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

Innovation, Science and Economic Development Canada
Canadian Intellectual Property Office

CA 3027448 A1 2017/12/21

(21) 3 027 448

(12) DEMANDE DE BREVET CANADIEN
CANADIAN PATENT APPLICATION

(13) A1

(86) Date de dépôt PCT/PCT Filing Date: 2017/06/13
(87) Date publication PCT/PCT Publication Date: 2017/12/21
(85) Entrée phase nationale/National Entry: 2018/12/12
(86) N° demande PCT/PCT Application No.: EP 2017/064353
(87) N° publication PCT/PCT Publication No.: 2017/216136
(30) Priorité/Priority: 2016/06/14 (EP EP16174382.8)

(51) Cl.Int./Int.Cl. *C12N 1/19* (2006.01), *C12N 15/81* (2006.01), *C12N 9/12* (2006.01), *C12N 9/88* (2006.01), *C12P 7/06* (2006.01)

(71) Demandeur/Applicant:
DSM IP ASSETS B.V., NL

(72) Inventeurs/Inventors:
PAPAPETRIDIS, IOANNIS, NL;
VAN MARIS, ANTONIUS JEROEN ADRIAAN, SE;
PRONK, JACOBUS THOMAS, NL;
DE WAAL, PAULUS PETRUS, NL;
LAURET, NIELS, NL;
VERHOEVEN, MAARTEN D, NL;
SCHMITZ, JOZEF PETRUS JOHANNES, NL

(74) Agent: SMART & BIGGAR

(54) Titre : CELLULE DE LEVURE RECOMBINEE
(54) Title: RECOMBINANT YEAST CELL

(57) Abrégé/Abstract:
The present invention describes a recombinant yeast cell functionally expressing one or more heterologous nucleic acid sequences encoding for ribulose-1,5-phosphate carboxylase/oxygenase (EC4.1.1.39; Rubisco), and optionally one or more molecular chaperones for Rubisco, and one or more phosphoribulokinase (EC2.7.1.19; PRK), wherein the phosphoribulokinase is under control of a promoter (the "PRK promoter") that enables higher expression under anaerobic conditions than under aerobic conditions.

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
21 December 2017 (21.12.2017)

WIPO | PCT

(10) International Publication Number
WO 2017/216136 A1

(51) International Patent Classification:
C12N 1/19 (2006.01) *C12N 15/81* (2006.01)
C12N 9/12 (2006.01) *C12P 7/06* (2006.01)
C12N 9/88 (2006.01)

(21) International Application Number:
PCT/EP2017/064353

(22) International Filing Date:
13 June 2017 (13.06.2017)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
EP16174382.8 14 June 2016 (14.06.2016) EP

(71) Applicant: **DSM IP ASSETS B.V.** [NL/NL]; Het Overloon 1, 6411 TE HEERLEN (NL).

(72) Inventors: **PAPAPETRIDIS, Ioannis**; c/o Technische Universiteit Delft - Sectie Industriële Microbiologie Afd. Biotechnologie - P.O.Box 5, 2600 AA Delft (NL). **VAN MARIS, Antonius Jeroen Adriaan**; Division of Industrial Biotechnology, School of Biotechnology KTH Royal Institute of Technology, AlbaNova University Center, SE 10691 Stockholm (SE). **PRONK, Jacobus Thomas**; c/o Technische Universiteit Delft, Sectie Industriële Microbiologie Afd. Biotechnologie, Van der Maasweg 9, 2629 HZ Delft (NL). **DE WAAL, Paulus Petrus**; P.O. Box 4, 6100 AA ECHT (NL). **LAURET, Niels**; 6100 AA ECHT (NL). **VERHOEVEN, Maarten D**; c/o Technische Universiteit Delft, Sectie Industriële Microbiologie Afd. Biotechnologie, Van der Maasweg 9, 2629 HZ DELFT (NL). **SCHMITZ, Jozef Petrus Johannes**; P.O.Box 4, 6100 AA ECHT (NL).

(74) Agent: **ARENDSEN, Alexander Ferdinand**; P.O. Box 4, 6100 AA ECHT (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, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, 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, ST, 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))*
— *with sequence listing part of description (Rule 5.2(a))*

WO 2017/216136 A1

(54) Title: RECOMBINANT YEAST CELL

(57) Abstract: The present invention describes a recombinant yeast cell functionally expressing one or more heterologous nucleic acid sequences encoding for ribulose-1,5-phosphate carboxylase/oxygenase (EC4.1.1.39; Rubisco), and optionally one or more molecular chaperones for Rubisco, and one or more phosphoribulokinase (EC2.7.1.19; PRK), wherein the phosphoribulokinase is under control of a promoter (the "PRK promoter") that enables higher expression under anaerobic conditions than under aerobic conditions.

RECOMBINANT YEAST CELL

Field of the invention

The invention relates to a recombinant yeast cell having the ability to produce a desired fermentation product, to the functional expression of heterologous peptides in a yeast cell, and to a method for producing a fermentation product wherein said yeast cell is used. The invention is further related to a use of CO_2 in the yeast cell.

Background of the invention

Microbial fermentation processes are applied to 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 also 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. 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. WO2014/129898 describes a recombinant cell functionally heterologous nucleic acid sequences encoding for ribulose-1,5-phosphate carboxylase/oxygenase (EC 4.1.1.39; herein abbreviated as "Rubisco"), and optionally molecular chaperones for Rubisco, and phosphoribulokinase (EC 2.7.1.19; herein abbreviated as "PRK"). WO2015107496 describes a recombinant cell functionally heterologous nucleic acid sequences encoding for ribulose-1,5-phosphate carboxylase/oxygenase units RbcL, RbcS and RcbX, molecular chaperones for Rubisco GroEL and GroES. In the examples PRK is expressed with a tetracyclin-iinducible promoter TetO7, see table 5. Thereby, a process aid is needed for this promoter i.e. the additions of a compound to the propagation which adds to the cost and complexity of the process. The said compound is doxycycline, an antibiotic, which is not preferred as an additive in the ethanol fermentation process. Although the described process in WO2014/129898 is advantageous, there is a continuing need for improvement, in particular improved production of a useful organic compound, such as ethanol. Also, it would be desirable to provide a microorganism wherein NADH-dependent side-products are further reduced. Also a process is desirable wherein no additives, such as antibiotic, are needed. Further, it is desirable that the propagation characteristics of the strain are improved. These are among objects of the invention.

Summary of the invention

One or more of the aforementioned objects is realized according to the present invention that provides a recombinant yeast cell functionally expressing one or more heterologous nucleic acid sequences encoding for ribulose-1,5-phosphate carboxylase/oxygenase (EC4.1.1.39; Rubisco), and optionally one or more molecular chaperones for Rubisco, and one or more phosphoribulokinase (EC2.7.1.19; PRK), wherein the phosphoribulokinase is under control of a promoter which has a PRK expression ratio $_{anaerobic/aerobic}$ of 2 or more. Advantageously, such recombinant yeast cell has improved product yields and/or reduced side-product formation and/or improved propagation characteristics and/or absence of additives, such as antibiotic, to the fermentation process, so that the conventional fermentation process does not need to be changed.

Description of the Figures

Fig. 1 PRK activity in cell-free extracts of IME324 (left in fig.1) and IMX774 (right in fig. 1), harvested during exponential growth phase of anaerobic shake-flask cultures in synthetic medium (20 g L^{-1} glucose). Values represent the averages and the standard deviations of activity when 30, 50 or 100 µl of cell-free extract were used. Data were collected from single cultures.

Fig. 2 Yields (Y) of glycerol, biomass and ethanol on glucose and the ratio of glycerol formation to biomass formation in anaerobic bioreactor batch cultures of *S. cerevisiae* strains IME324 and IMX774. Cultures were grown on synthetic medium containing 20 g L^{-1} glucose (pH 5) and sparged with a gas mixture of N_2/CO_2 (90%/10%). Yields and ratios were calculated from the exponential growth phase. The ethanol yield on glucose was corrected for evaporation. Values represent average and mean deviation of data from independent duplicate cultures.

Fig. 3 Sum of peak area of selected unique peptides from *S. oleracea prk* among samples from both strains against prk. The peak areas give an indication of protein amount.

Fig. 4 Illustration of integration of anaerobic promoter-PRK cassettes at INT1 intergenic locus.

Fig. 5 Average glycerol values end of MTP batch fermentation expressed in arbitrary units (AU) for control strains IME324, IMX774, IMX765 and transformants of IMX765 with expression cassette anaerobic promoter-PRK introduced. Error bars indicate standard error of the mean.

Fig. 6 Average ethanol values expressed in arbitrary units (AU) end of MTP batch fermentation for control strains IME324, IMX774, IMX765 and transformants of IMX765 with expression cassette anaerobic promoter-PRK introduced. Error bars indicate standard error of the mean.

Table 1 - Description of the sequence listing

SEQ ID NO:	Description	Purpose
1		pUDR240 construction
2		pUDR240 construction
3		pUDR240 construction
4		pUDR119 and pUDR164 construction
5	Plasmid construction	pUDR119 and pUDR164 construction
6		pUDR119 and pUDR164 construction
7		pUDR119 construction
8		pUDR164 construction
9		pUDR164 diagnostic PCR
10		Addition of 20 bp primer-binding sequence to cbbM
11		Addition of 20 bp primer-binding sequence to cbbM
12		cbbM cassette construction - D tag addition (single copy cbbm-prk-chaperone integration)
13		cbbM cassette construction - J tag addition (single copy cbbm-prk-chaperone integration)
14		cbbM cassette construction - SGA1 tag addition
15		cbbM cassette construction - G tag addition
16		cbbM cassette construction - A tag addition
17		cbbM cassette construction - G tag addition
18		cbbM cassette construction - B tag addition
19		cbbM cassette construction - A tag addition
20	cbbM cassettes construction	cbbM cassette construction - C tag addition
21		cbbM cassette construction - B tag addition

22		cbbM cassette construction - D tag addition
23		cbbM cassette construction - C tag addition
24		cbbM cassette construction - D tag addition
25		cbbM cassette construction - M tag addition
26		cbbM cassette construction - M tag addition
27		cbbM cassette construction - N tag addition
28		cbbM cassette construction - N tag addition
29		cbbM cassette construction - O tag addition
30		cbbM cassette construction - O tag addition
31		Diagnostic primer cbbM integration
32		Diagnostic primer cbbM integration
33		Diagnostic primer cbbM integration
34		Diagnostic primer cbbM integration
35		Diagnostic primer cbbM integration
36		Diagnostic primer cbbM integration
37		Diagnostic primer cbbM integration
38		Diagnostic primer cbbM integration
39		Diagnostic primer cbbM integration
40		Diagnostic primer cbbM integration
41		Diagnostic primer cbbM integration
42	cbbM integration diagnostic primers	Diagnostic primer cbbM integration
43		Diagnostic primer cbbM integration
44		Diagnostic primer cbbM integration
45		Diagnostic primer cbbM integration
46		Diagnostic primer cbbM integration
47		Diagnostic primer cbbM integration
48		Diagnostic primer cbbM integration
49		Diagnostic primer cbbM integration
50		Diagnostic primer cbbM integration
51		Diagnostic primer cbbM integration
52		Diagnostic primer cbbM integration
53		Diagnostic primer cbbM integration
54		Diagnostic primer cbbM integration
55		groEL cassette construction - J tag addition
56	groES, groEL cassette construction	groEL cassette construction - H tag addition
57		groES cassette construction - H tag addition
58		groES cassette construction - SGA1 tag addition
59		LYS1p prk cassette construction
60		LYS1p prk cassette construction
61		UBC6p prk cassette construction
62		UBC6p prk cassette construction
63		YEN1p prk cassette construction
64		YEN1p prk cassette construction
65		DAN1p prk cassette construction
66		DAN1p prk cassette construction
67	prk cassettes construction	prk cassette construction (PGK1t)
68		prk cassette construction (PGK1t) - D tag addition (single copy cbbm-prk-chaperone integration)
69		prk cassette construction (PGK1t) - X-2 tag addition
70		prk amplification (LYS1p cassette)
71		prk amplification

72		prk amplification (UBC6p cassette)
73		prk amplification (YEN1p cassette)
74		prk amplification (DAN1p cassette)
75	URA3 cassette construction	URA3 amplification - SGA1 tag addition (single copy cbbm-prk-chaperone integration)
76		URA3 amplification - C tag addition (single copy cbbm-prk-chaperone integration)
77		Diagnostic primer single copy cbbm-prk-chaperone integration
78		Diagnostic primer single copy cbbm-prk-chaperone integration
79	Misc. diagnostic primers	Diagnostic primer prk integration in X-2
80		Diagnostic primer prk integration in X-2
81		Diagnostic primer prk integration in X-2
82		Diagnostic primer prk integration in X-2
83		Repair fragment GPD1
84		Repair fragment GPD1
85		Repair fragment GPD2
86	IMX675 construction and verification	Repair fragment GPD2
87		Diagnostic primer GPD1 deletion
88		Diagnostic primer GPD1 deletion
89		Diagnostic primer GPD2 deletion
90		Diagnostic primer GPD2 deletion
91	Promoter ANB1, S. cervisiae	
92	Promoter DAN1, S. cerevisiae	
93	Sc_DAN1 promoter	
94	Sc_DIP5 promoter	
95	Sc_TIR3 promoter	
96	Sc_TIR2 promoter	
97	Sc_HEM13 promoter	
98	Sc_YHK8 promoter	
99	Sc_FET4 promoter	
100	Sc_TIR4 promoter	
101	Sc_AAC3 promoter	
102	Sc_PGK1 terminator	
103	Oligononucleotide primer Fw-PRK (DBC-15631)	
104	Oligononucleotide primer Rv-PRK (DBC-15632)	
105	gBLOCK bearing sequence of SNR52p-gRNA.INT1-SUP4t	
106	pRN599	
107	Forward primer INT1-5'flank (BOZ-783)	
108	reverse primer INT1-5 'flank with connector c (DBC-19944)	
109	forward primer connector c to amplify anaerobic promoter-PRK-PGK1 terminator expression cassette (DBC-5799)	
110	reverse primer connector d to amplify PRK expression cassette (DBC-5800)	
111	forward primer to amplify URA3 marker with flank connector d (DBC-19947)	

112	reverse primer to amplify URA3 locus with flank connector e (DBC-19949)	
113	forward primer INT1-3'flank with connector d (DBC-19946)	
114	reverse primer INT1-3 'flank (BoZ-788)	
115	forward primer on backbone pRN599 (DBC-13775)	
116	reverse primer on backbone pRN599 (DBC-13776)	
117	forward primer on gRNA cassette INT1 (DBC-13773)	
118	reverse primer on gRNA cassette INT1 (DBC-13774)	
119	forward primer 7933 YEN1p prk cassette construction	

Detailed description of the invention

The invention thus relates to a recombinant yeast cell functionally expressing one or more heterologous nucleic acid sequences encoding for ribulose-1,5-phosphate carboxylase/oxygenase (EC4.1.1.39; Rubisco), and optionally one or more molecular chaperones for Rubisco, and one or more phosphoribulokinase (EC2.7.1.19; PRK), wherein the phosphoribulokinase is under control of a promoter (herein “the PRK promoter”) which has a PRK expression ratio $_{anaerobic/aerobic}$ of 2 or more.

In an embodiment the PRK promoter is ROX1 repressed. ROX1 is herein Heme-dependent repressor of hypoxic gene(s); that mediates aerobic transcriptional repression of hypoxia induced genes such as COX5b and CYC7; the repressor function is regulated through decreased promoter occupancy in response to oxidative stress; and contains an HMG domain that is responsible for DNA bending activity; involved in the hyperosmotic stress resistance. ROX1 is regulated by oxygen.

Though not to be limiting for the scope of the invention, we have herein found that the regulation of ROX1 may function as follows: According to Kwast et al [18]: “Although Rox1 functions in an O2-independent manner, its expression is oxygen (heme) dependent, activated by the heme-dependent transcription factor Hap1 [19]. Thus, as oxygen levels fall to those that limit heme biosynthesis [20], ROX1 is no longer transcribed [21], its protein levels fall [22], and the genes it regulates are de-repressed”.

In an embodiment, the PRK promoter has one or more ROX1 binding motif.

In an embodiment the PRK is under control of a promoter (herein “the PRK promoter”) which has a PRK expression ratio $_{anaerobic/aerobic}$ of 2 or more and the Rubisco is under a constitutive promotor.

In an embodiment, the PRK promoter comprises in its sequence one or more of the motif NNNATTGTTNNN.

In an embodiment, the PRK promoter is the native promoter of a gene selected from the list consisting of: FET4, ANB1, YHR048W, DAN1, AAC3, TIR2, DIP5, HEM13, YNR014W, YAR028W, FUN 57, COX5B, OYE2, SUR2, FRDS1, PIS1, LAC1, YGR035C, YAL028W, EUG1, HEM14, ISU2, ERG26, YMR252C and SML1; in particular FET4, ANB1, YHR048W, DAN1, AAC3, TIR2, DIP5 and HEM13.

In an embodiment, the PRK promoter comprises in its sequence one or more of the motif: TCGTTYAG and/or AAAAATTGTTGA.

In an embodiment, the PRK promoter is comprises in its sequence one or more sequence motif: TCGTTYAG and/or AAAAATTGTTG,

In particular such PRK promoter is native promoter of a DAN, TIR or PAU gene. In an embodiment, the PRK promoter is the native promoter of a gene selected from the list consisting of: TIR2, DAN1, TIR4, TIR3, PAU7, PAU5, YLL064C, YGR294W, DAN3, YIL176C, YGL261C, YOL161C, PAU1, PAU6, DAN2, YDR542W, YIR041W, YKL224C, PAU3, YLL025W, YOR394W, YHL046C, YMR325W, YAL068C, YPL282C, PAU2, and PAU4, in particular the PRK promoter is the native promoter of a gene selected from the list consisting of: TIR2, DAN1, TIR4, TIR3, PAU7, PAU5, YLL064C, YGR294W, DAN3, YIL176C, YGL261C, YOL161C, PAU1, PAU6, DAN2, YDR542W, YIR041W, YKL224C, PAU3, and YLL025W.

The PRK promoter has a PRK expression ratio $_{anaerobic/aerobic}$ of 2 or more, preferably of 3 or more, 4 or more, 5 or more, 6 or more, 7 or more, 8 or more, 9 or more, 10 or more, 20 or more or 50 or more. This is to say that the expression of PRK is at least a factor 2 higher under anaerobic conditions than under aerobic conditions.

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

In an embodiment the PRK expression ratio is determined by measuring the amount of PRK protein of cells grown under aerobic and anaerobic conditions. The amount of PRK protein can be determined by proteomics, as shown in the Examples.

In another embodiment the level or PRK expression ratio is determined by measuring the PRK activity of cells grown under aerobic and anaerobic conditions, e.g. in a cell-free extract. Methods to measure PRK activity are for instance described in Example 1.

In yet another embodiment the level or PRK expression ratio is determined by measuring the transcription level (e.g. as amount of mRNA) of the PRK gene of cells grown under aerobic and anaerobic conditions. The skilled person knows how to determine translation levels using methods commonly known in the art, e.g. Q-PCR, real-time PCR, northern blot, RNA-seq.

As used herein "promoter" is a DNA sequence that directs the transcription of a (structural) gene, in particular one or more phosphoribulokinase gene. The promoter enables higher expression during anaerobic conditions than under aerobic conditions.

In an embodiment, the PRK promoter may be a synthetic oligonucleotide. It may be a product of artificial oligonucleotide synthesis. Artificial oligonucleotide synthesis is a method in synthetic biology that is used to create artificial oligonucleotides, such as genes, in the laboratory.

Commercial gene synthesis services are now available from numerous companies worldwide, some of which have built their business model around this task. Current gene synthesis approaches are most often based on a combination of organic chemistry and molecular biological techniques and entire genes may be synthesized "de novo", without the need for precursor template DNA.

In an embodiment, the promoter is located in the 5'-region of a the PRK gene, In an embodiment it is located proximal to the transcriptional start site of PRK gene.

The invention further relates to a vector (as defined hereinafter) comprising PRK and a promoter which has a PRK expression ratio $_{anaerobic/aerobic}$ of 2 or more.

The invention further relates to a process for preparing an organic compound, in particular an alcohol, comprising converting a carbon source, in particular a carbohydrate or another organic carbon source using a yeast cell, thereby forming the organic compound, wherein the yeast cell is a yeast cell according to the invention.

The term “a” or “an” as used herein is defined as “at least one” unless specified otherwise.

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’.

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).

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 *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 14 June 2016) 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 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. In an embodiment, 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. In an embodiment, 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%.

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. In an embodiment there is no (external) inducer needed.

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.

"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 recombinant yeast cell is preferably selected from the group of *Saccharomycetaceae*, 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 Bruxellis* and *Dekkera anomala; Metschnikowia, Issatchenkia, such as Issatchenkia orientalis, Kloeckera such as Kloeckera apiculata; Aureobasisium* such as *Aureobasidium pullulans*.

In an embodiment, the yeast cell is selected from the group of *Saccharomycetaceae*. 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 Rubiscos. 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. A suitable Rubisco in accordance with the invention is encoded by the *cbbM* gene from *Thiobacillus denitrificans*. An 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 the *cbbM* gene from *Thiobacillus denitrificans* . Suitable natural Rubisco polypeptides are given in Table 2, with identity to the *cbbM* gene from *Thiobacillus denitrificans*.

Table 2: Natural Rubisco polypeptides suitable for expression

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

Source	Accession no.	MAX ID (%)
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. . In an embodiment, 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. The conditions for an assay for determining this Rubisco activity are as found in the Examples.

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

$$ATP + D\text{-}ribulose\ 5\text{-}phosphate \rightleftharpoons ADP + D\text{-}ribulose\ 1,5\text{-}bisphosphate \quad (1)$$

Thus, the two substrates of this enzyme are ATP and D-ribulose 5-phosphate; 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 an 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 the PRK from *Spinacia*. Suitable natural PRK polypeptides are given in Table 3.

Table 3: Natural PRK polypeptides suitable for expression with identity to PRK from *Spinacia*

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 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 an 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. In an embodiment, 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 recombinant yeast 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. Good results have been achieved with a recombinant yeast cell comprising both the heterologous chaperones GroEL and GroES. As an 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 GroES. Suitable natural chaperones polypeptide homologous to GroES are given in Table 4.

Table 4: Natural chaperones homologous to GroES 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]
>gi|401842164|gb|EJT44422.1|:1-92 HSP10-like protein [Saccharomyces kudriavzevii IFO 1802]
>gi|402084027|gb|EJT79045.1|:1-102 hsp10-like protein [Gaeumannomyces graminis var. triti
>gi|403215209|emb|CCK69709.1|:1-104 hypothetical protein KNAG_0C06130 [Kazachstania naganishii CBS 8797]
>gi|406604629|emb|CCH43969.1|:4-100 hypothetical protein BN7_3524 [Wickerhamomyces ciferrii]

>gi|406867021|gb|EKD20060.1|:56-156 hypothetical protein MBM_02012 [Marssonina brunnea f.sp. 'multigermtubi' MB_m1]
>gi|407926227|gb|EKG19196.1|:74-174 GroES-like protein [Macrophomina phaseolina MS6]
>gi|408398157|gb|EKJ77291.1|:11-111 hypothetical protein FPSE_02566 [Fusarium pseudograminearum CS3096]
>gi|410082063|ref|XP_003958610.1|:1-103 hypothetical protein KAFR_0H00660 [Kazachstania africana CBS2517]
>gi|425777664|gb|EKV15823.1|:58-157 Chaperonin, putative [Penicillium digitatum Pd1]
>gi|440639680|gb|ELR09599.1|:1-102 chaperonin GroES [Geomyces destructans 20631-21]
>gi|444323906|ref|XP_004182593.1|:1-105 hypothetical protein TBLA_0J00760 [Tetrapisisporablattae CBS 6284]
>gi|448083208|ref|XP_004195335.1|:2-101 Piso0_005888 [Millerozyma farinosa CBS 7064]
>gi|448087837|ref|XP_004196425.1|:2-102 Piso0_005888 [Millerozyma farinosa CBS 7064]
>gi|448534948|ref|XP_003870866.1|:1-106 Hsp10 protein [Candida orthopsilosis Co 90-125]
>gi|449295977|gb|EMC91998.1|:1-102 hypothetical protein BAUCODRAFT_39148 [Baudoinia compn
>gi|46123659|ref|XP_386383.1|:3-103 hypothetical protein FG06207.1 [Gibberella zeae PH-1]
>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 an 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 of GroEL. Suitable natural chaperones polypeptides homologous to GroEL are given in Table 5.

Table 5: Natural chaperones homologous to GroEL 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 118893] >gi|311341587|gb|EFR00790.1| chaperonin GroL [Arthroderma gypseum CBS 118893]
>gi|320580028|gb|EFW94251.1| Tetradecameric mitochondrial chaperonin [Ogataea parapolymorpha DL-1]
>gi|320586014|gb|EFW98693.1| heat shock protein mitochondrial precursor [Grosmannia clavigera kw1407]
>gi|322692465|gb|EFY84374.1| Heat shock protein 60 precursor (Antigen HIS-62) [Metarhizium acridum CQMa 102]
>gi|322705285|gb|EFY96872.1| Heat shock protein 60 (Antigen HIS-62) [Metarhizium anisopliae ARSEF 23]
>gi|323303806|gb|EGA57589.1| Hsp60p [Saccharomyces cerevisiae FostersB]
>gi|323307999|gb|EGA61254.1| Hsp60p [Saccharomyces cerevisiae FostersO]
>gi|323332364|gb|EGA73773.1| Hsp60p [Saccharomyces cerevisiae AWRI796]
>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]
>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]
>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]
>gi|336468236|gb|EGO56399.1| hypothetical protein NEUTE1DRAFT_122948 [Neurospora tetrasperma FGSC 2508]

>gi|340522598|gb|EGR52831.1| hsp60 mitochondrial precursor-like protein [Trichoderma reesei QM6a]
>gi|341038907|gb|EGS23899.1| mitochondrial heat shock protein 60-like protein [Chaetomium thermophilum var. thermophilum DSM 1495]
>gi|342886297|gb|EGU86166.1| hypothetical protein FOXB_03302 [Fusarium oxysporum Fo5176]
>gi|344230084|gb|EGV61969.1| chaperonin GroL [Candida tenuis ATCC 10573]
>gi|344303739|gb|EGW33988.1| hypothetical protein SPAPADRAFT_59397 [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 4 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 4 and 5 similarly made with same organism source are also available to the skilled person for expression. Furthermore, one may combine a chaperone from Table 4 from one organism with a chaperone from Table 5 from another organism, or one may combine GroES with a chaperone from Table 4, or one may combine GroEL with a chaperone from Table 5. 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 an embodiment a fermentation process is provided, 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, and carbon

dioxide is present in the gas mixture above the fermentation broth and/or dissolved in the fermentation broth. 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*.

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.

As shown in the Examples below, the invention is in suitable 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 embodiments of the use of carbon dioxide as an electron acceptor in accordance with the invention.

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. Other embodiments of the invention are now described in more detail.

In an embodiment the invention relates to the use of the recombinant yeast cell as described herein in fermentation in the biofuel industry. The recombinant yeast cell may contain genes of a pentose metabolic pathway non-native to the recombinant yeast cell and/or that allow the recombinant yeast cell to convert pentose(s). In one embodiment, the recombinant yeast cell may comprise one or two or more copies of one or more xylose isomerases and/or one or two or more copies of one or more xylose reductase and xylitol dehydrogenase genes, allowing the recombinant yeast cell to convert xylose. In an embodiment thereof, these genes may be integrated into the recombinant yeast cell genome. In another embodiment, the recombinant yeast cell comprises the genes *araA*, *araB* and *araD*. It is then able to ferment arabinose. In one embodiment

of the invention the recombinant yeast cell comprises *xylA*-gene, *XYL1* gene and *XYL2* gene and/or *XKS1*-gene, to allow the recombinant yeast cell to ferment xylose; deletion of the aldose reductase (GRE3) gene; and/or overexpression of GAL2 and/or deletion of GAL80. Thus though inclusion of the above genes, suitable pentose or other metabolic pathway(s) may be introduced in the recombinant yeast cell that were non-native in the (wild type) recombinant yeast cell. According to an embodiment, the following genes may be introduced in the recombinant yeast cell by introduction into a host cell:

1) a set consisting of PPP-genes *TAL1*, *TKL1*, *RPE1* and *RKI1*, optionally under control of strong constitutive promoter;
2) a set consisting of a x*ylA*-gene under under control of strong constitutive promoter;
3) a set comprising a *XKS1*-gene under control of strong constitutive promoter,
4) a set consisting of the genes *araA*, *araB* and *araD* under control of a strong constitutive promoter
5) deletion of an aldose reductase gene

The above cells may be constructed using known recombinant expression techniques. The co-factor modification may be effected before, simultaneous or after any of the modifications 1) to 5). The recombinant yeast cell according to the invention may be subjected to evolutionary engineering to improve its properties. Evolutionary engineering processes are known processes. Evolutionary engineering is a process wherein industrially relevant phenotypes of a microorganism, herein the recombinant yeast cell, can be coupled to the specific growth rate and/or the affinity for a nutrient, by a process of rationally set-up natural selection. Evolutionary Engineering is for instance described in detail in Kuijper, M, et al, FEMS, Eukaryotic cell Research 5(2005) 925-934, WO2008/041840 and WO2009/112472. After the evolutionary engineering the resulting pentose fermenting recombinant yeast cell is isolated. The isolation may be executed in any known manner, e.g. by separation of cells from a recombinant yeast cell broth used in the evolutionary engineering, for instance by taking a cell sample or by filtration or centrifugation.

In an embodiment, the recombinant yeast cell is marker-free. As used herein, the term "marker" refers to a gene encoding a trait or a phenotype which permits the selection of, or the screening for, a host cell containing the marker. Marker-free means that markers are essentially absent in the recombinant yeast cell. Being marker-free is particularly advantageous when antibiotic markers have been used in construction of the recombinant yeast cell and are removed thereafter. Removal of markers may be done using any suitable prior art technique, e.g. intramolecular recombination.

In one embodiment, the industrial recombinant yeast cell is constructed on the basis of an inhibitor tolerant host cell, wherein the construction is conducted as described hereinafter. Inhibitor tolerant host cells may be selected by screening strains for growth on inhibitors containing materials, such as illustrated in Kadar et al, Appl. Biochem. Biotechnol. (2007), Vol. 136-140, 847-858, wherein an inhibitor tolerant *S. cerevisiae* strain ATCC 26602 was selected.

The recombinant yeast cell further may comprise those enzymatic activities required for conversion of pyruvate to a desired fermentation product, such as ethanol, butanol (e.g. n-butanol, 2-butanol and isobutanol), lactic acid, 3 -hydroxy- propionic acid, acrylic acid, acetic acid, succinic acid, citric acid, fumaric acid, malic acid, itaconic acid, an amino acid, 1,3- propane-diol, ethylene, glycerol, a ß-lactam antibiotic or a cephalosporin.

In an embodiment, the recombinant yeast cell is derived from an industrial recombinant yeast cell. An industrial cell and industrial recombinant yeast cell may be defined as follows. The living environments of (recombinant yeast cell) cells in industrial processes are significantly different from that in the laboratory. Industrial recombinant yeast cells must be able to perform well under multiple environmental conditions which may vary during the process. Such variations include change in nutrient sources, pH, ethanol concentration, temperature, oxygen concentration, etc., which together have potential impact on the cellular growth and ethanol production of *Saccharomyces cerevisiae*. Under adverse industrial conditions, the environmental tolerant strains should allow robust growth and production. Industrial recombinant yeast cell strains are generally more robust towards these changes in environmental conditions which may occur in the applications they are used, such as in the baking industry, brewing industry, wine making and the biofuel ethanol industry. In one embodiment, the industrial recombinant yeast cell is constructed on the basis of an industrial host cell, wherein the construction is conducted as described hereinafter. Examples of industrial yeast cell (*S. cerevisiae*) are Ethanol Red® (Fermentis) Fermiol® (DSM) and Thermosacc® (Lallemand).

The recombinant yeast cells according to the invention are preferably inhibitor tolerant, i.e. they can withstand common inhibitors at the level that they typically have with common pretreatment and hydrolysis conditions, so that the recombinant yeast cells can find broad application, i.e. it has high applicability for different feedstock, different pretreatment methods and different hydrolysis conditions. In an embodiment the recombinant yeast cell is inhibitor tolerant. Inhibitor tolerance is resistance to inhibiting compounds. The presence and level of inhibitory compounds in lignocellulose may vary widely with variation of feedstock, pretreatment method hydrolysis process. Examples of categories of inhibitors are carboxylic acids, furans and/or phenolic compounds. Examples of carboxylic acids are lactic acid, acetic acid or formic acid. Examples of furans are furfural and hydroxy- methylfurfural. Examples or phenolic compounds are vannilin, syringic acid, ferulic acid and coumaric acid. The typical amounts of inhibitors are for carboxylic acids: several grams per liter, up to 20 grams per liter or more, depending on the feedstock, the pretreatment and the hydrolysis conditions. For furans: several hundreds of milligrams per liter up to several grams per liter, depending on the feedstock, the pretreatment and the hydrolysis conditions. For phenolics: several tens of milligrams per liter, up to a gram per liter, depending on the feedstock, the pretreatment and the hydrolysis conditions.

In an embodiment, the recombinant yeast cell is a cell that is naturally capable of alcoholic fermentation, preferably, anaerobic alcoholic fermentation. A recombinant yeast cell preferably has a high tolerance to ethanol, a high tolerance to low pH (i.e. capable of growth at a pH lower than

about 5, about 4, about 3, or about 2.5) and towards organic and/or a high tolerance to elevated temperatures.

In an embodiment one or more genes of the non-oxidative branch of the pentose phosphate pathway of the recombinant yeast of the invention are overexpressed, and/or a glycerol-3-phosphate dehydrogenase (GPD) gene is deleted or disrupted. In another embodiment a glycerol-3-phosphate dehydrogenase (GPD) gene is deleted or disrupted. In yet another embodiment one or more genes of the non-oxidative branch of the pentose phosphate pathway of the recombinant yeast of the invention are overexpressed and a glycerol-3-phosphate dehydrogenase (GPD) gene is deleted or disrupted. The GPD gene may be a *GPD1* and/or a *GPD2* gene. Both *GPD1* and *GPD2* genes may be deleted or disrupted, although it is preferred that *GPD2*, but not *GPD1* is deleted or disrupted. The GPD gene encodes for an enzyme having at least EC number 1.1.1.8. WO2011/010923 describes methods to delete or disrupt a glycerol-3-phosphate dehydrogenase. In an embodiment the one or more genes of the pentose phosphate pathway that is overexpressed encodes for an enzyme selected from the list of a transaldolase (EC 2.2.1.2), a transketolase (EC 2.2.1.1), a ribose-5-phosphate isomerase (EC 5.3.1.6) and a D-ribulose-5-phosphate 3-epimerase (EC 5.1.3.1). In another embodiment the one or more genes of the pentose phosphate pathway that is overexpressed is selected from the list of TAL1, TAL2, NQM1, TKL1, TKL2, RPE1 and RKI1.

The invention also relates to a process for the fermentation of a substrate to produce a fermentation product with a recombinant yeast cell as described herein, in the biofuel industry, wherein the glycerol yield is at least 5%, at least 10% or at least 10%, at least 20% or at least 30% lower than that of a process with the corresponding wild-type recombinant yeast cell. In an embodiment of such process, the ethanol yield is not increased or decreased, compared to that of a process with the corresponding wild-type recombinant yeast cell.

Any of the above characteristics or activities of a recombinant yeast cell may be naturally present in the cell or may be introduced or modified by genetic modification.

Recombinant expression

The recombinant yeast cell is a recombinant cell. That is to say, a recombinant yeast cell comprises, or is transformed with or is genetically modified with a nucleotide sequence that does not naturally occur in the cell in question. Techniques for the recombinant expression of enzymes in a cell, as well as for the additional genetic modifications of a recombinant yeast cell are well known to those skilled in the art. Typically such techniques involve transformation of a cell with nucleic acid construct comprising the relevant sequence. Such methods are, for example, known from standard handbooks, such as Sambrook and Russel (2001) "Molecular Cloning: A Laboratory Manual (3rd edition), Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, or F. Ausubel *et al.*, eds., "Current protocols in molecular biology", Green Publishing and Wiley Interscience, New York (1987). Methods for transformation and genetic modification of fungal host cells are known from e.g. EP-A-0635574, WO98/46772, WO 99/60102, WO00/37671, WO90/14423, EP-A-0481008, EP-A-0635574 and US6,265,186.

Bioproducts production

Over the years suggestions have been made for the introduction of various organisms for the production of bio-ethanol from crop sugars. In practice, however, all major bio-ethanol production processes have continued to use the recombinant yeast cells of the genus *Saccharomyces* as ethanol producer. This is due to the many attractive features of *Saccharomyces* species for industrial processes, i. e. a high acid-, ethanol tolerance and osmotolerance, capability of anaerobic growth, and of course its high alcoholic fermentative capacity. In an embodiment, recombinant yeast cell species as host cells include *S. cerevisiae*, *S. bulderi*, *S. barnetti*, *S. exiguus*, *S. uvarum*, *S. diastaticus*, *K. lactis*, *K. marxianus* or *K. fragilis*. A recombinant yeast cell may be a cell suitable for the production of ethanol. A recombinant yeast cell may, however, be suitable for the production of fermentation products other than ethanol. Such non-ethanolic fermentation products include in principle any bulk or fine chemical that is producible by a eukaryotic microorganism such as a recombinant yeast cell or a filamentous fungus.

In an embodiment, recombinant yeast cell for production of non-ethanolic fermentation products is a host cell that contains a genetic modification that results in decreased alcohol dehydrogenase activity.

Lignocellulose

Lignocellulose, which may be considered as a potential renewable feedstock, generally comprises the polysaccharides cellulose (glucans) and hemicelluloses (xylans, heteroxylans and xyloglucans). In addition, some hemicellulose may be present as glucomannans, for example in wood-derived feedstocks. The enzymatic hydrolysis of these polysaccharides to soluble sugars, including both monomers and multimers, for example glucose, cellobiose, xylose, arabinose, galactose, fructose, mannose, rhamnose, ribose, galacturonic acid, glucuronic acid and other hexoses and pentoses occurs under the action of different enzymes acting in concert. In addition, pectins and other pectic substances such as arabinans may make up considerably proportion of the dry mass of typically cell walls from non-woody plant tissues (about a quarter to half of dry mass may be pectins).

Pretreatment

Before enzymatic treatment, the lignocellulosic material may be pretreated. The pretreatment may comprise exposing the lignocellulosic material to an acid, a base, a solvent, heat, a peroxide, ozone, mechanical shredding, grinding, milling or rapid depressurization, or a combination of any two or more thereof. This chemical pretreatment is often combined with heat-pretreatment, e.g. between 150-220 °C for 1 to 30 minutes.

Enzymatic hydrolysis

The pretreated material is commonly subjected to enzymatic hydrolysis to release sugars

that may be fermented according to the invention. This may be executed with conventional methods, e.g. contacting with cellulases, for instance cellobiohydrolase(s), endoglucanase(s), beta-glucosidase(s) and optionally other enzymes, The conversion with the cellulases may be executed at ambient temperatures or at higher temperatures, at a reaction time to release sufficient amounts of sugar(s). The result of the enzymatic hydrolysis is hydrolysis product comprising C5/C6 sugars, herein designated as the sugar composition.

The sugar composition

The sugar composition used according to the invention comprises glucose and one or more pentose, e.g. arabinose and/or xylose. Any sugar composition may be used in the invention that suffices those criteria. Optional sugars in the sugar composition are galactose and mannose. In an embodiment, the sugar composition is a hydrolysate of one or more lignocellulosic material. Lignocelllulose herein includes hemicellulose and hemicellulose parts of biomass. Also lignocellulose includes lignocellulosic fractions of biomass. Suitable lignocellulosic materials may be found in the following list: orchard primings, chaparral, mill waste, urban wood waste, municipal waste, logging waste, forest thinnings, short-rotation woody crops, industrial waste, wheat straw, oat straw, rice straw, barley straw, rye straw, flax straw, soy hulls, rice hulls, rice straw, corn gluten feed, oat hulls, sugar cane, corn stover, corn stalks, corn cobs, corn husks, switch grass, miscanthus, sweet sorghum, canola stems, soybean stems, prairie grass, gamagrass, foxtail; sugar beet pulp, citrus fruit pulp, seed hulls, cellulosic animal wastes, lawn clippings, cotton, seaweed, trees, softwood, hardwood, poplar, pine, shrubs, grasses, wheat, wheat straw, sugar cane bagasse, corn, corn husks, corn hobs, corn kernel, fiber from kernels, products and by-products from wet or dry milling of grains, municipal solid waste, waste paper, yard waste, herbaceous material, agricultural residues, forestry residues, municipal solid waste, waste paper, pulp, paper mill residues, branches, bushes, canes, corn, corn husks, an energy crop, forest, a fruit, a flower, a grain, a grass, a herbaceous crop, a leaf, bark, a needle, a log, a root, a sapling, a shrub, switch grass, a tree, a vegetable, fruit peel, a vine, sugar beet pulp, wheat midlings, oat hulls, hard or soft wood, organic waste material generated from an agricultural process, forestry wood waste, or a combination of any two or more thereof. The conversion of glucose, xylose, arabinose and galactose to fermentation product is of great economic importance. Also mannose is present in some lignocellulose materials be it usually in lower amounts than the previously mentioned sugars. Advantageously therefore also mannose is converted by the recombinant yeast cell. It is expected that recombinant yeast cells of the present invention can be further manipulated to achieve other desirable characteristics, or even higher overall ethanol yields. Selection of improved recombinant yeast cells by passaging the recombinant yeast cells on medium containing hydrolysate has resulted in improved recombinant yeast cell with enhanced fermentation rates. Using the teachings of the present invention, one could readily such improved strains. By pentose-containing material, it is meant any medium comprising pentose, whether liquid or solid. Suitable pentose-containing materials include hydrolysates of polysaccharide or lignocellulosic biomass such as corn hulls,

wood, paper, agricultural byproducts, and the like.

By a "hydrolysate" as used herein, it is meant a polysaccharide that has been depolymerized through the addition of water to form mono and oligosaccharide sugars. Hydrolysates may be produced by enzymatic or acid hydrolysis of the polysaccharide-containing material.

Preferably, the recombinant yeast cell is able to grow under conditions similar to those found in industrial sources of pentose. The method of the present invention would be most economical when the pentose-containing material can be inoculated with the recombinant yeast cell variant without excessive manipulation. By way of example, the pulping industry generates large amounts of cellulosic waste. Saccharification of the cellulose by acid hydrolysis yields hexoses and pentoses that can be used in fermentation reactions. However, the hydrolysate or sulfite liquor contains high concentrations of sulfite and phenolic inhibitors naturally present in the wood which inhibit or prevent the growth of most organisms. The examples below describe the fermentation of pentose in acid hydrolysates (or sulfite waste liquor) of hard woods and soft woods by the recombinant yeast cells of the present invention. It is reasonably expected that recombinant yeast cell strains capable of growing in sulfite waste liquor could grow be expected grow in virtually any other biomass hydrolysate.

Propagation

The invention further relates to a process for aerobic propagation of the recombinant yeast cell, in particular aerobic propagation of the recombinant yeast cell strain. Propagation is herein any process of recombinant yeast cell growth that leads to increase of an initial recombinant yeast cell population. Main purpose of propagation is to increase a recombinant yeast cell population using the recombinant yeast cell's natural reproduction capabilities as living organisms. There may be other reasons for propagation, for instance, in case dry recombinant yeast cell is used, propagation is used to rehydrate and condition the recombinant yeast cell, before it is grown. Fresh recombinant yeast cell, whether active dried recombinant yeast cell or wet cake may be added to start the propagation directly. The conditions of propagation are critical for optimal recombinant yeast cell production and subsequent fermentation, such as for example fermentation of lignocellulosic hydrolysate into ethanol. They include adequate carbon source, aeration, temperature and nutrient additions. Tank size for propagation and is normally between 2 percent and 5 percent of the (lignocellulosic hydrolysate to ethanol) fermentor size. In the propagation the recombinant yeast cell needs a source of carbon. The source of carbon may herein comprise glycerol, ethanol, acetate and/or sugars (C6 and C5 sugars). Other carbon sources may also be used. The carbon source is needed for cell wall biosynthesis and protein and energy production. Propagation is an aerobic process, thus the propagation tank must be properly aerated to maintain a certain level of dissolved oxygen. Adequate aeration is commonly achieved by air inductors installed on the piping going into the propagation tank that pull air into the propagation mix as the tank fills and during recirculation. The capacity for the propagation mix to retain dissolved oxygen

is a function of the amount of air added and the consistency of the mix, which is why water is often added at a ratio of between 50:50 to 90:10 mash to water. "Thick" propagation mixes (80:20 mash-to-water ratio and higher) often require the addition of compressed air to make up for the lowered capacity for retaining dissolved oxygen. The amount of dissolved oxygen in the propagation mix is also a function of bubble size, so some ethanol plants add air through spargers that produce smaller bubbles compared to air inductors. Along with lower glucose, adequate aeration is important to promote aerobic respiration, which differs from the comparably anaerobic environment of fermentation. One sign of inadequate aeration or high glucose concentrations is increased ethanol production in the propagation tank. Generally during propagation, recombinant yeast cell requires a comfortable temperature for growth and metabolism, for instance the temperature in the propagation reactor is between 25-40°C. Generally lower temperatures result in slower metabolism and reduced reproduction, while higher temperatures can cause production of stress compounds and reduced reproduction. In an embodiment the propagation tanks are indoors and protected from the insult of high summer or low winter temperatures, so that maintaining optimum temperatures of between within the range of 30-35 degrees C is usually not a problem. Further propagation may be conducted as propagation of recombinant yeast cell is normally conducted.

Fermentation

The invention provides a process for the fermentation of a recombinant yeast cell according to the invention e.g. ethanol, that is advantageous in the biofuel industry.

In an embodiment, the recombinant yeast cell according to the invention may be a pentose and glucose fermenting recombinant yeast cell, including but not limited to such cells that are capable of anaerobic simultaneous pentose and glucose consumption. In an embodiment of the process the pentose-containing material comprises a hydrolysate of ligno-cellulosic material. The hydrolysate may be an enzymatic hydrolysate of ligno-cellulosic material.

The fermentation process may be an aerobic or an anaerobic fermentation process. An anaerobic fermentation process is herein defined as a fermentation process run in the absence of oxygen or in which substantially no oxygen is consumed, preferably less than about 5, about 2.5 or about 1 mmol/L/h, more preferably 0 mmol/L/h is consumed (i.e. oxygen consumption is not detectable), and wherein organic molecules serve as both electron donor and electron acceptors. In the absence of oxygen, NADH produced in glycolysis and biomass formation, cannot be oxidised by oxidative phosphorylation. To solve this problem many microorganisms use pyruvate or one of its derivatives as an electron and hydrogen acceptor thereby regenerating NAD^+.

Thus, in an embodiment, anaerobic fermentation process pyruvate is used as an electron (and hydrogen acceptor) and is reduced to fermentation products such as ethanol, butanol, lactic acid, 3 -hydroxy-propionic acid, acrylic acid, acetic acid, succinic acid, malic acid, fumaric acid, an amino acid and ethylene.

The fermentation process is preferably run at a temperature that is optimal for the cell. Thus, for most recombinant yeast cells, the fermentation process is performed at a temperature

which is less than about 50°C, less than about 42°C, or less than about 38°C. For recombinant yeast cell or filamentous fungal host cells, the fermentation process is preferably performed at a temperature which is lower than about 35, about 33, about 30 or about 28°C and at a temperature which is higher than about 20, about 22, or about 25°C.

The ethanol yield on xylose and/or glucose in the process preferably is at least about 50, about 60, about 70, about 80, about 90, about 95 or about 98%. The ethanol yield is herein defined as a percentage of the theoretical maximum yield.

The invention also provides a process for producing a fermentation product. The fermentation process according to the present invention may be run under aerobic and anaerobic conditions. In an embodiment, the process is carried out under micro-aerophilic or oxygen limited conditions. An anaerobic fermentation process is herein defined as a fermentation process run in the absence of oxygen or in which substantially no oxygen is consumed, preferably less than about 5, about 2.5 or about 1 mmol/L/h, and wherein organic molecules serve as both electron donor and electron acceptors. An oxygen-limited fermentation process is a process in which the oxygen consumption is limited by the oxygen transfer from the gas to the liquid. The degree of oxygen limitation is determined by the amount and composition of the ingoing gasflow as well as the actual mixing/mass transfer properties of the fermentation equipment used. Preferably, in a process under oxygen-limited conditions, the rate of oxygen consumption is at least about 5.5, more preferably at least about 6, such as at least 7 mmol/L/h. A process of the invention may comprise recovery of the fermentation product. In an embodiment of the process, the cell ferments both the xylose and glucose, preferably simultaneously in which case preferably a cell is used which is insensitive to glucose repression to prevent diauxic growth. In addition to a source of xylose (and glucose) as carbon source, the fermentation medium will further comprise the appropriate ingredient required for growth of the cell. Compositions of fermentation media for growth of microorganisms such as recombinant yeast cells are well known in the art. The fermentation processes may be carried out in batch, fed-batch or continuous mode. A separate hydrolysis and fermentation (SHF) process or a simultaneous saccharification and fermentation (SSF) process may also be applied. A combination of these fermentation process modes may also be possible for optimal productivity. These processes are described hereafter in more detail.

SSF mode

For Simultaneous Saccharification and Fermentation (SSF) mode, the reaction time for liquefaction/hydrolysis or presaccharification step is dependent on the time to realize a desired yield, i.e. cellulose to glucose conversion yield. Such yield is preferably as high as possible, preferably 60% or more, 65% or more, 70% or more, 75% or more 80% or more, 85% or more,90% or more, 95% or more, 96% or more, 97% or more, 98% or more, 99% or more, even 99.5% or more or 99.9% or more. According to the invention very high sugar concentrations in SHF mode and very high product concentrations (e.g. ethanol) in SSF mode are realized. In SHF operation the glucose concentration is 25g/L or more, 30 g/L or more, 35g/L or more, 40 g/L or more, 45 g/L

or more, 50 g/L or more, 55 g/L or more, 60 g/L or more, 65 g/L or more, 70 g/L or more , 75 g/L or more, 80 g/L or more, 85 g/L or more, 90 g/L or more, 95 g/L or more, 100 g/L or more, 110 g/L or more, 120g/L or more or may e.g. be 25g/L-250 g/L, 30gl/L-200g/L, 40g/L-200 g/L, 50g/L-200g/L, 60g/L-200g/L, 70g/L-200g/L, 80g/L-200g/L, 90 g/L-200g/L.

Product concentration in SSF mode

In SSF operation, the product concentration (g/L) is dependent on the amount of glucose produced, but this is not visible since sugars are converted to product in the SSF, and product concentrations can be related to underlying glucose concentration by multiplication with the theoretical maximum yield (Yps max in gr product per gram glucose). The theoretical maximum yield (Yps max in gr product per gram glucose) of a fermentation product can be derived from textbook biochemistry. For ethanol, 1 mole of glucose (180 gr) yields according to normal glycolysis fermentation pathway in recombinant yeast cell 2 moles of ethanol (=2x46 = 92 gr ethanol. The theoretical maximum yield of ethanol on glucose is therefore 92/180 = 0.511 gr ethanol/gr glucose). For n-butanol (MW 74 gr/mole) or iso butanol, the theoretical maximum yield is 1 mole of butanol per mole of glucose. So Yps max for (iso-)butanol = 74/180 = 0.411 gr (iso-)butanol/gr glucose. For lactic acid the fermentation yield for homolactic fermentation is 2 moles of lactic acid (MW = 90 gr/mole) per mole of glucose. According to this stoichiometry, the Yps max = 1 gr lactic acid/gr glucose. Similar calculation may be made for C5/C6 fermentations, in which in addition to glucose also pentoses are included e.g. xylose and/or arabinose. For other fermentation products a similar calculation may be made.

SSF mode

In SSF operation the product concentration is 25g * Yps g/L /L or more, 30 * Yps g/L or more, 35g * Yps /L or more, 40 * Yps g/L or more, 45 * Yps g/L or more, 50 * Yps g/L or more, 55 * Yps g/L or more, 60 * Yps g/L or more, 65 * Yps g/L or more, 70 * Yps g/L or more , 75 * Yps g/L or more, 80 * Yps g/L or more, 85 * Yps g/L or more, 90 * Yps g/L or more, 95 * Yps g/L or more, 100 * Yps g/L or more, 110 * Yps g/L or more, 120g/L * Yps or more or may e.g. be 25 * Yps g/L-250 * Yps g/L, 30 * Yps gl/L-200 * Yps g/L, 40 * Yps g/L-200 * Yps g/L, 50 * Yps g/L-200 * Yps g/L, 60 * Yps g/L-200 * Yps g/L, 70 * Yps g/L-200 * Yps g/L, 80 * Yps g/L-200 * Yps g/L, 90 * Yps g/L , 80 * Yps g/L-200 * Yps g/L. Accordingly, the invention provides a method for the preparation of a fermentation product, which method comprises:

a. degrading lignocellulose using a method as described herein; and
b. fermenting the resulting material,

thereby to prepare a fermentation product.

Fermentation product

The fermentation product of the invention may be any useful product. In one embodiment, it is a product selected from the group consisting of ethanol, n-butanol, 2-butanol, isobutanol, lactic

acid, 3-hydroxy-propionic acid, acrylic acid, acetic acid, succinic acid, fumaric acid, malic acid, itaconic acid, maleic acid, citric acid, adipic acid, an amino acid, such as lysine, methionine, tryptophan, threonine, and aspartic acid, 1,3-propane-diol, ethylene, glycerol, a β-lactam antibiotic and a cephalosporin, vitamins, pharmaceuticals, animal feed supplements, specialty chemicals, chemical feedstocks, plastics, solvents, fuels, including biofuels and biogas or organic polymers, and an industrial enzyme, such as a protease, a cellulase, an amylase, a glucanase, a lactase, a lipase, a lyase, an oxidoreductases, a transferase or a xylanase.

Recovery of the fermentation product

For the recovery of the fermentation product existing technologies are used. For different fermentation products different recovery processes are appropriate. Existing methods of recovering ethanol from aqueous mixtures commonly use fractionation and adsorption techniques. For example, a beer still can be used to process a fermented product, which contains ethanol in an aqueous mixture, to produce an enriched ethanol-containing mixture that is then subjected to fractionation (e.g., fractional distillation or other like techniques). Next, the fractions containing the highest concentrations of ethanol can be passed through an adsorber to remove most, if not all, of the remaining water from the ethanol. In an embodiment in addition to the recovery of fermentation product, the yeast may be recycled. The following non-limiting examples are intended to be purely illustrative.

EXAMPLES

Example 1: Expression of *DAN1p-prk* in a Rubisco expressing strain from the CEN.PK lineage

Maintenance of strains

Strains originating from the CEN.PK lineage were used in this example (Table 6) [1,2]. Cultures were propagated in synthetic medium [3], supplemented with 20 g L^{-1} glucose. Propagation of *E. coli DH5a* stock cultures was performed in LB medium (5 g L^{-1} Bacto yeast extract, 10 g L^{-1} Bacto tryptone, 5 g L^{-1} NaCl), supplemented with 100 µg mL^{-1} ampicillin or 50 µg mL^{-1} kanamycin. Frozen stocks were prepared by addition of glycerol (30 % v/v final concentration) to growing cultures and subsequent storage at -80°C.

Table 6. *S. cerevisiae* strains used in this example.

Strain name	Relevant Genotype	Origin
CEN.PK113-5D	*MATa ura3-52*	[1,2]
IMX585	*MATa URA3 can1::cas9-natNT2*	[4]
IMX581	*MATa ura3-52 can1::cas9-natNT2*	[4]
IMX675	*MATa ura3-52 can1::cas9-natNT2 gpd1Δ gpd2Δ*	Examples
IME324	*MATa ura3-52 can1::cas9-natNT2* p426-*TEF* (empty)	Examples

IMX765	*MATa ura3-52 can1::cas9-natNT2 sga1:: cbbM* (9 copies), *groES, groEL*	Examples
IMX773	*MATa ura3-52 can1::cas9-natNT2 sga1:: cbbM* (9 copies), *groES, groEL X-2::YEN1p-prk* pUDR164	Examples
IMX774	*MATa ura3-52 can1::cas9-natNT2 sga1:: cbbM* (9 copies), *groES, groEL X-2::DAN1p-prk* pUDR164	Examples
IMX1443	IMX774 *gpd2::RPE1, TKL1, TAL1, TAL2, RKI1, NQM1*	Examples

Plasmid and cassette construction

A list of plasmids used in this example is given in Table 7. CRISPR/Cas9 genome editing was used to perform genetic modifications in all constructed strains [5]. Unique CRISPR/Cas9 sequences targeting *GPD1*, *GPD2*, *SGA1* or *X-2* were identified using a publicly available list [5]. For markerless deletion of *GPD1* and *GPD2*, plasmid pUDR240 was constructed. The plasmid backbone was PCR amplified using primer combination 5793-5793 (double-binding) and pROS10 as template. The plasmid insert, containing the expression cassettes coding for the unique 20-bp gRNA sequences targeting *GPD1* and *GPD2*, was obtained using primer combination 6965-6966 and plasmid pROS10 as template. For markerless genomic integration of gene cassettes, plasmids expressing unique gRNAs targeting the *SGA1* locus or the intergenic region *X-2* [6] were constructed. The plasmid backbones of puDR119 and pURD164 were obtained by PCR amplification using the primer combination 5792-5980 and plasmids pMEL11 and pMEL10, respectively, as templates. The plasmid inserts of pUDR119 and pUDR164, containing the expression cassettes coding for the unique 20-bp gRNA sequences targeting *SGA1* and *X-2* respectively, were obtained by PCR amplification using the primer combinations 5979-7023 for *SGA1* and 5979-7374 for *X-2* and plasmids pMEL11 and pMEL10, respectively, as templates. Phusion® Hot Start II High Fidelity DNA Polymerase (Thermo Scientific, Waltham, MA, USA) was used for construction of plasmids and expression cassettes in all cases, according to the manufacturer's guidelines. The assembly of plasmids pUDR119, pUDR164 and pUDR240 was performed in vitro using the Gibson Assembly® Cloning kit (New England Biolabs, Ipswich, MA, USA) following the supplier's guidelines. The assembly was enabled by homologous sequences present at the 5' and 3' ends of the PCR-amplified plasmid backbones and inserts. In each case, 1 ul of the Gibson-assembly mix was used for *E. coli DH5a* transformation by electroporation, performed in a Gene PulserXcell Electroporation System (Biorad, Hercules, CA, USA). Correct assembly of plasmids was confirmed by diagnostic PCR (Dreamtaq®, Thermo Scientific) or restriction digestion. The constructed plasmids pUDR119, pUDR164 and pUDR240 were isolated from transformed *E. coli* cultures using a Sigma GenElute Plasmid kit (Sigma-Aldrich, St. Louis, MO, USA) and used for transformation of *S. cerevisiae*.

A yeast expression cassette of *cbbM* was obtained by PCR amplification using plasmid pBTWW002 as template and primer combination 7549-7550. The resulting fragment was ligated to a pJET/1.2 blunt vector (Thermo-Scientific) following the supplier's protocol and cloned to *E. coli*. The resulting plasmid was used as PCR template to generate integration *cbbM* cassettes, using primer combinations 7074-7075 (integration at the *SGA1* locus along with *prk*, *groES*, *groEL*), 7548-6285, 6280-6273, 6281-6270, 6282-6271, 6284-6272, 6283-6275, 6287-6276, 6288-6277, 6289-

7075 (multiple-*cbbM*-copy integration at the *SGA1* locus). The expression cassettes of *cbbM* were genetically identical, except for different overhangs present at the 5' and 3' ends of the fragments to allow for *in vivo* homologous recombination. Yeast expression cassettes of *groEL* and *groES* were obtained using plasmids pUD232 and pUD233 as templates and primer combinations 7076-7077 and 7078-7079 respectively. The genomic sequences corresponding to the constitutive promoters of *LYS1*, *UBC6*, *YEN1* and the anaerobically active promoter of *DAN1* [7] were obtained by PCR amplification with primer combinations 7082-7083, 7292-7294, 7293-7295 and 7930-7931 respectively, using genomic DNA of IMX585 as template. In the case of integration at the *X-2* locus, primer combination 7933-7295 was used for amplification of the *YEN1* promoter region. The terminator of *PGK1* was obtained by PCR amplification with genomic DNA of IMX585 as template using primer combinations 7084-7085 (integration at the *SGA1* locus along with *cbbM*, *groES*, *groEL*) and 7084-7934 (individual integration of *prk* at the *X-2* locus). The ORF of *prk* was obtained by PCR amplification using primer combinations 7080-7081 (*LYS1p* cassette construction), 7296-7081 (*UBC6p* cassette construction), 7297-7081 (*YEN1p* cassette construction), 7932-7081 (*DAN1p* cassette construction) and plasmid pUDE046 as template. The various primer combinations resulted in *prk*-ORF fragments with homologous overhangs to the different promoter sequences and the terminator of *PGK1*. The expression cassettes *LYS1p-prk-PGK1t*, *UBC6p-prk-PGK1t* and *YEN1p-prk-PGK1t* were assembled *in vitro* using fusion PCR by combining the respective promoter/*prk*/*PGK1t* fragments as templates and primer combinations 7082-7085, 7292-7085 and 7293-7085 respectively, in the case of aimed integration at the *SGA1* locus of strain IMX675 (along with a KlURA3 fragment, *cbbM* cassete and *groEL*, *groES* chaperones. When *prk* cassettes were integrated individually (integration at the *X-2* locus), the complete expression cassettes (*YEN1p-prk-PGK1t* and *DAN1p-prk-PGK1t*) was assembled by *in vivo* homologous recombination after transformation to yeast and correct assembly was verified by diagnostic PCR. Primer combination 7086-7087 was used to obtain a *KlURA3* fragment using plasmid pUG72 as template.

Table 7. Plasmids used in this example.

Name	Characteristics	Origin
p426-*TEF*	2 µm ori, *URA3*, empty vector	[8]
pUG72	*KlURA3,* PCR template	[9]
pMEL10	2 µm ori, *URA3*, *SNR52*p-gRNA.*CAN1-SUP4*t	[4]
pMEL11	2 µm ori, *amdS*, *SNR52*p-gRNA.*CAN1-SUP4*t	[4]
pROS10	URA3, gRNA.*CAN1*-2 µm ori-gRNA.*ADE2*	[4]
pUD232	Delivery vector, *TEF1*p-*groEL-ACT1*t	[10]
pUD233	Delivery vector, *TPI1*p-*groES-PGI1*t	[10]
pUDE046	2 µm ori, *GAL1p-prk-CYC1t*	[10]
pBTWW002	2 µm ori, *URA3*, *TDH3p-cbbM-CYC1t*	[10]
pUDR119	2 µm ori, *amdS*, *SNR52*p-gRNA.*SGA1-SUP4*t	This example

pUDR164	2 µm ori, *URA3*, *SNR52*p-gRNA.*X-2*-*SUP4*t	This example
pUDR240	*URA3*, gRNA.*GPD1*-2 µm ori-gRNA.*GPD2*	This example

Strain construction

The lithium-acetate transformation protocol was used for yeast transformations [11]. Transformation mixtures were plated on synthetic medium agar plates [3] (2 % Bacto Agar, BD, Franklin Lakes, NJ, USA), supplemented with 20 g L^{-1} glucose in the case of transformations performed with puDR164 and pUDR240. In transformations performed with plasmid pUDR119, the agar plates were prepared as described previously [12]. For the construction of strain IMX765 uracil was additionally supplemented to the agar plates (150 mg L^{-1}) (Sigma-Aldrich). Confirmation of the desired genotypes in each case was performed by diagnostic colony PCR. Recycling of pUDR240 was performed using 5-fluoorotic acid (Zymo Research, Irvine, CA, USA) counter-selection, following the supplier's guidelines. Recycling of pUDR119 was performed as described previously [12]. Strain IMX675 was constructed by co-transformation of the double-gRNA-expressing, *GPD1*/*GPD2* targeting plasmid pUDR240 and the repair-oligonucleotide combinations 6967-6968 and 6969-6970 to the Cas9-expressing strain IMX581 (after plasmid recycling from the correct mutant). The expression cassettes *LYS1p-prk-PGK1t*, *UBC6p-prk-PGK1t* and *YEN1p-prk-PGK1t* were respectively co-transformed to strain IMX675 along with a single copy of the *cbbM* cassette, *groEL*, *groES*, the *URA3* fragment and the gRNA-expressing, *SGA1*-targeting plasmid pUDR119. Overhangs present at the 5' and 3' ends of the molecules were designed to allow for complete assembly of the pathways in the *SGA1* locus. Strain IMX765 was obtained by co-transformation of pUDR119, 9 copies of the expression cassette of *cbbM* and the expression cassettes of *groEL* and *groES* to IMX581 (after plasmid recycling from the correct mutant). Overhangs present at the 5' and 3' ends of the molecules allowed for in vivo assembly of the entire construct (11 fragments) and integration in the *SGA1* locus. Strain IMX774 was obtained by transformation of strain IMX765 with the gRNA-expressing, *X-2* targeting plasmid pUDR164 and the *DAN1p*, *prk ORF*, *PGK1t* fragments which were assembled *in vivo* into the complete construct and subsequently integrated in the *X-2* locus. Strain IMX773 was obtained by transformation of strains IMX765 with pUDR164 and the *YEN1p, prk ORF, PGK1t* fragments which were similarly assembled *in vivo* and subsequently integrated in the *X-2* locus. The control strain IME324 was obtained by transformation of IMX581 with the empty vector p426-*TEF*.

Cultivation media and analytical methods

Physiological characterization of *S. cerevisiae* strains was performed in anaerobic batch cultivations in 2-L bioreactors (Applikon, Delft, The Netherlands), with 1-L working volume. Salt solutions were sterilized by autoclaving at 120°C for 20 min. Glucose solutions were autoclaved separately at 110°C for 20 min and subsequently added to the sterile salt solutions. All fermentations were performed in synthetic medium [3] (20 g L^{-1} glucose), supplemented with sterile solutions of the anaerobic growth factors ergosterol (10 mg L^{-1}) and Tween 80 (420 mg L^{-1}), as well

as with 0.2 g L^{-1} sterile antifoam C (Sigma-Aldrich). Anaerobic conditions were maintained by sparging of a gas mixture of N_2 / CO_2 (90%/10%, <10 ppm oxygen) at a rate of 0.5 L min^{-1} and culture pH was maintained at 5 by automatic addition of 2M KOH. All cultivations were performed at a stirrer speed of 800 rpm and at a temperature of 30°C. Oxygen diffusion in the bioreactors was minimized by equipping them with Norprene tubing and Viton O-rings. Pre-culture shake flask cultivations were performed aerobically in 500-mL flasks containing 100 mL synthetic medium (20 g L^{-1} glucose). Initial flask pH was adjusted to 6 by addition of KOH. Cultures were grown at 30°C and shaken at 200 rpm. In each case, pre-culture flasks were inoculated from frozen *S. cerevisiae* stock cultures. After incubation for 8-12 h, cultures from these flasks were used to inoculate fresh pre-culture flasks for bioreactor inoculum propagation. Bioreactors were inoculated to a starting OD660 of ca. 0.2. Off-gas analysis, biomass dry weight measurements, HPLC analysis of culture supernatants and correction for ethanol evaporation in bioreactor experiments were performed as described previously [13]. Optical density was determined at 660 nm, using a Libra S11 spectrophotometer (Biochrom, Cambridge, UK). Yields of products in each cultivation were calculated from samples taken at mid-exponential phase (minimum of five samples), as described previously [14]. For the calculation of the degree of reduction (electron) balances in performed fermentations, reported degree of reduction values for biomass, CO_2, NH_4^+ and extracellular metabolites (glucose, ethanol, glycerol, succinate, pyruvate, lactate, acetate) were used [15]. For determination of *in vitro* enzymatic activity of PRK, cells from exponentially growing, anaerobic shake-flask cultures in synthetic-medium were harvested and cell-free extracts were prepared as previously described [16]. The harvesting and sonification buffer contained 100 mM Tris-HCl, 20 mM $MgCl_2 \cdot 6H_2O$ and 5mM DTT (pH 8.2). The PRK assay contained 50 mM Tris-HCl (pH 8.2), 40 mM KCl, 10 mM $MgCl_2 \cdot 6H_2O$, 0.15 mM NADH, 1 mM ATP, 3 mM phosphoenolpyruvate, 1 mM 1,4-dithiothreitol, 5 U of pyruvate kinase (EC 2.7.1.40), 6 U of L-lactate dehydrogenase (EC 1.1.1.27) and 30, 50 or 100 µl cell-free extract in 1 ml total volume. Reactions were started by addition of D-ribulose-5-phosphate (2.5 mM final concentration) and PRK activity was measured at 30°C on a Hitachi 100-60 spectrophotometer by monitoring NADH oxidation at 340 nm over time. Protein content determination in cell-free extracts was performed as previously described (Lowrey protein assay) [17].

Physiological characterization of strains

Expression of *cbbM* and *prk* in *S. cerevisiae* has previously been shown to result in decreased formation of the by-product glycerol under anaerobic conditions that are relevant for industrial ethanol production [10]. However, in this previous research, *prk* was expressed under the control of the galactose-inducible *GAL1*-promoter. The requirement for the presence of galactose and low levels of glucose are a drawback of the use of this promoter. This example investigates the expression of *prk* under the control of different promoters in a strain that co-expresses Rubisco. Expression cassettes were constructed based on three constitutive promoter sequences of varied expression strengths, with *LYS1p* being the strongest and *YEN1p* being the weakest, as well as a weak promoter that is only active in anaerobic conditions (*DAN1p*) [7]. Initially, transformations were

performed with *prk* cassettes under the control of *LYS1p*, *UBC6p* and *YEN1p* and copies of *cbbM*, *groEL*, *groES*, according to [10]. It was not possible to obtain correct mutant colonies in the case where the *LYS1p-prk-PGK1t* cassette was used. In the case where *UBC6p-prk-PGK1t* was used, colonies were obtained but growth was inconsistent and severely impacted, and a stable strain was not obtained, see Table 9. However, in the cases where *YEN1p-prk-PGK1t* (weakest constitutive promoter tested) it was possible to obtain correct mutant colonies (Table 9). These results could be an indication that high & medium-constitutive expression of *prk* in *S. cerevisiae* under aerobic conditions (biomass propagation phase) results in inhibitory effects and are in agreement to data available in literature [10]. Based on the above, a multi-copy, *cbbM*-expressing strain was constructed with *prk* under the control of *YEN1p* (IMX773, low-*prk*-expression strain). Additionally, a strain was constructed in which a *DAN1p-prk-PGK1t* cassette was used instead. The promoter of *DAN1* (or any other similar promoter) is of particular interest, because it is active in the process conditions of bioethanol production (anaerobic conditions in this case) and does not require the use of specific carbon sources (like *GAL1p* does) or any other change to the commonly used production process. This promoter (or any other similar promoter) should alleviate the toxicity of *prk*-expression under aerobic conditions (no transcription). The strain was designated IMX774 (high-*prk*-expression strain). Figure 1 shows the PRK activity in cell-free extracts of exponentially growing shake-flask cultures on synthetic medium containing 20 g L^{-1} glucose. Left: IME324, Right: IMX774. To determine whether the promoter of *DAN1* could drive the expression of *prk* in *S. cerevisiae*, PRK enzymatic activity determination was performed *in vitro*, using cell-free extracts of anaerobically-grown cultures of strains IME324 (reference) and IMX774 (9**cbbM*, *DAN1p-prk*). PRK activity in IMX774 was ca. 0.8 µmol (mg protein)$^{-1}$ min^{-1} (Fig 1). To quantitatively analyze the impact of expression of the introduced Rubisco pathway, strains IME324 (reference), IMX773 (9**cbbM*, *YEN1p-prk*) and IMX774 (9**cbbM*, *DAN1p-prk*) were compared in glucose-grown, anaerobic batch cultures in bioreactors. The engineered strains IMX773 and IMX774 grew at 82% and 61%, respectively, of the specific growth rate of the reference strain IME324. The glycerol yield on glucose of strains IMX773 and IMX774 was 0.098 and 0.058 g/g, respectively. This corresponds to 96% of the glycerol yield for IMX773 compared to IME324 indicating a limited impact of the Rubisco pathway in a design in which *prk* is expressed under control of the weak constitutive promoter *YEN1p* (Table 8). In contrast, for IMX774 a 43% decrease in glycerol yield compared to the reference strain IME324 was observed (Table 8, Fig 2). Furthermore, the ratio of glycerol production per g biomass formed of strain IMX774 decreased by 41% compared to the reference strain, whereas for IMX773 this was maintained to 97% of IME324 levels indicating a limited effect. A decrease of glycerol production can be expected when NAD^+ is being regenerated via the Rubisco pathway. Therefore, these findings are in agreement with results reported elsewhere on galactose-grown Rubisco-expressing strains [10] and show that the effect of the pathway on galactose-grown cultures can be replicated in glucose-grown ones, when *prk* is expressed under the control of the promoter of *DAN1* but not under control of a weak constitutive promoter such as *YEN1p*. Even more, *prk* expression in IMX765 under control of stronger constitutive promoters did

not even yield viable colonies, indicating cellular toxicity as a result of transformation of these *prk* specific expression cassettes. Strains IME324 and IMX774 showed an ethanol yield on glucose of 0.356 and 0.409 g/g respectively (corrected for evaporation). This means that the combination of the decrease in glycerol production, CO_2 fixation via the Rubisco pathway and decrease in biomass yield of the engineered, Rubisco-expressing strain IMX774 (Fig 2) resulted in a ca. 14% increase in the ethanol yield on glucose in the experiments performed in this example.

Table 8. Maximum specific growth rate (μ), yields (Y) of glycerol, biomass and ethanol on glucose and the ratio of glycerol formation to biomass formation in anaerobic bioreactor batch cultures of *S. cerevisiae* strains IME324, IMX773 and IM3X774.

Strain	**IME324**		**IMX773**		**IMX774**	
Relevant genotype	reference		*9*cbbM*, *YEN1p-prk*, *groES*, *groEL*		*9*cbbM*, *DAN1p-prk*, *groES*, *groEL*	
μ (h^{-1})	0.33 ± 0.01		0.27 ± 0.01		0.20 ± 0.03	
Y glycerol/glucose (g g^{-1})	0.102 ± 0.001		0.098 ± 0.000		0.058 ± 0.005	
Y biomass/glucose (g_x g^{-1})	0.091 ± 0.000		0.089 ± 0.000		0.087 ± 0.007	
Y EtOH/glucose (g g^{-1})	0.356 ± 0.004		0.385 ± 0.002		0.409 ± 0.001	
Ratio glycerol produced/biomass (mmol g_x^{-1})	12.262 ± 0.122		11.879 ± 0.008		7.272 ± 0.115	
*Degree of reduction (γ_D) balance	0.95 ± 0.01	0.95 ± 0.01	0.99 ± 0.01	X1.01 ± 0.01	1.03 ± 0.01	1.01 ± 0.01

*Degree of reduction balances are given for each individual experiment of independent duplicate cultures. Balances calculated over the exponential sampling phase. Averages and standard deviation values of the balances over the sampling points are given.

Cultures were grown on synthetic medium containing 20 g L^{-1} glucose (pH 5) and sparged with a gas mixture of N_2/CO_2 (90%/10%). Yields and ratios were calculated from the exponential growth phase. The ethanol yield on glucose was corrected for evaporation. Values represent average and mean deviation of data from independent duplicate cultures.

Table 9. Aerobic growth properties and glycerol reduction

Promoter fro the PRK	Strength promoter	Conditions active	Aerobic propagation	Glycerol reduction
LYS1p	strong	Constitutive	No colonies formed	n.a.
UBC6p	medium	Constitutive	No colonies formed	n.a.
YEN1p	Weak	Constitutive	Growth	-3%
DAN1p	weak	Anaerobic	Growth	-41%

Example 2: *S. oleracea prk* protein expressed exclusively under anaerobic conditions in IMX774

Shake flask cultivation strains

IME324 and IMX774 were cultivated in duplicate in mineral medium (according to Luttik et al., 2000) supplemented with 20 g L^{-1} glucose and 0.05 g L^{-1} uracil in shake flasks under aerobic and anaerobic conditions. After overnight aerobic propagation on YePhD, 75 mg L^{-1} of yeast was inoculated to the above described medium in either a 100 mL shake flask filled with 25 mL medium closed afterwards with a cotton plug to recreate aerobic cultivation conditions, or a 25 mL shake flask filled with 25 mL medium (leaving limited head space for aeration) closed afterwards with a water lock to recreate conditions which shortly after inoculation and closing off become anaerobic in the vessel. After 24 hours of cultivation at 32°C and 250 rpm (for aerobic cultures) and 150 rpm (for anaerobic cultures), 10 OD600 units/ mL of cells were harvested from each of the eight shake flasks by centrifugation and cells were washed with ice cold demineralized water. Cell pellets were stored at -80 °C for further processing.

Protein extraction and proteomics

Frozen cells were lysed using mechanical based disruption approach via VK05 glass beads and Precellys 24 homogeniser (Bertin Technologies) in the environment of cold Methanol (Sigma). Protein concentration of the disrupted cell suspension was measured using the Qubit 2.0 fluorometer (Invitrogen, Life Technologies). Two hundred fifty ug of total protein was taken from each methanol suspension and 10 ug BSA was spiked to all the samples for quality control. Proteins were extracted from the disrupted cell suspension using chloroform (Sigma) and 20% TCA (Sigma). The obtained protein pellet was dissolved in 100 mM NH4HCO3 buffer at pH 7 (Sigma) to a final concentration of 0.5 ug/uL. Proteins were reduced through the addition of 5 ul of 500 mM Tris(2-carboxyethyl)phosphine hydrochloride solution (TCEP, sigma) and incubated at 55 °C for 30 minutes in a thermocycler to facilitate disulfide reduction. Alkylation was performed through the addition of 5 ul of 550mM iodoactamide and incubated at 25°C in the dark for 30 minutes. Proteolysis was carried out overnight in a thermomixer at 37°C with Trypsin Gold (Promega) at an enzyme to substrate ration of 1:25, which specifically cleaves C-terminally of Lysine and Arginine. Tryptic digests were analyzed on an Ultimate3000 coupled to a QExactive Plus (Thermo Scientific). Gradient elution of peptide was performed on a C18 (Acquity UPLC CSH C18 Column, 130Å, 1.7 µm, 2.1 mm X 100 mm). Twenty uL of injected peptides were separated with a gradient of mobile phase A (99.9% water and 0.1% formic acid; VWR) to 20% B (99.9% acetonitrile and 0.1% formic acid; VWR) over 20 minutes, and to 50% B over 30 minutes, for a final length of 60 minutes. Data acquisition on the qExactive MS was carried out using a data-dependent method. The top 15 precursors were selected for tandem-MS/MS (MS2) analysis after HCD fragmentation. Full MS scans covering the mass range of 400 to 1600 were acquired at a resolution of 70,000 (at m/z 200), with a maximum fill time of 75 milliseconds, and an automatic gain control (AGC) target value of 3e6. MS2 scans were acquired at a resolution of 17,500 (at m/z 200), with a maximum fill time of

75 milliseconds, and an AGC target value of 1e5. An isolation window of 2.0 m/z with a fixed first mass of 110.0 m/z was applied in all experiments. HCD fragmentation was induced with a normalized collision energy of (NCE) of 27 for all peptides. Charge state exclusion was set to ignore unassigned 1 charge. Isotope exclusion was enabled and peptide match was preferred. All LC-MS/MS results were searched against the *S. cerevisiae* protein database to which the amino acid sequences of the heterologous introduced enzymes were manually added, using Sequest HT on Proteome Discover 1.4 Sequest HT (Thermo Fisher Scientific, San Jose, CA, USA). The cleavage preference of trypsin was used, allowing up to 2 missed cleavages (C-Term K/R restrict P). Dynamic modifications were set to carbamidomethyl (C), deamidation (N/Q) and oxidation (M). Precursor mass tolerance was set to 10 ppm and fragment mass tolerance 0.6 Da. Following peptide identification, their q-values were calculated based on target decoy approach with a 1% false discovery rate (FDR) and filtered in the Percolator.

Analysis results

Around 1000 unique proteins were identified in each sample. To quantify the amount of *S. oleracea* PRK, among all samples, most abundant unique peptides were selected. Peak areas of all peptides were summed up. By normalizing against beta-actin (*S. cerevisiae* Act1p) amount determined by LC-MS/MS, this value finally gave an indication concerning the protein amount among all samples. As shown in Figure 3, under both aerobic and anerobic conditions no unique *prk* peptides are detected for negative control strain IME324 which does not express *prk*. For IMX774 prk peptides are detected solely under anaerobic conditions indicating that the *DAN1p* is only inducing *prk* expression.under anaerobic conditions.

Example 3: Expression of *S. oleracea prk* under control of more anaerobically upregulated promoters results in glycerol reduction in IMX765

Introduction of *S. oleracea prk* expressed under control of the anaerobically upregulated *DAN1* promoter was found to be beneficial to the reduction of glycerol byproduct formation and ethanol yield improvement (Example 1). In constrast to *DAN1p*, *prk* introduction with other, constitutively active promoters (*UBC6, LYS1, YEN1*) placed upstream of *S. oleracea* prk did not yield viable, correct transformants (*UBC6, LYS1*) or did not show an impact on glycerol reduction (*YEN1p*) by implementing the Rubisco pathway. This led us to the conclusion that anaerobically induced / derepressed *prk* is beneficial to Rubisco pathway flux and results in ethanol yield improvement by reduction of glycerol. *DAN1* is regulated by repressor *ROX1* which is absent in anaerobic conditions thereby relieving repression on the *DAN1* promoter. More such promoters are regulated by *ROX1* which display a differential expression pattern being expressed at a higher level under anaerobic conditions than under aerobic conditions (Kwast et al., 1998). In this example, several of these promoters were placed upstream of *S. oleracea prk* and introduced to strain IMX765 after which resulting strains were subjected to anaerobic growth cultivations. In Table 10, the promoters tested in Example 3 are listed.

Table 10. Anaerobic PRK promotors

Promoter	SEQ ID NO	Length (bp)
Sc_DAN1_long	93	1000
Sc_DIP5	94	967
Sc_TIR3	95	1000
Sc_TIR2	96	1000
Sc_HEM13	97	1000
Sc_YHK8	98	1000
Sc_FET4	99	1000
Sc_TIR4	100	1000
Sc_AAC3	101	600
Sc_ANB1	91	601

Material and Methods

Expression cassette construction

The promoters (listed in Table 10; SEQ ID NOs: 91 and 93-101) and terminator, namely Sc_PGK1.ter (SEQ ID NO: 102) sequences were synthesized at DNA2.0 (Menlo Park, CA 94025, USA). The *S. oleracea prk* ORF sequence (Sole_PRK.orf) was obtained by PCR amplification using primer combinations DBC-15631 (SEQ ID NO: 103) and DBC-15632 (SEQ ID NO: 104) using pUDE046 as template. The promoter, ORF and terminator sequences were recombined by using the Golden Gate technology, as described by Engler et al (2011) and references therein. The expression cassettes were cloned into a standard subcloning vector. The promoters listed in Table 10 were ligated to Sole_PRK ORF and Sc_PGK1 terminator resulting in expression cassettes listed in Table 11.

Table 11: List of expression (promoter-ORF-terminator) cassettes resulting from Golden Gate Cloning of containing promoter variants, *S. oleracea* PRK ORF and S. cerevisiae PGK1 terminator; on the 5' end and 3'end connector sequences compatable to neighbouring pathway brick are listed.

Cassette	5'connector	promoter	ORF	Terminator	3'-connector
cDAN1	C	Sc_DAN1	Sole_PRK	Sc_PGK1	D
cDIP5	C	Sc_DIP5	Sole_PRK	Sc_PGK1	D
cTIR3	C	Sc_TIR3	Sole_PRK	Sc_PGK1	D
cTIR2	C	Sc_TIR2	Sole_PRK	Sc_PGK1	D
cHEM13	C	Sc_HEM13	Sole_PRK	Sc_PGK1	D
cYHK8	C	Sc_YHK8	Sole_PRK	Sc_PGK1	D
cFET4	C	Sc_FET4	Sole_PRK	Sc_PGK1	D
cTIR4	C	Sc_TIR4	Sole_PRK	Sc_PGK1	D

cAAC3	C	Sc_AAC3	Sole_PRK	Sc_PGK1	D
cANB1	C	Sc_ANB1	Sole_PRK	Sc_PGK1	D

Strain construction

Approach

The followed strain construction approach is described in patent application PCT/EP2013/056623 and PCT/EP2016/050136. PCT/EP2013/056623 describes the techniques enabling the construction of expression cassettes from various genes of interest in such a way, that these cassettes are combined into a pathway and integrated in a specific locus of the yeast genome upon transformation of this yeast. PCT/EP2016/050136 describes the use of a CRISPR-Cas9 system for integration of expression cassettes into the genome of a host cell, in this case *S. cerevisiae*. In the construction of IMX765 a *S. pyogenes* Cas9 expression cassette was already integrated at the *CAN1* locus. Upon introduction of an *in vivo* assembled gRNA-expressing plasmid and repair DNA fragments the intended modifications were made. Firstly, an integration site in the yeast genome was selected. DNA fragments of approximately 500 bp of the up- and downstream parts of the integration locus were amplified by PCR using primers introducing connectors to the generated PCR products. These connectors (50 bp in size) allow for correct *in vivo* recombination of the pathway upon transformation in yeast. Secondly, the genes of interest, are amplified by PCR, incorporating a different connector (compatible with the connector on the of the neighbouring biobrick) at each flank. Upon transformation of yeast cells with the DNA fragments, *in vivo* recombination and integration into the genome takes place at the desired location. This technique facilitates parallel testing of multiple genetic designs, as one or more genes from the pathway can be replaced with (an)other gene(s) or genetic element(s), as long as that the connectors that allow for homologous recombination remain constant and compatible with the preceding and following biobrick in the design (patent application PCT/EP2013/056623).

gRNA expression plasmid

Integration site: the expression cassettes were targeted at the INT1 locus. The INT1 integration site is a non-coding region between *NTR1* (YOR071c) and *GYP1* (YOR070c) located on chromosome XV of *S. cerevisiae*. The guide sequence to target INT1 was designed with a gRNA designer tool (https://www.dna20.com/eCommerce/cas9/input). The gRNA expression cassette (as described by DiCarlo et al., Nucleic Acids Res. 2013; pp.1-8) was ordered as synthetic DNA cassette (gBLOCK) at Integrated DNA Technologies (Leuven, Belgium) (INT1 gBLOCK; SEQ ID NO: 105). In *vivo* assembly of the gRNA expression plasmid is then completed by co-transforming a linear fragment derived from yeast vector pRN599. pRN599 is a multi-copy yeast shuttling vector that contains a functional kanMX marker cassette conferring resistance against G418. The backbone of this plasmid is based on pRS305 (Sikorski and Hieter, Genetics 1989, vol. 122, pp.19-

27), including a functional 2 micron ORI sequence and a functional kanMX marker cassette (SEQ ID NO: 106).

Transformation of IMX765 with specified DNA fragments upon assembly comprising anaerobic promoter PRK cassette

Strain IMX765 was transformed with the following fragments resulting in the assembly of anaerobic promoter PRK as depicted in Figure 4:

1) a PCR fragment (5'-INT1) generated with primers BoZ-783 (SEQ ID NO: 107) and DBC-19944 (SEQ ID NO: 108) with genomic DNA of strain CEN.PK113-7D as template;
2) a PCR fragment (Anaerobic.pro-PRK) generated with primers DBC-5799 (SEQ ID NO: 109 and DBC-5800 (SEQ ID NO: 110) using either one of the plasmids listed in Table 7 as template;
3) a PCR fragment generated with primers DBC-19947 (SEQ ID NO: 111) and DBC-19949 (SEQ ID NO.112) using genomic DNA of strain CEN.PK113-7D as template; this PCR resulted in the 1.2 kb marker cassette conD-*URA3*-conE (*URA3* marker flanked by connectors D and E).
4) a PCR fragment (3'-INT1) generated with primers DBC-19946 (SEQ ID NO: 113) and BoZ-788 (SEQ ID NO: 114) using genomic DNA of strain CEN.PK113-7D as template;
5) a PCR fragment (BB-599) generated with primers DBC-13775 (SEQ ID NO: 115) and DBC-13776 (SEQ ID NO: 116) using pRN599 (SEQ ID NO: 106) as template;
5) a PCR fragment (gRNA-INT1) generated with primers DBC-13773 (SEQ ID NO: 117) and DBC-13774 (SEQ ID NO: 118) using INT1 gRNA (SEQ ID NO: 105) as template;

Transformants were selected on mineral medium (according to recipe Luttik et al., 2000, Journal of Bacteriology 182, 24: 501-517) supplemented with 1.5% bactoagar supplemented with 20 g L^{-1} glucose and 0.2 mg G418 mL^{-1}. Diagnostic PCR was performed to confirm the correct assembly and integration at the INT1 locus of the PRK expression cassettes.

Microtiterplate batch fermentation experiment

Six to nine correct transformants per transformation design and controls strains IME324, IMX765 and IMX774 in nine fold were inoculated to 240µl mineral medium (according to Luttik *et al.*, 2000) supplemented with 20 g L^{-1} glucose and 0.05 g L^{-1} uracil in microtiterplate format. Inoculated microtiterplates were cultivated for 2 days at 30°C. Subsequently, grown cultures on agar medium were transferred to liquid 350µl mineral medium (according to Luttik *et al.*, 2000) supplemented with 20 g L^{-1} glucose and 0.05 g L^{-1} uracil with the aid of a pintool to the second microtiterplate for biomass propagation. The microtiterplates were sealed with a gas permeable seal enabling aerobic cultivation conditions and were incubated at 32°C, shaking at 750 rpm at 80% humidity for 2 days. After 2 days, 5µl grown liquid culture was transferred to 270µl mineral medium (according to Luttik *et al.*, 2000) supplemented with 20 g L^{-1} glucose and 0.05 g L^{-1} uracil to the third microtiterplate for main fermentation. A high volume of 270 µl leaving little head space volume in the well and sealing microtiterplates with aluminium seal allowed for screening mainly under

anaerobic conditions. Cultures were grown for 48 hours at 32°C, 250 rpm at 80% humidity. After 48 hours. samples were taken in order to measure residual glucose, glycerol and ethanol.

Analysis of glucose, ethanol, and glycerol

For the quantification of glucose, ethanol and glycerol, 150 µl of the supernatant sample was transferred accurately into a suitable vial. Subsequently 100 µl internal standard solution, containing maleic acid (20 g/l), EDTA (40 g/l), DSS (4,4-dimethyl-4-silapentane-1-sulfonic acid) (0.5 g/L), adjusted to pH 6.40 with NaOH, in D_2O was added. The sample was further diluted with 400 µl D_2O. 1D ^1H NMR spectra of the clear solution were recorded on a Bruker Avance III HD spectrometer, operating at a proton frequency of 500MHz, equipped with a He-cooled cryo probe, using a pulse program with water suppression (ZGCPPR), solvent suppression power of 8Hz, at a temperature of 300K, 90 degree excitation pulse was used and acquisition time of 2.0 seconds and a relaxation delay of 5 seconds. The number of scans was set at 8, dummy scans were not used. The analyte concentrations [in g/L] were calculated based on the following signals (δ relative to DSS):

- Glucose: α-H1 glucose signal (d, 5.21 ppm, 0.38 H, J =4Hz)
- Ethanol: (t, 1.17 ppm, 3H, J =7 Hz)
- Glycerol H1/H3 signals: (dd, 3.55 ppm, 2H, J =7 Hz, 12 Hz)
- The signal used for the standard:
- Maleic acid: (s, 6.05 ppm, 2H)

Results microtiterplate fermentation experiment

Glucose was completely exhausted in the medium by all replicates of transformants and control strains. The average glycerol levels (expressed as arbitrary units) detected for the transformants with different anaerobic promoter – PRK cassettes and control strains IME324, IMX765 and IMX774 are listed in Table 11 and illustrated in Figure 5. In this MTP fermentation set up, again IMX774 displayed reduced glycerol levels (87%, reduction of 13%) compared to reference strain IME324 as shown in Example 1. Transformants of cDAN1, which are a reconstruction of IMX774, showed a similar glycerol reduction (85% of IME324 glycerol levels) as IMX774. All tested anaerobic promoter designs displayed reduced glycerol levels compare to IME324. Also one can appreciate designs with anaerobic promoters resulting in transformants with on average even larger reductions of glycerol levels than cDAN1, such as cTIR3, cHEM13, cAAC3, and cANB1 yielding on average 76%, 68%, 82%, and 80%, respectively, of IME324 glycerol levels. As can be observed in Table 12 and Figure 6 these reductions in glycerol yield were reflected in at least similar, often higher ethanol titers compared to IME324 as seen for cDAN1 (102%), cTIR3 (103%), cHEM13 (104%), cAAC3 (101%), and cANB1 (103%).

Table 12. Average glycerol yields expressed as arbitrary units per g glucose of transformants per design anaerobic promoter design and reference strains IME324, IMX774, IMX765 (transformation control) determined in microtiterplate fermentation

experiment on mineral medium (according to Luttik) supplemented with 20 g L^{-1} glucose and 0.05 g L^{-1} uracil. N.A., not applicable.

Background Strain	Cassette	Promoter PRK	Glycerol (AU)	Standard error	Relative to IME324 (%)	Replicates (n=x)
IME324	N.A.	N.A.	1.59	0.01	100	9
IMX765	URA3	N.A.	1.60	0.02	101	9
IMX774	DAN1 (Ex1)	Sc_DAN1 (Ex 1)	1.39	0.01	87	9
IMX765	cDAN1	Sc_DAN1	1.35	0.03	85	8
IMX765	cDIP5	Sc_DIP5	1.53	0.02	96	6
IMX765	cTIR3	Sc_TIR3	1.21	0.01	76	9
IMX765	cTIR2	Sc_TIR2	1.50	0.01	94	9
IMX765	cHEM13	Sc_HEM13	1.08	0.01	68	9
IMX765	cYHK8	Sc_YHK8	1.52	0.05	96	9
IMX765	cFET4	Sc_FET4	1.46	0.02	92	9
IMX765	cTIR4	Sc_TIR4	1.4	0.02	88	9
IMX765	cAAC3	Sc_AAC3	1.31	0.01	82	9
IMX765	cANB1	Sc_ANB1	1.27	0.02	80	6

Table 13. Average ethanol yields expressed as arbitrary units per g glucose of transformants per design anaerobic promoter design and reference strains IME324, IMX774, IMX765 (transformation control) determined in microtiterplate fermentation experiment on mineral medium (according to Luttik) supplemented with 20 g L^{-1} glucose and 0.05 g L^{-1} uracil. N.A., not applicable.

Background Strain	Cassette	Promoter PRK	Ethanol (AU))	Standard error	Relative to IME324 (%)	Replicates (n=x)
IME324	N.A.	N.A.	7.43	0.03	100	9
IMX765	URA3	N.A.	7.37	0.06	99	9
IMX774	DAN1 (Ex 1)	Sc_DAN1 (Ex 1)	7.46	0.08	100	9
IMX765	cDAN1	Sc_DAN1 (long)	7.56	0.03	102	8
IMX765	cDIP5	Sc_DIP5	7.48	0.05	101	6
IMX765	cTIR3	Sc_TIR3	7.66	0.03	103	9
IMX765	cTIR2	Sc_TIR2	7.52	0.03	101	9
IMX765	cHEM13	Sc_HEM13	7.70	0.03	104	9

IMX765	cYHK8	Sc_YHK8	7.49	0.04	101	9
IMX765	cFET4	Sc_FET4	7.48	0.04	101	9
IMX765	cTIR4	Sc_TIR4	7.51	0.03	101	9
IMX765	cAAC3	Sc_AAC3	7.51	0.03	101	9
IMX765	cANB1	Sc_ANB1	7.67	0.03	103	6

Example 4. Overexpression of PPP genes and deletion of GPD2 gene.

Genes of the non-oxidative branch of the pentose-phosphate pathway (TKL1, NQM1, TKL1, TKL2, RPE1, RKI1) were overexpressed by transforming IMX774 with expression cassettes of the abovementioned genes under control of constitutive promoters as described in pending European Patent Application EP16194660.3. The expression casettes were integrated at the GPD2 locus by co-transforming the guide RNA expression plasmid with GPD2 targeting sequence, thereby abbrogating the coding sequence of gpd2. The resulting strain was named IMX1443. Strain IMX1443 was compared with IME324 and IMX774 in a batch fermentation experiment as described in Example 1. Results are listed in Table 14.

Table 14. Results using yeast strain with PPP genes and deletion of GPD2 gene

Strain	IME324	IMX774	IMX1443
Relevant genotype	reference	9**cbbM*, *DAN1p-prk*, *groES*, *groEL*	9**cbbM*, *DAN1p-prk*, *groES*, *groEL* *gpd2::RPE1, TKL1, TAL1, TAL2, RKI1, NQM1*
μ (h^{-1})	0.33 ± 0.01	0.20 ± 0.03	0.30 ± 0.03
Y glycerol/glucose (g g^{-1})	0.102 ± 0.001	0.058 ± 0.005	0.014 ± 0.001
Y biomass/glucose (g_x g^{-1})	0.091 ± 0.000	0.087 ± 0.007	0.096 ± 0.004
Y EtOH/glucose (g g^{-1})	0.356 ± 0.004	0.409 ± 0.001	0.420 ± 0.001
Ratio glycerol produced/biomass (mmol g_x^{-1})	12.262 ± 0.122	7.272 ± 0.115	1.557 ± 0.003

Reference List

1. Entian KD, Kötter P. Yeast genetic strain and plasmid collections. Method Microbiol. 2007;629-66.

2. Nijkamp JF, van den Broek M, Datema E, de Kok S, Bosman L, Luttik MA, Daran-Lapujade P, Vongsangnak W, Nielsen J, Heijne WHM, Klaassen P, Paddon CJ, Platt D, Kötter P, van Ham RC, Reinders MJT, Pronk JT, de Ridder D, Daran J-M. De novo sequencing, assembly and analysis of the genome of the laboratory strain *Saccharomyces cerevisiae* CEN.PK113-7D, a model for modern industrial biotechnology. Microb Cell Fact. 2012;11:36.
3. Verduyn C, Postma E, Scheffers WA, van Dijken JP. Effect of benzoic acid on metabolic fluxes in yeasts: A continuous-culture study on the regulation of respiration and alcoholic fermentation. Yeast. 1992;8:501-17.
4. Mans R, van Rossum HM, Wijsman M, Backx A, Kuijpers NG, van den Broek M, Daran-Lapujade P, Pronk JT, van Maris AJA, Daran J-M. CRISPR/Cas9: a molecular Swiss army knife for simultaneous introduction of multiple genetic modifications in *Saccharomyces cerevisiae*. FEMS Yeast Res. 2015;15:fov004.
5. DiCarlo JE, Norville JE, Mali P, Rios X, Aach J, Church GM. Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems. Nucleic Acids Res. 2013;1-8.
6. Mikkelsen MD, Buron LD, Salomonsen B, Olsen CE, Hansen BG, Mortensen UH, Halkier BA. Microbial production of indolylglucosinolate through engineering of a multi-gene pathway in a versatile yeast expression platform. Metab Eng. 2012;14:104-11.
7. Knijnenburg TA, Daran JM, van den Broek MA, Daran-Lapujade PA, de Winde JH, Pronk JT, Reinders MJ, Wessels LF. Combinatorial effects of environmental parameters on transcriptional regulation in *Saccharomyces cerevisiae*: A quantitative analysis of a compendium of chemostat-based transcriptome data. BMC Genomics. 2009;10:53.
8. Mumberg D, Müller R, Funk M. Yeast vectors for the controlled expression of heterologous proteins in different genetic backgrounds. Gene. 1995;156:119-22.
9. Gueldener U, Heinisch J, Koehler GJ, Voss D, Hegemann JH. A second set of loxP marker cassettes for Cre-mediated multiple gene knockouts in budding yeast. Nucleic Acids Res. 2002;30:e23.
10. Guadalupe-Medina V, Wisselink H, Luttik M, de Hulster E, Daran J-M, Pronk JT, van Maris AJA. Carbon dioxide fixation by Calvin-Cycle enzymes improves ethanol yield in yeast. Biotechnol Biofuels. 2013;6:125.
11. Daniel Gietz R, Woods RA: Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. Methods Enzymol. 2002:87-96.
12. Solis-Escalante D, Kuijpers NGA, Bongaerts N, Bolat I, Bosman L, Pronk JT, Daran J-M, Daran-Lapujade P. amdSYM, a new dominant recyclable marker cassette for *Saccharomyces cerevisiae*. FEMS Yeast Res. 2013;13:126-39.
13. Guadalupe-Medina V, Almering MJH, van Maris AJA, Pronk JT. Elimination of glycerol production in anaerobic cultures of a *Saccharomyces cerevisiae* strain engineered to use acetic acid as an electron acceptor. Appl Environ Microb. 2010;76:190-5.

14. Papapetridis I, van Dijk M, Dobbe AP, Metz B, Pronk JT, van Maris AJA. Improving ethanol yield in acetate-reducing *Saccharomyces cerevisiae* by cofactor engineering of 6-phosphogluconate dehydrogenase and deletion of *ALD6*. Microb Cell Fact. 2016;15:1-16.
15. Heijnen JJ, van Dijken JP. In search of a thermodynamic description of biomass yields for the chemotrophic growth of microorganisms. Biotechnol Bioeng. 1992;39:833-58.
16. Postma E, Verduyn C, Scheffers WA, van Dijken JP. Enzymic analysis of the crabtree effect in glucose-limited chemostat cultures of *Saccharomyces cerevisiae*. Appl Environ Microbiol. 1989;55:468-77.
17. Verduyn C, Postma E, Scheffers WA, van Dijken JP. Physiology of *Saccharomyces cerevisiae* in anaerobic glucose-limited chemostat cultures. J Gen Microbiol. 1990;136:395-403.
18. Kwast et al. Genomic Analysis of Anaerobically induced genes in Saccharomyces cerevisiae: Functional roles of ROX1 and other factors in mediating the anoxic response, 2002, Journal of bacteriology vol 184, no1 p250-265.
19. Keng, T. 1992. HAP1 and ROX1 form a regulatory pathway in the repression of HEM13 transcription in Saccharomyces cerevisiae. Mol. Cell. Biol. 12: 2616–2623.
20. Labbe-Bois, R., and P. Labbe. 1990. Tetrapyrrole and heme biosynthesis in the yeast Saccharomyces cerevisiae, p. 235–285. In H. A. Dailey (ed.), Biosynthesis of heme and chlorophylls. McGraw-Hill, New York, N.Y.
21. Zitomer, R. S., and C. V. Lowry. 1992. Regulation of gene expression by oxygen in Saccharomyces cerevisiae. Microbiol. Rev. 56:1–11.
22. Zitomer, R. S., P. Carrico, and J. Deckert. 1997. Regulation of hypoxic gene expression in yeast. Kidney Int. 51:507–513.
23. Cohen et al., Induction and repression of DAN1 and the family of anaerobic mannoprotein genes in Saccharomyces cerevisiae occurs through a complex array of regulatory sites. Nucleic Acid Research, 2001 Vol. 29, No3, 799-808
24. Ter Kinde and de Steensma, A microarray-assisted screen for potential Hap1 and Rox1 target genes in Saccharomyces cerevisiae, 2002, Yeast 19: 825-840.
25. Sertil et al. The DAN1 gene of S cerevisiae is regulated in parallel with the hypoxic gene , but by a different mechanism, 1997, Gene Vol 192, pag 199-205.

CLAIMS

1. A recombinant yeast cell functionally expressing one or more heterologous nucleic acid sequences encoding for ribulose-1,5-phosphate carboxylase/oxygenase (EC4.1.1.39; Rubisco), and optionally one or more molecular chaperones for Rubisco, and one or more phosphoribulokinase (EC2.7.1.19; PRK), wherein the phosphoribulokinase is under control of a promoter (the "PRK promoter") which has a PRK expression ratio $_{anaerobic/aerobic}$ of 2 or more.
2. A recombinant yeast cell according to claim 1, wherein the PRK promoter is ROX1-repressed.
3. A recombinant yeast cell according to claim 1 or 2, wherein the PRK promoter has one or more ROX1 binding motif.
4. A recombinant yeast cell according to any of claims 1 to 3, wherein PRK promoter comprises in its sequence one or more of the motif NNNATTGTTNNN.
5. A recombinant yeast cell according to any of claims 1 to 4, wherein the PRK promoter is the native promoter of a gene selected from the list consisting of: FET4, ANB1, YHR048W, DAN1, AAC3, TIR2, DIP5, HEM13, YNR014W, YAR028W, FUN 57, COX5B, OYE2, SUR2, FRDS1, PIS1, LAC1, YGR035C, YAL028W, EUG1, HEM14, ISU2, ERG26, YMR252C and SML1.
6. A recombinant yeast cell according to claim 5, wherein the PRK promoter is the native promoter of a gene selected from the list consisting of: FET4, ANB1, YHR048W, DAN1, AAC3, TIR2, DIP5 and HEM13.
7. A recombinant yeast cell according to claim 1, wherein the PRK promoter comprises in its sequence one or more of the motif: TCGTTYAG and/or AAAAATTGTTGA.
8. A recombinant yeast cell according to claim 7, wherein the PRK promoter is the native promoter of a gene selected from the list consisting of: TIR2, DAN1, TIR4, TIR3, PAU7, PAU5, YLL064C, YGR294W, DAN3, YIL176C, YGL261C, YOL161C, PAU1, PAU6, DAN2, YDR542W, YIR041W, YKL224C, PAU3, YLL025W, YOR394W, YHL046C, YMR325W, YAL068C, YPL282C, PAU2, PAU4
9. A recombinant yeast cell according to claim 8, wherein the PRK promoter is the native promoter of a gene selected from the list consisting of: TIR2, DAN1, TIR4, TIR3, PAU7, PAU5, YLL064C, YGR294W, DAN3, YIL176C, YGL261C, YOL161C, PAU1, PAU6, DAN2, YDR542W, YIR041W, YKL224C, PAU3, YLL025W.
10. A recombinant yeast strain according to any of claims 1 to 9, wherein the PRK promoter is a synthetic oligonucleotide.
11. A recombinant yeast cell according to any of claims 1 to 10, wherein the PRK promoter enables expression only during anaerobic conditions.
12. A recombinant yeast cell according to any of claims 1 to 11, wherein the PRK promoter is the promoter of ANB1 and/or DAN1.

13. A recombinant yeast cell according to any of claims 1 to 12 whererin the Rubisco is under a constitutive promotor.
14. A recombinant yeast cell according to any of claims 1 to 13 in which one or more genes of the non-oxidative branch of the pentose phosphate pathway are overexpressed and/or which yeast cell comprises a deletion or disruption of a glycerol-3-phosphate dehydrogenase (GPD) gene.
15. A vector comprising PRK and a PRK promoter that enables higher expression during anaerobic conditions than under aerobic conditions and/or is a PRK promoter mentioned in any of claims 1 to 14.
16. A process for preparing an organic compound, in particular an alcohol, comprising converting a carbon source, in particular a carbohydrate or another organic carbon source using a recombinant yeast cell according to any of claims 1 to 14, thereby forming the organic compound.

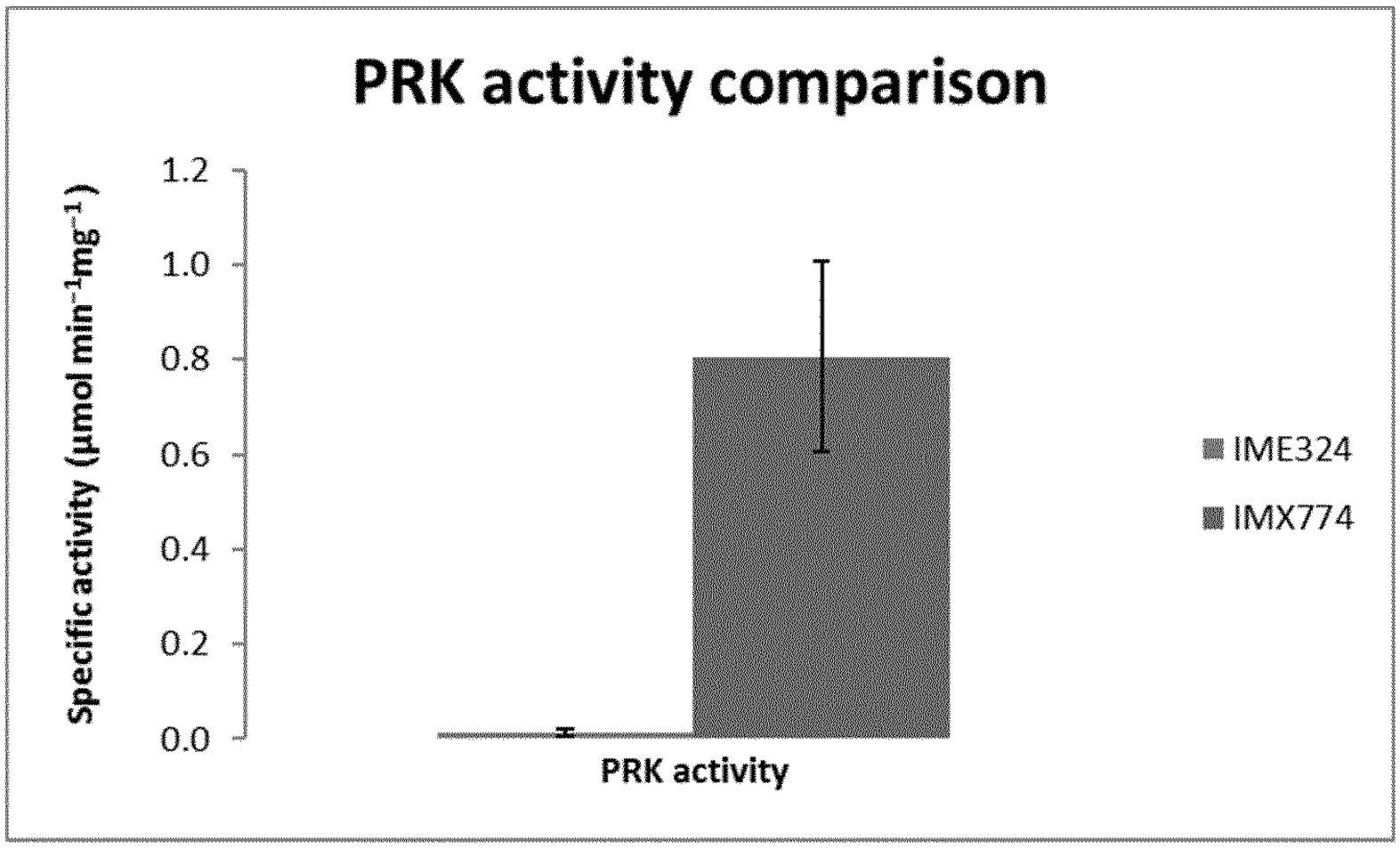
PRK activity comparison
Specific activity (µmol min⁻¹mg⁻¹)
1.2
1.0
0.8
0.6
0.4
0.2
0.0
PRK activity
IME324
IMX774

Fig. 1

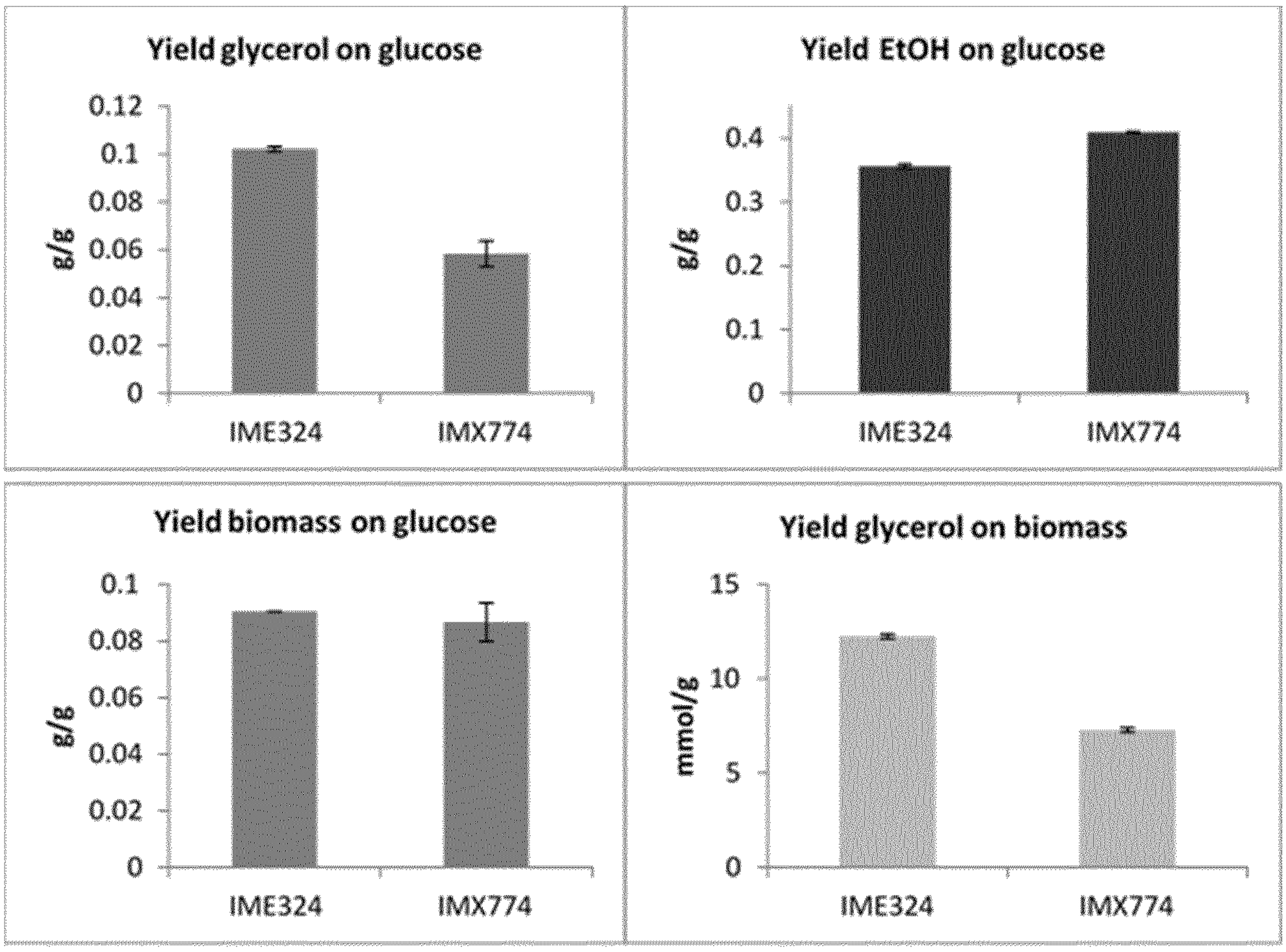
Yield glycerol on glucose
g/g
0.12
0.1
0.08
0.06
0.04
0.02
0
IME324
IMX774
Yield EtOH on glucose
g/g
0.4
0.3
0.2
0.1
0
IME324
IMX774
Yield biomass on glucose
g/g
0.1
0.08
0.06
0.04
0.02
0
IME324
IMX774
Yield glycerol on biomass
mmol/g
15
10
5
0
IME324
IMX774

Fig. 2

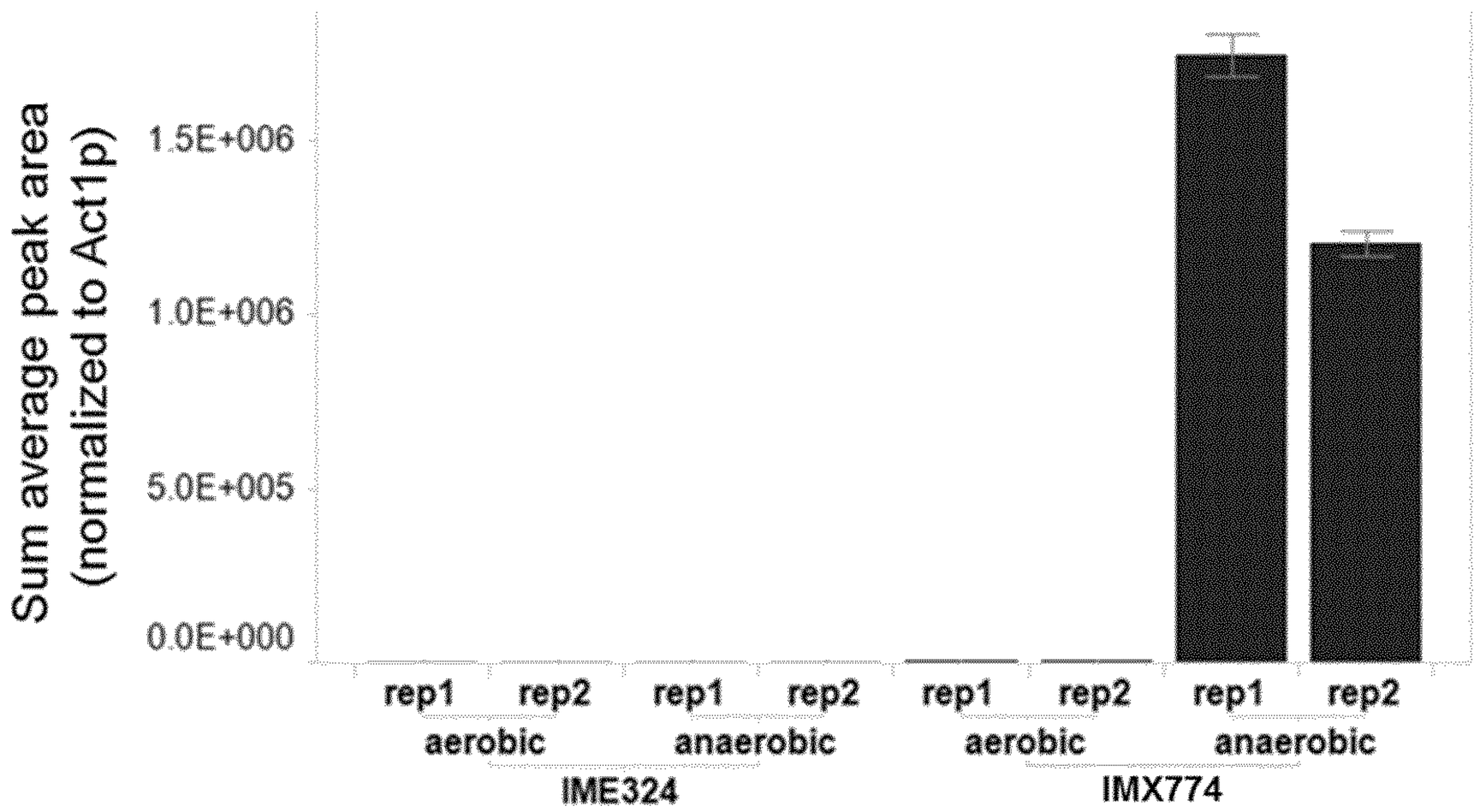
Sum average peak area
(normalized to Act1p)
1.5E+006
1.0E+006
5.0E+005
0.0E+000
rep1
rep2
rep1
rep2
rep1
rep2
rep1
rep2
aerobic
anaerobic
aerobic
anaerobic
IME324
IMX774

Fig. 3

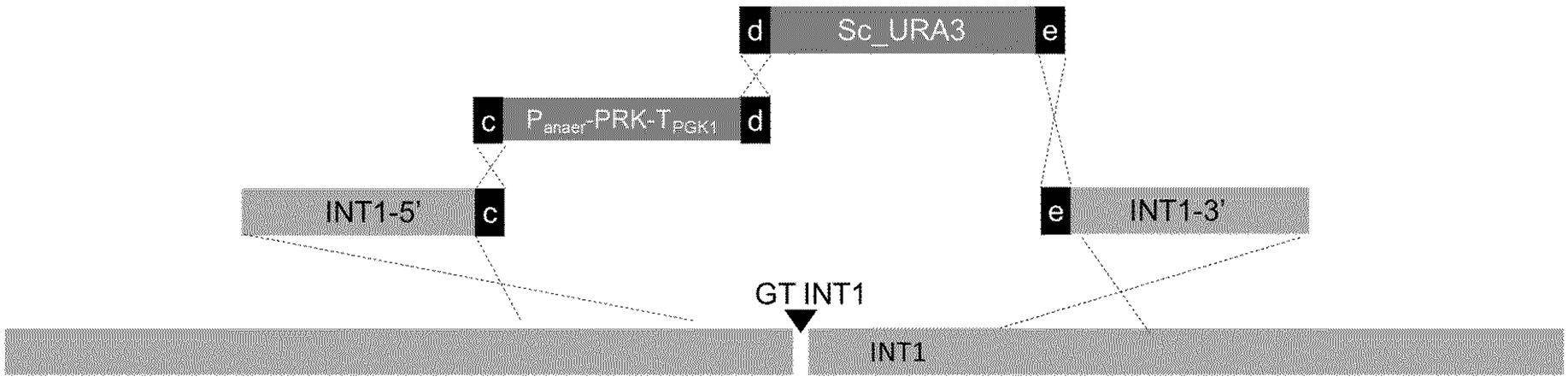

Fig. 4

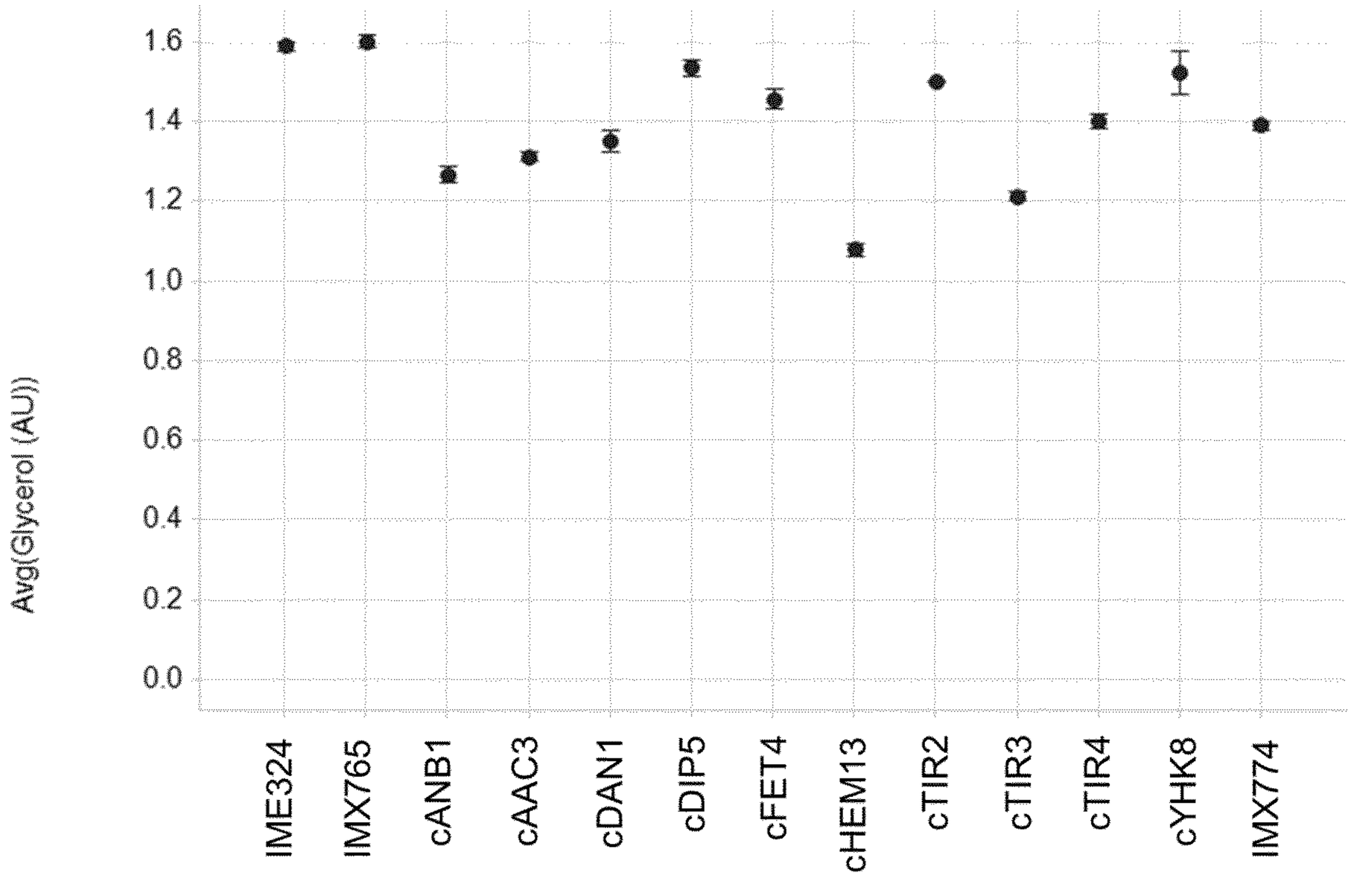
Avg(Glycerol (AU))
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
IME324
IMX765
cANB1
cAAC3
cDAN1
cDIP5
cFET4
cHEM13
cTIR2
cTIR3
cTIR4
cYHK8
IMX774

Fig. 5

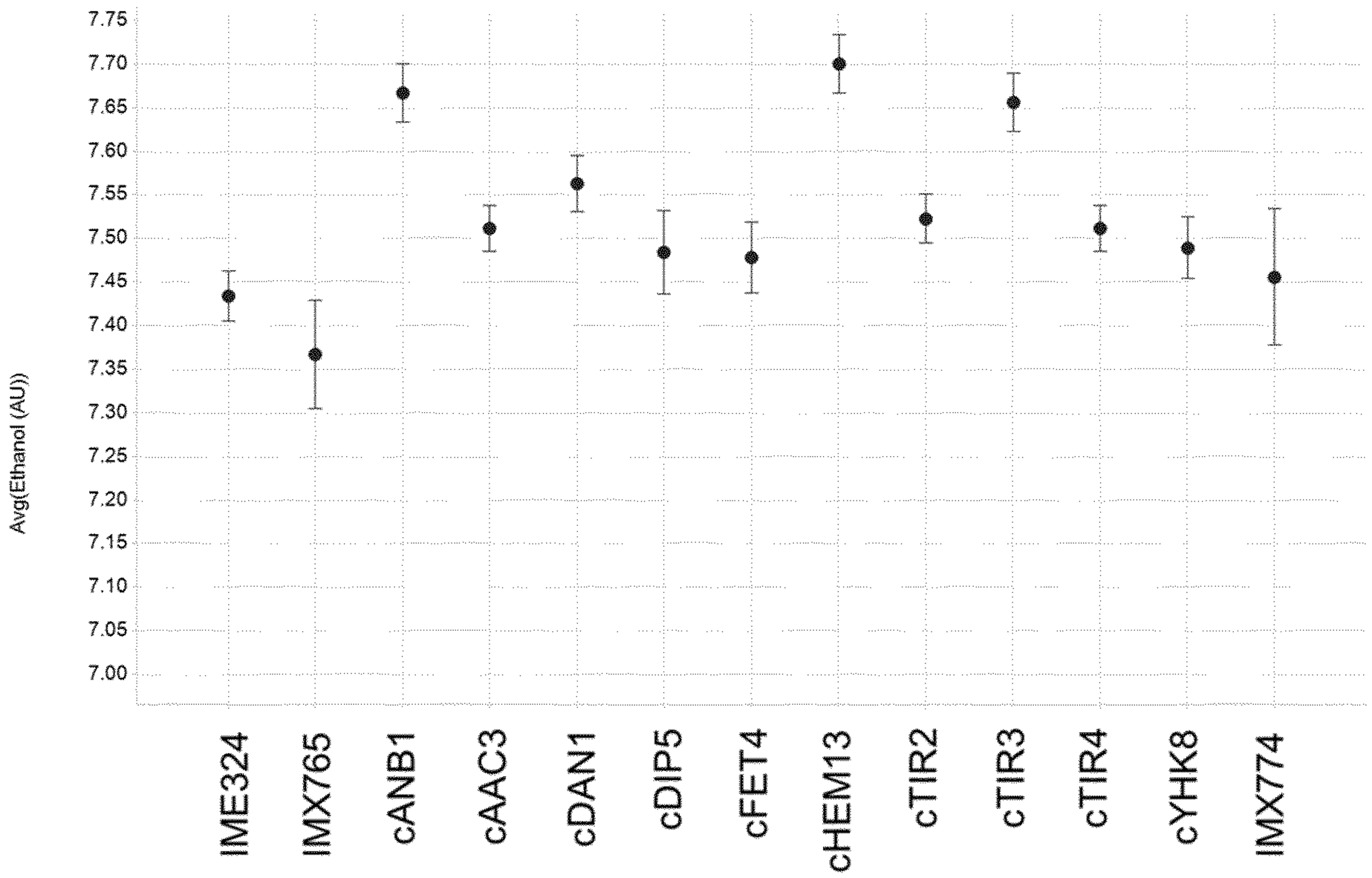
Avg(Ethanol (AU))
7.75
7.70
7.65
7.60
7.55
7.50
7.45
7.40
7.35
7.30
7.25
7.20
7.15
7.10
7.05
7.00
IME324
IMX765
cANB1
cAAC3
cDAN1
cDIP5
cFET4
cHEM13
cTIR2
cTIR3
cTIR4
cYHK8
IMX774

Fig. 6