Innovation, Sciences et Développement économique Canada
Office de la Propriété Intellectuelle du Canada

Innovation, Science and Economic Development Canada
Canadian Intellectual Property Office

CA 2902149 C 2022/07/26

(11)(21) **2 902 149**

(12) **BREVET CANADIEN**
CANADIAN PATENT

(13) **C**

(86) Date de dépôt PCT/PCT Filing Date: 2014/02/21
(87) Date publication PCT/PCT Publication Date: 2014/08/28
(45) Date de délivrance/Issue Date: 2022/07/26
(85) Entrée phase nationale/National Entry: 2015/08/21
(86) N° demande PCT/PCT Application No.: NL 2014/050106
(87) N° publication PCT/PCT Publication No.: 2014/129898
(30) Priorité/Priority: 2013/02/22 (EP13156448.6)

(51) Cl.Int./Int.Cl. *C12N 1/16* (2006.01), *C12N 9/12* (2006.01), *C12N 9/88* (2006.01), *C12P 5/00* (2006.01), *C12P 7/00* (2006.01), *C12P 7/06* (2006.01)

(72) Inventeurs/Inventors:
VAN MARIS, ANTONIUS JEROEN ADRIAAN, NL;
PRONK, JACOBUS THOMAS, NL;
GUADALUPE MEDINA, VICTOR GABRIEL, NL;
WISSELINK, HENDRIK WOUTER, NL

(73) Propriétaire/Owner:
DSM IP ASSETS B.V., NL

(74) Agent: SMART & BIGGAR LLP

(54) Titre : MICROORGANISME RECOMBINE DESTINE A ETRE UTILISE DANS UN PROCEDE AVEC UN RENDEMENT EN PRODUIT ACCRU
(54) Title: RECOMBINANT MICRO-ORGANISM FOR USE IN METHOD WITH INCREASED PRODUCT YIELD

(57) Abrégé/Abstract:
The invention relates to a recombinant yeast cell, in particular a transgenic yeast cell, functionally expressing one or more recombinant, in particular heterologous, nucleic acid sequences encoding ribulose-1,5-biphosphate carboxylase oxygenase (Rubisco) and phosphoribulokinase (PRK). The invention further relates to the use of carbon dioxide as an electron acceptor in a recombinant chemotrophic micro-organism, in particular a eukaryotic micro-organism.

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
28 August 2014 (28.08.2014)

WIPO | PCT

(10) International Publication Number
WO 2014/129898 A3

(51) International Patent Classification:
C12N 1/16 (2006.01) *C12P 5/00* (2006.01)
C12N 9/12 (2006.01) *C12P 7/00* (2006.01)
C12N 9/88 (2006.01) *C12P 7/06* (2006.01)

(21) International Application Number:
PCT/NL2014/050106

(22) International Filing Date:
21 February 2014 (21.02.2014)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
13156448.6 22 February 2013 (22.02.2013) EP

(71) Applicant: **TECHNISCHE UNIVERSITEIT DELFT** [NL/NL]; Stevinweg 1, NL-2628 CN Delft (NL).

(72) Inventors: **VAN MARIS, Antonius Jeroen Adriaan**; c/o TU Delft Valorisation Centre, Postbus 5, NL-2600 AA Delft (NL). **PRONK, Jacobus Thomas**; c/o TU Delft Valorisation Centre, Postbus 5, NL-2600 AA Delft (NL). **GUADALUPE MEDINA, Victor Gabriel**; c/o TU Delft Valorisation Centre, Postbus 5, NL-2600 AA Delft (NL). **WISSELINK, Hendrik Wouter**; c/o TU Delft Valorisation Centre, Postbus 5, NL-2600 AA Delft (NL).

(74) Agent: **JANSEN, C.M.**; V.O., Johan de Wittlaan 7, NL-2517 JR Den Haag (NL).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report (Art. 21(3))*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

(88) Date of publication of the international search report:
18 December 2014

WO 2014/129898 A3

(54) Title: RECOMBINANT MICRO-ORGANISM FOR USE IN METHOD WITH INCREASED PRODUCT YIELD

(57) Abstract: The invention relates to a recombinant yeast cell, in particular a transgenic yeast cell, functionally expressing one or more recombinant, in particular heterologous, nucleic acid sequences encoding ribulose-1,5-biphosphate carboxylase oxygenase (Rubisco) and phosphoribulokinase (PRK). The invention further relates to the use of carbon dioxide as an electron acceptor in a recombinant chemotrophic micro-organism, in particular a eukaryotic micro-organism.

Title: RECOMBINANT MICRO-ORGANISM FOR USE IN METHOD WITH INCREASED PRODUCT YIELD

The invention relates to a recombinant micro-organism having the ability to produce a desired fermentation product, to the functional expression of heterologous peptides in a micro-organism, and to a method for producing a fermentation product wherein said microorganism is used. In a preferred embodiment the micro-organism is a yeast. The invention is further related to a use of CO_2 in micro-organisms.

Microbial fermentation processes are applied for industrial production of a broad and rapidly expanding range of chemical compounds from renewable carbohydrate feedstocks.

Especially in anaerobic fermentation processes, redox balancing of the cofactor couple $NADH/NAD^+$ can cause important constraints on product yields. This challenge is exemplified by the formation of glycerol as major by-product in the industrial production of – for instance - fuel ethanol by *Saccharomyces cerevisiae*, a direct consequence of the need to reoxidize NADH formed in biosynthetic reactions.

Ethanol production by *Saccharomyces cerevisiae* is currently, by volume, the single largest fermentation process in industrial biotechnology, but various other compounds, including other alcohols, carboxylic acids, isoprenoids, amino acids *etc*, are currently produced in industrial biotechnological processes.

Various approaches have been proposed to improve the fermentative properties of organisms used in industrial biotechnology by genetic modification.

WO 2008/028019 relates to a method for forming fermentation products utilizing a microorganism having at least one heterologous gene sequence, the method comprising the steps of converting at least one carbohydrate to 3 -phosphoglycerate and fixing carbon dioxide, wherein at least one of said steps is catalyzed by at least one exogenous enzyme. Further, it relates to a microorganism for forming fermentation products through fermentation of at least one sugar, the microorganism comprising at least one heterologous gene sequence encoding at least one enzyme selected from the group consisting of phosphopentose epimerase, phosphoribulokinase, and ribulose bisphosphate carboxylase.

In an example, a yeast is mentioned wherein a heterologous PRK and a heterologous Rubisco gene are incorporated. In an embodiment the yeast is used for ethanol production. The results (Figure 24) show concentrations for transgenic controls and the modified strains. Little difference is noticeable between modified yeast and its corresponding control. No information is apparent regarding product yield, sugar conversion, yeast growth, evaporation rates of ethanol. Thus, it is apparent that results are not conclusive with respect to an improvement in ethanol yield.

Further, WO 2008/028019 is silent on the problem of glycerol side-product formation.

A major challenge relating to the stoichiometry of yeast-based production of ethanol, but also of other compounds, is that substantial amounts of NADH-dependent side-products (in particular glycerol) are generally formed as a by-product, especially under anaerobic and oxygen-limited conditions or under conditions where respiration is otherwise constrained or absent. It has been estimated that, in typical industrial ethanol processes, up to about 4 wt.% of the sugar feedstock is converted into glycerol (Nissen *et al.* Yeast 16 (2000) 463-474). Under conditions that are ideal for anaerobic growth, the conversion into glycerol may even be higher, up to about 10 %.

Glycerol production under anaerobic conditions is primarily linked to redox metabolism. During anaerobic growth of *S. cerevisiae,* sugar dissimilation occurs via alcoholic fermentation. In this process, the NADH formed in the glycolytic glyceraldehyde-3-phosphate dehydrogenase reaction is reoxidized by converting acetaldehyde, formed by decarboxylation of pyruvate to ethanol via NAD^+-dependent alcohol dehydrogenase. The fixed stoichiometry of this redox-neutral dissimilatory pathway causes problems when a net reduction of NAD^+ to NADH occurs elsewhere in metabolism. Under anaerobic conditions, NADH reoxidation in *S. cerevisiae* is strictly dependent on reduction of sugar to glycerol. Glycerol formation is initiated by reduction of the glycolytic intermediate dihydroxyacetone phosphate (DHAP) to glycerol 3-phosphate (glycerol-3P), a reaction catalyzed by NAD^+-dependent glycerol 3-phosphate dehydrogenase. Subsequently, the glycerol 3-phosphate formed in this reaction is hydrolysed by glycerol-3-phosphatase to yield glycerol and inorganic phosphate. Consequently, glycerol is a major by-product during anaerobic production of ethanol by *S. cerevisiae,* which is undesired as it reduces overall conversion of sugar to ethanol. Further, the presence of glycerol in effluents of ethanol production plants may impose costs for waste-water treatment.

In WO 2011/010923, the NADH-related side-product (glycerol) formation in a process for the production of ethanol from a carbohydrate containing feedstock – in particular a carbohydrate feedstock derived from lignocellulosic biomass - glycerol side-production problem is addressed by providing a recombinant yeast cell comprising one or more recombinant nucleic acid sequences encoding an NAD^+-dependent acetylating acetaldehyde dehydrogenase (EC 1.2.1.10) activity, said cell either lacking enzymatic activity needed for the NADH-dependent glycerol synthesis or the cell having a reduced enzymatic activity with respect to the NADH-dependent glycerol synthesis compared to its corresponding wild-type yeast cell. A cell is described that is effective in essentially eliminating glycerol production. Also, the cell uses acetate to reoxidise NADH, whereby ethanol yield can be increased if an acetate-containing feedstock is used.

Although the described process in WO 2011/010923 is advantageous, there is a continuing need for alternatives, in particular alternatives that also allow the production of a useful organic compound, such as ethanol, without needing acetate or other organic electron acceptor molecules in order to eliminate or at least reduce NADH-dependent side-product synthesis. It would in particular be desirable to provide a microorganism wherein NADH-dependent side-product synthesis is reduced and which allows increased product yield, also in the absence of acetate.

The inventors realised that it may be possible to reduce or even eliminate NADH-dependent side-product synthesis by functionally expressing a recombinant enzyme in a heterotrophic, chemotrophic microorganism cell, in particular a yeast cell, using carbon dioxide as a substrate.

Accordingly, the present invention relates to the use of carbon dioxide as an electron acceptor in a recombinant chemoheterotrophic micro-organism, in particular a eukaryotic micro-organism. Chemotrophic, (chemo)heterotrophic and autotrophic and other classifications of a microorganism are herein related to the micro-organism before recombination, this organism is herein also referred to as the host. For instance, through recombination as disclosed herein a host micro-organism that is originally (chemo)heterotroph and not autotrophic may become autotrophic after recombination, since applying what is disclosed herein causes that the recombined organism may assimilate carbon dioxide, thus resulting in (partial) (chemo)autotrophy.

Advantageously, the inventors have found a way to incorporate the carbon dioxide as a co-substrate in metabolic engineering of heterotrophic industrial microorganisms that can be used to improve product yields and/or to reduce side-product formation.

In particular, the inventors found it to be possible to reduce or even eliminate NADH-dependent side-product synthesis by functionally expressing at least two recombinant enzyme from two specific groups in a eukaryotic microorganism, in particular a yeast cell, wherein one of the enzymes catalysis a reaction wherein carbon dioxide is used and the other uses ATP as a cofactor.

Accordingly, the invention further relates to a recombinant, in a particular transgenic, eukaryotic microorganism, in particular a yeast cell, said microorganism functionally expressing one or more recombinant, in particular heterologous, nucleic acid sequences encoding a ribulose-1,5-biphosphate carboxylase oxygenase (Rubisco) and a phosphoribulokinase (PRK).

A microorganism according to the invention has in particular been found advantageous in that in the presence of Rubisco and the PRK NADH-dependent side-product formation (glycerol) is reduced considerably or essentially completely eliminated and production of the desired product can be increased. It is thought that the carbon dioxide acts as an electron acceptor for NADH whereby less NADH is available for the reaction towards the side-product (such as glycerol).

The invention further relates to a method for preparing an organic compound, in particular an alcohol, organic acid or amino acid, comprising converting a carbon source, in particular a carbohydrate or another organic carbon source using a microorganism, thereby forming the organic compound, wherein the microorganism is a microorganism according to the invention or wherein carbon dioxide is used as an electron acceptor in a recombinant chemotrophic or chemoheterotrophic micro-organism.

The invention further relates to a vector for the functional expression of a heterologous polypeptide in a yeast cell, wherein said vector comprises a heterologous nucleic acid sequence encoding Rubisco and PRK, wherein said Rubisco exhibits activity of carbon fixation. The term "a" or "an" as used herein is defined as "at least one" unless specified otherwise.

In an embodiment, there is provided a recombinant yeast cell, functionally expressing one or more heterologous nucleic acid sequences encoding ribulose-1,5-biphosphate carboxylase oxygenase (Rubisco) and phosphoribulokinase (PRK), said yeast cell comprising one or more chaperones selected from the group consisting of GroEL, GroES, functional homologues of GroEL and functional homologues of GroES , wherein said functional homologues of a chaperone selected from the group consisting of GroEL and GroES have more than 50%, sequence identity with GroEL or GroES; and wherein the ribulose-1,5-biphosphate carboxylase oxygenase (Rubisco) is a Rubisco comprising the protein sequence of SEQUENCE ID NO: 2 or a functional homologue thereof comprising a sequence having at least 80% sequence identity with SEQUENCE ID NO:2.

In another embodiment, there is provided one or more vectors for the functional expression of a heterologous polypeptide in a yeast cell, wherein said vector or vectors comprise one or more heterologous nucleic acid sequences encoding Rubisco and PRK, wherein said Rubisco exhibits activity of carbon fixation, said vector or vectors further comprising one or more heterologous nucleic acid sequences encoding a chaperone selected from the group consisting of GroEL, GroES, functional homologues of GroEL and functional homologues of GroES wherein said functional homologues of a chaperone selected from the group consisting of GroEL and GroES have more than 50%, sequence identity with GroEL or GroES.

In still another embodiment, there is provided a method for preparing an alcohol comprising fermenting a carbohydrate, with a yeast cell as described herein, thereby forming the alcohol, wherein the yeast cell is present in a reaction medium.

When referring to a noun (*e.g.* a compound, an additive, *etc.*) in the singular, the plural is meant to be included. Thus, when referring to a specific moiety, *e.g.* "compound", this means "at least one" of that moiety, *e.g.* "at least one compound", unless specified otherwise.

The term 'or' as used herein is to be understood as 'and/or'.

Date Reçue/Date Received 2020-05-13

When referring to a compound of which several isomers exist (*e.g.* a D and an L enantiomer), the compound in principle includes all enantiomers, diastereomers and cis/trans isomers of that compound that may be used in the particular method of the invention; in particular when referring to such as compound, it includes the natural isomer (s).

For the purpose of clarity and a concise description features are described herein as part of the same or separate embodiments, however, it will be appreciated that the scope of the invention may include embodiments having combinations of all or some of the features described". In view of this passage it is evident to the skilled reader that the variants of claim 1 as filed may be combined with other features described in the application as filed, in particular with features disclosed in the dependent claims, such claims usually relating to the most preferred embodiments of an invention.

The term 'fermentation', 'fermentative' and the like is used herein in a classical sense, *i.e.* to indicate that a process is or has been carried out under anaerobic conditions. Anaerobic conditions are herein defined as conditions without any oxygen or in which essentially no oxygen is consumed by the yeast cell, in particular a yeast cell, and usually corresponds to an oxygen consumption of less than 5 mmol/l.h, in particular to an oxygen consumption of less than 2.5 mmol/l.h, or less than 1 mmol/l.h. More preferably 0 mmol/L/h is consumed (i.e. oxygen consumption is not detectable. This usually corresponds to a dissolved oxygen concentration in the culture broth of less than 5 % of air saturation, in particular to a dissolved oxygen concentration of less than 1 % of air saturation, or less than 0.2 % of air saturation.

The term "yeast" or "yeast cell" refers to a phylogenetically diverse group of single-celled fungi, most of which are in the division of *Ascomycota* and

Date Reçue/Date Received 2020-05-13

Basidiomycota. The budding yeasts ("true yeasts") are classified in the order *Saccharomycetales*, with *Saccharomyces cerevisiae* as the most well known species.

The term “recombinant (cell)” or “recombinant micro-organism”as used herein, refers to a strain (cell) containing nucleic acid which is the result of one or more genetic modifications using recombinant DNA technique(s) and/or another mutagenic technique(s). In particular a recombinant cell may comprise nucleic acid not present in a corresponding wild-type cell, which nucleic acid has been introduced into that strain (cell) using recombinant DNA techniques (a transgenic cell), or which nucleic acid not present in said wild-type is the result of one or more mutations – for example using recombinant DNA techniques or another mutagenesis technique such as UV-irradiation – in a nucleic acid sequence present in said wild-type (such as a gene encoding a wild-type polypeptide) or wherein the nucleic acid sequence of a gene has been modified to target the polypeptide product (encoding it) towards another cellular compartment. Further, the term “recombinant (cell)” in particular relates to a strain (cell) from which DNA sequences have been removed using recombinant DNA techniques.

The term “transgenic (yeast) cell” as used herein, refers to a strain (cell) containing nucleic acid not naturally occurring in that strain (cell) and which has been introduced into that strain (cell) using recombinant DNA techniques, *i.e.* a recombinant cell).

The term "mutated" as used herein regarding proteins or polypeptides means that at least one amino acid in the wild-type or naturally occurring protein or polypeptide sequence has been replaced with a different amino acid, inserted or deleted from the sequence via mutagenesis of nucleic acids encoding these amino acids. Mutagenesis is a well-known method in the art, and includes, for example, site-directed mutagenesis by means of PCR or via oligonucleotide-mediated mutagenesis as described in Sambrook et al., Molecular Cloning-A Laboratory Manual, 2nd ed., Vol. 1-3 (1989). The term "mutated" as used herein regarding genes means that at least one nucleotide in the nucleic acid sequence of that gene or a regulatory sequence thereof, has been replaced with a different nucleotide, or has been deleted from the sequence via mutagenesis, resulting in the transcription of a protein sequence with a qualitatively of quantitatively altered function or the knock-out of that gene.

The term “gene”, as used herein, refers to a nucleic acid sequence containing a template for a nucleic acid polymerase, in eukaryotes, RNA polymerase II. Genes are transcribed into mRNAs that are then translated into protein.

The term "nucleic acid" as used herein, includes reference to a deoxyribonucleotide or ribonucleotide polymer, *i.e.* a polynucleotide, in either single or double-stranded form, and unless otherwise limited, encompasses known analogues having the essential nature of natural nucleotides in that they hybridize to single-stranded nucleic acids in a manner similar to naturally occurring nucleotides (e. g., peptide nucleic acids). A polynucleotide can be full-length or a subsequence of a native or heterologous structural or regulatory gene. Unless otherwise indicated, the term includes reference to the specified sequence as well as the complementary sequence thereof. Thus, DNAs or RNAs with backbones modified for stability or for other reasons are "polynucleotides" as that term is intended herein. Moreover, DNAs or RNAs comprising unusual bases, such as inosine, or modified bases, such as tritylated bases, to name just two examples, are polynucleotides as the term is used herein. It will be appreciated that a great variety of modifications have been made to DNA and RNA that serve many useful purposes known to those of skill in the art. The term polynucleotide as it is employed herein embraces such chemically, enzymatically or metabolically modified forms of polynucleotides, as well as the chemical forms of DNA and RNA characteristic of viruses and cells, including among other things, simple and complex cells.

The terms "polypeptide", "peptide" and "protein" are used interchangeably herein to refer to a polymer of amino acid residues. The terms apply to amino acid polymers in which one or more amino acid residue is an artificial chemical analogue of a corresponding naturally occurring amino acid, as well as to naturally occurring amino acid polymers. The essential nature of such analogues of naturally occurring amino acids is that, when incorporated into a protein, that protein is specifically reactive to antibodies elicited to the same protein but consisting entirely of naturally occurring amino acids. The terms "polypeptide", "peptide" and "protein" are also inclusive of modifications including, but not limited to, glycosylation, lipid attachment, sulphation, gamma-carboxylation of glutamic acid residues, hydroxylation and ADP-ribosylation.

When an enzyme is mentioned with reference to an enzyme class (EC), the enzyme class is a class wherein the enzyme is classified or may be classified, on the basis of the Enzyme Nomenclature provided by the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology (NC-IUBMB), which nomenclature may be found at http://www.chem.qmul.ac.uk/iubmb/enzyme/. Other suitable enzymes that have not (yet) been classified in a specified class but may be classified as such, are meant to be included.

If referred herein to a protein or a nucleic acid sequence, such as a gene, by reference to a accession number, this number in particular is used to refer to a protein or nucleic acid sequence (gene) having a sequence as can be found via www.ncbi.nlm.nih.gov/, (as available on 13 July 2009) unless specified otherwise.

Every nucleic acid sequence herein that encodes a polypeptide also, by reference to the genetic code, describes every possible silent variation of the nucleic acid. The term "conservatively modified variants" applies to both amino acid and nucleic acid sequences. With respect to particular nucleic acid sequences, conservatively modified variants refers to those nucleic acids which encode identical or conservatively modified variants of the amino acid sequences due to the degeneracy of the genetic code. The term "degeneracy of the genetic code" refers to the fact that a large number of functionally identical nucleic acids encode any given protein. For instance, the codons GCA, GCC, GCG and GCU all encode the amino acid alanine. Thus, at every position where an alanine is specified by a codon, the codon can be altered to any of the corresponding codons described without altering the encoded polypeptide. Such nucleic acid variations are "silent variations" and represent one species of conservatively modified variation.

The term “functional homologue” (or in short “homologue”) of a polypeptide having a specific sequence (*e.g.* SEQ ID NO: X), as used herein, refers to a polypeptide comprising said specific sequence with the proviso that one or more amino acids are substituted, deleted, added, and/or inserted, and which polypeptide has (qualitatively) the same enzymatic functionality for substrate conversion. This functionality may be tested by use of an assay system comprising a recombinant yeast cell comprising an expression vector for the expression of the homologue in yeast, said expression vector comprising a heterologous nucleic acid sequence operably linked to a promoter functional in the yeast and said heterologous nucleic acid sequence encoding the homologous polypeptide of which enzymatic activity for converting acetyl-Coenzyme A to acetaldehyde in the yeast cell is to be tested, and assessing whether said conversion occurs in said cells. Candidate homologues may be identified by using *in silico* similarity analyses. A detailed example of such an analysis is described in Example 2 of WO2009/013159. The skilled person will be able to derive there from how suitable candidate homologues may be found and, optionally upon codon(pair) optimization, will be able to test the required functionality of such candidate homologues using a suitable assay system as described above. A suitable homologue represents a polypeptide having an amino acid sequence similar to a specific polypeptide of more than 50%, preferably of 60 % or more, in particular of at least 70 %, more in particular of at least 80 %, at least 90 %, at least 95 %, at least 97 %, at least 98 % or at least 99 % and having the required enzymatic functionality. With respect to nucleic acid sequences, the term functional homologue is meant to include nucleic acid sequences which differ from another nucleic acid sequence due to the degeneracy of the genetic code and encode the same polypeptide sequence.

Sequence identity is herein defined as a relationship between two or more amino acid (polypeptide or protein) sequences or two or more nucleic acid (polynucleotide) sequences, as determined by comparing the sequences. Usually, sequence identities or similarities are compared over the whole length of the sequences compared. In the art, "identity" also means the degree of sequence relatedness between amino acid or nucleic acid sequences, as the case may be, as determined by the match between strings of such sequences.

Amino acid or nucleotide sequences are said to be homologous when exhibiting a certain level of similarity. Two sequences being homologous indicate a common evolutionary origin. Whether two homologous sequences are closely related or more distantly related is indicated by “percent identity” or “percent similarity”, which is high or low respectively. Although disputed, to indicate “percent identity” or “percent similarity”, “level of homology” or “percent homology” are frequently used interchangeably. A comparison of sequences and determination of percent identity between two sequences can be accomplished using a mathematical algorithm. The skilled person will be aware of the fact that several different computer programs are available to align two sequences and determine the homology between two sequences (Kruskal, J. B. (1983) An overview of sequence comparison In D. Sankoff and J. B.

Kruskal, (ed.), Time warps, string edits and macromolecules: the theory and practice of sequence comparison, pp. 1-44 Addison Wesley). The percent identity between two amino acid sequences can be determined using the Needleman and Wunsch algorithm for the alignment of two sequences. (Needleman, S. B. and Wunsch, C. D. (1970) J. Mol. Biol. 48, 443-453). The algorithm aligns amino acid sequences as well as nucleotide sequences. The Needleman-Wunsch algorithm has been implemented in the computer program NEEDLE. For the purpose of this invention the NEEDLE program from the EMBOSS package was used (version 2.8.0 or higher, EMBOSS: The European Molecular Biology Open Software Suite (2000) Rice,P. Longden,I. and Bleasby,A. Trends in Genetics 16, (6) pp276—277, http://emboss.bioinformatics.nl/). For protein sequences, EBLOSUM62 is used for the substitution matrix. For nucleotide sequences, EDNAFULL is used. Other matrices can be specified. The optional parameters used for alignment of amino acid sequences are a gap-open penalty of 10 and a gap extension penalty of 0.5. The skilled person will appreciate that all these different parameters will yield slightly different results but that the overall percentage identity of two sequences is not significantly altered when using different algorithms.

Global Homology Definition

The homology or identity is the percentage of identical matches between the two full sequences over the total aligned region including any gaps or extensions. The homology or identity between the two aligned sequences is calculated as follows: Number of corresponding positions in the alignment showing an identical amino acid in both sequences divided by the total length of the alignment including the gaps. The identity defined as herein can be obtained from NEEDLE and is labelled in the output of the program as “IDENTITY”.

Longest Identity Definition

The homology or identity between the two aligned sequences is calculated as follows: Number of corresponding positions in the alignment showing an identical amino acid in both sequences divided by the total length of the alignment after subtraction of the total number of gaps in the alignment. The identity defined as herein can be obtained from NEEDLE by using the NOBRIEF option and is labeled in the output of the program as “longest-identity”.

A variant of a nucleotide or amino acid sequence disclosed herein may also be defined as a nucleotide or amino acid sequence having one or several substitutions, insertions and/or deletions as compared to the nucleotide or amino acid sequence specifically disclosed herein (e.g. in de the sequence listing).

Optionally, in determining the degree of amino acid similarity, the skilled person may also take into account so-called "conservative" amino acid substitutions, as will be clear to the skilled person. Conservative amino acid substitutions refer to the interchangeability of residues having similar side chains. For example, a group of amino acids having aliphatic side chains is glycine, alanine, valine, leucine, and isoleucine; a group of amino acids having aliphatic-hydroxyl side chains is serine and threonine; a group of amino acids having amide-containing side chains is asparagine and glutamine; a group of amino acids having aromatic side chains is phenylalanine, tyrosine, and tryptophan; a group of amino acids having basic side chains is lysine, arginine, and histidine; and a group of amino acids having sulphur-containing side chains is cysteine and methionine. Preferred conservative amino acids substitution groups are: valine-leucine-isoleucine, phenylalanine-tyrosine, lysine-arginine, alanine-valine, and asparagine-glutamine. Substitutional variants of the amino acid sequence disclosed herein are those in which at least one residue in the disclosed sequences has been removed and a different residue inserted in its place. Preferably, the amino acid change is conservative. Preferred conservative substitutions for each of the naturally occurring amino acids are as follows: Ala to ser; Arg to lys; Asn to gln or his; Asp to glu; Cys to ser or ala; Gln to asn; Glu to asp; Gly to pro; His to asn or gln; Ile to leu or val; Leu to ile or val; Lys to arg; gln or glu; Met to leu or ile; Phe to met, leu or tyr; Ser to thr; Thr to ser; Trp to tyr; Tyr to trp or phe; and, Val to ile or leu.

Nucleotide sequences of the invention may also be defined by their capability to hybridise with parts of specific nucleotide sequences disclosed herein, respectively, under moderate, or preferably under stringent hybridisation conditions. Stringent hybridisation conditions are herein defined as conditions that allow a nucleic acid sequence of at least about 25, preferably about 50 nucleotides, 75 or 100 and most preferably of about 200 or more nucleotides, to hybridise at a temperature of about 65°C in a solution comprising about 1 M salt, preferably 6 x SSC or any other solution having a comparable ionic strength, and washing at 65°C in a solution comprising about 0.1 M salt, or less, preferably 0.2 x SSC or any other solution having a comparable ionic strength. Preferably, the hybridisation is performed overnight, i.e. at least for 10 hours and preferably washing is performed for at least one hour with at least two changes of the washing solution. These conditions will usually allow the specific hybridisation of sequences having about 90% or more sequence identity.

Moderate conditions are herein defined as conditions that allow a nucleic acid sequences of at least 50 nucleotides, preferably of about 200 or more nucleotides, to hybridise at a temperature of about 45°C in a solution comprising about 1 M salt, preferably 6 x SSC or any other solution having a comparable ionic strength, and washing at room temperature in a solution comprising about 1 M salt, preferably 6 x SSC or any other solution having a comparable ionic strength. Preferably, the hybridisation is performed overnight, i.e. at least for 10 hours, and preferably washing is performed for at least one hour with at least two changes of the washing solution. These conditions will usually allow the specific hybridisation of sequences having up to 50% sequence identity. The person skilled in the art will be able to modify these hybridisation conditions in order to specifically identify sequences varying in identity between 50% and 90%.

"Expression" refers to the transcription of a gene into structural RNA (rRNA, tRNA) or messenger RNA (mRNA) with subsequent translation into a protein.

As used herein, "heterologous" in reference to a nucleic acid or protein is a nucleic acid or protein that originates from a foreign species, or, if from the same species, is substantially modified from its native form in composition and/or genomic locus by deliberate human intervention. For example, a promoter operably linked to a heterologous structural gene is from a species different from that from which the structural gene was derived, or, if from the same species, one or both are substantially modified from their original form. A heterologous protein may originate from a foreign species or, if from the same species, is substantially modified from its original form by deliberate human intervention.

The term “heterologous expression” refers to the expression of heterologous nucleic acids in a host cell. The expression of heterologous proteins in eukaryotic host cell systems such as yeast are well known to those of skill in the art. A polynucleotide comprising a nucleic acid sequence of a gene encoding an enzyme with a specific activity can be expressed in such a eukaryotic system. In some embodiments, transformed/transfected yeast cells may be employed as expression systems for the expression of the enzymes. Expression of heterologous proteins in yeast is well known. Sherman, F., et al., Methods in Yeast Genetics, Cold Spring Harbor Laboratory (1982) is a well recognized work describing the various methods available to express proteins in yeast. Two widely utilized yeasts are *Saccharomyces cerevisiae* and *Pichia pastoris*. Vectors, strains, and protocols for expression in *Saccharomyces* and *Pichia* are known in the art and available from commercial suppliers (*e.g.*, Invitrogen). Suitable vectors usually have expression control sequences, such as promoters, including 3-phosphoglycerate kinase or alcohol oxidase, and an origin of replication, termination sequences and the like as desired.

As used herein "promoter" is a DNA sequence that directs the transcription of a (structural) gene. Typically, a promoter is located in the 5'-region of a gene, proximal to the transcriptional start site of a (structural) gene. Promoter sequences may be constitutive, inducible or repressible. If a promoter is an inducible promoter, then the rate of transcription increases in response to an inducing agent.

The term “vector” as used herein, includes reference to an autosomal expression vector and to an integration vector used for integration into the chromosome.

The term "expression vector" refers to a DNA molecule, linear or circular, that comprises a segment encoding a polypeptide of interest under the control of (*i.e.* operably linked to) additional nucleic acid segments that provide for its transcription. Such additional segments may include promoter and terminator sequences, and may optionally include one or more origins of replication, one or more selectable markers, an enhancer, a polyadenylation signal, and the like. Expression vectors are generally derived from plasmid or viral DNA, or may contain elements of both. In particular an expression vector comprises a nucleic acid sequence that comprises in the 5' to 3' direction and operably linked: (a) a yeast-recognized transcription and translation initiation region, (b) a coding sequence for a polypeptide of interest, and (c) a yeast-recognized transcription and translation termination region. “Plasmid" refers to autonomously replicating extrachromosomal DNA which is not integrated into a microorganism's genome and is usually circular in nature.

An “integration vector” refers to a DNA molecule, linear or circular, that can be incorporated in a microorganism's genome and provides for stable inheritance of a gene encoding a polypeptide of interest. The integration vector generally comprises one or more segments comprising a gene sequence encoding a polypeptide of interest under the control of (*i.e.* operably linked to) additional nucleic acid segments that provide for its transcription. Such additional segments may include promoter and terminator sequences, and one or more segments that drive the incorporation of the gene of interest into the genome of the target cell, usually by the process of homologous recombination. Typically, the integration vector will be one which can be transferred into the target cell, but which has a replicon which is nonfunctional in that organism. Integration of the segment comprising the gene of interest may be selected if an appropriate marker is included within that segment.

By "host cell" is meant a cell which contains a vector and supports the replication and/or expression of the vector. Host cells may be prokaryotic cells such as *E. coli*, or eukaryotic cells such as yeast, insect, amphibian, or mammalian cells. Preferably, host cells are eukaryotic cells of the order of *Actinomycetales*.

"Transformation" and "transforming", as used herein, refers to the insertion of an exogenous polynucleotide into a host cell, irrespective of the method used for the insertion, for example, direct uptake, transduction, f-mating or electroporation. The exogenous polynucleotide may be maintained as a non-integrated vector, for example, a plasmid, or alternatively, may be integrated into the host cell genome.

The microorganism, preferably is selected from the group of *Saccharomyceraceae*, such as *Saccharomyces cerevisiae*, *Saccharomyces pastorianus, Saccharomyces beticus, Saccharomyces fermentati, Saccharomyces paradoxus, Saccharomyces uvarum and Saccharomyces bayanus;_Schizosaccharomyces such as Schizosaccharomyces pombe, Schizosaccharomyces japonicus, Schizosaccharomyces octosporus* and *Schizosaccharomyces cryophilus; Torulaspora such as Torulaspora delbrueckii; Kluyveromyces* such as *Kluyveromyces marxianus*; *Pichia* such as *Pichia stipitis*, *Pichia pastoris* or *pichia angusta, Zygosaccharomyces* such as *Zygosaccharomyces bailii*; *Brettanomyces* such as *Brettanomyces intermedius,.Brettanomyces bruxellensis, Brettanomyces anomalus, Brettanomyces custersianus, Brettanomyces naardenensis, Brettanomyces nanus, Dekkera* *bruxellensis* and *Dekkera anomala; Metschnikowia, Issatchenkia, such as Issatchenkia orientalis, Kloeckera such as Kloeckera apiculata; Aureobasidium* such as *Aureobasidium pullulans.*

In a highly preferred embodiment, the microorganism is a yeast cell is selected from the group of *Saccharomyceraceae*. In particular, good results have been achieved with a *Saccharomyces cerevisiae* cell. It has been found possible to use such a cell according to the invention in a method for preparing an alcohol (ethanol) wherein the NADH-dependent side-product formation (glycerol) was reduced by about 90 %, and wherein the yield of the desired product (ethanol) was increase by about 10 %, compared to a similar cell without Rubisco and PRK.

The Rubisco may in principle be selected from eukaryotic and prokaryotic Rubisco's.

The Rubisco is preferably from a non-phototrophic organism. In particular, the Rubisco may be from a chemolithoautotrophic microorganism.

Good results have been achieved with a bacterial Rubisco. Preferably, the bacterial Rubisco originates from a *Thiobacillus,* in particular, *Thiobacillus denitrificans,* which is chemolithoautotrophic.

The Rubisco may be a single-subunit Rubisco or a Rubisco having more than one subunit. In particular, good results have been achieved with a single-subunit Rubisco.

In particular, good results have been achieved with a form-II Rubisco, more in particular CbbM.

SEQUENCE ID NO: 2 shows the sequence of a particularly preferred Rubisco in accordance with the invention. It is encoded by the *cbbM* gene from *Thiobacillus denitrificans*. A preferred alternative to this Rubisco, is a functional homologue of this Rubisco, in particular such functional homologue comprising a sequence having at least 80% , 85%, 90 % or 95% sequence identity with SEQUENCE ID NO: 2. Suitable natural Rubisco polypeptides are given in Table 1.

Table 1: Rubisco polypeptides

Source	**Accession no.**	**MAX ID (%)**
Thiobacillus denitrificans	AAA99178.2	100
Sideroxydans lithotrophicus ES-1	YP_003522651.1	94

Thiothrix nivea DSM 5205	ZP_10101642.1	91
Halothiobacillus neapolitanus c2	YP_003262978.1	90
Acidithiobacillus ferrooxidans ATCC 53993	YP_002220242.1	88
Rhodoferax ferrireducens T118	YP_522655.1	86
Thiorhodococcus drewsii AZ1	ZP_08824342.1	85
uncultured prokaryote	AGE14067.1	82

In accordance with the invention, the Rubisco is functionally expressed in the microorganism, at least during use in an industrial process for preparing a compound of interest.

To increase the likelihood that herein enzyme activity is expressed at sufficient levels and in active form in the transformed (recombinant) host cells of the invention, the nucleotide sequence encoding these enzymes, as well as the Rubisco enzyme and other enzymes of the invention (see below), are preferably adapted to optimise their codon usage to that of the host cell in question. The adaptiveness of a nucleotide sequence encoding an enzyme to the codon usage of a host cell may be expressed as codon adaptation index (CAI). The codon adaptation index is herein defined as a measurement of the relative adaptiveness of the codon usage of a gene towards the codon usage of highly expressed genes in a particular host cell or organism. The relative adaptiveness (w) of each codon is the ratio of the usage of each codon, to that of the most abundant codon for the same amino acid. The CAI index is defined as the geometric mean of these relative adaptiveness values. Non-synonymous codons and termination codons (dependent on genetic code) are excluded. CAI values range from 0 to 1, with higher values indicating a higher proportion of the most abundant codons (see Sharp and Li , 1987, Nucleic Acids Research 15: 1281-1295; also see: Jansen et al., 2003, Nucleic Acids Res. 31(8):2242-51). An adapted nucleotide sequence preferably has a CAI of at least 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8 or 0.9. Most preferred are the sequences which have been codon optimised for expression in the fungal host cell in question such as e.g. *S. cerevisiae* cells.

Preferably, the functionally expressed Rubisco has an activity, defined by the rate of ribulose-1,5-bisphosphate- dependent ^{14}C-bicarbonate incorporation by cell extracts of at least 1 $nmol.min^{-1}.(mg\ protein)^{-1}$, in particular an activity of at least 2 $nmol.min^{-1}.(mg\ protein)^{-1}$, more in particular an activity of at least 4 $nmol.min^{-1}$.(mg protein)$^{-1}$. The upper limit for the activity is not critical. In practice, the activity may be about 200 nmol.min^{-1}.(mg protein)$^{-1}$ or less, in particular 25 nmol.min^{-1}.(mg protein)$^{-1}$, more in particular 15 nmol.min^{-1}.(mg protein)$^{-1}$ or less, *e.g.* about 10 nmol.min^{-1}.(mg protein)$^{-1}$ or less. When referred herein to the activity of Rubisco, in particular the activity at 30 °C is meant. The conditions for an assay for determining this Rubisco activity are as found in the Examples, below (Example 4).

A functionally expressed phosphoribulokinase (PRK, (EC 2.7.1.19)) according to the invention is capable of catalysing the chemical reaction:

ATP + D-ribulose 5-phosphate $\rightleftharpoons$ *ADP + D-ribulose 1,5-bisphosphate (1)*

Thus, the two substrates of this enzyme are ATP and D-ribulose 5-phosphate, whereas its two products are ADP and D-ribulose 1,5-bisphosphate.

PRK belongs to the family of transferases, specifically those transferring phosphorus-containing groups (phosphotransferases) with an alcohol group as acceptor. The systematic name of this enzyme class is ATP:D-ribulose-5-phosphate 1-phosphotransferase. Other names in common use include phosphopentokinase, ribulose-5-phosphate kinase, phosphopentokinase, phosphoribulokinase (phosphorylating), 5-phosphoribulose kinase, ribulose phosphate kinase, PKK, PRuK, and PRK. This enzyme participates in carbon fixation.

The PRK can be from a prokaryote or a eukaryote. Good results have been achieved with a PRK originating from a eukaryote. Preferably the eukaryotic PRK originates from a plant selected from *Caryophyllales* , in particular from *Amaranthaceae,* more in particular from *Spinacia.*

As a preferred alternative to PRK from *Spinacia* a functional homologue of PRK from *Spinacia* may be present, in particular a functional homologue comprising a sequence having at least 70%, 75%, 80%. 85%, 90 % or 95% sequence identity with SEQUENCE ID NO 4.

Suitable natural PRK polypeptides are given in Table 2.

Table 2: Natural PRK polypeptides suitable for expression

Source	**Accession no.**	**MAX ID (%)**
Spinacia oleracea	P09559.1	100
Medicago truncatula	XP_003612664.1	88

Arabidopsis thaliana	NP_174486.1	87
Vitis vinifera	XP_002263724.1	86
Closterium peracerosum	BAL03266.1	82
Zea mays	NP_001148258.1	78

In an advantageous embodiment, the recombinant microorganism further comprises a nucleic acid sequence encoding one or more heterologous prokaryotic or eukaryotic molecular chaperones, which – when expressed – are capable of functionally interacting with an enzyme in the microorganism, in particular with at least one of Rubisco and PRK.

Chaperonins are proteins that provide favourable conditions for the correct folding of other proteins, thus preventing aggregation. Newly made proteins usually must fold from a linear chain of amino acids into a three-dimensional form. Chaperonins belong to a large class of molecules that assist protein folding, called molecular chaperones. The energy to fold proteins is supplied by adenosine triphosphate (ATP). A review article about chaperones that is useful herein is written by Yébenes (2001); “Chaperonins: two rings for folding”; Hugo Yébenes et al. Trends in Biochemical Sciences, August 2011, Vol. 36, No. 8.

In a preferred embodiment, the chaperone or chaperones are from a bacterium, more preferably from *Escherichia,* in particular *E. coli* GroEL and GroEs from E. coli may in particular encoded in a microorganism according to the invention. Other preferred chaperones are chaperones from Saccharomyces, in particular Saccharomyces cerevisiae Hsp10 and Hsp60. If the chaperones are naturally expressed in an organelle such as a mitochondrion (examples are Hsp60 and Hsp10 of Saccharomyces cerevisiae) relocation to the cytosol can be achieved e.g. by modifying the native signal sequence of the chaperonins.

In eukaryotes the proteins Hsp60 and Hsp10 are structurally and functionally nearly identical to GroEL and GroES, respectively. Thus, it is contemplated that Hsp60 and Hsp10 from any eukaryotic cell may serve as a chaperone for the Rubisco. See Zeilstra-Ryalls J, Fayet O, Georgopoulos C (1991). "The universally conserved GroE (Hsp60) chaperonins". Annu Rev Microbiol. 45: 301–25. doi:10.1146/annurev.mi.45.100191.001505. PMID 1683763 and Horwich AL, Fenton WA, Chapman E, Farr GW (2007). "Two Families of Chaperonin: Physiology and Mechanism". Annu Rev Cell Dev Biol. 23: 115–45. doi:10.1146/annurev.cellbio.23.090506.123555. PMID 17489689.

Particularly good results have been achieved with a recombinant yeast cell comprising both the heterologous chaperones GroEL and GroES.

As a preferred alternative to GroEL a functional homologue of GroEL may be present, in particular a functional homologue comprising a sequence having at least 70%, 75%, 80%, 85%, 90 % or 95% sequence identity with SEQUENCE ID NO: 10.

Suitable natural chaperones polypeptide homologous to SEQUENCE ID NO: 10 are given in Table 3.

Table 3: Natural chaperones homologous to SEQUENCE ID NO: 10 polypeptides suitable for expression

>gi|115388105|ref|XP_001211558.1|:2-101 10 kDa heat shock protein, mitochondrial [Aspergillus terreus NIH2624]
>gi|116196854|ref|XP_001224239.1|:1-102 conserved hypothetical protein [Chaetomium globosum CBS 148.51]
>gi|119175741|ref|XP_001240050.1|:3-102 hypothetical protein CIMG_09671 [Coccidioides immitis RS]
>gi|119471607|ref|XP_001258195.1|:12-111 chaperonin, putative [Neosartorya fischeri NRRL181]
>gi|121699818|ref|XP_001268174.1|:8-106 chaperonin, putative [Aspergillus clavatus NRRL 1]
>gi|126274604|ref|XP_001387607.1|:2-102 predicted protein [Scheffersomyces stipitis CBS 6054]
>gi|146417701|ref|XP_001484818.1|:5-106 conserved hypothetical protein [Meyerozyma guilliermondii ATCC 6260]
>gi|154303611|ref|XP_001552212.1|:1-102 10 kDa heat shock protein, mitochondrial [Botryotinia fuckeliana B05.10]
>gi|156049571|ref|XP_001590752.1|:1-102 hypothetical protein SS1G_08492 [Sclerotinia sclerotiorum 1980]
>gi|156840987|ref|XP_001643870.1|:1-103 hypothetical protein Kpol_495p10

[Vanderwaltozyma polyspora DSM 70924]
>gi|169608295|ref|XP_001797567.1|:1-101 hypothetical protein SNOG_07218 [Phaeosphaeria nodorum SN15]
>gi|171688384|ref|XP_001909132.1|:1-102 hypothetical protein [Podospora anserina S mat+]
>gi|189189366|ref|XP_001931022.1|:71-168 10 kDa chaperonin [Pyrenophora tritici-repentis Pt-1C-BFP]
>gi|19075598|ref|NP_588098.1|:1-102 mitochondrial heat shock protein Hsp10 (predicted) [Schizosaccharomyces pombe 972h-]
>gi|212530240|ref|XP_002145277.1|:3-100 chaperonin, putative [Talaromyces marneffei ATCC 18224]
>gi|212530242|ref|XP_002145278.1|:3-95 chaperonin, putative [Talaromyces marneffei ATCC 18224]
>gi|213404320|ref|XP_002172932.1|:1-102 mitochondrial heat shock protein Hsp10 [Schizosaccharomyces japonicus yFS275]
>gi|225557301|gb|EEH05587.1|:381-478 pre-mRNA polyadenylation factor fip1 [Ajellomyces capsulatus G186AR]
>gi|225684092|gb|EEH22376.1|:3-100 heat shock protein [Paracoccidioides brasiliensis Pb03
>gi|238490530|ref|XP_002376502.1|:2-104 chaperonin, putative [Aspergillus flavus NRRL3357
>gi|238878220|gb|EEQ41858.1|:1-106 10 kDa heat shock protein, mitochondrial [Candida albicans WO-1]
>gi|240280207|gb|EER43711.1|:426-523 pre-mRNA polyadenylation factor fip1 [Ajellomyces capsulatus H143]
>gi|241950445|ref|XP_002417945.1|:1-103 10 kda chaperonin, putative; 10 kda heat shock protein mitochondrial (hsp10), putative [Candida dubliniensis CD36]
>gi|242819222|ref|XP_002487273.1|:90-182 chaperonin, putative [Talaromyces stipitatus ATC
>gi|254566327|ref|XP_002490274.1|:1-102 Putative protein of unknown function [Komagataella pastoris GS115]

>gi|254577241|ref|XP_002494607.1|:1-103 ZYRO0A05434p [Zygosaccharomyces rouxii]
>gi|255717999|ref|XP_002555280.1|:1-103 KLTH0G05588p [Lachancea thermotolerans]
>gi|255956581|ref|XP_002569043.1|:2-101 Pc21g20560 [Penicillium chrysogenum Wisconsin 54-1255]
>gi|258572664|ref|XP_002545094.1|:16-108 chaperonin GroS [Uncinocarpus reesii 1704]
>gi|261190594|ref|XP_002621706.1|:3-100 chaperonin [Ajellomyces dermatitidis SLH14081]
>gi|295664909|ref|XP_002793006.1|:3-100 10 kDa heat shock protein, mitochondrial [Paracoccidioides sp. 'lutzii'Pb01]
>gi|296412657|ref|XP_002836039.1|:76-177 hypothetical protein [Tuber melanosporum Mel28]
>gi|302307854|ref|NP_984626.2|:2-102 AEL235Wp [Ashbya gossypii ATCC 10895]
>gi|302894117|ref|XP_003045939.1|:1-102 predicted protein [Nectria haematococca mpVI 77-13-4]
>gi|303318351|ref|XP_003069175.1|:3-100 10 kDa heat shock protein, mitochondrial , putative [Coccidioides posadasii C735 delta SOWgp]
>gi|310795300|gb|EFQ30761.1|:1-102 chaperonin 10 kDa subunit [Glomerella graminicola M1.001]
>gi|315053085|ref|XP_003175916.1|:12-109 chaperonin GroS [Arthroderma gypseum CBS 118893]
>gi|317032114|ref|XP_001394060.2|:334-433 heat shock protein [Aspergillus niger CBS 513.88]
>gi|317032116|ref|XP_001394059.2|:2-101 heat shock protein [Aspergillus niger CBS 513.88]
>gi|320583288|gb|EFW97503.1|:6-106 chaperonin, putative heat shock protein, putative [Ogataea parapolymorpha DL-1]
>gi|320591507|gb|EFX03946.1|:1-102 heat shock protein [Grosmannia clavigera kw1407]

>gi|322700925|gb|EFY92677.1|:1-102 chaperonin [Metarhizium acridum CQMa 102]
>gi|325096696|gb|EGC50006.1|:409-506 pre-mRNA polyadenylation factor fip1 [Ajellomyces capsulatus H88]
>gi|326471604|gb|EGD95613.1|:14-111 chaperonin 10 Kd subunit [Trichophyton tonsurans CBS112818]
>gi|327293056|ref|XP_003231225.1|:3-100 chaperonin [Trichophyton rubrum CBS 118892]
>gi|330942654|ref|XP_003306155.1|:37-136 hypothetical protein PTT_19211 [Pyrenophora teres f. teres 0-1]
>gi|336268042|ref|XP_003348786.1|:47-147 hypothetical protein SMAC_01809 [Sordaria macrospora khell]
>gi|340519582|gb|EGR49820.1|:1-109 predicted protein [Trichoderma reesei QM6a]
>gi|340960105|gb|EGS21286.1|:3-103 putative mitochondrial 10 kDa heat shock protein [Chaetomium thermophilum var. thermophilum DSM 1495]
>gi|342883802|gb|EGU84224.1|:1-102 hypothetical protein FOXB_05181 [Fusarium oxysporum Fo5176]
>gi|344302342|gb|EGW32647.1|:2-102 hypothetical protein SPAPADRAFT_61712 [Spathaspora passalidarum NRRL Y-27907]
>gi|345570750|gb|EGX53571.1|:1-102 hypothetical protein AOL_s00006g437 [Arthrobotrys oligospora ATCC 24927]
>gi|346321154|gb|EGX90754.1|:1-102 chaperonin [Cordyceps militaris CM01]
>gi|346970393|gb|EGY13845.1|:1-102 heat shock protein [Verticillium dahliae VdLs.17]
>gi|354548296|emb|CCE45032.1|:1-106 hypothetical protein CPAR2_700360 [Candida parapsilosis]
>gi|358385052|gb|EHK22649.1|:1-102 hypothetical protein TRIVIDRAFT_230640 [Trichoderma virens Gv 29-8]
>gi|358393422|gb|EHK42823.1|:1-101 hypothetical protein TRIATDRAFT_258186 [Trichoderma atroviride IMI 206040]
>gi|361126733|gb|EHK98722.1|:1-97 putative 10 kDa heat shock protein,

mitochondrial [Glare lozoyensis 74030]
>gi|363753862|ref|XP_003647147.1|:2-102 hypothetical protein Ecym_5593 [Eremothecium cymbalariae DBVPG#7215]
>gi|365758401|gb|EHN00244.1|:1-106 Hsp10p [Saccharomyces cerevisiae x Saccharomyces kudriavzevii VIN7]
>gi|365987664|ref|XP_003670663.1|:1-103 hypothetical protein NDAI_0F01010 [Naumovozyma dairenensis CBS 421]
>gi|366995125|ref|XP_003677326.1|:1-103 hypothetical protein NCAS_0G00860 [Naumovozyma castellii CBS 4309]
>gi|366999797|ref|XP_003684634.1|:1-103 hypothetical protein TPHA_0C00430 [Tetrapisispora phaffii CBS 4417]
>gi|367009030|ref|XP_003679016.1|:1-103 hypothetical protein TDEL_0A04730 [Torulaspora delbruekii]
>gi|367023138|ref|XP_003660854.1|:1-104 hypothetical protein MYCTH_59302 [Myceliophthora thermophila ATCC 42464]
>gi|367046344|ref|XP_003653552.1|:1-102 hypothetical protein THITE_2116070 [Thielavia terrestris NRRL8126]
>gi|378726440|gb|EHY52899.1|:9-109 chaperonin GroES [Exophiala dermatitidis NIH/UT8656]
>gi|380493977|emb|CCF33483.1|:1-102 chaperonin 10 kDa subunit [Colletotrichum higginsianu
>gi|385305728|gb|EIF49680.1|:1-102 10 kda heat shock mitochondrial [Dekkera bruxellensis AWRI1499]
>gi|389628546|ref|XP_003711926.1|:1-102 hsp10-like protein [Magnaporthe oryzae 70-15]
>gi|396462608|ref|XP_003835915.1|:1-101 similar to 10 kDa heat shock protein [Leptosphaeria maculans JN3]
>gi|398392541|ref|XP_003849730.1|:1-102 hypothetical protein MYCGRDRAFT_105721 [Zymoseptoria tritici IPO323]
>gi|400597723|gb|EJP65453.1|:24-124 chaperonin 10 kDa subunit [Beauveria bassiana ARSEF 2860]
>gi|401623646|gb|EJS41738.1|:1-106 hsp10p [Saccharomyces arboricola H-6]

<table>
<tr><td>>gi|401842164|gb|EJT44422.1|:1-92 HSP10-like protein [Saccharomyces kudriavzevii IFO 1802]</td></tr>
<tr><td>>gi|402084027|gb|EJT79045.1|:1-102 hsp10-like protein [Gaeumannomyces graminis var. triti</td></tr>
<tr><td>>gi|403215209|emb|CCK69709.1|:1-104 hypothetical protein KNAG_0C06130 [Kazachstania naganishii CBS 8797]</td></tr>
<tr><td>>gi|406604629|emb|CCH43969.1|:4-100 hypothetical protein BN7_3524 [Wickerhamomyces ciferrii]</td></tr>
<tr><td>>gi|406867021|gb|EKD20060.1|:56-156 hypothetical protein MBM_02012 [Marssonina brunnea f.sp. 'multigermtubi' MB_m1]</td></tr>
<tr><td>>gi|407926227|gb|EKG19196.1|:74-174 GroES-like protein [Macrophomina phaseolina MS6]</td></tr>
<tr><td>>gi|408398157|gb|EKJ77291.1|:11-111 hypothetical protein FPSE_02566 [Fusarium pseudograminearum CS3096]</td></tr>
<tr><td>>gi|410082063|ref|XP_003958610.1|:1-103 hypothetical protein KAFR_0H00660 [Kazachstania africana CBS2517]</td></tr>
<tr><td>>gi|425777664|gb|EKV15823.1|:58-157 Chaperonin, putative [Penicillium digitatum Pd1]</td></tr>
<tr><td>>gi|440639680|gb|ELR09599.1|:1-102 chaperonin GroES [Geomyces destructans 20631-21]</td></tr>
<tr><td>>gi|444323906|ref|XP_004182593.1|:1-105 hypothetical protein TBLA_0J00760 [Tetrapisisporablattae CBS 6284]</td></tr>
<tr><td>>gi|448083208|ref|XP_004195335.1|:2-101 Piso0_005888 [Millerozyma farinosa CBS 7064]</td></tr>
<tr><td>>gi|448087837|ref|XP_004196425.1|:2-102 Piso0_005888 [Millerozyma farinosa CBS 7064]</td></tr>
<tr><td>>gi|448534948|ref|XP_003870866.1|:1-106 Hsp10 protein [Candida orthopsilosis Co 90-125]</td></tr>
<tr><td>>gi|449295977|gb|EMC91998.1|:1-102 hypothetical protein BAUCODRAFT_39148 [Baudoinia compn</td></tr>
<tr><td>>gi|46123659|ref|XP_386383.1|:3-103 hypothetical protein FG06207.1 [Gibberella zeae PH-1]</td></tr>
</table>

>gi|50289455|ref|XP_447159.1|:1-103 hypothetical protein [Candida glabrata CBS 138]
>gi|50308731|ref|XP_454370.1|:1-103 hypothetical protein [Kluyveromyces lactis NRRL Y-1140]
>gi|50411066|ref|XP_457014.1|:1-106 DEHA2B01122p [Debaryomyces hansenii CBS767]
>gi|50545998|ref|XP_500536.1|:1-102 YALI0B05610p [Yarrowia lipolytica]
>gi|51013895|gb|AAT93241.1|:1-106 YOR020C [Saccharomyces cerevisiae]
>gi|6324594|ref|NP_014663.1|:1-106 Hsp10p [Saccharomyces cerevisiae S288c]
>gi|67523953|ref|XP_660036.1|:2-101 hypothetical protein AN2432.2 [Aspergillus nidulans FGSC A4]
>gi|70992219|ref|XP_750958.1|:12-106 chaperonin [Aspergillus fumigatus Af293]
>gi|85079266|ref|XP_956315.1|:1-104 hypothetical protein NCU04334 [Neurospora crassa OR74A]

As a preferred alternative to GroES a functional homologue of GroES may be present, in particular a functional homologue comprising a sequence having at least 70%, 75%, 80%, 85%, 90 % or 95% sequence identity with SEQUENCE ID NO: 12.

Suitable natural chaperones polypeptides homologous to SEQUENCE ID NO: 12 are given in Table 4.

Table 4: Natural chaperones homologous to SEQUENCE ID NO: 12 polypeptides suitable for expression

>gi|115443330|ref|XP_001218472.1| heat shock protein 60, mitochondrial precursor [Aspergillus terreus NIH2624] >gi|114188341|gb|EAU30041.1| heat shock protein 60, mitochondrial precursor [Aspergillus terreus NIH2624]
>gi|119480793|ref|XP_001260425.1| antigenic mitochondrial protein HSP60, putative [Neosartorya fischeri NRRL 181] >gi|119408579|gb|EAW18528.1|

antigenic mitochondrial protein HSP60, putative [Neosartorya fischeri NRRL 181]
>gi|126138730|ref|XP_001385888.1| hypothetical protein PICST_90190 [Scheffersomyces stipitis CBS 6054] >gi|126093166|gb|ABN67859.1| mitochondrial groEL-type heat shock protein [Scheffersomyces stipitis CBS 6054]
>gi|145246630|ref|XP_001395564.1| heat shock protein 60 [Aspergillus niger CBS 513.88] >gi|134080285|emb|CAK46207.1| unnamed protein product [Aspergillus niger] >gi|350636909|gb|EHA25267.1| hypothetical protein ASPNIDRAFT_54001 [Aspergillus niger ATCC 1015]
>gi|146413148|ref|XP_001482545.1| heat shock protein 60, mitochondrial precursor [Meyerozyma guilliermondii ATCC 6260]
>gi|154277022|ref|XP_001539356.1| heat shock protein 60, mitochondrial precursor [Ajellomyces capsulatus NAm1] >gi|150414429|gb|EDN09794.1| heat shock protein 60, mitochondrial precursor [Ajellomyces capsulatus NAm1]
>gi|154303540|ref|XP_001552177.1| heat shock protein 60 [Botryotinia fuckeliana B05.10] >gi|347840915|emb|CCD55487.1| similar to heat shock protein 60 [Botryotinia fuckeliana]
>gi|156063938|ref|XP_001597891.1| heat shock protein 60, mitochondrial precursor [Sclerotinia sclerotiorum 1980] >gi|154697421|gb|EDN97159.1| heat shock protein 60, mitochondrial precursor [Sclerotinia sclerotiorum 1980 UF-70]
>gi|156844469|ref|XP_001645297.1| hypothetical protein Kpol_1037p35 [Vanderwaltozyma polyspora DSM 70294] >gi|156115957|gb|EDO17439.1| hypothetical protein Kpol_1037p35 [Vanderwaltozyma polyspora DSM 70294]
>gi|16416029|emb|CAB91379.2| probable heat-shock protein hsp60 [Neurospora crassa] >gi|350289516|gb|EGZ70741.1| putative heat-shock protein hsp60 [Neurospora tetrasperma FGSC 2509]
>gi|169626377|ref|XP_001806589.1| hypothetical protein SNOG_16475 [Phaeosphaeria nodorum SN15] >gi|111055053|gb|EAT76173.1| hypothetical protein SNOG_16475 [Phaeosphaeria nodorum SN15]
>gi|169783766|ref|XP_001826345.1| heat shock protein 60 [Aspergillus oryzae RIB40] >gi|238493601|ref|XP_002378037.1| antigenic mitochondrial protein HSP60, putative [Aspergillus flavus NRRL3357] >gi|83775089|dbj|BAE65212.1| unnamed protein product [Aspergillus oryzae RIB40]

>gi|220696531|gb|EED52873.1| antigenic mitochondrial protein HSP60, putative [Aspergillus flavus NRRL3357] >gi|391869413|gb|EIT78611.1| chaperonin, Cpn60/Hsp60p [Aspergillus oryzae 3.042]
>gi|189190432|ref|XP_001931555.1| heat shock protein 60, mitochondrial precursor [Pyrenophora tritici-repentis Pt-1C-BFP] >gi|187973161|gb|EDU40660.1| heat shock protein 60, mitochondrial precursor [Pyrenophora tritici-repentis Pt-1C-BFP]
>gi|190348913|gb|EDK41467.2| heat shock protein 60, mitochondrial precursor [Meyerozyma guilliermondii ATCC 6260]
>gi|225554633|gb|EEH02929.1| hsp60-like protein [Ajellomyces capsulatus G186AR]
>gi|238880068|gb|EEQ43706.1| heat shock protein 60, mitochondrial precursor [Candida albicans WO-1]
>gi|239613490|gb|EEQ90477.1| chaperonin GroL [Ajellomyces dermatitidis ER-3]
>gi|240276977|gb|EER40487.1| hsp60-like protein [Ajellomyces capsulatus H143]
>gi|241958890|ref|XP_002422164.1| heat shock protein 60, mitochondrial precursor, putative [Candida dubliniensis CD36] >gi|223645509|emb|CAX40168.1| heat shock protein 60, mitochondrial precursor, putative [Candida dubliniensis CD36]
>gi|254572906|ref|XP_002493562.1| Tetradecameric mitochondrial chaperonin [Komagataella pastoris GS115] >gi|238033361|emb|CAY71383.1| Tetradecameric mitochondrial chaperonin [Komagataella pastoris GS115]
>gi|254579947|ref|XP_002495959.1| ZYRO0C07106p [Zygosaccharomyces rouxii] >gi|238938850|emb|CAR27026.1| ZYRO0C07106p [Zygosaccharomyces rouxii]
>gi|255712781|ref|XP_002552673.1| KLTH0C10428p [Lachancea thermotolerans] >gi|238934052|emb|CAR22235.1| KLTH0C10428p [Lachancea thermotolerans CBS 6340]
>gi|255721795|ref|XP_002545832.1| heat shock protein 60, mitochondrial precursor [Candida tropicalis MYA-3404] >gi|240136321|gb|EER35874.1| heat shock protein 60, mitochondrial precursor [Candida tropicalis MYA-3404]
>gi|255941288|ref|XP_002561413.1| Pc16g11070 [Penicillium chrysogenum Wisconsin 54-1255] >gi|211586036|emb|CAP93777.1| Pc16g11070 [Penicillium

chrysogenum Wisconsin 54-1255]
>gi|259148241|emb|CAY81488.1| Hsp60p [Saccharomyces cerevisiae EC1118]
>gi|260950325|ref|XP_002619459.1| heat shock protein 60, mitochondrial precursor [Clavispora lusitaniae ATCC 42720] >gi|238847031|gb|EEQ36495.1| heat shock protein 60, mitochondrial precursor [Clavispora lusitaniae ATCC 42720]
>gi|261194577|ref|XP_002623693.1| chaperonin GroL [Ajellomyces dermatitidis SLH14081] >gi|239588231|gb|EEQ70874.1| chaperonin GroL [Ajellomyces dermatitidis SLH14081] >gi|327355067|gb|EGE83924.1| chaperonin GroL [Ajellomyces dermatitidis ATCC 18188]
>gi|296422271|ref|XP_002840685.1| hypothetical protein [Tuber melanosporum Mel28] >gi|295636906|emb|CAZ84876.1| unnamed protein product [Tuber melanosporum]
>gi|296809035|ref|XP_002844856.1| heat shock protein 60 [Arthroderma otae CBS 113480] >gi|238844339|gb|EEQ34001.1| heat shock protein 60 [Arthroderma otae CBS 113480]
>gi|302308696|ref|NP_985702.2| AFR155Wp [Ashbya gossypii ATCC 10895] >gi|299790751|gb|AAS53526.2| AFR155Wp [Ashbya gossypii ATCC 10895] >gi|374108933|gb|AEY97839.1| FAFR155Wp [Ashbya gossypii FDAG1]
>gi|302412525|ref|XP_003004095.1| heat shock protein [Verticillium albo-atrum VaMs.102] >gi|261356671|gb|EEY19099.1| heat shock protein [Verticillium albo-atrum VaMs.102]
>gi|302505585|ref|XP_003014499.1| hypothetical protein ARB_07061 [Arthroderma benhamiae CBS 112371] >gi|291178320|gb|EFE34110.1| hypothetical protein ARB_07061 [Arthroderma benhamiae CBS 112371]
>gi|302656385|ref|XP_003019946.1| hypothetical protein TRV_05992 [Trichophyton verrucosum HKI 0517] >gi|291183723|gb|EFE39322.1| hypothetical protein TRV_05992 [Trichophyton verrucosum HKI 0517]
>gi|302915513|ref|XP_003051567.1| predicted protein [Nectria haematococca mpVI 77-13-4] >gi|256732506|gb|EEU45854.1| predicted protein [Nectria haematococca mpVI 77-13-4]
>gi|310794550|gb|EFQ30011.1| chaperonin GroL [Glomerella graminicola M1.001]
>gi|315048491|ref|XP_003173620.1| chaperonin GroL [Arthroderma gypseum CBS

<table>
<tr><td>118893] >gi|311341587|gb|EFR00790.1| chaperonin GroL [Arthroderma gypseum CBS 118893]</td></tr>
<tr><td>>gi|320580028|gb|EFW94251.1| Tetradecameric mitochondrial chaperonin [Ogataea parapolymorpha DL-1]</td></tr>
<tr><td>>gi|320586014|gb|EFW98693.1| heat shock protein mitochondrial precursor [Grosmannia clavigera kw1407]</td></tr>
<tr><td>>gi|322692465|gb|EFY84374.1| Heat shock protein 60 precursor (Antigen HIS-62) [Metarhizium acridum CQMa 102]</td></tr>
<tr><td>>gi|322705285|gb|EFY96872.1| Heat shock protein 60 (Antigen HIS-62) [Metarhizium anisopliae ARSEF 23]</td></tr>
<tr><td>>gi|323303806|gb|EGA57589.1| Hsp60p [Saccharomyces cerevisiae FostersB]</td></tr>
<tr><td>>gi|323307999|gb|EGA61254.1| Hsp60p [Saccharomyces cerevisiae FostersO]</td></tr>
<tr><td>>gi|323332364|gb|EGA73773.1| Hsp60p [Saccharomyces cerevisiae AWRI796]</td></tr>
<tr><td>>gi|326468648|gb|EGD92657.1| heat shock protein 60 [Trichophyton tonsurans CBS 112818] >gi|326479866|gb|EGE03876.1| chaperonin GroL [Trichophyton equinum CBS 127.97]</td></tr>
<tr><td>>gi|330915493|ref|XP_003297052.1| hypothetical protein PTT_07333 [Pyrenophora teres f. teres 0-1] >gi|311330479|gb|EFQ94847.1| hypothetical protein PTT_07333 [Pyrenophora teres f. teres 0-1]</td></tr>
<tr><td>>gi|336271815|ref|XP_003350665.1| hypothetical protein SMAC_02337 [Sordaria macrospora k-hell] >gi|380094827|emb|CCC07329.1| unnamed protein product [Sordaria macrospora k-hell]</td></tr>
<tr><td>>gi|336468236|gb|EGO56399.1| hypothetical protein NEUTE1DRAFT_122948 [Neurospora tetrasperma FGSC 2508]</td></tr>
<tr><td>>gi|340522598|gb|EGR52831.1| hsp60 mitochondrial precursor-like protein [Trichoderma reesei QM6a]</td></tr>
<tr><td>>gi|341038907|gb|EGS23899.1| mitochondrial heat shock protein 60-like protein [Chaetomium thermophilum var. thermophilum DSM 1495]</td></tr>
<tr><td>>gi|342886297|gb|EGU86166.1| hypothetical protein FOXB_03302 [Fusarium oxysporum Fo5176]</td></tr>
<tr><td>>gi|344230084|gb|EGV61969.1| chaperonin GroL [Candida tenuis ATCC 10573]</td></tr>
<tr><td>>gi|344303739|gb|EGW33988.1| hypothetical protein SPAPADRAFT_59397</td></tr>
</table>

[Spathaspora passalidarum NRRL Y-27907]
>gi|345560428|gb|EGX43553.1| hypothetical protein AOL_s00215g289 [Arthrobotrys oligospora ATCC 24927]
>gi|346323592|gb|EGX93190.1| heat shock protein 60 (Antigen HIS-62) [Cordyceps militaris CM01]
>gi|346975286|gb|EGY18738.1| heat shock protein [Verticillium dahliae VdLs.17]
>gi|354545932|emb|CCE42661.1| hypothetical protein CPAR2_203040 [Candida parapsilosis]
>gi|358369894|dbj|GAA86507.1| heat shock protein 60, mitochondrial precursor [Aspergillus kawachii IFO 4308]
>gi|358386867|gb|EHK24462.1| hypothetical protein TRIVIDRAFT_79041 [Trichoderma virens Gv29-8]
>gi|358399658|gb|EHK48995.1| hypothetical protein TRIATDRAFT_297734 [Trichoderma atroviride IMI 206040]
>gi|363750488|ref|XP_003645461.1| hypothetical protein Ecym_3140 [Eremothecium cymbalariae DBVPG#7215] >gi|356889095|gb|AET38644.1| Hypothetical protein Ecym_3140 [Eremothecium cymbalariae DBVPG#7215]
>gi|365759369|gb|EHN01160.1| Hsp60p [Saccharomyces cerevisiae x Saccharomyces kudriavzevii VIN7]
>gi|365764091|gb|EHN05616.1| Hsp60p [Saccharomyces cerevisiae x Saccharomyces kudriavzevii VIN7]
>gi|365985626|ref|XP_003669645.1| hypothetical protein NDAI_0D00880 [Naumovozyma dairenensis CBS 421] >gi|343768414|emb|CCD24402.1| hypothetical protein NDAI_0D00880 [Naumovozyma dairenensis CBS 421]
>gi|366995970|ref|XP_003677748.1| hypothetical protein NCAS_0H00890 [Naumovozyma castellii CBS 4309] >gi|342303618|emb|CCC71399.1| hypothetical protein NCAS_0H00890 [Naumovozyma castellii CBS 4309]
>gi|367005154|ref|XP_003687309.1| hypothetical protein TPHA_0J00520 [Tetrapisispora phaffii CBS 4417] >gi|357525613|emb|CCE64875.1| hypothetical

protein TPHA_0J00520 [Tetrapisispora phaffii CBS 4417]
>gi|367017005|ref|XP_003683001.1| hypothetical protein TDEL_0G04230 [Torulaspora delbrueckii] >gi|359750664|emb|CCE93790.1| hypothetical protein TDEL_0G04230 [Torulaspora delbrueckii]
>gi|367035486|ref|XP_003667025.1| hypothetical protein MYCTH_2097570 [Myceliophthora thermophila ATCC 42464] >gi|347014298|gb|AEO61780.1| hypothetical protein MYCTH_2097570 [Myceliophthora thermophila ATCC 42464]
>gi|367055018|ref|XP_003657887.1| hypothetical protein THITE_127923 [Thielavia terrestris NRRL 8126] >gi|347005153|gb|AEO71551.1| hypothetical protein THITE_127923 [Thielavia terrestris NRRL 8126]
>gi|378728414|gb|EHY54873.1| heat shock protein 60 [Exophiala dermatitidis NIH/UT8656]
>gi|380494593|emb|CCF33032.1| heat shock protein 60 [Colletotrichum higginsianum]
>gi|385305893|gb|EIF49836.1| heat shock protein 60 [Dekkera bruxellensis AWRI1499]
>gi|389638386|ref|XP_003716826.1| heat shock protein 60 [Magnaporthe oryzae 70-15] >gi|351642645|gb|EHA50507.1| heat shock protein 60 [Magnaporthe oryzae 70-15] >gi|440474658|gb|ELQ43388.1| heat shock protein 60 [Magnaporthe oryzae Y34] >gi|440480475|gb|ELQ61135.1| heat shock protein 60 [Magnaporthe oryzae P131]
>gi|393243142|gb|EJD50658.1| chaperonin GroL [Auricularia delicata TFB-10046 SS5]
>gi|396494741|ref|XP_003844378.1| similar to heat shock protein 60 [Leptosphaeria maculans JN3] >gi|312220958|emb|CBY00899.1| similar to heat shock protein 60 [Leptosphaeria maculans JN3]
>gi|398393428|ref|XP_003850173.1| chaperone ATPase HSP60 [Zymoseptoria tritici IPO323] >gi|339470051|gb|EGP85149.1| hypothetical protein MYCGRDRAFT_75170 [Zymoseptoria tritici IPO323]
>gi|401624479|gb|EJS42535.1| hsp60p [Saccharomyces arboricola H-6]
>gi|401842294|gb|EJT44530.1| HSP60-like protein [Saccharomyces kudriavzevii

IFO 1802]
>gi|402076594|gb|EJT72017.1| heat shock protein 60 [Gaeumannomyces graminis var. tritici R3-111a-1]
>gi|403213867|emb|CCK68369.1| hypothetical protein KNAG_0A07160 [Kazachstania naganishii CBS 8797]
>gi|406606041|emb|CCH42514.1| Heat shock protein 60, mitochondrial [Wickerhamomyces ciferrii]
>gi|406863285|gb|EKD16333.1| heat shock protein 60 [Marssonina brunnea f. sp. 'multigermtubi' MB_m1]
>gi|407922985|gb|EKG16075.1| Chaperonin Cpn60 [Macrophomina phaseolina MS6]
>gi|408399723|gb|EKJ78816.1| hypothetical protein FPSE_00959 [Fusarium pseudograminearum CS3096]
>gi|410083028|ref|XP_003959092.1| hypothetical protein KAFR_0I01760 [Kazachstania africana CBS 2517] >gi|372465682|emb|CCF59957.1| hypothetical protein KAFR_0I01760 [Kazachstania africana CBS 2517]
>gi|444315528|ref|XP_004178421.1| hypothetical protein TBLA_0B00580 [Tetrapisispora blattae CBS 6284] >gi|387511461|emb|CCH58902.1| hypothetical protein TBLA_0B00580 [Tetrapisispora blattae CBS 6284]
>gi|448090588|ref|XP_004197110.1| Piso0_004347 [Millerozyma farinosa CBS 7064] >gi|448095015|ref|XP_004198141.1| Piso0_004347 [Millerozyma farinosa CBS 7064] >gi|359378532|emb|CCE84791.1| Piso0_004347 [Millerozyma farinosa CBS 7064] >gi|359379563|emb|CCE83760.1| Piso0_004347 [Millerozyma farinosa CBS 7064]
>gi|448526196|ref|XP_003869293.1| Hsp60 heat shock protein [Candida orthopsilosis Co 90-125] >gi|380353646|emb|CCG23157.1| Hsp60 heat shock protein [Candida orthopsilosis]
>gi|46123737|ref|XP_386422.1| HS60_AJECA Heat shock protein 60, mitochondrial precursor (Antigen HIS-62) [Gibberella zeae PH-1]
>gi|50292099|ref|XP_448482.1| hypothetical protein [Candida glabrata CBS 138] >gi|49527794|emb|CAG61443.1| unnamed protein product [Candida glabrata]
>gi|50310975|ref|XP_455510.1| hypothetical protein [Kluyveromyces lactis NRRL

Y-1140] >gi|49644646|emb|CAG98218.1| KLLA0F09449p [Kluyveromyces lactis]
>gi|50422027|ref|XP_459575.1| DEHA2E05808p [Debaryomyces hansenii CBS767] >gi|49655243|emb|CAG87802.1| DEHA2E05808p [Debaryomyces hansenii CBS767]
>gi|50555023|ref|XP_504920.1| YALI0F02805p [Yarrowia lipolytica] >gi|49650790|emb|CAG77725.1| YALI0F02805p [Yarrowia lipolytica CLIB122]
>gi|6323288|ref|NP_013360.1| Hsp60p [Saccharomyces cerevisiae S288c] >gi|123579|sp|P19882.1|HSP60_YEAST RecName: Full=Heat shock protein 60, mitochondrial; AltName: Full=CPN60; AltName: Full=P66; AltName: Full=Stimulator factor I 66 kDa component; Flags: Precursor >gi|171720|gb|AAA34690.1| heat shock protein 60 (HSP60) [Saccharomyces cerevisiae] >gi|577181|gb|AAB67380.1| Hsp60p: Heat shock protein 60 [Saccharomyces cerevisiae] >gi|151941093|gb|EDN59473.1| chaperonin [Saccharomyces cerevisiae YJM789] >gi|190405319|gb|EDV08586.1| chaperonin [Saccharomyces cerevisiae RM11-1a] >gi|207342889|gb|EDZ70518.1| YLR259Cp-like protein [Saccharomyces cerevisiae AWRI1631] >gi|256271752|gb|EEU06789.1| Hsp60p [Saccharomyces cerevisiae JAY291] >gi|285813676|tpg|DAA09572.1| TPA: chaperone ATPase HSP60 [Saccharomyces cerevisiae S288c] >gi|323353818|gb|EGA85673.1| Hsp60p [Saccharomyces cerevisiae VL3] >gi|349579966|dbj|GAA25127.1| K7_Hsp60p [Saccharomyces cerevisiae Kyokai no. 7] >gi|392297765|gb|EIW08864.1| Hsp60p [Saccharomyces cerevisiae CEN.PK113-7D] >gi|226279|prf||1504305A mitochondrial assembly factor
>gi|68485963|ref|XP_713100.1| heat shock protein 60 [Candida albicans SC5314] >gi|68486010|ref|XP_713077.1| heat shock protein 60 [Candida albicans SC5314] >gi|6016258|sp|O74261.1|HSP60_CANAL RecName: Full=Heat shock protein 60, mitochondrial; AltName: Full=60 kDa chaperonin; AltName: Full=Protein Cpn60; Flags: Precursor >gi|3552009|gb|AAC34885.1| heat shock protein 60 [Candida albicans] >gi|46434552|gb|EAK93958.1| heat shock protein 60 [Candida albicans SC5314] >gi|46434577|gb|EAK93982.1| heat shock protein 60 [Candida albicans SC5314]
>gi|71001164|ref|XP_755263.1| antigenic mitochondrial protein HSP60 [Aspergillus fumigatus Af293] >gi|66852901|gb|EAL93225.1| antigenic

mitochondrial protein HSP60, putative [Aspergillus fumigatus Af293] >gi|159129345|gb|EDP54459.1| antigenic mitochondrial protein HSP60, putative [Aspergillus fumigatus A1163]
>gi|90970323|gb|ABE02805.1| heat shock protein 60 [Rhizophagus intraradices]

In an embodiment, a 10 kDa chaperone from Table 3 is combined with a matching 60kDa chaperone from table 4 of the same organism genus or species for expression in the host.

For instance: >gi|189189366|ref|XP_001931022.1|:71-168 10 kDa chaperonin [Pyrenophora tritici-repentis] expressed together with matching >gi|189190432|ref|XP_001931555.1| heat shock protein 60, mitochondrial precursor [Pyrenophora tritici-repentis Pt-1C-BFP].

All other combinations from Table 3 and 4 similarly made with same organism source are also available to the skilled person for expression.

Further, one may combine a chaperone from Table 3 from one organism with a chaperone from Table 4 from another organism, or one may combine GroES with a chaperone from Table 3, or one may combine GroEL with a chaperone from Table 4.

As follows from the above, the invention further relates to a method for preparing an organic compound comprising converting a carbon source, using a microorganism, thereby forming the organic compound. The method may be carried out under aerobic, oxygen-limited or anaerobic conditions.

The invention allows in particular a reduction in formation of an NADH dependent side-product, especially glycerol, by up to 100 %, up to 99 %, or up to 90 %, compared to said production in a corresponding reference strain. The NADH dependent side-product formation is preferably reduced by more than 10 % compared to the corresponding reference strain, in particular by at least 20 %, more in particular by at least 50 %. NADH dependent side-product production is preferably reduced by 10-100 %, in particular by 20-95 %, more in particular by 50-90 %.

In preferred method wherein Rubisco, or another enzyme capable of catalysing the formation of an organic compound from CO_2 (and another substrate) or another enzyme that catalyses the function of CO_2 as an electron acceptor, is used, the carbon dioxide concentration in the reaction medium is at least 5 % of the CO_2

saturation concentration under the reaction conditions, in particular at least 10 % of said CO_2 saturation concentration, more in particular at least 20 % of said CO_2 saturation concentration. This is in particular advantageous with respect to product yield. The reaction medium may be oversaturated in CO_2 concentration, saturated in CO_2 concentration or may have a concentration below saturation concentration. In a specific embodiment, the CO_2 concentration is 75 % of the saturation concentration or less, in particular 50 % of said saturation concentration or less, more in particular is 25 % of the CO_2 saturation concentration or less.

In a specific embodiment, the carbon dioxide or part thereof is formed *in situ* by the microorganism. If desired, the method further comprises the step of adding external CO_2 to the reaction system, usually by aeration with CO_2 or a gas mixture containing CO_2, for instance a CO_2 /nitrogen mixture. Adding external CO_2 in particular is used to (increase or) maintain the CO_2 within a desired concentration range, if no or insufficient CO_2 is formed *in situ*.

Determination of the CO_2 concentration in a fluid is within the routine skills of the person skilled in the art. In practice, one may routinely determine the CO_2 concentration in the gas phase above a culture of the yeast (practically the off-gas if the medium is purged with a gas). This can routinely be measured using a commercial gas analyser, such as a RosemountNGA200000 gas analyser (Rosemount Analytical,Orrvile,USA). The concentration in the liquid phase (relative to the saturation concentration), can then be calculated from the measured value in the gas, from the CO_2 saturation concentration and Henri coefficients of under the existing conditions in the method. These parameters are available from handbooks or can be routinely determined.

As a carbon source, in principle any carbon source that the microorganism can use as a substrate can be used. In particular an organic carbon source may be used, selected from the group of carbohydrates and lipids (including fatty acids). Suitable carbohydrates include monosaccharides, disaccharides, and hydrolysed polysaccharides (e.g. hydrolysed starches, lignocellulosic hydrolysates). Although a carboxylic acid may be present, it is not necessary to include a carboxylic acid such as acetic acid, as a carbon source.

It is in particular an advantage of the present invention that an improved ethanol yield and a reduced glycerol production is feasible compared to, e.g., a wild type yeast cell, without needing to intervene in the genome of the cell by inhibition of a glycerol 3-phosphate phosphohydrolase and/or encoding a glycerol 3-phosphate dehydrogenase gene.

Still, in a specific embodiment, a yeast cell according to the invention may comprise a deletion or disruption of one or more endogenous nucleotide sequence encoding a glycerol 3-phosphate phosphohydrolase and/or encoding a glycerol 3-phosphate dehydrogenase gene:

Herein in the cell, enzymatic activity needed for the NADH-dependent glycerol synthesis is reduced or deleted. The reduction or deleted of this enzymatic activity can be achieved by modifying one or more genes encoding a NAD-dependent glycerol 3-phosphate dehydrogenase activity (GPD) or one or more genes encoding a glycerol phosphate phosphatase activity (GPP), such that the enzyme is expressed considerably less than in the wild-type or such that the gene encoded a polypeptide with reduced activity.

Such modifications can be carried out using commonly known biotechnological techniques, and may in particular include one or more knock-out mutations or site-directed mutagenesis of promoter regions or coding regions of the structural genes encoding GPD and/or GPP. Alternatively, yeast strains that are defective in glycerol production may be obtained by random mutagenesis followed by selection of strains with reduced or absent activity of GPD and/or GPP. *S. cerevisiae GPD1*, *GPD2*, *GPP1* and *GPP2* genes are shown in WO 2011/010923, and are disclosed in SEQ ID NO: 24-27 of that application. The contents of this application are incorporated by reference, in particular the contents relating to GPD and/or GPP.

As shown in the Examples below, the invention is in particular found to be advantageous in a process for the production of an alcohol, notably ethanol. However, it is contemplated that the insight that CO_2 can be used as an electron acceptor in microorganisms that do not naturally allow this, has an industrial benefit for other biotechnological processes for the production of organic molecules, in particular organic molecules of a relatively low molecular weight, particularly organic molecules with a molecular weight below 1000 g/mol. The following items are mentioned herein as preferred embodiments of the use of carbon dioxide as an electron acceptor in accordance with the invention.

1. Use of carbon dioxide as an electron acceptor in a recombinant chemotrophic micro-organism is a non-phototrophic eukaryotic micro-organism.

2. Use of carbon dioxide as an electron acceptor in a recombinant chemotrophic micro-organism , wherein the micro-organism produces an organic compound under anaerobic conditions.

3. Use according to item 1 or 2, wherein the carbon dioxide serves as an electron acceptor in a process with NADH as an electron donor.

5. Use according to any of the preceding items, wherein the micro-organism produces an organic compound in a process with an excess production of ATP and/or NADH.

6. Use according to any of the preceding items, wherein the micro-organism comprises a heterologous nucleic acid sequence encoding a polypeptide from a (naturally) autotrophic organism.

7. Use according to item 6, wherein the micro-organism comprises a heterologous nucleic acid sequence encoding a first prokaryotic chaperone for said polypeptide and preferably a nucleic acid sequence encoding a second prokaryotic chaperone - different from the first - for said polypeptide.

8. Use according to item 7, wherein the chaperones are GroEL and GroES.

9. Use according to any of the preceding items, wherein the micro-organism produces an organic compound selected from the group consisting of alcohols (such as methanol, ethanol, propanol, butanol, phenol, polyphenol), ribosomal peptides, antibiotics (such as penicillin), bio-diesel, alkynes, alkenes, isoprenoids, esters, carboxylic acids (such as succinic acid, citric acid, adipic acid, lactic acid), amino acids, polyketides, lipids, and carbohydrates.

10. Use according to any of the preceding items, wherein the microorganism comprises a heterologous nucleic acid sequence functionally expressing a polypeptide selected from the group consisting of carbonic anhydrases, carboxylases, oxygenases, hydrogenases, dehydrogenases, isomerases, aldolases, transketolases, transaldolases, phosphatases, epimerases, kinases, carboxykinases, oxidoreductases, aconitases, fumarases, reductases, lactonases, phosphoenolpyruvate (PEP) carboxylases, phosphoglycerate kinases, glyceraldehyde 3-phosphate dehydrogenases, triose phosphate isomerases, fructose-1,6-bisphosphatases, sedoheptulose-1,7-

bisphosphatases, phosphopentose isomerases, phosphopentose epimerase, phosphoribulokinases (PRK), glucose 6-phosphate dehydrogenases, 6-phosphogluconolactonases, 6-phosphogluconate dehydrogenases, ribulose 5-phosphate isomerases, ribulose 5-phosphate 3-epimerases, Ribulose-1,5-bisphosphate carboxylase oxygenases, lactate dehydrogenases, malate synthases, isocitrate lyases, pyruvate carboxylases, phosphoenolpyruvate carboxykinases, fructose-1,6-bisphosphatases, phosphoglucoisomerases, glucose-6-phosphatases, hexokinases, glucokinases, phosphofructokinases, pyruvate kinases, succinate dehydrogenases, citrate synthases, isocitrate dehydrogenases, α-ketoglutarate dehydrogenases, succinyl-CoA synthetases, malate dehydrogenases, nucleoside-diphosphate kinases, xylose reductases, xylitol dehydrogenases, xylose isomerases, isoprenoid synthases, and xylonate dehydratases.

11. Use according to item 10, wherein the microorganism comprises a heterologous nucleic acid sequence functionally expressing Ribulose-1,5-bisphosphate carboxylase oxygenase (Rubisco) and/or a heterologous nucleic acid sequence functionally expressing a phosphoribulokinase (PRK).

12. Use according to any of the preceding items, wherein the microorganism is selected from the group of is selected from the group consisting of *Saccharomyceraceae, Penicillium, Yarrowia* and *Aspergillus.*

13. Use according to any of the preceding items, wherein the carbon dioxide is used as an electron acceptor to reduce production of an NAD+-dependent side-product or NADH-dependent side-product, such as glycerol, in a process for preparing another organic compound, such as another alcohol or a carboxylic acid.

14. Recombinant micro-organism, in particular a eukaryotic micro-organism, having an enzymatic system allowing the micro-organism to use carbon dioxide as an electron acceptor under chemotrophic (non-phototrophic) conditions., wherein the microorganism is preferably as defined in the prevision items.

15. Recombinant micro-organism according to item 14, wherein the micro-organism has an enzymatic system for producing an organic compound in a process with an excess production of ATP and/or NADH.

The production of the organic compound of interest may take place in a organism known for it usefulness in the production of the organic compound of interest, with the proviso that the organism has been genetically modified to enable the use of carbon dioxide as an electron acceptor in the organism.

Although it is contemplated that the invention is interesting for the production of a variety of industrially relevant organic compounds, a method or use according the invention is in particular considered advantageous for the production of an alcohol, in particular an alcohol selected from the group of ethanol, n-butanol and 2,3-butanediol; or in the production of an organic acid/carboxylate, in particular a carboxylate selected from the group of L-lactate, 3-hydroxypropionate, D-malate, L-malate, succinate, citrate, pyruvate and itaconate.

Regarding the production of ethanol, details are found herein above, when describing the yeast cell comprising PRK and Rubisco and in the examples. The ethanol or another alcohol is preferably produced in a fermentative process.

For the production of several organic acids (carboxylates), *e.g.* citric acid, an aerobic process is useful. For citric acid production for instance *Aspergillus niger, Yarrowia lipolytica*, or another known citrate producing organism may be used.

An example of an organic acid that is preferably produced anaerobically is lactic acid. Various lactic acid producing bacterial strains and yeast strains that have been engineered for lactate production are generally known in the art.

EXAMPLES

Example 1. Construction of the expression vector

Phosphoribulokinase (PRK) cDNA from *Spinacia oleracea* (spinach) (EMBL accession number: X07654.1) was PCR-amplified using Phusion Hot-start polymerase (Finnzymes, Landsmeer, the Netherlands) and the oligonucleotides XbaI_prk-FW2 and RV1_XhoI_prk (Table 5), and was ligated in pCR®-Blunt II-TOPO® (Life Technologies Europe BV, Bleiswijk, the Netherlands).

Table 5 Oligonucleotides

Number	Name	Sequence (5' to 3')	Purpose
	Cloning		
1	XbaI_prk_FW2	TGACATCTAGATGTCACAACAACAAACAATTG	cloning of PRK into pUDE046.
2	RV1 XhoI prk	TGACATCTAGATGTCACAACAACAAACAATTG	cloning of PRK into pUDE046.
	Primers used for in vivo plasmid assembly		
3	HR-cbbM-FW-65	TTGTAAAACGACGGCCAGTGAGCGCGCGTAATACGACTCACTATAGGGCGAATTGGGTACAGCTGGAGCTCAGTTTATCATTATC	Rubisco *cbbM* cassette for plasmids pUDC075, pUDC099, and pUDC100.
4	HR-cbbM-RV-65	GGAATCTGTGTAGTATGCCTGGAATGTCTGCCGTGCCATAGCCATGTATGCTGATATGTCGGTACCGGCCGCAAATTAAAG	Rubisco *cbbM* cassette for plasmids pUDC075, pUDC099, and pUDC100
5	linker-cbbO2-pRS416	ATCACTCTTACCAGGCTAGGACGACCCTACTCATGTATTGAGATCGACGAGATTTCTAGGCCAGCTTTTGTTCCCTTTAGTGAGGGTTAATTGCGCGCTTGGCGTAATCATGGTCATAGC	Linker fragment for assembly of plasmid pUDC099.
6	linker-cbbM-GroEL	GACATATCAGCATACATGGCTATGGCACGGCAGACATTCCAGGCATACTACACAGATTCCATCACTCTTACCAGGCTAGGACGACCCTACTCATGTATTGAGATCGACGAGATTTCTAGG	Linker fragment for assembly of plasmid pUDC100.

	Primers used for in vivo integration assembly		
7	FW pTDH3- HR-CAN1up	GTTGGATCCAGTTTTTAATCTGTCGTCAATCGAAAGTTTATTTCAGAGTTCTTCAGACTTCTTAACTCCTGTAAAAACAAAAAAAAAAAAAGGCATAGCAAGCTGGAGCTCAGTTTATC	1st cloning expression cassette linker fragment between *CAN1* upstream and PRK expression cassette (IMI229), and CAN1up-linker and *KlLEU2* expression cassette (IMI232).
8	RV linker-iHR2B	AGATATACTGCAAAGTCCGGAGCAACAGTCGTATAACTCGAGCAGCCCTCTACTTTGTTGTTGCGCTAAGAGAATGGACC	1st cloning fragment: linker fragment between CAN1up-linker and PRK expression cassette (IMI229).
9	RV linker-iHR6	GCTATGACCATGATTACGCCAAGCGCGCAATTAACCCTCACTAAAGGGAACAAAAGCTGGTTGCGCTAAGAGAATGGACC	1st cloning fragment: linker fragment between CAN1up-linker and *KlLEU2* expression cassette (IMI232).
10	FW pGAL1-prk HR2B	CAACAAAGTAGAGGGCTGCTCGAGTTATACGACTGTTGCTCCGGACTTTGCAGTATATCTGCTGGAGCTCTAGTACGGATT	2nd cloning fragment: *GAL1*$_p$-*PRK*-*CYC1*$_t$ expression cassette (IMI229) from pUDE046.
11	RV CYC1t-prk HR2	GGAATCTGTGTAGTATGCCTGGAATGTCTGCCGTGCCATAGCCATGTATGCTGATATGTCGTACCGGCCGCAAATTAAAG	2nd cloning fragment: *GAL1*p-*PRK*-*CYC1*$_t$ expression cassette (IMI229) from pUDE046.
12	FW HR2-cbbQ2-HR3	GACATATCAGCATACATGGCTATGG	3rdI cloning fragment: *PGI1*$_p$-*cbbQ2*-*TEF2*t cassette (IMI229).

13	RV HR2-cbbQ2-HR3	GGACACGCTTGACAGAATGTCAAAGG	3rd cloning fragment: *PGI1*$_p$-*cbbQ2*-*TEF2*t cassette (IMI229).
14	FW HR3-cbbO2-HR4	CGTCCGATATGATCTGATTGG	4th TARI cloning fragment: *PGK1*$_p$-*cbbO2*-*ADH1*$_t$ cassette (IMI229).
15	RV HR3-cbbO2-HR4	CCTAGAAATCTCGTCGATCTC	4th cloning fragment: *PGK1*$_p$-*cbbO2*-*ADH1*$_t$ cassette (IMI229).
16	FW HR4-GroEL-HR5	ATCACTCTTACCAGGCTAGG	5th cloning fragment:*TEF1*$_p$-*groEL*-*ACT1*$_t$ cassette (IMI229).
17	RV HR4-GroEL-HR5	CTGGACCTTAATCGTGTGCGCATCCTC	5th cloning fragment: *TEF1*$_p$-*groEL*-*ACT1*$_t$ cassette (IMI229).
18	FW HR5-GroES-HR6	CCGTATAGCTTAATAGCCAGCTTTATC	6th cloning fragment: *TPI1*$_p$-*groES*-*PGI1*$_t$ cassette (IMI229).
19	RV HR5-GroES-HR6	GCTATGACCATGATTACGCCAAGC	6th cloning fragment: *TPI1*$_p$-*groES*-*PGI1*$_t$ cassette (IMI229).
20	FW HR6-LEU2-CAN1dwn	CCAGCTTTTGTTCCCTTTAGTGAGGGTTAATTGCGCGCTTGGCGTAATCATGGTCATAGCCTGTGAAGATCCCAGCAAAG	7th (IMI229) or 2nd (IMI232) cloning fragment: *KlLEU2* cassette from pUG73.
21	RV LEU2 HR-CAN1	AGCTCATTGATCCCTTAAACTTTCTTTTCGGTGTATGACTTATGAGGGTGAGAATGCGAAATGGCGTGGAAATGTGATCAAAGGTAATAAAACGTCATATATCCGCAGGCTAACCGGAAC	7th (IMI229) or 2nd (IMI232) cloning fragment: *KlLEU2* cassette from pUG73.

	Primers used for verification of the in vivo assembled constructs		
22	m-PCR-HR1-FW	GGCGATTAAGTTGGGTAACG	Diagnostic for assembly of plasmids pUDC075, pUDC099, and pUDC100,.
23	m-PCR-HR1-RV	AACTGAGCTCCAGCTGTACC	Diagnostic for assembly of plasmids pUDC075, pUDC099, pUDC100, and integration in strain IMI229.
24	m-PCR-HR2-FW	ACGCGTGTACGCATGTAAC	Diagnostic for assembly of pUDC075, pUDC099, pUDC100, and integration in strain IMI229
25	m-PCR-HR2-RV	GCGCGTGGCTTCCTATAATC	Diagnostic for assembly of pUDC075, pUDC099, pUDC100, and integration in strain IMI229
26	m-PCR-HR3-FW	GTGAATGCTGGTCGCTATAC	Diagnostic for assembly of pUDC075, pUDC099, pUDC100, and integration in strain IMI229.
27	m-PCR-HR3-RV	GTAAGCAGCAACACCTTCAG	Diagnostic for assembly of pUDC075, pUDC099, pUDC100, and integration in strain IMI229.
28	m-PCR-HR4-FW	ACCTGACCTACAGGAAAGAG	Diagnostic for assembly of pUDC075, pUDC099, pUDC100, and integration in

			strain IMI229.
29	m-PCR-HR4-RV	TGAAGTGGTACGGCGATGC	Diagnostic for assembly of pUDC075, pUDC099, pUDC100, and integration in strain IMI229.
30	m-PCR-HR5-FW	ATAGCCACCCAAGGCATTTC	Diagnostic for assembly of pUDC075, pUDC099, pUDC100, and integration in strain IMI229.
31	m-PCR-HR5-RV	CCGCACTTTCTCCATGAGG	Diagnostic for assembly of pUDC075, pUDC099, pUDC100, and integration in strain IMI229.
32	m-PCR-HR6-FW	CGACGGTTACGGTGTTAAG	Diagnostic for assembly of pUDC075, pUDC099, pUDC100, and integration in strain IMI229.
33	m-PCR-HR6-RV	CTTCCGGCTCCTATGTTGTG	Diagnostic for assembly of pUDC075, pUDC099, pUDC100, and integration in strain IMI229.

After restriction by *Xba*I and *Xho*I, the PRK-containing fragment was ligated into pTEF424. The *TEF1*p was later replaced by *GAL1*p from plasmid pSH47 by *Xba*I and *Sac*I restriction/ligation, creating plasmid pUDE046 (see Table 6).

Table 6: Plasmids

Name	Relevant genotype	Source/reference
pFL451	pAOX1-prk (Spinach)-AOX1t (pHIL2-D2 HIS4 Amp centromeric)	Brandes *et al.* 1996.[14]
pCR®-Blunt II-TOPO	*bla*	Life Technologies Europe BV
pTEF424_TEF	*TRP1* 2μ *bla*	Mumberg *et al.*1995[25].
pSH47	*URA3 CEN6 ARS4* $GAL1_p$-*cre*-$CYC1_t$ *bla*	Güldener et al 1996[26]
pUD0E46	*TRP1* 2μ *GAL1*p-prk-$CYC1_t$ *bla*	This study.
pPCR-Script	*bla*	Life Technologies Europe BV
pGPD_426	*URA3* 2μ *bla*	Mumberg *et al.*1995[25].
pRS416	*URA3 CEN6 ARS4 bla*	Mumberg *et al.*1995[25].
pBTWW002	*URA3* 2μ $TDH3_p$-*cbbM*-$CYC1_t$ *bla*	This study.
pUDC098	*URA3 CEN6 ARS4* $TDH3_p$-*cbbM*-$CYC1_t$ *bla*	This study.
pMK-RQ	*nptII*	Life Technologies Europe BV
pUD230	$PGI1_p$-*cbbQ2*-$TEF2_t$ *nptII*	Life Technologies Europe BV
pUD231	$PGK1_p$-*cbbO2*-$ADH1_t$ *nptII*	Life Technologies

		Europe BV
pUD232	*TEF1*$_{p}$-*groEL-ACT1*$_{t}$ *nptII*	Life Technologies Europe BV
pUD233	*TPI1*$_{p}$-*groES-PGI1*$_{t}$ *nptII*	Life Technologies Europe BV
pUDC075	*URA3 CEN6 ARS4 TDH3*$_{p}$-*cbbM-CYC1*$_{t;}$*PGI1*$_{p}$-*cbbQ2-TEF2*$_{t}$;*PGK1*$_{p}$-*cbbO2-ADH1*$_{t}$;*TEF1*$_{p}$-*groEL-ACT1*$_{t}$;*TPI1*$_{p}$-*groES-PGI1*$_{t}$ *bla*	This study.
pUDC099	*URA3 CEN6 ARS4 TDH3*$_{p}$-*cbbM-CYC1*$_{t;}$*PGI1*$_{p}$-*cbbQ2-TEF2*$_{t}$;*PGK1*$_{p}$-*cbbO2-ADH1*$_{t}$ *bla*	This study.
pUDC100	*URA3 CEN6 ARS4 TDH3*$_{p}$-*cbbM-CYC1*$_{t;}$ *TEF1*$_{p}$-*groEL-ACT1*$_{t}$;*TPI1*$_{p}$-*groES-PGI1*$_{t}$ *bla*	This study.

Rubisco form II gene *cbbM* from *Thiobacillus denitrificans* (*T. denitrificans*) flanked by *Kpn*I and *Sac*I sites was codon optimized synthesized at GeneArt (Life Technologies Europe BV), and ligated into pPCR-Script., the plasmid was then digested by *Bam*HI and *Sac*I. The *cbbM*-containing fragment was ligated into the *Bam*HI and *Sac*I restricted vector pGPD_426 creating plasmid pBTWW002.The *cbbM* expression cassette was transferred into pRS416 using *Kpn*I and *Sac*I, yielding pUDC098.

Expression cassette of the specific Rubisco form II cheparones from *T. denitrificans cbbQ2* and *cbbO2*, and chaperones *gro*EL and *gro*ES from *E. coli.* were condon optimized. The expression cassettes contained a yeast constitutive promoters and terminator, flanking the codon optimized gene. The cassette was flanked by unique 60 bp regions obtained by randomly combining bar-code sequences used in the *Saccharomyces* Genome Deletion Project and an *Eco*RV site (GeneArt). The expression cassettes were inserted in plasmid pMK-RQ (GeneArt) using the *Sfi*I cloning sites yielding pUB230 (PGI1p-cbbQ2-TEF2t), pUD231 (PGK1p-cbbO2-ADH1t), pUD232(TEF1p-groEL-ACT1t), and pUDE233 (TPI1p-groES-PGI1t) Table 6). The expression cassette TDH3p-cbbM-CYC1t was PCR-amplified from plasmid pBTWW002 using Phusion Hot-Start Polymerase

(Finnzymes) and primers HR-cbbM-FW-65 and HR-cbbM-RV-65 in order to incorporate the 60-bp region for recombination cloning.

Example 2. Strain construction, isolation and maintenance

All *Saccharomyces cerevisiae* strains used (Table 7)belong to the CEN.PK family. All strains were grown in 2% w/v glucose synthetic media supplemented with 150 mg L^{-1} uracil when required until they reached end exponential phase, then sterile glycerol was added up to ca. 30% v/v and aliquots of 1 ml were stored at -80° C.

Table 7: *Saccharomyces cerevisiae* strains

Strain	Relevant genotype	Source/reference
CEN.PK113-5D	*MATa ura3-52*	Euroscarf
CEN.PK102-3A	*MATa ura3-52 leu2-3, 112*	Euroscarf
IMC014	*MATa ura3-52* pUDC075 (*CEN6 ARS4 URA3 TDH3*$_p$-*cbbM-CYC1*$_t$ *PGI1*$_p$-*cbbQ2-TEF2*$_t$ *PGK1*$_p$-*cbbO2-ADH1*$_t$ *TEF1*$_p$-*groEL-ACT1*$_t$ *TPI1*$_p$-*groES-PGI1*$_t$)	This study.
IMC033	*MATa ura3-52* pUDC098 (*CEN6 ARS4 URA3 TDH3*$_p$-*cbbM-CYC1*$_t$)	This study.
IMC034	*MATa ura3-52* pUDC099 (*CEN6 ARS4 URA3 TDH3*$_p$-*cbbM-CYC1*$_t$ *PGI1*$_p$-*cbbQ2-TEF2*$_t$ *PGK1*$_p$-*cbbO2-ADH1*$_t$ *cbbO2*-pRS416 linker)	This study.
IMC035	*MATa ura3-52* pUDC100 (*CEN6 ARS4 URA3 TEF1*$_p$-*groEL-ACT1*$_t$ *TPI1*$_p$-*groES-PGI1*$_t$ *cbbM-GroEL* linker)	This study.
IMI229	*MATa ura3-52 leu2-3, 112 can1Δ::GAL1*$_p$-*prk-CYC1*$_t$ *PGI1*$_p$-*cbbQ2-TEF2*$_t$,*PGK1*$_p$-*cbbO2-ADH1*$_t$,*TEF1*$_p$-*groEL-ACT1*$_t$,*TPI1*$_p$-*groES-PGI1*$_t$ *KlLEU2*	This study.
IMI232	*MATa ura3-52 leu2-3, 112 can1Δ::KlLEU2*	This study.
IMU032	IMI232 p426_GPD (2μ *URA3*)	This study.
IMU033	IMI229 pUDC100 (*CEN6 ARS4 URA3 TEF1*$_p$-*groEL-ACT1*$_t$ *TPI1*$_p$-*groES-PGI1*$_t$ *cbbM-GroEL* linker)	This study.

The strain IMC014 that co-expressed the Rubisco form II *ccbM* and the four chaperones *cbbQ2, cbbO2, groEL,* and *groES* was constructed using *in vivo* transformation associated recombination. 200 fmol of each expression cassette were pooled with 100 fmol of the *Kpn*I/*Sac*I linearized pRS416 backbone in a final volume of 50µl and transformed in CEN.PK 113-5D using the lithium acetate protocol (Gietz, et al., Yeast Transformation by the LiAc/SS Carrier DNA/PEG Method in Yeast Protocol, *Humana press*, 2006). Cells were selected on synthetic medium. Correct assembly of the fragment of pUDC075 was performed by multiplex PCR on transformant colonies using primers enabling amplification over the regions used for homologous recombination (Table 5) and by restriction analysis after re-transformation of the isolated plasmid in *E. coli* DH5α. PUDC075 was sequenced by Next-Generation Sequencing (Illumina, San Diego, California, U.S.A.) (100br reads paired-end, 50Mb) and assembled with Velvet (Zerbino, et al., Velvet: Algorithms for De Novo Short Read Assembly Using De Bruijn Graphs, *Genome Research*, 2008). The assembled sequence did not contain mutations in any of the assembled expression cassettes. The strains IMC034 and IMC035 that expressed *ccbM/ccbQ2/ccbO2* and *ccbM/groEL/groES* respectively were constructed using the same *in vivo* assembly method with the following modification. To construct plasmids pUDC099 and pUDC100, 120 bp cbbO2-pRS416 linker and cbbM-GroEL linker were used to close the assembly respectively (Table 5), 100 fmol of each of complementary 120 bp oligonucleotides were added to the transformation. The strain IMC033 that only expressed the cbbM gene was constructed by transforming CEN.PK113-5D with pUDC098.

To construct the strain IMU033 that co-expressed PRK, *ccbM, ccbQ2, ccbO2, GroEL, GroES,* the intermediate strain IMI229 was constructed by integrating PRK, the four chaperones and *KlLEU2* (Güldener, et al., A second set of *loxP* marker cassettes for Cre-mediated multiple gene knockouts in budding yeast, *Nucleic Acids Research*, 2002) at the CAN1 locus by *in vivo* homologous integration in CEN.PK102-3A. The expression cassettes were PCR amplified using Phusion Hot-Start Polymerase (Finnzymes, Thermo Fisher Scientific Inc. Massachusetts, U.S.A.), the corresponding oligonucleotides and DNA templates (Table 5). Finally, the strain IMI229 was transformed with pUDC100 that carries the Rubisco form II *ccbM* and the two *E. coli* chaperones *groEL* and *groES*.

Strain IMI232 was constructed by transforming CEN.PK102-3A with the *KlLEU2* cassette. IMI232 was finally transformed with the plasmid p426GPD to restore prototrophy resulting in the reference strain IMU032.

Example 3. Experimental set-up of chemostat and batch experiments

Anaerobic chemostat cultivation was performed essentially as described (Basso, et al., Engineering topology and kinetics of sucrose metabolism in *Saccharomyces cerevisiae* for improved ethanol yield, *Metabolic Engineering* 13:694-703, 2011), but with 12.5 g l-1 glucose and 12.5 g l-1 galactose as the carbon source and where indicated, a mixture of 10% CO_2/90% N_2 replaced pure nitrogen as the sparging gas. Residual glucose and galactose concentrations were determined after rapid quenching (Mashego, et al., Critical evaluation of sampling techniques for residual glucose determination in carbon-limited chemostat culture of *Saccharomyces cerevisiae, Biotechnology and Bioengineering* 83:395-399, 2003) using commercial enzymatic assays for glucose (Boehringer, Mannheim, Germany) and D-galactose (Megazyme, Bray, Ireland). Anaerobic bioreactor batch cultures were grown essentially as described (Guadalupe Medina, et al., Elimination of glycerol production in anaerobic cultures of a *Saccharomyces cerevisiae* strain engineered to use acetic acid as an electron acceptor. *Applied and Environmental Microbiology* 76:190-195, 2010), but with 20 g L^{-1} galactose and a sparging gas consisting of 10% CO_2 and 90% N_2. Biomass and metabolite concentrations in batch and chemostat and batch cultures were determined as described by Guadalupe et al. (Guadalupe Medina, et al., Elimination of glycerol production in anaerobic cultures of a Saccharomyces cerevisiae strain engineered to use acetic acid as an electron acceptor. Appl. Environ. Microbiol. 76, 190-195, 2010). In calculations of ethanol fluxes and yields, ethanol evaporation was corrected for based on a first-order evaporation rate constant of 0.008 h^{-1} in the bioreactor set-ups and under the conditions used in this study.

Example 4. Enzyme assays for phosphoribulokinase (PRK) and Rubisco

Cell extracts for analysis of phosphoribulokinase (PRK) activity were prepared as described previously (Abbott, et al., Catalase Overexpression reduces lactic acid-induced oxidative stress in *Saccharomyces cerevisiae, Applied and Environmental Microbiology* 75:2320-2325, 2009). PRK activity was measured at 30°C by a coupled spectrophotometric assay (MacElroy, et al., Properties of Phosphoribulokinase from *Thiobacillus neapolitanus*, *Journal of Bacteriology* 112:532-538, 1972). Reaction rates were proportional to the amounts of cell extract added. Protein concentrations were determined by the Lowry method (Lowry, et al., Protein measurement with the Folin phenol reagent, *The Journal of Biological Chemistry* 193:265-275, 1951) using bovine serum albumin as a standard.

Cell extracts for Rubisco activity assays were prepared as described in Abbott, D. A. et al. Catalase overexpression reduces lactic acid-induced oxidative stress in Saccharomyces cerevisiae. Appl. Environ. Microbiol. 75:2320-2325, 2009, with two modifications: Tris-HCl (1 mM, pH 8.2) containing 20 mM $MgCl_2 \cdot 6H_2O$, 5 mM of DTT 5 mM $NaHCO_3$ was used as sonication buffer and Tris-HCl (100 mM, pH 8.2), 20 mM $MgCl_2 \cdot 6H_2O$ and 5 mM of DTT as freezing buffer. Rubisco activity was determined by measuring $^{14}CO_2$-fixation (PerkinElmer, Groningen, The Netherlands) as described (Beudeker, et al., Relations between d-ribulose-1,5-biphosphate carboxylase, carboxysomes and CO_2 fixing capacity in the obligate chemolithotroph *Thiobacillus neapolitanus* grown under different limitations in the chemostat, *Archives of Microbiology* 124:185-189, 1980) and measuring radioactive counts in a TRI-CARB® 2700TR Series liquid scintillation counter (PerkinElmer, Groningen, The Netherlands), using Ultima Gold™ scintillation cocktail (PerkinElmer, Groningen, The Netherlands). Protein concentrations were determined by the Lowry method (Lowry, O. H., Rosebrough, N. J., Farr, A. L., & Randall, R. J. Protein measurement with the Folin phenol reagent. J. Biol. Chem. 193:265-275, 1951) using standard solutions of bovine serum albumin dissolved in 50 mM Tris-HCl (pH 8.2).

Example 5. The activity of Rubisco and the activity of PRK in cell extracts,

In order to study a possible requirement of heterologous chaperones of Rubisco in *S. cerevisiae*, the form-II Rubisco-encoding *cbbM* gene from *T. denitrificans* was codon-optimised and expressed from a centromeric vector, both alone and in combination with expression cassettes for the codon-optimised *E. coli groEL/groES* and/or *T. denitrificans cbbO2/cbbQ2* genes. Analysis of ribulose-1,5-biphosphate-dependent CO_2 fixation by yeast cell extracts demonstrated that functional expression of *T. denitrificans* Rubisco in *S. cerevisiae* was observed upon co-expression of *E. coli* GroEL/GroES. Rubisco activity increased from <0.2 $nmol.min^{-1}.(mg\ protein)^{-1}$ to more than 6 $nmol.min^{-1}.(mg\ protein)^{-1}$. Results of these experiments are visualised in Figure 1, showing specific ribulose-1,5-bisphosphate carboxylase (Rubisco) activity in cell extracts of *S. cerevisiae* expressing Rubisco form II CbbM from T. denitrificans either alone (IMC033) or in combination with the E. coli chaperones GroEL/GroES (IMC035), The *T. denitrificans* chaperones CbbO2/CbbQ2 [20] (IMC034) or all four chaperones (IMC014). Heterologously expressed genes were codon optimised for expression in yeast and expressed from a single centromeric vector. Biomass samples were taken from anaerobic batch cultures on synthetic media (pH 5.0, 30°C), sparged with nitrogen and containing 20 g l-1 glucose as carbon source. Rubisco activities, measured as 14CO2-fixation in cell extracts, in a wild-type reference strain and in S. cerevisiae strains expressing cbbM and cbbM-cbbQ2-cbbO2 were below the detection limit of the enzyme assay (0.2 nmol CO2 min-1 mg protein-1

Co-expression of CbbO2/cbbQ2 did not result in a significant further increase of Rubisco activity. The positive effect of GroEL/GroES on Rubisco expression in *S. cerevisiae* demonstrates the potential value of this approach for metabolic engineering, especially when prokaryotic enzymes need to be functionally expressed in the cytosol of eukaryotes.

The *Spinach oleracea* PRK gene was integrated together with *E. coli groEL/groES* and *T. denitrificans cbbO2/cbbQ2* into the *S. cerevisiae* genome at the *CAN1* locus, under control of the galactose-inducible *GAL1* promoter. This induced in high PRK activities in cell extracts of *S. cerevisiae* strain IMU033, which additionally carried the centromeric expression cassette for *T. denitrificans* Rubisco . This engineered yeast strain was used to quantitatively analyze the physiological impacts of the expression of Rubisco and PRK.

Table 8

	IMU032 (reference strain)		**IMU033 (expressing PRK and Rubisco)**	
CO_2 in inlet gas (%)	0	10	0	10
CO_2 in outlet gas (%)	0.89 ± 0.03	10.8 ± 0.0	1.02 ± 0.00	10.8 ± 0.1
Phosphoribulokinase (μmol mg protein^{-1} min^{-1})	0.58 ± 0.09	0.51 ± 0.12	14.4 ± 1.5	15.2 ± 1.0
Rubisco (nmol mg protein^{-1} min^{-1})	< 0.2*	< 0.2	4.59 ± 0.30	2.67 ± 0.28
Biomass yield on sugar (g g^{-1})	0.083 ± 0.000[a]	0.084 ± 0.000[b]	0.093 ± 0.001[a]	0.095 ± 0.000[b]
Ethanol yield on sugar (mol mol^{-1})	1.56 ± 0.03[c]	1.56 ± 0.02[d]	1.73 ± 0.02[c]	1.73 ± 0.01[d]
Glycerol yield on sugar (mol mol^{-1})	0.14 ± 0.00[e]	0.12 ± 0.00[f]	0.04 ± 0.00[e, g]	0.01 ± 0.00[f, g]

Table 8 show increased ethanol yields on sugar of an *S. cerevisiae* strain expressing phosphoribulokinase (PRK) and Rubisco. Physiological analysis of S. cerevisiae IMU033 expressing PRK and Rubisco and the isogenic reference strain IMU032 in anaerobic chemostat cultures, grown at a dilution rate of 0.05 h-1 on a synthetic medium (pH 5) supplemented with 12.5 g l-1 glucose and 12.5 g l-1 galactose as carbon sources. To assess the impact of CO2 concentration, chemostat cultures were run sparged either with pure nitrogen gas or with a blend of 10% CO2 and 90% nitrogen. Results are represented as average ± mean deviations of data from independent duplicate chemostat experiments. Data pairs labelled with the same subscripts (a,a, b,b, etc.) are considered statistically different in a standard t-test ($p < 0.02$).

Expression of Rubisco and the four chaperones without co-expression of PRK (strain IMC014) did not result in decreased glycerol yield (0.13 mol mol^{-1}) compared to the reference strain IMU032 (0.12 mol mol^{-1}) in carbon-limited chemostat cultures supplemented with CO_2, indicating that expression of a phosphoribulokinase (PRK) gene is required for the functional pathway in *S. cerevisiae* to decrease glycerol production. The physiological impact of expression of PRK and Rubisco on growth, substrate consumption and product formation in galactose-grown anaerobic batch cultures of *S. cerevisiae* was also investigated and compared with an isogenic reference strain. Growth conditions: T = 30 °C, pH 5.0, 10% CO_2 in inlet gas. Two independent replicate experiments were carried out, whose growth kinetic parameters differed by less than 5%. Ethanol yield on galactose was 8 % higher and glycerol production was reduced by 60 % in the yeast cell in which PRK and Rubisco were functionally expressed, compared to the yeast cell lacking these enzymes. The differences were statistically significant (standard t-test (p value < 0.02). The activities of phosphoribulokinase and of Rubisco in cell extracts of the engineered strain IMU033 (table 7) enable the use of CO_2 as an electron acceptor. The ethanol yields and glycerol yields of strain IMU033 relative to the reference strain IMU032 (table 8) show that this is possible in an anaerobic fermentation with increased ethanol production.

SEQUENCES

SEQUENCE ID NO 1:

Rubisco *cbbM* gene (synthetic; based on *cbbM* gene from *Thiobacillus denitrificans* – pBTWW002, codon optimized Source: Hernandez *et al* 1996, GenBank ID: L37437.2)

ATGGATCAATCTGCAAGATATGCTGACTTGTCTTTAAAGGAAGAAGATTTGAT
TAAAGGTGGTAGACATATTTTGGTTGCTTACAAAATGAAACCAAAATCTGGTT
ATGGTTATTTGGAAGCTGCTGCTCATTTTGCTGCTGAATCTTCTACAGGTACAA
ATGTTGAAGTTTCTACTACAGATGATTTTACAAAAGGTGTTGATGCTTTAGTTT
ACTACATCGATGAAGCTTCAGAAGATATGAGAATTGCTTATCCATTGGAATTAT
TCGACAGAAATGTTACTGACGGAAGATTCATGTTAGTTTCTTTTTTGACTTTGG
CTATTGGTAACAATCAAGGAATGGGAGATATAGAACATGCAAAAATGATAGAT
TTTTACGTTCCAGAAAGATGTATTCAAATGTTTGATGGTCCAGCTACAGATATT
TCTAATTTGTGGAGAATTTTGGGTAGACCAGTAGTTAATGGTGGTTATATTGCT
GGTACTATTATTAAGCCAAAATTGGGTTTAAGACCAGAACCATTTGCTAAAGC
TGCTTATCAATTTTGGTTGGGTGGAGATTTTATCAAGAATGACGAACCACAAG
GTAATCAAGTTTTTTGTCCATTGAAAAAAGTTTTGCCATTGGTTTACGATGCTA
TGAAAAGAGCACAAGATGATACTGGTCAAGCAAAATTGTTTTCTATGAATATT
ACTGCAGACGATCATTATGAAATGTGTGCAAGAGCTGATTATGCTTTGGAAGT
TTTCGGTCCAGATGCAGATAAATTGGCTTTTTTGGTAGATGGTTACGTTGGAG
GTCCAGGAATGGTTACTACTGCTAGAAGGCAATATCCTGGTCAATATTTGCAT
TATCATAGAGCAGGTCACGGTGCTGTTACTTCTCCATCTGCTAAAAGAGGTTA
TACTGCTTTTGTTTTGGCTAAAATGTCTAGATTGCAAGGCGCTTCAGGTATTCA
TGTTGGTACTATGGGTTATGGAAAAATGGAAGGAGAAGGCGACGATAAGATTA
TTGCTTATATGATAGAAAGGGACGAATGTCAAGGTCCAGTTTATTTTCAAAAAT
GGTACGGTATGAAACCAACTACTCCAATTATCTCCGGAGGAATGAATGCTTTG
AGATTGCCTGGTTTTTTCGAAAATTTGGGTCATGGTAACGTTATTAATACTGCA
GGTGGTGGTTCTTACGGTCATATTGATTCTCCTGCTGCTGGTGCTATTTCTTTG
AGACAATCTTACGAATGTTGGAAACAAGGTGCAGATCCAATTGAATTTGCTAA
GGAACATAAGGAATTTGCAAGAGCTTTTGAATCTTTTCCAAAAGATGCTGATA
AGTTATTTCCAGGATGGAGAGAAAAATTGGGAGTTCATTCTTAA

SEQUENCE ID NO 2:

Translated protein sequence of *cbbM* gene from *Thiobacillus denitrificans*

MDQSARYADLSLKEEDLIKGGRHILVAYKMKPKSGYGYLEAAAHFAAESSTGT
NVEVSTTDDFTKGVDALVYYIDEASEDMRIAYPLELFDRNVTDGRFMLVSFLTL
AIGNNQGMGDIEHAKMIDFYVPERCIQMFDGPATDISNLWRILGRPVVNGGYIA
GTIIKPKLGLRPEPFAKAAYQFWLGGDFIKNDEPQGNQVFCPLKKVLPLVYDA
MKRAQDDTGQAKLFSMNITADDHYEMCARADYALEVFGPDADKLAFLVDGYV
GGPGMVTTARRQYPGQYLHYHRAGHGAVTSPSAKRGYTAFVLAKMSRLQGAS
GIHVGTMGYGKMEGEGDDKIIAYMIERDECQGPVYFQKWYGMKPTTPIISGGM
NALRLPGFFENLGHGNVINTAGGGSYGHIDSPAAGAISLRQSYECWKQGADPIE
FAKEHKEFARAFESFPKDADKLFPGWREKLGVHS

SEQUENCE ID NO 3:

prk gene from *Spinacea oleracea* – pBTWW001, plasmid constructed using restriction and ligation. Source: Milanez and Mural 1988, GenBank ID: M21338.1

ATGTCACAACAACAAACAATTGTGATTGGTTTAGCAGCAGATTCAGGTTGTGG
TAAGAGTACATTCATGAGGAGGTTAACAAGTGTTTTCGGTGGCGCGGCCGAGC
CACCAAAGGGTGGTAACCCAGATTCAAACACATTGATTAGTGACACTACTACT
GTTATCTGTTTGGATGATTTTCATTCCCTTGATAGAAATGGCAGGAAAGTGGA
AAAAGTTACTGCTTTAGACCCAAAAGCTAATGATTTTGATCTTATGTATGAACA
AGTTAAGGCTTTGAAAGAAGGTAAAGCTGTTGATAAACCTATTTATAATCATGT
TTCTGGTTTGTTGGACCCTCCTGAGCTTATTCAACCTCCTAAGATCTTGGTCAT
TGAAGGGTTACACCCCATGTATGACGCACGTGTGAGGGAATTGCTAGACTTCA
GCATCTACTTGGACATTAGCAATGAAGTTAAATTTGCCTGGAAAATTCAGAGA
GACATGAAAGAAAGAGGACACAGTCTTGAAAGCATCAAAGCCAGTATTGAATC
CAGAAAGCCAGATTTTGATGCTTACATTGACCCACAAAAGCAGCATGCTGATG
TAGTGATTGAAGTATTGCCAACTGAACTCATTCCTGATGATGATGAAGGCAAA
GTGTTGAGAGTAAGGATGATTCAGAAAGAAGGAGTCAAGTTTTTCAACCCAGT
TTACTTGTTTGATGAAGGATCTACCATTTCATGGATTCCATGTGGTAGAAAATT
AACATGTTCTTACCCTGGTATCAAATTTTCCTATGGCCCAGACACCTTCTATGG
CAACGAGGTGACAGTAGTAGAGATGGATGGGATGTTTGACAGATTAGACGAA
CTAATCTACGTCGAAAGCCATTTGAGCAATCTATCAACCAAGTTTTATGGTGAA

GTCACTCAACAAATGTTGAAGCACCAAAATTTCCCAGGAAGCAACAATGGAAC
TGGTTTCTTCCAAACCATAATTGGATTGAAGATCAGAGACTTGTTCGAGCAGC
TCGTTGCTAGCAGGTCTACAGCAACTGCAACAGCTGCTAAAGCCTAG

SEQUENCE ID NO 4:

Translated protein sequence of *prk* gene from *Spinacea oleracea*

MSQQQTIVIGLAADSGCGKSTFMRRLTSVFGGAAEPPKGGNPDSNTLISDTTTVI
CLDDFHSLDRNGRKVEKVTALDPKANDFDLMYEQVKALKEGKAVDKPIYNHV
SGLLDPPELIQPPKILVIEGLHPMYDARVRELLDFSIYLDISNEVKFAWKIQRDM
KERGHSLESIKASIESRKPDFDAYIDPQKQHADVVIEVLPTELIPDDDEGKVLRV
RMIQKEGVKFFNPVYLFDEGSTISWIPCGRKLTCSYPGIKFSYGPDTFYGNEVTV
VEMDGMFDRLDELIYVESHLSNLSTKFYGEVTQQMLKHQNFPGSNNGTGFFQ
TIIGLKIRDLFEQLVASRSTATATAAKA

SEQUENCE ID NO 5:

cbbQ2 gene (synthetic, based on cbbQ2 gene from *Thiobacillus denitrificans* –codon optimized, original sequence obtained from Beller *et al* 2006, GenBank Gene ID: 3672366, Protein ID: AAZ98590.1

ATGACTACTAACAAGGAACAATACAAGGTTCACCAAGAACCATACTACCAAGC
TCAAGGTAGAGAAGTTCAATTGTACGAAGCTGCTTACAGAAACAGATTGCCAG
TTATGGTTAAGGGTCCAACTGGTTGTGGTAAGTCTAGATTCGTTGAATACATG
GCTTGGAAGTTGAACAAGCCATTGATCACTGTTGCTTGTAACGAAGACATGAC
TGCTTCTGACTTGGTTGGTAGATACTTGTTGGAAGCTAACGGTACTAGATGGT
TGGACGGTCCATTGACTACTGCTGCTAGAATCGGTGCTATCTGTTACTTGGAC
GAAGTTGTTGAAGCTAGACAAGACACTACTGTTGTTATCCACCCATTGACTGA
CCACAGAAGAACTTTGCCATTGGACAAGAAGGGTGAATTGATCGAAGCTCACC
CAGACTTCCAATTGGTTATCTCTTACAACCCAGGTTACCAATCTTTGATGAAGG
ACTTGAAGCAATCTACTAAGCAAAGATTCGCTGCTTTCGACTTCGACTACCCA
GACGCTGCTTTGGAAACTACTATCTTGGCTAGAGAAACTGGTTTGGACGAAAC
TACTGCTGGTAGATTGGTTAAGATCGGTGGTGTTGCTAGAAACTTGAAGGGTC

ACGGTTTGGACGAAGGTATCTCTACTAGATTGTTGGTTTACGCTGCTACTTTG
ATGAAGGACGGTGTTGACGCTGGTGACGCTTGTAGAATGGCTTTGGTTAGACC
AATCACTGACGACGCTGACATCAGAGAAACTTTGGACCACGCTATCGACGCTA
CTTTCGCTTAA

SEQUENCE ID NO 6:
Translated protein sequence of *cbbQ2* gene from *Thiobacillus denitrificans*

MTTNKEQYKVHQEPYYQAQGREVQLYEAAYRNRLPVMVKGPTGCGKSRFVEY
MAWKLNKPLITVACNEDMTASDLVGRYLLEANGTRWLDGPLTTAARIGAICYL
DEVVEARQDTTVVIHPLTDHRRTLPLDKKGELIEAHPDFQLVISYNPGYQSLMK
DLKQSTKQRFAAFDFDYPDAALETTILARETGLDETTAGRLVKIGGVARNLKGH
GLDEGISTRLLVYAATLMKDGVDAGDACRMALVRPITDDADIRETLDHAIDATF
A

SEQUENCE ID NO 7:
cbbO2 gene (Synthetic, based on cbbO2 gene from *Thiobacillus denitrificans* –codon optimized, original sequence obtained from Beller *et al* 2006, GenBank Gene ID: 3672365, Protein ID: YP_316394.1

ATGGCTGCTTACTGGAAGGCTTTGGACACTAGATTCGCTCAAGTTGAAGAAGT
TTTCGACGACTGTATGGCTGAAGCTTTGACTGTTTTGTCTGCTGAAGGTGTTG
CTGCTTACTTGGAAGCTGGTAGAGTTATCGGTAAGTTGGGTAGAGGTGTTGAA
CCAATGTTGGCTTTCTTGGAAGAATGGCCATCTACTGCTCAAGCTGTTGGTGA
AGCTGCTTTGCCAATGGTTATGGCTTTGATCCAAAGAATGCAAAAGTCTCCAA
ACGGTAAGGCTATCGCTCCATTCTTGCAAACTTTGGCTCCAGTTGCTAGAAGA
TTGCAATCTGCTGAACAATTGCAACACTACGTTGACGTTACTTTGGACTTCATG
ACTAGAACTACTGGTTCTATCCACGGTCACCACACTACTTTCCCATCTCCAGGT
TTGCCAGAATTCTTCGCTCAAGCTCCAAACTTGTTGAACCAATTGACTTTGGCT
GGTTTGAGAAACTGGGTTGAATACGGTATCAGAAACTACGGTACTCACCCAGA
AAGACAACAAGACTACTTCTCTTTGCAATCTGCTGACGCTAGAGCTGTTTTGC
AAAGAGAAAGACACGGTACTTTGTTGGTTGACGTTGAAAGAAAGTTGGACTTG
TACTTGAGAGGTTTGTGGCAAGACCACGACCACTTGGTTCCATACTCTACTGC

TTTCGACGAAATCAGAAAGCCAGTTCCATACTACGACAAGTTGGGTATGAGAT
TGCCAGACGTTTACGACGACTTGGTTTTGCCATGTCCAGCTGGTAGAGGTGGT
GCTGGTGGTGAAGACGTTTTGGTTTCTGGTTTGGACAGATACAGAGCTACTTT
GGCTCACATGGTTGGTCACAGAAGATGGTCTGAAGCTCAAATCGCTGACAACT
GGTCTCCATTCCAAAGAATGGCTGTTGAATTCTTCGAAGACTGTAGAGTTGAA
ACTTTGTTGATGAGAGAATACCCAGGTTTGGCTAGAATCTTCAGAGCTTTGCA
CCCAAAGCCAGTTGAAGCTGCTTGTGACGGTGAAACTACTTCTTGTTTGAGAC
ACAGATTGGCTATGTTGTCTAGAGCTTTCATCGACCCAGACCACGGTTACGCT
GCTCCAGTTTTGAACGACTTCGTTGCTAGATTCCACGCTAGATTGGCTGACGG
TACTTCTTCTACTTCTGAAATGGCTGACTTGGCTTTGTCTTACGTTGCTAAGAC
TAGAAGACCATCTGACCAATTCGCTAAGGTTCACTTCGACGACACTGTTGTTG
ACTACAGAGACGACAACAGACAATTGTGGAAGTTCATCGAAGAAGGTGACGAA
GAAGAAGCTTTCGACGCTAAGAGAAAGATCGAACCAGGTGAAGAAATCCAAG
GTTTGCCACCAAGACACTACCCAGAATGGGACTACACTTCTCAAACTTACAGA
CCAGACTGGGTTTCTGTTTACGAAGGTTTGCACAGATCTGGTAACGCTGGTGA
CATCGACAGATTGTTGGCTAAGCACGCTGCTTTGGCTAAGAGATTGAAGAAGA
TGTTGGACTTGTTGAAGCCACAAGACAAGGTTAGAGTTAGATACCAAGAAGAA
GGTTCTGAATTGGACTTGGACGTTGCTATCAGATCTTTGATCGACTTCAAGGG
TGGTGCTACTCCAGACCCAAGAATCAACATGTCTCACAGATCTGACGGTAGAG
ACATCGCTGTTATGTTGTTGTTGGACTTGTCTGAATCTTTGAACGAAAAGGCT
GCTGGTGCTGGTCAAACTATCTTGGAATTGTCTCAAGAAGCTGTTTCTTTGTTG
GCTTGGTCTATCGAAAAGTTGGGTGACCCATTCGCTATCGCTGGTTTCCACTC
TAACACTAGACACGACGTTAGATACTTCCACATCAAGGGTTACTCTGAAAGAT
GGAACGACGACGTTAAGGCTAGATTGGCTGCTATGGAAGCTGGTTACTCTACT
AGAATGGGTGCTGCTATGAGACACGCTGCTCACTACTTGTCTGCTAGACCAGC
TGACAAGAAGTTGATGTTGATCTTGACTGACGGTAGACCATCTGACGTTGACG
CTGCTGACGAAAGATTGTTGGTTGAAGACGCTAGACAAGCTGTTAAGGAATTG
GACAGACAAGGTATCTTCGCTTACTGTATCTCTTTGGACGCTCAATTGAAGGC
TGGTGCTGACGACTACGTTGCTGAAATCTTCGGTAGACAATACACTGTTATCG
ACAGAGTTGAAAGATTGCCAGAAAGATTGCCAGAATTGTTCATGGCTTTGACT
AAGTAA

SEQUENCE ID NO: 8

Translated protein sequence of *cbbO2* gene from *Thiobacillus denitrificans*

MAAYWKALDTRFAQVEEVFDDCMAEALTVLSAEGVAAYLEAGRVIGKLGRGVE
PMLAFLEEWPSTAQAVGEAALPMVMALIQRMQKSPNGKAIAPFLQTLAPVARR
LQSAEQLQHYVDVTLDFMTRTTGSIHGHHTTFPSPGLPEFFAQAPNLLNQLTLA
GLRNWVEYGIRNYGTHPERQQDYFSLQSADARAVLQRERHGTLLVDVERKLDL
YLRGLWQDHDHLVPYSTAFDEIRKPVPYYDKLGMRLPDVYDDLVLPCPAGRGG
AGGEDVLVSGLDRYRATLAHMVGHRRWSEAQIADNWSPFQRMAVEFFEDCRV
ETLLMREYPGLARIFRALHPKPVEAACDGETTSCLRHRLAMLSRAFIDPDHGYA
APVLNDFVARFHARLADGTSSTSEMADLALSYVAKTRRPSDQFAKVHFDDTVV
DYRDDNRQLWKFIEEGDEEEAFDAKRKIEPGEEIQGLPPRHYPEWDYTSQTYR
PDWVSVYEGLHRSGNAGDIDRLLAKHAALAKRLKKMLDLLKPQDKVRVRYQE
EGSELDLDVAIRSLIDFKGGATPDPRINMSHRSDGRDIAVMLLLDLSESLNEKA
AGAGQTILELSQEAVSLLAWSIEKLGDPFAIAGFHSNTRHDVRYFHIKGYSERW
NDDVKARLAAMEAGYSTRMGAAMRHAAHYLSARPADKKLMLILTDGRPSDVD
AADERLLVEDARQAVKELDRQGIFAYCISLDAQLKAGADDYVAEIFGRQYTVID
RVERLPERLPELFMALTK

SEQUENCE ID NO: 9

GroEL gene (synthetic, based on GroEL from *E. coli* – codon optimized, original sequence obtained from Durfee *et al* 2008, Gene ID: 6061450, Protein ID: YP_001732912.1

ATGGCTGCTAAGGACGTTAAGTTCGGTAACGACGCTAGAGTTAAGATGTTGAG
AGGTGTTAACGTTTTGGCTGACGCTGTTAAGGTTACTTTGGGTCCAAAGGGTA
GAAACGTTGTTTTGGACAAGTCTTTCGGTGCTCCAACTATCACTAAGGACGGT
GTTTCTGTTGCTAGAGAAATCGAATTGGAAGACAAGTTCGAAAACATGGGTGC
TCAAATGGTTAAGGAAGTTGCTTCTAAGGCTAACGACGCTGCTGGTGACGGTA
CTACTACTGCTACTGTTTTGGCTCAAGCTATCATCACTGAAGGTTTGAAGGCT
GTTGCTGCTGGTATGAACCCAATGGACTTGAAGAGAGGTATCGACAAGGCTGT
TACTGCTGCTGTTGAAGAATTGAAGGCTTTGTCTGTTCCATGTTCTGACTCTAA

GGCTATCGCTCAAGTTGGTACTATCTCTGCTAACTCTGACGAAACTGTTGGTA
AGTTGATCGCTGAAGCTATGGACAAGGTTGGTAAGGAAGGTGTTATCACTGTT
GAAGACGGTACTGGTTTGCAAGACGAATTGGACGTTGTTGAAGGTATGCAATT
CGACAGAGGTTACTTGTCTCCATACTTCATCAACAAGCCAGAAACTGGTGCTG
TTGAATTGGAATCTCCATTCATCTTGTTGGCTGACAAGAAGATCTCTAACATCA
GAGAAATGTTGCCAGTTTTGGAAGCTGTTGCTAAGGCTGGTAAGCCATTGTTG
ATCATCGCTGAAGACGTTGAAGGTGAAGCTTTGGCTACTTTGGTTGTTAACAC
TATGAGAGGTATCGTTAAGGTTGCTGCTGTTAAGGCTCCAGGTTTCGGTGACA
GAAGAAAGGCTATGTTGCAAGACATCGCTACTTTGACTGGTGGTACTGTTATC
TCTGAAGAAATCGGTATGGAATTGGAAAAGGCTACTTTGGAAGACTTGGGTCA
AGCTAAGAGAGTTGTTATCAACAAGGACACTACTACTATCATCGACGGTGTTG
GTGAAGAAGCTGCTATCCAAGGTAGAGTTGCTCAAATCAGACAACAAATCGAA
GAAGCTACTTCTGACTACGACAGAGAAAAGTTGCAAGAAAGAGTTGCTAAGTT
GGCTGGTGGTGTTGCTGTTATCAAGGTTGGTGCTGCTACTGAAGTTGAAATGA
AGGAAAAGAAGGCTAGAGTTGAAGACGCTTTGCACGCTACTAGAGCTGCTGTT
GAAGAAGGTGTTGTTGCTGGTGGTGGTGTTGCTTTGATCAGAGTTGCTTCTAA
GTTGGCTGACTTGAGAGGTCAAAACGAAGACCAAAACGTTGGTATCAAGGTTG
CTTTGAGAGCTATGGAAGCTCCATTGAGACAAATCGTTTTGAACTGTGGTGAA
GAACCATCTGTTGTTGCTAACACTGTTAAGGGTGGTGACGGTAACTACGGTTA
CAACGCTGCTACTGAAGAATACGGTAACATGATCGACATGGGTATCTTGGACC
CAACTAAGGTTACTAGATCTGCTTTGCAATACGCTGCTTCTGTTGCTGGTTTGA
TGATCACTACTGAATGTATGGTTACTGACTTGCCAAAGAACGACGCTGCTGAC
TTGGGTGCTGCTGGTGGTATGGGTGGTATGGGTGGTATGGGTGGTATGATGT
AA

SEQUENCE ID NO: 10

Translated protein sequence of *GroEL* gene from *E. coli*

MAAKDVKFGNDARVKMLRGVNVLADAVKVTLGPKGRNVVLDKSFGAPTITKD
GVSVAREIELEDKFENMGAQMVKEVASKANDAAGDGTTTATVLAQAIITEGLK
AVAAGMNPMDLKRGIDKAVTAAVEELKALSVPCSDSKAIAQVGTISANSDETVG
KLIAEAMDKVGKEGVITVEDGTGLQDELDVVEGMQFDRGYLSPYFINKPETGA

VELESPFILLADKKISNIREMLPVLEAVAKAGKPLLIIAEDVEGEALATLVVNTM
RGIVKVAAVKAPGFGDRRKAMLQDIATLTGGTVISEEIGMELEKATLEDLGQAK
RVVINKDTTTIIDGVGEEAAIQGRVAQIRQQIEEATSDYDREKLQERVAKLAGGV
AVIKVGAATEVEMKEKKARVEDALHATRAAVEEGVVAGGGVALIRVASKLADL
RGQNEDQNVGIKVALRAMEAPLRQIVLNCGEEPSVVANTVKGGDGNYGYNAA
TEEYGNMIDMGILDPTKVTRSALQYAASVAGLMITTECMVTDLPKNDAADLGA
AGGMGGMGGMGGMM

SEQUENCE ID NO: 11

GroES gene (synthetic, based on GroES *E. coli* – codon optimized, original sequence obtained from Durfee et al 2008, Gene ID: 6061370, Protein ID: YP_001732911.1

ATGAACATCAGACCATTGCACGACAGAGTTATCGTTAAGAGAAAGGAAGTTGA
AACTAAGTCTGCTGGTGGTATCGTTTTGACTGGTTCTGCTGCTGCTAAGTCTA
CTAGAGGTGAAGTTTTGGCTGTTGGTAACGGTAGAATCTTGGAAAACGGTGAA
GTTAAGCCATTGGACGTTAAGGTTGGTGACATCGTTATCTTCAACGACGGTTA
CGGTGTTAAGTCTGAAAAGATCGACAACGAAGAAGTTTTGATCATGTCTGAAT
CTGACATCTTGGCTATCGTTGAAGCTTAA

SEQUENCE ID NO: 12

Translated protein sequence of *GroES* gene from *E. coli*

MNIRPLHDRVIVKRKEVETKSAGGIVLTGSAAAKSTRGEVLAVGNGRILENGEV
KPLDVKVGDIVIFNDGYGVKSEKIDNEEVLIMSESDILAIVEA

CLAIMS:

1. A recombinant yeast cell, functionally expressing one or more heterologous nucleic acid sequences encoding ribulose-1,5-biphosphate carboxylase oxygenase (Rubisco) and phosphoribulokinase (PRK), said yeast cell comprising one or more chaperones selected from the group consisting of GroEL, GroES, functional homologues of GroEL and functional homologues of GroES, wherein said functional homologues of a chaperone selected from the group consisting of GroEL and GroES have more than 50%, sequence identity with GroEL or GroES; and
wherein the ribulose-1,5-biphosphate carboxylase oxygenase (Rubisco) is a Rubisco comprising the protein sequence of SEQUENCE ID NO: 2 or a functional homologue thereof comprising a sequence having at least 80% sequence identity with SEQUENCE ID NO:2.

2. A recombinant yeast cell according to claim 1, wherein said one or more chaperones are prokaryotic molecular chaperones.

3. A recombinant yeast cell according to claim 2, wherein said one or more chaperones originate from a bacterium.

4. A recombinant yeast cell according to claim 2, wherein said one or more chaperones originate from Escherichia coli (E. coli).

5. A recombinant yeast cell according to any one of claims 1 to 4, wherein said functional homologue of a chaperone selected from the group consisting of GroEL and GroES has
60 % or more sequence identity with GroEL or GroES.

Date Reçue/Date Received 2021-03-31

6. A recombinant yeast cell according to any one of claims 1 to 5, wherein said Rubisco is a single subunit Rubisco.

7. A recombinant yeast cell according to any one of claims 1 to 6, wherein said Rubisco is a prokaryotic form-II Rubisco.

8. A recombinant yeast cell according to any one of claims 1 to 7, wherein said yeast cell is selected from the group consisting of *Saccharomyceraceae*; *Schizosaccharomyces*; *Torulaspora*; *Kluyveromyces*; *Pichia*, *Zygosaccharomyces*; *Brettanomyces*; *Metschnikowia*, *Issatchenkia*, *Kloeckera*; and *Aureobasidium*.

9. A recombinant yeast cell according to any one of claims 1 to 8, wherein said yeast cell is a *Saccharomyces cerevisiae* cell.

10. A recombinant yeast cell according to any one of claims 1 to 9, wherein the PRK is a PRK originating from a eukaryote.

11. A recombinant yeast cell according to claim 10, wherein the PRK is originating from Spinacia.

12. One or more vectors for the functional expression of a heterologous polypeptide in a yeast cell, wherein said vector or vectors comprise one or more heterologous nucleic acid sequences encoding Rubisco and PRK, wherein said Rubisco exhibits activity of carbon fixation, said vector or vectors further comprising one or more heterologous nucleic acid sequences encoding a chaperone selected from the group consisting of GroEL, GroES, functional homologues of GroEL and functional homologues of GroES wherein said functional homologues of a chaperone selected from the group consisting of GroEL and GroES have more than 50%, sequence identity with GroEL or GroES.

Date Reçue/Date Received 2021-03-31

13. A method for preparing an alcohol comprising fermenting a carbohydrate, with a yeast cell according to any one of claims 1-11, thereby forming the alcohol, wherein the yeast cell is present in a reaction medium.

14. A method according to claim 13, wherein the reaction medium comprises carbon dioxide wherein the carbon dioxide concentration in the reaction medium is at least 5 % of the carbon dioxide saturation concentration.

15. Method according to claim 13 or 14, wherein ethanol is formed.

Date Reçue/Date Received 2021-03-31

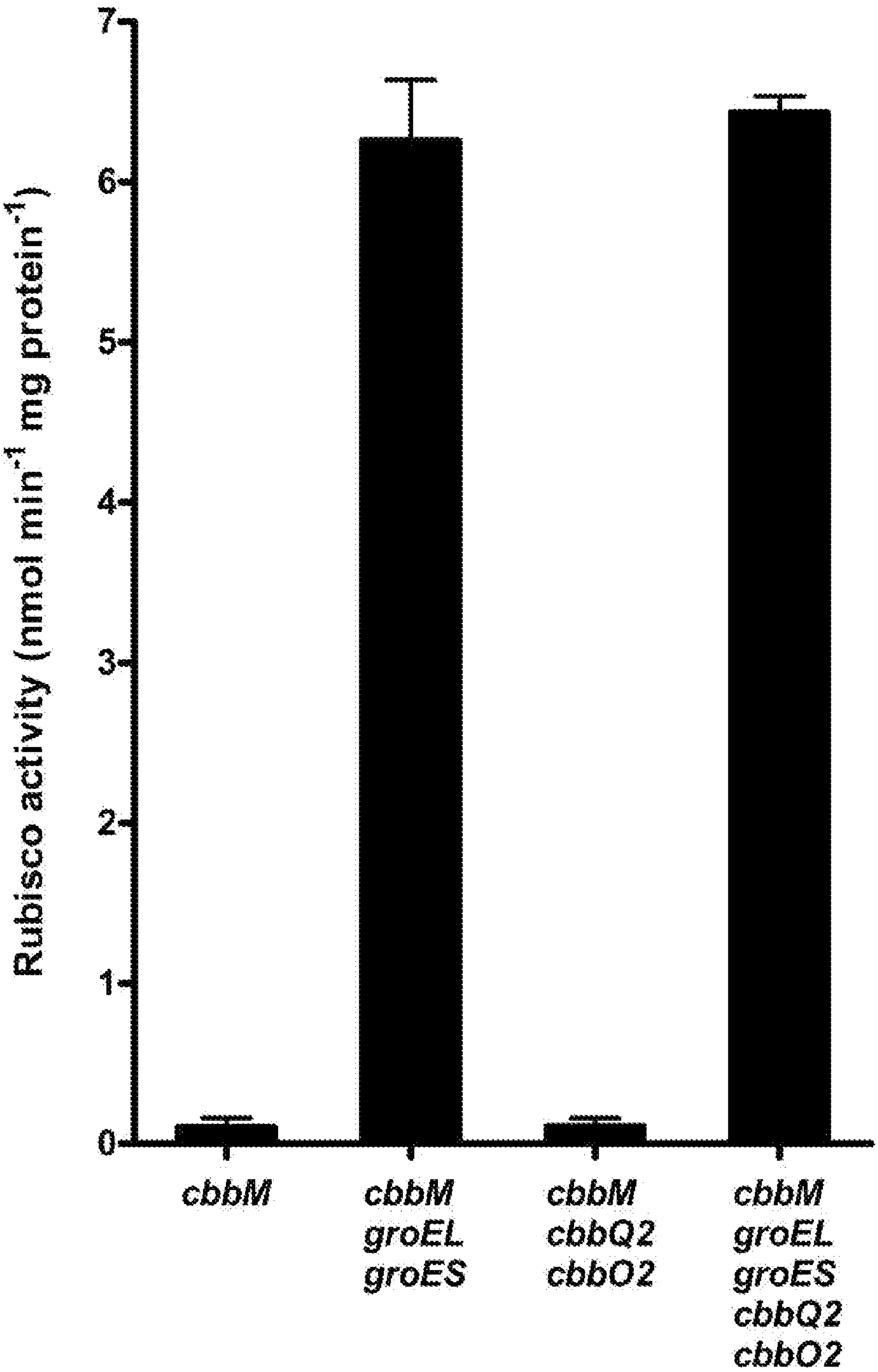
Rubisco activity (nmol min-1 mg protein-1)
7
6
5
4
3
2
1
0
cbbM
cbbM groEL groES
cbbM cbbQ2 cbbO2
cbbM groEL groES cbbQ2 cbbO2

Fig. 1