr Asn Tyr Thr Glu Asp 85 90 95
Gly Leu Ile Glu Tyr Leu Asn His Tyr Asp Ser Glu Thr Cys Lys Asp 100 105 110
Ile Ile Thr Gln Tyr Asn Val Val Asp Thr Ser Asn Cys Ile Ser 115 120 125
Asn Thr Thr Asp Gln Ala Arg Leu Gln Arg Arg
130
135

<210> SEQ ID NO 63
<211> LENGTH: 587
<212> TYPE: DNA
<213> ORGANISM: Zygosaccharomyces bailii

<400> SEQUENCE: 63
ttatagctct aaaaataaat aataattt acctaatctca ggttcccccc caaatctctt 60
atatcaata ataagctact ctctcccccc cggagatttt tttaggggga gggggaccct 120
taatcagag gggtgatttt ttttagctca ctggygggagt attataaatg atagtagcct 180
tttgtttta aaaaagcct ccaaaaaata atatctttaa ggttattaat aagttataa 240
tatatatat gttatatcgttt ttagtacttta aagtaattga agatatctctt 300
atagtagctt cgtttctca caaggtacat atatatagc cttatataag tccttctca 360
gaatattga atatatgtg atatataggc aagggctctt ataccttctg cccctaaaaa 420
catatatattt ctaataatttt cttttcatc ttcatttta ctttttacta ctctctctct 480
atagtaaagtc atacaccagt tttttctaaa caacaaaaac cggaggggttc catttttat 540
atatatatcattttctca cttttatatt aaaaacctaa ccgttta 587

<210> SEQ ID NO 64
<211> LENGTH: 435
<212> TYPE: DNA
<213> ORGANISM: Zygosaccharomyces bailii

<400> SEQUENCE: 64
ttatattct atgctgctaa ataataagt tttatatatt ataataatttt caaccaacca 60
cococococca aaaaagaggt cctttctgtg gctgctcocca accacccctt cgggttctgtg 120
cococatag gtttctctat taaataataa aaaaactttta tggagggaco taaaagaaat 180
aataggaag taaaattataa taaggctcc ccttttttct cttctctoc caacccctaa 240
aatatcctct ggagggggag ggaggaagtt tattagtgtt ggagggtgta agttataaat 300
agagattt ataatgattat aaaaatcata ataatcgtt cttattaata ataataattat 360
taacagatat aagccacaag gtcgaggagct ttcttgggga gaaagatgta gtttaattga 420
atcatctta cttga 435

<210> SEQ ID NO 65
<211> LENGTH: 299
<212> TYPE: DNA
<213> ORGANISM: Zygosaccharomyces bailii

<400> SEQUENCE: 65
tttggggtct ggggtctgtg aaaaacctaa atgggacccc taagaggggc cccctatta 60
gagtctgagc ggggaaagcc cggagcctgg cgggaaggga aggggaaaaa cggagggag 120
cggggctag gggggctgca gttgagtgatg ttaagctcgc ctttaaccoc ccaccggcgg 180
cgagatagc gcggctgcag gggggccagt ggggtcttttt cgggggat tggggggtga 240
cccctattt gttatatccc taaaaataat cctggtactg agagatct 299

<210> SEQ ID NO 66
<211> LENGTH: 153

---continued---
-continued

<212> TYPE: DNA
<213> ORGANISM: Zygosaccharomyces bailii

<400> SEQUENCE: 66

cctcacgtga acattccccg tctgcctttt ccacattttc attttacaat gagaagtttt
60
cggaaaaaa aattcaaatg gctgatggc aacccttccc atgttctgtg agccctcaat
120
gttttccto tctccagagt caggtttctg tgt
153

<210> SEQ ID NO: 67
<211> LENGTH: 231
<212> TYPE: DNA
<213> ORGANISM: Zygosaccharomyces bailii

<400> SEQUENCE: 67

aattcaatag atgagctgtt aacctgcttttt gtagatggag attaaggaatttctggggtata
60
tttgaagt tggatcataaa accagttataa ttatctagtt tattttctaa toctcctaactc
120
aattgatcct aataaatata gatgatgataa agaattgaaa cttgtgcaat aatttgcctcagc
180
attgtgtcct aatattttcct aagagggaa gctagaggtg ggtcaggtcag t
231

<210> SEQ ID NO: 68
<211> LENGTH: 52
<212> TYPE: DNA
<213> ORGANISM: Zygosaccharomyces bailii

<400> SEQUENCE: 52

cctcgatggt ggctgcctaggt ctgctgctgg tctagcaggg gatagctggc
52

<210> SEQ ID NO: 69
<211> LENGTH: 116
<212> TYPE: DNA
<213> ORGANISM: Zygosaccharomyces bailii

<400> SEQUENCE: 69

catctgccgctt ggtgggactcctt gtgctgcttgg gggagtgtgc gggacgcagc cggatctgcggc
60
gttgtggctt ggtggggcgtg cggatctgc tttgtgcagat tctactgggg cggggctgg
116

<210> SEQ ID NO: 70
<211> LENGTH: 268
<212> TYPE: DNA
<213> ORGANISM: Zygosaccharomyces bailii

<400> SEQUENCE: 70

cggcgtcctt ttggcgcctt gggtggtca gggcgcgtgag tttcgtcttg cggcgtgcgtg
60
cggggcata tggggagaaa cgggtgctttg ggtttcgtcttg gggtggtca gggcgcgtgag
120
tgcgtaataa cgggtggtcct ggggtggtca ggggtggtca ggggtggtcct ggggtggtcct
180
cgggtgtctt ggggtggtca ggggtggtca ggggtggtca ggggtggtca ggggtggtca
240
cgggtgtctt ggggtggtca ggggtggtca
268

<210> SEQ ID NO: 71
<211> LENGTH: 869
<212> TYPE: DNA
<213> ORGANISM: Zygosaccharomyces bailii

<400> SEQUENCE: 71

cggcgtcctt ttggcgcctt gggtggtca gggcgcgtgag tttcgtcttg cggcgtgcgtg
60
-continued

tcgtaaacc gcttgtgca gacgcgtatc gtatgggatc cagatattgt gatattatct 120
tggttattgc ctctgtatcg acggtgtaaa ttcggtgaat ctgtgttgg ccctaatgac 180
cggttttgg aatatttgctc tattgacctg cctttggatg ttcagcgtga aatatggtt 240
ggagtcggat tgttcgttct ccctgtgctga gatctcccgg eacgtatgag ccacggaag 300
atatototo gtagcttttggt cgttaagaga agcaatagac gocaatgca ttagaagac 360
tgtgtacac cggagaattg gagtctttgct gtctgatgtta ccagtattatg tattgctgcc 420
acccotctgc tcttatcccc tttttcggaga gacgagcata cagccacgccc aatacaggg 480
agccacattta gttcaggtg cacagagac acagagatgg gttcagctttc gaggcttca 540
atagaggtg ggttgcgtagg gtttcctcg ttgattgact cttcagaaata 600
aacaggytt ggcgctctt agtggagtt cttgaggagc tattggtctc atctgctgg 660
ctactctgta cagccagttta aggggactc ttaagcgatt ctcotctgctc ttcagacttt 720
atatatttgc aacggcaat tttgaggttt ctccgtgtat cagtaaaagcg ctcagtgccc 780
ttccagcggt aacagaagcg ttcagacgat ctccagacat atctctcg aaatctcag 840
caccaagag aagacacgta gttctcattt 869

<210> SEQ ID NO: 72
<211> LENGTH: 1425
<212> TYPE: DNA
<213> ORGANISM: Syngnathozoon bali
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (1)...(1425)
<400> SEQUENCE: 72

atg tct gac ggg gag cag ctt gtc aag aca ttt cca tta gac cag gtt
Met Ser Glu Phe Ser Glu Leu Val Arg Ile Leu Pro Leu Asp Gin Val
1  5 10 15 48

gaa gaa gat tgg tgt tgg gac gcc gag cct gaa ctt ata ctt taa
Ala Glu Ile Lys Arg Ile Leu Ser Arg Gly Asp Pro Ile Pro Leu Gin
20 25 30 35 96

agc tta got tct cta ata act atg tgt atc ata aca ggt gcc acc atg tca
Arg Leu Ala Ala Leu Ser Leu Thr Met Val Ile Leu Thr Val Asn Met Ser
35 40 45 50 144

aag aag aag tgt tgt gag agc ccc atg ccc acc ctt ttt act aca tat
Lys Lys Arg Lys Ser Ser Pro Ile Lys Leu Ser Ser Phe Thr Lys Tyr
50 55 60 65 192

cct ggg agg gtc gcc aag tca ttt tga aca cag tgg tgg gac aca aca
Arg Arg Asn Val Ala Lys Leu Ser Tyr Asp Met Ser Ser Lys Thr
65 70 75 80 240

gta ttc gaa tgg gag ctc aac aga aag ctt cca aat cag gaa gaa
Val Phe Phe Glu Tyr His Leu Lys Asn Thr Gin Asp Leu Gin Gly Val
85 90 95 100 288

tcc gag cca gcc att gcc ccc tac aat ttc tgt gta aag tgt cac aag
Leu Glu Gln Ala Ile Ala Pro Tyr Aan Phe Val Val Lys Val His Lys
100 105 110 115 336

aag cca att gat tgg cag aag cag ctc tga aag ctc gtt gat gag aga
Leu Pro Ile Asp Trp Gln Lys Gin Leu Ser Ser Val His Glu Arg Lys
115 120 125 130 384

ggg cag gac aag att ctc aag aac aat gtt ggc gcc gag aag atc tct
Ala Gly His Arg Ser Ile Leu Ser Aan Asn Val Gly Ala Glu Ile Ser
130 135 140 145 432
-continued

tgg acc ggc tcs tgg acc ccc ttg acc tgg act ggt ccc aeg aca ctc
Ser Thr Gly Ser Ser Thr Pro Leu Thr Trp Thr Ala Pro Lys Thr Leu 
450 455 460

tcc act ggt cta atg aca cct gaa gas gag tag
Ser Thr Gly Leu Met Thr Pro Gly Glu Glu *
465 470

<210> SEQ ID NO 73
<211> LENGTH: 474
<212> TYPE: PRT
<213> ORGANISM: Zygosaccharomyces bailii

<400> SEQUENCE: 73

Met Ser Glu Phe Ser Glu Leu Val Arg Ile Leu Pro Leu Asp Gln Val 
1  5   10  15

Ala Glu Ile Lys Arg Ile Leu Ser Arg Gly Asp Pro Ile Pro Leu Gln 
20  25  30

Arg Leu Ala Ser Leu Leu Thr Met Val Ile Leu Thr Val Asn Met Ser 
35  40  45

Lys Lys Arg Lys Ser Ser Pro Ile Lys Leu Ser Thr Phe Thr Lys Tyr 
50  55  60

Arg Arg Asn Val Ala Lys Ser Leu Tyr Tyr Asp Met Ser Ser Lys Thr 
65  70  75  80

Val Phe Phe Glu Tyr His Leu Lys Asn Thr Gln Asp Leu Gln Glu Gly 
85  90  95

Leu Glu Gln Ala Ile Ala Pro Tyr Asn Phe Val Val Lys Val His Lys 
100 105 110

Lys Pro Ile Asp Trp Gln Lys Glu Leu Ser Ser Val His Glu Arg Lys 
115 120 125

Ala Gly His Arg Ser Ile Leu Ser Asn Asn Val Gly Ala Glu Ile Ser 
130 135 140

Lys Leu Ala Glu Thr Lys Asp Ser Thr Trp Ser Phe Ile Glu Arg Thr 
145 150 155 160

Met Asp Leu Ile Glu Ala Arg Thr Arg Gln Pro Thr Thr Arg Val Ala 
165 170 175

Tyr Arg Phe Leu Leu Gln Leu Thr Phe Met Asn Cys Cys Arg Ala Asn 
180 185 190

Asp Leu Lys Asn Ala Asp Pro Ser Thr Phe Glu Ile Ile Ala Asp Pro 
195 200 205

His Leu Gly Arg Ile Leu Arg Ala Val Pro Glu Thr Lys Thr Ser 
210 215 220

Ile Glu Arg Phe Ile Tyr Phe Phe Pro Cys Lys Gly Arg Cys Asp Pro 
225 230 235 240

Leu Leu Ala Leu Asp Ser Tyr Leu Leu Trp Val Gly Pro Val Pro Lys 
245 250 255

Thr Gln Thr Thr Asp Glu Glu Thr Gln Tyr Asp Tyr Gln Leu Leu Gln 
260 265 270

Asp Thr Leu Leu Ile Ser Tyr Asp Arg Phe Ile Ala Lys Glu Ser Lys 
275 280 285

Glu Asn Ile Phe Lys Ile Pro Asn Gly Pro Lys Ala His Leu Gly Arg 
290 295 300

His Leu Met Ala Ser Tyr Leu Gly Asn Asn Ser Leu Lys Ser Glu Ala 
305 310 315 320
Thr Leu Tyr Gly Asn Trp Ser Val Glu Arg Gln Glu Gly Val Ser Lys
325 330 335
Met Ala Aep Ser Arg Tyr Met His Thr Val Lys Ser Pro Pro Ser
340 345 350
Tyr Leu Phe Ala Phe Leu Ser Gly Tyr Tyr Lys Ser Asn Gln Gly
355 360 365
Glu Tyr Val Leu Ala Glu Thr Leu Tyr Asn Pro Leu Asp Tyr Asp Lys
370 375 380
Thr Leu Pro Ile Thr Thr Asn Glu Lys Leu Ile Cys Arg Arg Tyr Gly
385 390 395 400
Lys Asn Ala Lys Val Ile Pro Lys Asp Ala Leu Tyr Leu Tyr Thr
405 410 415
Tyr Ala Gln Gln Lys Arg Lys Gln Leu Ala Asp Pro Asn Glu Gln Asn
420 425 430
Arg Leu Phe Ser Ser Glu Ser Pro Ala His Pro Phe Leu Thr Pro Gln
435 440 445
Ser Thr Gly Ser Ser Ser Thr Pro Leu Thr Trp Thr Ala Pro Lys Thr Leu
450 455 460
Ser Thr Gly Leu Met Thr Pro Gly Glu Glu
465 470

<210> SEQ ID NO 74
<211> LENGTH: 1075
<212> TYPE: DNA
<213> ORGANISM: Zygosaccharomyces bailii
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (1) ...(1074)

<400> SEQUENCE: 74
atg ttc tcc agg gaa gag gtt agg gcc tcc agg ccc act aea gag atg
Met Phe Ser Arg Glu Val Arg Ala Ser Arg Pro Thr Lys Glu Met
1  5  10  15
aag atg atc ttt gat gtg ctt atg aca ttt cct tac ttc gcg gta cat
Lys Met Ile Phe Asp Val Asp Leu Met Thr Pro Tyr Phe Ala Val His
20  25  30
96
<400> SEQUENCE: 74

gtt cct tcc aag aat ata ctt atc aca cca aea ggc aca gtt gag ata
Val Pro Ser Lys Asn Ile Leu Ile Thr Pro Lys Gly Thr Val Ile
35  40
144

cct gaa aac tat cca aat tat ccc ats tgg gcc ato ttc tac gtc aea
Pro Glu Asn Tyr Glu Asn Leu Ala Phe Thr Val Lys Glu
50  55  60
192

tat tta atg aag aea aat cog tac cat ctt cca aca acc ggt gac
Tyr Met Lys Lys Asn Pro Tyr Asp Leu Pro Ser Thr Val Asn
65  70  75  80
240

ttg cgc gaa ccc tat gta gtt gat aac ato act aac cog ttc cag
Trp Pro Glu Pro Tyr Val Val Val Val Thr Ile Thr Lys Arg Phe Gln
85  90  95
288

gac cat aea cta ttt gca aac aea aat gtt gat gtc tac gtt gaa aea
Asp His Lys Leu Phe Ala Asn Lys Ala Asp Val Lys Glu Arg
100  105 110
336

cat cca aat gca att gcc tgg ggt att aag att cct gac ttc cag
Leu Glu Asn Ala Ile Ala Ser Gly Ile Lys Ile Pro Glu Ser Lys Lys
115  120  125
384

eat gaa cga tca ggg cag cca aea aag acg aea aat gtt aca aea gga
Asn Glu Arg Leu Gly Gln Pro Lys Thr Lys Asn Val Thr Lys Glu
130  135  140
---continued---

```
at tga gga gac ctt taa ctc taa ctc cac taa tgc gag aaa aga att gga
Aan * Gly Aep Leu Tyr * Cys His * Cys Glu Lys Arg Ile Gly
145 150 155

tga gta ctt cag aaa act tca gga tgg tac att aac cgg aga ttt gga
  Val Leu Gln Lys Thr Ser Gly Trp Tyr Ile Asn Arg Arg Phe Gly
  160 165 170

agg tgg ctt gtt cgg gaa gac ggt cct ctc gaa ttg ccg aag ttc
Gly Trp Leu Val Gln Gly Gln Asn Ala His Ile Val * Ser Phe Val
  175 180 185

cgg agg ccc cca gaa act cca gtt tat ggc cac caa tgt tgc taa
Arg Arg Thr His Pro Arg Thr Pro Val Tyr Gly His Glu Cys Ser *
  190 195 200

agt gat gaa gaa gat ggt gct ggt cat tgg cct cca taa aag taa
Ser Leu Asp Arg Gly Asp Ser Val Arg His Gly Phe Glu * Lys Cys
  205 210 215

aat tga cga taa tga tct cga gga aga aga ggc taa tgc atc ggg cga
Aan * Arg * * Ser Arg Gly Arg Arg Ala * Cys Ile Gly Arg
  220 225

aac aac tac gac gac cgc gaa aga tga aag gag tca tgg cac ctc
Thr Asn Tyr Asp Ser Pro Arg Arg Gly Ile Arg Gly Ser Gly Tyr His Ile
  230 235 240 245

cac tgg ttt ggg ggc ctt gta acc tca aat taa tgc cat aga aac gga
Gln Trp Phe Gly Arg Ser Glu Tyr Ser Aan * Cys His Asn Arg Gly
  250 255 260

sga gcc gcc gcc gcc gct gcc gcc gcc cca cga taa gta gct cca ccc
Gly Ile Ile Leu Gly Ser Tyr Glu Gly Ala Pro * * Ala Thr His
  265 270

cct tcc aac cca gtt aga ggg cgg gag gcc ggt gct ttt cct
Leu Ser Asn Thr Val Arg Arg Val Glu Ser Gly Ser Phe Phe Thr
  275 280 285 290

sga cca taa aag aat tct cca aac att tac aa aag tca gcc cca cca
Gly Pro * Lys Aan Thr Pro Asn Ile Tyr Lys Ala Lys Gly Arg Cys
  295 300 305

cgc gcc tcc tgg ttt taa cct cag gca gtt cgc gcc gcc gcc gct
Pro Arg Ser Phe Leu Tyr Lys Ser Glu Arg Val Ser Gly Asn Glu Leu
  310 315 320

gaa cta cgg ctc ggg cag tga tct gcg cct gaa ctg gct cgg cgg cct
Glu Phe Arg Ile Tyr Arg Gly Ile Arg Cys Lys Aan Gly Aan
  325 330 335

agg tga gct tca aag tta a
Arg * Ala Ser Lys Leu
340
```

<210> SEQ ID NO 75
<211> LENGTH: 341
<212> TYPE: FRT
<213> ORGANISM: Zygosaccharomyces bailii

<400> SEQUENCE: 75

Met Phe Ser Arg Glu Glu Val Arg Ala Ser Arg Pro Thr Lys Glu Met
  1  5  10  15

Lys Met Ile Phe Asp Val Leu Met Thr Phe Pro Tyr Phe Ala Val His
  20 25 30

Val Pro Ser Lys Asn Ile Leu Thr Pro Lys Gly Thr Val Glu Ile
  35 40 45

Pro Glu Asn Tyr Glu Asn Tyr Pro Ile Leu Ala Ile Phe Tyr Val Lys
  50 55 60
Tyr Leu Met Lys Lys Asn Pro Tyr Asp Leu Leu Pro Ser Thr Val Asn 65 70 75 80
Trp Pro Glu Pro Tyr Val Val Val Asn Thr Ile Thr Lys Arg Phe Glu 85 90 95
Asp His Lys Leu Phe Ala Asn Lys Asn Ala Asp Val Tyr Val Glu Arg 100 105 110
Leu Glu Asn Ala Ile Ala Ser Gly Ile Lys Ile Pro Glu Ser Lys Lys 115 120 125
Asn Glu Arg Leu Gly Gln Pro Lys Thr Lys Asn Val Thr Lys Glu 130 135 140
Asn Gly Asp Leu Tyr Cys His Cys Glu Lys Arg Ile Gly Val Leu Gln 145 150 155 160
Lys Thr Ser Gly Trp Tyr Ile Asn Arg Arg Phe Gly Gly Trp Leu Val 165 170 175
Gln Gly Gln Asn Ala His Ile Val Ser Phe Val Arg Arg Thr His Pro 180 185 190
Arg Thr Pro Val Tyr Gly His Gin Cys Ser Ser Leu Asp Arg Gly Asp 195 200 205
Ser Val Arg His Gly Phe Glu Lys Cys Asn Arg Ser Arg Gly Arg Arg 210 215 220
Asp Cys Ile Gly Arg Thr Asn Tyr Asp Ser Pro Arg Gly Ile Arg Gly 225 230 235 240
Ser Gly Tyr His Ile Gln Trp Phe Gly Arg Ser Glu Tyr Ser Asn Cys 245 250 255
His Arg Asn Gly Gln Ile Ile Leu Gly Ser Tyr Gln Gly Ala Pro Ala 260 265 270
Thr His Leu Ser Asn Thr Val Arg Val Gin Glu Ser Gly Ser Phe 275 280 285
Phe Thr Gly Pro Lys Asn Thr Pro Asn Ile Tyr Lys Ala Gly Tyr 290 295 300
Cys Pro Arg Ser Phe Leu Tyr Lys Ser Gin Arg Val Ser Gly Asn Gin 305 310 315 320
Leu Glu Phe Arg Ile Tyr Arg Gly Ile Arg Cys Lys Lys Asn Gin 325 330 335
Arg Ala Ser Lys Leu 340

<210> SEQ ID NO 76
<211> LENGTH: 750
<212> TYPE: DNA
<213> ORGANISM: Zygosaccharomyces bailii
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (1)...(750)

<400> SEQUENCE: 76
atg aac tca ggg ttc tct ctt gcc tac gga aat gtt gac tct gat tet
Met Asn Ser Glu Phe Ser Leu Ala Tyr Gly Asn Val Asp Ser Asp Tyr 1 5 10 15

96

96
gct ctc gat ctc gaa cgg ctt gat tcc aac tgg aag ggc acc ggc
Ala Leu Asp Leu Leu Glu Arg Leu Asp Ser Asn Trp Lys Gly Thr Gln 20 25 30

144

cc ttt act cat ata ccc gaa acc ttt cca att ggc ctt ggc aat gtt
Leu Phe Thr His Ile Arg Glu Thr Phe Glu Ile Gly Leu Gly Asn Val
<210> SEQ ID NO: 77
<211> LENGTH: 249
<212> TYPE: PRT
<213> ORGANISM: Zygosaccharomyces bailii

<400> SEQUENCE: 77

Met Aan Ser Glu Phe Ser Leu Ala Tyr Gly Asn Val Asp Ser Asp Tyr 1 5 10 15
Ala Leu Asp Leu Leu Glu Arg Leu Aas Ser Aas Thr Tyr Gly Thr Leu 20 25 30
Leu Phe Thr His Ile Arg Gly Thr Phe Gln Ile Gly Leu Gly Asn Val 35 40 45
Ile Ile Val Ser Glu Gln Ser Glu Ser Leu Arg Ile Pro Pro Ser Leu 50 55 60
Leu Gly Ser Ser Ser Pro Ala Asp Ser Aas Ser Pro Pro Gly Thr 65 70 75 80

35 40 45

atc ata gta gca gaa cag aat gaa agc ctt aga ata ccc ccc ctc cta cta
Ile Ile Val Ser Glu Gln Ser Glu Ser Leu Arg Ile Pro Pro Ser Leu
50 55 60

cct ggt gtc gat agt cca goa gat agc gac aat aat gtt cca gga aca
Leu Gly Ser Ser Ser Ser Pro Ala Asp Ser Aas Ser Pro Pro Gly Thr
65 70 75 80

cct gat aat gaa gac gca ccc tgg ttt att tgt gaa gat ctc tgg aca
Pro Thr Asn Glu Ala Gln Pro Thr Pro Leu Ser Glu Ser Leu Ser Leu
85 90 95

ggc cct ttc acg gaa gcc cac tca act cca tca tct att gag aca ctc
Gly Pro Phe Thr Glu Ala Gln Ser Thr Glu Ser Ser Ile Glu Thr Leu
100 105 110

gaa ggt gag ccc cta gct gtc tct tgt tga ctc cag ctc aag cta aat ggc
Glu Gly Glu His His Ala Val Ser Ser Leu His Leu Lys Leu Aas Gly
115 120 125

cct ccc tgt att gga cgg cgt gta tgg cgg gct act cgc aas atg gat
Leu Ser Cys Ile Gly Arg Ala Val Trp Arg Ala Thr Arg Lys Met Aas
130 135 140

cgc ata gaa gaa gta gac ata taa aac tca ata aca gaa ccc aca
Thr Arg Thr Glu Val Asp Ile Leu Aas Ser Ile Thr Glu Pro Arg
145 150 155 160

gaa ctc cca tta ccc ggt atc aac aag atg cgt cca taa tgt gct
Arg Leu Thr Leu Pro Gly Ile Aas Met Arg Glu Gly Cys Ile Val Arg
165 170 175

cga tta cct cct gta ccc tca gaa cta cgq gaa gag att cct cca gcc
Leu Leu Leu Val Pro Ile Glu Glu Ile Leu Ser Phe
180 185 190

gcc ata ggt tgg gta ccc tca gaa aca ata gaa gat att cga tct
Ala Ile Ala Ser Gly Ile Pro Ser Glu Thr Ile Glu Asp Ile Arg Ser
195 200 205

tca cca aat att tca gct ggt gat gac aat gcc aca aca gca cca cat
Ser Thr Asn Ile Ser Gly Asp Aas Thr Gln Val Asp Thr Cys Ser Ala His
210 215 220

cgc gcc cct ccc tga ggg cta ccc tca gaa aca taa gat att cta cct
Arg Arg Ile Ile Arg Gly Asp Ser Glu Ser Leu Arg Ala Gly Gly
225 230 235 240

cgc gct cgg aag cgg aca aag gcc aca gct gaa gat tga gct
Arg Arg Ile Ile Arg Gly His Thr Gln +
245

<210> SEQ ID NO: 77
<211> LENGTH: 249
<212> TYPE: PRT
<213> ORGANISM: Zygosaccharomyces bailii

<400> SEQUENCE: 77

Met Aan Ser Glu Phe Ser Leu Ala Tyr Gly Asn Val Asp Ser Asp Tyr 1 5 10 15
Ala Leu Asp Leu Leu Glu Arg Leu Aas Ser Aas Thr Tyr Gly Thr Leu 20 25 30
Leu Phe Thr His Ile Arg Gly Thr Phe Gln Ile Gly Leu Gly Asn Val 35 40 45
Ile Ile Val Ser Glu Gln Ser Glu Ser Leu Arg Ile Pro Pro Ser Leu 50 55 60
Leu Gly Ser Ser Ser Pro Ala Asp Ser Aas Ser Pro Pro Gly Thr 65 70 75 80
-continued

Pro Thr Asn Glu Ala Gln Pro Trp Phe Ile Ser Glu Asp Leu Ser Lys
  85   90   95
Gly Pro Phe Thr Glu Ala Gln Ser Thr Asp Ser Ile Glu Thr Leu
 100  105  110
Glu Gly Glu His His Ala Val Ser Ser Leu His Leu Lys Leu Asn Gly
 115  120  125
Leu Ser Cys Ile Gly Arg Ala Val Trp Arg Ala Thr Arg Lys Met Asp
 130  135  140
Thr Arg Thr Glu Val Asp Asp Ile Leu Asn Ser Ile Thr Glu Pro Arg
 145  150  155  160
Arg Leu Thr Leu Pro Gly Ile Asn Lys Met Arg Glu Cys Ile Val Arg
 165  170  175
Leu Leu Leu Leu Val Pro Ile Gln Val Arg Glu Ile Leu Ser Phe
 180  185  190
Ala Ile Ala Ser Gly Ile Pro Ser Glu Thr Ile Glu Asp Ile Arg Ser
 195  200  205
Ser Thr Asn Ile Ser Ala Val Arg Thr Asn Gly Arg Gly Ile Ala His
 210  215  220
Asn Ser Lys Arg Ser Leu Ala Pro Thr Gln Asp Ser Arg Asn Leu
 225  230  235  240
Arg Arg Arg Ile Arg Gly His Thr Gln
 245

<210> SEQ ID NO 78
<211> LENGTH: 453
<212> TYPE: DNA
<213> ORGANISM: Saccharomyces cerevisiae
<220> FEATURE:
<221> NAME/KEY: promoter
<222> LOCATION: (1)...(450)
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (451)...(453)
<223> OTHER INFORMATION: start codon

<400> SEQUENCE: 78
ctactattc ctcctgagat tatatctgg asccacatcg gttgttggaas gattacccgt
  60
tttaagacct ttctagttcc tttattgtat ttacacctgg asccacatctt tttggatcct
 120
agttttttaat ctcctgagat atgtgatatt ccctgaaacctat ctaacacatctt
 180
tttctttgtt acccctgcag tttgatcaag gatgtgtgatt ttaccccttt gtaataatctt
 240
ggagctata actctttgaa cttgctctgc tctgagttatt atttttctttaaaacagtt
 300
agaggtttag tgaatttca acatataca ttttttttttttttctttt cttaaatttctttttt
 360
gtaatatttttcttatattttc ttctttttttttcttttcttttctttttc tttaattttttttt
 420
aatatatattttaatctatcc ttatatctta atatatcatat tttttttttttttttttttttttt
 485

<210> SEQ ID NO 79
<211> LENGTH: 489
<212> TYPE: DNA
<213> ORGANISM: Zygosaccharomyces bailii
<220> FEATURE:
<221> NAME/KEY: promoter
<222> LOCATION: (1)...(496)
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (497)...(499)
<223> OTHER INFORMATION: start codon

<400> SEQUENCE: 79

    ggtgcgatt gctctcaacct ttctttttgt attgagcgctt attgacatcttctt  60
    tttaacgctc gcctcaaat atctgtctca gaagaaggtg ataaatactgccaa 120
    ctggtatca cttctaaact gctggtggtac aagcaacactg ctttacacg acatcctctat 180
    cttctcgag tgggcgaaacc gctcagctgg ttgattatca atgtcagtgcc tggatagca 240
    cattatctt cccatatatc tcgctgaacg cgggyctca acagtctttttgtt ctctcaacttt 300
    cttaagacct gcgctatcct ctaaaagcaac aggatactca aagacccctc gcgtgcaacg 360
    gcgtgctagc gctggttcaca tatctattgg tcaacctttcc tctgtctaaaa gatggtaagc 420
    aagctggtata cttccatctct tctagatctaa ctaacacagacctacctgcgg 480
    tatataaact acaaaatg 499

<210> SEQ ID NO 80
<211> LENGTH: 27
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR Primer

<400> SEQUENCE: 80

    aagagactcc aacgctcgoc aacctgta 27

<210> SEQ ID NO 81
<211> LENGTH: 32
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR Primer

<400> SEQUENCE: 81

    aagagattgc gaagaccaca atctgcatggt gac 32

<210> SEQ ID NO 82
<211> LENGTH: 29
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR Primer

<400> SEQUENCE: 82

    atctgatcct gcctctctct ctttctgta 29

<210> SEQ ID NO 83
<211> LENGTH: 29
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR Primer

<400> SEQUENCE: 83

    ttggtttatg tttaacctga tgtagcttc 29

<210> SEQ ID NO 84
<211> LENGTH: 27
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
tagctactctctccaggtgctcattag 27

SEQ ID NO: 85
LENGTH: 25
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: PCR Primer

ctatagcagagtttaggacgttg 25

SEQ ID NO: 86
LENGTH: 25
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: PCR Primer

agatgaaactcagagttctctcttgg 25

SEQ ID NO: 87
LENGTH: 22
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: PCR Primer

atctcatattggtagttccccctg 22

SEQ ID NO: 88
LENGTH: 30
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: PCR Primer

gttttaatttttgagttccotttaattg 30

SEQ ID NO: 89
LENGTH: 26
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: PCR Primer

attgtttctccaggagaaggttag 26

SEQ ID NO: 90
LENGTH: 27
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: PCR Primer

SEQ ID NO: 90
agastcastc atttacgttg gcagggag

<210> SEQ ID NO 91
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR Primer

<400> SEQUENCE: 91
tasaaaatgc ccgcaatatt tctgtc

<210> SEQ ID NO 92
<211> LENGTH: 708
<212> TYPE: DNA
<213> ORGANISM: Zygoascuscheromyces rouxii

<400> SEQUENCE: 92
cgtcagaag gcttcagtag tggctccctgt tggctcgctc gcagcccttt ggaaatctt
60
gtgcgcgcgc tcagcagcga atggcccttt agacacacat ccaaccttg caaactctg
120
ttcotcctcct ttcctctttt ttcctccctt cctctccctc ctgctctctc
180
cgcagatgt cggagaaagag acagctgca gcagggcgac caagtttttga cgctttggac
240
caggtcatcgt cgtcgcagac gcggactata gcgttacac gatcgctgtaa
300
gcgcgcgggc tggggtatata aagggctaca toctttctcc caagcagggc atacccgca
360
tctacacat gctgcctcttt ttgccttgcg ccattagccgc cggacactt gctacccgca
420
tgacacagagt ggctgtatctt ccggagaggat tgggctgtat ggagccaga tcctaccttt
480
ttcotcagct gcacagcgc ggacacacc gcgcgtttcc aagttggtta aagttggcag
540
ggcgctgggt gcgcagacat atgcttgatg tggatctgat ccggagag cagatgtgca
600
tagacagact gctggctctt gcttatcttt ttctttttct otataactc cagaacaat
660
cagagttccg ttggttcata tcgaatanaa agtacatcag agataaca
708

<210> SEQ ID NO 93
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR Primer

<400> SEQUENCE: 93
tgcgaagac cctaaagtc t

<210> SEQ ID NO 94
<211> LENGTH: 29
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR Primer

<400> SEQUENCE: 94
tgctgtgtat gtactttttta tttgatatg

<210> SEQ ID NO 95
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
1-21. (canceled)

22. A process for the production of a protein comprising the steps of:
   a) culturing a *Zygosaccharomyces bailii* strain,
   b) expressing and secreting a protein,
   c) isolating the protein.

23. The process of claim 22, wherein the *Z. bailii* strain is transformed with a vector comprising a DNA sequence coding for the protein, functionally linked to a signal sequence leading to the secretion of the protein and further functionally linked to a promoter.

24. The process of claim 23, wherein the vector is an extra-chromosomal plasmid.

25. The process of claim 24, wherein the plasmid is derived from an endogenous episomal plasmid from a *Z. bailii* strain.

26. The process of claim 23, wherein the plasmid comprises sequences for replication, stabilization, or plasmid copy number control, obtainable from *Z. bailii*.

27. The process of claim 25, wherein the plasmid comprises at least 35 bases of one of the sequences selected from the group consisting of SEQ ID No.: 63, SEQ ID No.: 64, SEQ ID No.: 65, SEQ ID No.: 66, SEQ ID No.: 67, SEQ ID No.: 68, SEQ ID No.: 69, SEQ ID No.: 70, and SEQ ID No.: 71.

28. The process of claim 23, wherein the promoter is a triose-phosphate isomerase promoter, obtainable from *Saccharomyces cerevisiae* or from *Z. bailii*.

29. The process of claim 23, wherein the promoter is a glyceraldehyde phosphate dehydrogenase promoter, obtainable from *Saccharomyces cerevisiae*, *Z. bailii* or *Z. rouxii*.

30. The process of claim 23, wherein the signal sequence is a continuous stretch of 15 to 60 amino acids, comprising one or more positively charged amino acid(s) followed by a stretch of about 5 to 10 hydrophobic amino acids, which are optionally interrupted by non-hydrophobic residues.


32. The process of claim 22, wherein the *Z. bailii* strain is transformed with a vector comprising the DNA sequence coding for the protein, functionally linked to the signalling pre-sequence of the alpha-subunit of the K1 killer toxin of *Kluyveromyces lactis* and further functionally linked to the triose-phosphate isomerase promoter from *S. cerevisiae*.

33. The process of claim 32, wherein the vector is the plasmid pZkl as shown in FIG. 1b.

34. The process of claim 22, wherein the *Z. bailii* strain is transformed with a vector comprising the DNA sequence coding for the protein, functionally linked to the signal sequence of the pre-pro alpha-factor of *S. cerevisiae* and further functionally linked to the triose-phosphate isomerase promoter from *S. cerevisiae*.

35. The process of claim 34, wherein the vector is the plasmid pZppta as shown in FIG. 1c.

36. The process of claim 23, wherein the DNA sequence coding for the protein is derived from animal, bacterial, fungal, plant, or viral sources.

37. The process of claim 23, wherein the *Z. bailii* strain that is transformed is selected from the list of: ATCC 36947, ATCC 60483, NCYC 1427 or ATCC 8766.

38. The process of claim 22, wherein the *Z. bailii* strain has been subjected to a selection process for improved secretion.

39. The process of claim 22, wherein the *Z. bailii* strain is cultivated in a chemically defined medium.

40. The process of claim 22, wherein the protein is isolated from the culture medium.

41. A *Z. bailii* strain, expressing and secreting a heterologous protein.

42. The *Z. bailii* strain of claim 41, wherein the cells are transformed with a vector comprising a DNA sequence coding for the heterologous protein, functionally linked to a signal sequence leading to the secretion of the protein and further functionally linked to a promoter.

* * * * *