

GLYCEROL 3- PHOSPHATE DEHYDROGENASE FOR BUTANOL PRODUCTION

CROSS REFERENCE TO RELATED APPLICATIONS

[0001] This application claims benefit of priority from United States Provisional Application No. 61/782,651, filed March 14, 2013, and United States Provisional Application No. 61/934,096, filed January 31, 2014, each of which is hereby incorporated by reference in its entirety.

FIELD OF THE INVENTION

[0002] The invention relates to the field of industrial microbiology and the fermentative production of butanol and isomers thereof. More specifically, the invention relates to glycerol-3-phosphate dehydrogenase (GPD) enzymes with a high K_M for NADH, substantially the same affinity for NADH and NADPH and/or GPD enzymes that are feedback inhibited, recombinant microorganisms comprising such enzymes, and methods of using such enzymes to produce butanol.

REFERENCE TO SEQUENCE LISTING SUBMITTED ELECTRONICALLY

[0003] The content of the electronically submitted sequence listing in ASCII text file (Name: 20140314_CL5707USNP_Sequence Listing; Size: 401,408 bytes, and Date of Creation: March 12, 2014) filed with the application is incorporated herein by reference in its entirety.

BACKGROUND

[0004] Butanol is an important industrial chemical, useful as a fuel additive, as a feedstock chemical in the plastics industry, and as a food grade extractant in the food and flavor industry. Each year 10 to 12 billion pounds of butanol are produced by petrochemical means and the need for this commodity chemical will likely increase in the future.

[0005] Methods for the chemical synthesis of isobutanol are known, such as oxo synthesis, catalytic hydrogenation of carbon monoxide (Ullmann's Encyclopedia of Industrial Chemistry, 6th edition, 2003, Wiley-VCH Verlag GmbH and Co., Weinheim, Germany, Vol. 5, pp. 716-719) and Guerbet condensation of methanol with n-propanol (Carlini *et al.*, *J. Molec. Catal. A. Chem.* 220:215-220, 2004). These processes use starting materials derived from

petrochemicals, are generally expensive, and are not environmentally friendly. The production of isobutanol from plant-derived raw materials would minimize greenhouse gas emissions and would represent an advance in the art.

[0006] Isobutanol is produced biologically as a by-product of yeast fermentation or by recombinantly engineered microorganisms modified to express a butanol biosynthetic pathway for producing butanol. *See e.g.*, U.S. Patent No. 7,851,188, which is incorporated herein by reference in its entirety. As a component of “fusel oil” that forms as a result of the incomplete metabolism of amino acids by fungi, isobutanol is specifically produced from catabolism of L-valine. After the amine group of L-valine is harvested as a nitrogen source, the resulting α -keto acid is decarboxylated and reduced to isobutanol by enzymes of the so-called Ehrlich pathway (Dickinson *et al.*, *J. Biol. Chem.* 273:25752-25756, 1998).

[0007] One of the key yield loss mechanisms in yeast butanol production is the loss of carbon and reducing equivalents that are diverted from glycolysis by the conversion of dihydroxyacetone phosphate to glycerol. The first step in this conversion is catalyzed by an enzyme called glycerol-3-phosphate dehydrogenase (GPD). Eliminating GPD, and therefore glycerol production, in butanol-producing yeast, has been proposed previously. However, glycerol is required for growth and is an osmoprotectant.

[0008] Accordingly, methods of increasing butanol yield and decreasing glycerol production represent an advance in the art.

BRIEF SUMMARY OF THE INVENTION

[0009] Provided herein are GPD enzymes, recombinant microorganisms, and methods for production of butanol.

[0010] Provided herein are recombinant microorganisms comprising (a) an engineered butanol biosynthetic pathway comprising at least one polypeptide that is heterologous to the recombinant microorganism; (b) a heterologous glycerol-3-phosphate dehydrogenase (GPD), wherein the heterologous GPD has a higher K_M for NADH as compared to the K_M of the endogenous GPD of the microorganism; and (c) a deletion or disruption in an endogenous gene encoding GPD. In some embodiments, the recombinant microorganism has improved or increased production of butanol as compared to a control recombinant microorganism that lacks the heterologous GPD. In some embodiments, the recombinant microorganism has reduced or decreased production of glycerol as compared to a control recombinant microorganism that lacks

the heterologous GPD. In some embodiments, the recombinant microorganism has an increased butanol to glycerol molar ratio as compared to a control recombinant microorganism that lacks the heterologous GPD. In some embodiments, the recombinant microorganism has an increased effective yield as compared to a control recombinant microorganism that lacks the heterologous GPD.

[0011] Also provided herein are recombinant microorganisms comprising (a) an engineered butanol biosynthetic pathway comprising at least one polypeptide that is heterologous to the recombinant microorganism; (b) a heterologous glycerol-3-phosphate dehydrogenase (GPD), wherein the heterologous GPD has substantially the same affinity for NADH and NADPH and/or is feedback inhibited; and (c) a deletion or disruption in an endogenous gene encoding GPD. In some embodiments, the recombinant microorganism has improved production of butanol as compared to a control recombinant microorganism that lacks the heterologous GPD. Optionally, the heterologous GPD is feedback inhibited by glycerol-3-phosphate. In some embodiments, the recombinant microorganism has reduced or decreased production of glycerol as compared to a control recombinant microorganism that lacks the heterologous GPD. In some embodiments, the recombinant microorganism has an increased butanol to glycerol molar ratio as compared to a control recombinant microorganism that lacks the heterologous GPD. In some embodiments, the recombinant microorganism has an increased effective yield as compared to a control recombinant microorganism that lacks the heterologous GPD.

[0012] In certain embodiments, the heterologous GPD is a naturally occurring GPD. In certain embodiments, the naturally occurring GPD is selected from EC number 1.1.1.8, 1.1.5.3, or 1.1.1.94. The naturally occurring GPD can be a GPD from an organism selected from the group consisting of *Leishmania mexicana*, *Dunaliella viridis*, *Jaculus orientalis*, *Archeoglobus fulgidus*, *Rickettsia prowazekii*, *Beggiatoa alba*, *Kangiella koreensis*, *Aspergillus oryzae*, *Candida versatilis*, *Escherichia coli*, and *Oryctolagus cuniculu*.

[0013] In certain embodiments, the heterologous GPD is an engineered GPD. The engineered GPD can comprise at least one substitution corresponding to position 42, 44, 45, 71, 73, 75, 95, 124, 126, 129, 151, 152, 183, 184, 185, 246, 310, 336, 337, or 339 of SEQ ID NO: 195. In certain embodiments the engineered GPD comprises at least one substitution at a residue corresponding to position 73 of SEQ ID NO:195. In certain embodiments the engineered GPD comprises at least one substitution at a residue corresponding to position 129 of SEQ ID NO:195. In certain embodiments the engineered GPD comprises at least one substitution at a residue

corresponding to position 73 of SEQ ID NO:195 and a substitution at a residue corresponding to position 129 of SEQ ID NO:195.

[0014] Also provided are engineered glycerol-3-phosphate dehydrogenase (GPD) enzymes. In certain embodiments, the engineered GPD enzyme has at least 85% identity to SEQ ID NO:195. In certain embodiments, the engineered GPD enzyme comprises at least one substitution at a residue corresponding to position 42, 44, 45, 71, 73, 75, 95, 124, 126, 129, 151, 152, 183, 184, 185, 246, 310, 336, 337, or 339 of SEQ ID NO: 195. In certain embodiments, the engineered GPD enzyme comprises at least one substitution corresponding to position 73 of SEQ ID NO:195. In certain embodiments, the engineered GPD enzyme comprises at least one substitution corresponding to position 129 of SEQ ID NO:195. In certain embodiments, the engineered GPD enzyme comprises at least one substitution corresponding to position 73 of SEQ ID NO:195 and a substitution corresponding to position 129 of SEQ ID NO:129. In certain embodiments, the engineered GPD enzyme has a K_M for NADH from about 0.01 mM to 1mM.

[0015] Also provided are recombinant microorganisms comprising any of the engineered GPD enzymes disclosed herein. Optionally, the recombinant microorganism can comprise an engineered butanol biosynthetic pathway that comprises at least one gene that is heterologous to the recombinant microorganism. The recombinant microorganism can, for example, comprise a deletion or disruption of an endogenous gene encoding GPD. In certain embodiments, the recombinant microorganism has improved or increased production of butanol compared to a microorganism that lacks the engineered GPD enzyme. In some embodiments, the recombinant microorganism has reduced or decreased production of glycerol as compared to a control recombinant microorganism that lacks the engineered GPD. In some embodiments, the recombinant microorganism has an increased butanol to glycerol molar ratio as compared to a control recombinant microorganism that lacks the engineered GPD. In some embodiments, the recombinant microorganism has an increased effective yield as compared to a control recombinant microorganism that lacks the engineered GPD.

[0016] Also provided are methods for the production of butanol. The methods comprise providing a recombinant microorganism comprising (i) an engineered butanol biosynthetic pathway, (ii) a deletion or disruption in an endogenous gene encoding GPD, and; (iii) at least one of (a) an engineered GPD enzyme; (b) a heterologous glycerol-3-phosphate dehydrogenase (GPD), wherein the heterologous GPD has a higher K_M for NADH as compared to the K_M of the microorganism's endogenous GPD; or (c) a heterologous GPD, wherein the heterologous GPD

has substantially the same affinity for NADH and NADPH and/or is feedback inhibited; and contacting the recombinant microorganism with at least one fermentable carbon substrate under conditions wherein butanol is produced. Optionally, the heterologous GPD is feedback inhibited by glycerol-3-phosphate. In certain embodiments, the recombinant microorganism is grown under anaerobic conditions.

[0017] The recombinant microorganism can comprise an engineered butanol biosynthetic pathway selected from the group consisting of (a) a 1-butanol biosynthetic pathway; (b) a 2-butanol biosynthetic pathway; and (c) an isobutanol biosynthetic pathway.

[0018] Optionally, the 1-butanol biosynthetic pathway comprises at least one polypeptide that performs one of the following substrate to product conversions: (a) acetyl-CoA to acetoacetyl-CoA, as catalyzed by acetyl-CoA acetyltransferase; (b) acetoacetyl-CoA to 3-hydroxybutyryl-CoA, as catalyzed by 3-hydroxybutyryl-CoA dehydrogenase; (c) 3-hydroxybutyryl-CoA to crotonyl-CoA, as catalyzed by crotonase; (d) crotonyl-CoA to butyryl-CoA, as catalyzed by butyryl-CoA dehydrogenase; (e) butyryl-CoA to butyraldehyde, as catalyzed by butyraldehyde dehydrogenase; and (f) butyraldehyde to 1-butanol, as catalyzed by 1-butanol dehydrogenase.

[0019] Optionally, the 2-butanol biosynthetic pathway comprises at least one polypeptide that performs one of the following substrate to product conversions: (a) pyruvate to alpha-acetolactate, as catalyzed by acetolactate synthase; (b) alpha-acetolactate to acetoin, as catalyzed by acetolactate decarboxylase; (c) acetoin to 2,3-butanediol, as catalyzed by butanediol dehydrogenase; (d) 2,3-butanediol to 2-butanone, as catalyzed by butanediol dehydratase; and (e) 2-butanone to 2-butanol, as catalyzed by 2-butanol dehydrogenase.

[0020] Optionally, the isobutanol biosynthetic pathway comprises at least one polypeptide that performs one of the following substrate to product conversions: (a) pyruvate to acetolactate, as catalyzed by acetolactate synthase; (b) acetolactate to 2,3-dihydroxyisovalerate, as catalyzed by ketol-acid reductoisomerase; (c) 2,3-dihydroxyisovalerate to α -ketoisovalerate, as catalyzed by dihydroxyacid dehydratase; (d) α -ketoisovalerate to isobutyraldehyde, as catalyzed by a branched chain keto acid decarboxylase; and (e) isobutyraldehyde to isobutanol, as catalyzed by branched-chain alcohol dehydrogenase.

[0021] In certain embodiments, the recombinant microorganism is from a genus selected from the group consisting of *Clostridium*, *Zymomonas*, *Escherichia*, *Salmonella*, *Serratia*, *Erwinia*, *Klebsiella*, *Shigella*, *Rhodococcus*, *Pseudomonas*, *Bacillus*, *Lactobacillus*,

Enterococcus, Alcaligenes, Paenibacillus, Arthrobacter, Corynebacterium, Brevibacterium, Schizosaccharomyces, Kluyveromyces, Yarrowia, Pichia, Zygosaccharomyces, Debaryomyces, Candida, Brettanomyces, Pachysolen, Hansenula, Issatchenkia, Trichosporon, Yamadazyma, and Saccharomyces.

[0022] Also provided are recombinant microorganisms comprising (a) a heterologous glycerol-3-phosphate dehydrogenase (GPD), wherein the heterologous GPD has a higher K_M for NADH as compared to the K_M of the endogenous GPD of the microorganism; and (b) a deletion or disruption in an endogenous gene encoding GPD. In some embodiments the microorganism has decreased production of glycerol as compared to a control recombinant microorganism that lacks the heterologous GPD.

[0023] Also provided are recombinant microorganisms comprising (a) a heterologous glycerol-3-phosphate dehydrogenase (GPD), wherein the heterologous GPD has substantially the same affinity for NADH and NADPH and/or is feedback inhibited; and (b) a deletion or disruption in an endogenous gene encoding GPD. Optionally, the heterologous GPD is feedback inhibited by glycerol-3-phosphate. In some embodiments the microorganism has decreased production of glycerol as compared to a control recombinant microorganism that lacks the heterologous GPD.

[0024] Also provided are recombinant microorganisms comprising a heterologous GPD, wherein the heterologous GPD has substantially the same affinity for NADH and NADPH and/or is feedback inhibited, and wherein the recombinant microorganism has decreased production of glycerol as compared to a control recombinant microorganism that lacks the heterologous GPD. Optionally, the heterologous GPD is feedback inhibited by glycerol-3-phosphate.

BRIEF DESCRIPTION OF THE DRAWINGS/FIGURES

[0025] Figure 1 shows a map of the plasmid used to express variant GPD proteins in *E. coli*.

[0026] Figure 2 shows a partial alignment of the GPD sequences of human truncated (SEQ ID NO:190), *Saccharomyces cerevisiae* (GPD1) (SEQ ID NO:191), *Rickettsia prowazekii* (SEQ ID NO:192), *Beggiatoa alba* (SEQ ID NO:193), and *Kangiella koreensis* (SEQ ID NO:194). Asterisk (*) indicates the positions of the phe41 and phe97 in the human truncated sequence.

[0027] Figure 3 shows a graph demonstrating 20 hour production data for the indicated GPD1 variant and control cell cultures. Two clones for each variant were tested in duplicate. 2145: isobutanologen control strain with WT GPD which is PNY2145 transformed with pLMH11-JM44; EC_1 and EC_2: *E. coli* optimized GPD; E3 and E8: *E. coli* optimized GPD1 variants; N3 and N8: yeast native codon-usage GPD variants; M3 and M8: Yeast codon optimized GPD variants. Variants E3, N3, and M3 has F73A substitution and variants E8, N8, and M8 have F73G/F129G substitutions.

[0028] Figure 4 shows a graph demonstrating a comparison of isobutanol (iBuOH)/glycerol (Gly) ratio with measured GPD activity (U/mg). The regression equation for the iBuOH/Gly ratio equals $2.93 - 234\text{GPD (U/mg)}$ ($R\text{-Sq} = 60.1\%$; $R\text{-Sq}_{(\text{pred})} = 25.04\%$).

[0029] Figure 5 shows a graph demonstrating a comparison of measured isobutanol titer to values calculated by the linear regression equation (FIT_2) ($S = 0.159277$; $R\text{-Sq} = 98.6\%$; $R\text{-Sq}_{(\text{adj})} = 97.6\%$; $\text{PRESS} = 0.415575$; $R\text{-Sq}_{(\text{pred})} = 94.32\%$). The constant and coefficients for the regression equation are provided in Table 12.

[0030] Figure 6 shows a graph demonstrating a comparison of isobutanol yield (grams isobutanol/gram glucose consumed) to values calculated by the linear regression equation (FIT_4) ($S = 0.0101071$; $R\text{-Sq} = 93.9\%$; $R\text{-Sq}_{(\text{adj})} = 91.4\%$; $\text{PRESS} = 0.00197878$; $R\text{-Sq}_{(\text{pred})} = 76.30\%$). The constant and coefficients for the regression equation are provided in Table 13.

[0031] Figure 7 shows a graph of the isobutanol yield (grams of isobutanol produced per gram of glucose consumed) at 28 and 42 hours for CPN97 and PNY2310 isobutanologen strains.

[0032] Figure 8 shows a graph of the isobutanol/glycerol ratio at 28 and 42 hours for CPN97 and PNY2310 isobutanologen strains.

[0033] Figure 9 shows a graph of glucose consumed as a function of time for CPN97 and PNY2310 isobutanologen strains.

DETAILED DESCRIPTION OF THE INVENTION

[0034] Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. In case of conflict, the present application including the definitions will control. Also, unless otherwise required by context, singular terms shall include pluralities and plural terms shall include the singular. All publications, patents and other references mentioned herein are incorporated by reference in their entireties for all purposes as if each individual publication or

patent application were specifically and individually indicated to be incorporated by reference, unless only specific sections of patents or patent publications are indicated to be incorporated by reference.

[0035] In order to further define this invention, the following terms, abbreviations and definitions are provided.

[0036] It will be understood that “derived from” with reference to polypeptides disclosed herein encompasses sequences synthesized based on the amino acid sequences of the GPDs, or other enzymes, present in the indicated organisms as well as those cloned directly from the genetic material of the organism.

[0037] As used herein, the terms “comprises,” “comprising,” “includes,” “including,” “has,” “having,” “contains” or “containing,” or any other variation thereof, will be understood to imply the inclusion of a stated integer or group of integers but not the exclusion of any other integer or group of integers and are intended to be non-exclusive or open-ended. For example, a composition, a mixture, a process, a method, an article, or an apparatus that comprises a list of elements is not necessarily limited to only those elements but can include other elements not expressly listed or inherent to such composition, mixture, process, method, article, or apparatus. Further, unless expressly stated to the contrary, “or” refers to an inclusive or and not to an exclusive or. For example, a condition A or B is satisfied by any one of the following: A is true (or present) and B is false (or not present), A is false (or not present) and B is true (or present), and both A and B are true (or present).

[0038] As used herein, the term “consists of,” or variations such as “consist of” or “consisting of,” as used throughout the specification and claims, indicate the inclusion of any recited integer or group of integers, but that no additional integer or group of integers can be added to the specified method, structure, or composition.

[0039] As used herein, the term “consists essentially of,” or variations such as “consist essentially of” or “consisting essentially of,” as used throughout the specification and claims, indicate the inclusion of any recited integer or group of integers, and the optional inclusion of any recited integer or group of integers that do not materially change the basic or novel properties of the specified method, structure or composition. *See* M.P.E.P. § 2111.03.

[0040] Also, the indefinite articles “a” and “an” preceding an element or component of the invention are intended to be nonrestrictive regarding the number of instances, *i.e.*, occurrences of the element or component. Therefore “a” or “an” should be read to include one or

at least one, and the singular word form of the element or component also includes the plural unless the number is obviously meant to be singular.

[0041] The term “invention” or “present invention” as used herein is a non-limiting term and is not intended to refer to any single embodiment of the particular invention but encompasses all possible embodiments as described in the claims as presented or as later amended and supplemented, or in the specification.

[0042] As used herein, the term “about” modifying the quantity of an ingredient or reactant of the invention employed refers to variation in the numerical quantity that can occur, for example, through typical measuring and liquid handling procedures used for making concentrates or solutions in the real world; through inadvertent error in these procedures; through differences in the manufacture, source, or purity of the ingredients employed to make the compositions or to carry out the methods; and the like. The term “about” also encompasses amounts that differ due to different equilibrium conditions for a composition resulting from a particular initial mixture. Whether or not modified by the term “about”, the claims include equivalents to the quantities. In one embodiment, the term “about” means within 10% of the reported numerical value, or within 5% of the reported numerical value.

[0043] The term “butanol biosynthetic pathway” as used herein refers to the enzymatic pathway to produce 1-butanol, 2-butanol, or isobutanol.

[0044] The term “1-butanol biosynthetic pathway” refers to an enzymatic pathway to produce 1-butanol. A “1-butanol biosynthetic pathway” can refer to an enzyme pathway to produce 1-butanol from acetyl-coenzyme A (acetyl-CoA). For example, 1-butanol biosynthetic pathways are disclosed in U.S. Patent Application Publication No. 2008/0182308 and International Publication No. WO 2007/041269, which are herein incorporated by reference in their entireties.

[0045] The term “2-butanol biosynthetic pathway” refers to an enzymatic pathway to produce 2-butanol. A “2-butanol biosynthetic pathway” can refer to an enzyme pathway to produce 2-butanol from pyruvate. For example, 2-butanol biosynthetic pathways are disclosed in U.S. Patent No. 8,206,970, U.S. Patent Application Publication No. 2007/0292927, International Publication Nos. WO 2007/130518 and WO 2007/130521, which are herein incorporated by reference in their entireties.

[0046] The term “isobutanol biosynthetic pathway” refers to an enzymatic pathway to produce isobutanol. An “isobutanol biosynthetic pathway” can refer to an enzyme pathway to

produce isobutanol from pyruvate. For example, isobutanol biosynthetic pathways are disclosed in U.S. Patent No. 7,851,188, U.S. Application Publication No. 2007/0092957, and International Publication No. WO 2007/050671, which are herein incorporated by reference in their entireties. From time to time “isobutanol biosynthetic pathway” is used synonymously with “isobutanol production pathway.”

[0047] The term “butanol” as used herein refers to the butanol isomers 1-butanol (1-BuOH), 2-butanol (2-BuOH), *tert*-butanol (*t*-BuOH), and/or isobutanol (iBuOH or *i*-BuOH, also known as 2-methyl-1-propanol), either individually or as mixtures thereof. From time to time, as used herein the terms “biobutanol” and “bio-produced butanol” may be used synonymously with “butanol.”

[0048] Uses for butanol can include, but are not limited to, fuels (e.g., biofuels), a fuel additive, an alcohol used for the production of esters that can be used as diesel or biodiesel fuel, as a chemical in the plastics industry, an ingredient in formulated products such as cosmetics, and a chemical intermediate. Butanol may also be used as a solvent for paints, coatings, varnishes, resins, gums, dyes, fats, waxes, resins, shellac, rubbers, and alkaloids.

[0049] As used herein, the term “bio-produced” means that the molecule (e.g., butanol) is produced from a renewable source (e.g., the molecule can be produced during a fermentation process from a renewable feedstock). Thus, for example, bio-produced isobutanol can be isobutanol produced by a fermentation process from a renewable feedstock. Molecules produced from a renewable source can further be defined by the $^{14}\text{C}/^{12}\text{C}$ isotope ratio. A $^{14}\text{C}/^{12}\text{C}$ isotope ratio in range of from 1:0 to greater than 0:1 indicates a bio-produced molecule, whereas a ratio of 0:1 indicates that the molecule is fossil derived.

[0050] A recombinant host cell comprising an “engineered alcohol production pathway” (such as an engineered butanol or isobutanol production pathway) refers to a host cell containing a modified pathway that produces alcohol in a manner different than that normally present in the host cell. Such differences include production of an alcohol not typically produced by the host cell, or increased or more efficient production.

[0051] The term “heterologous biosynthetic pathway” as used herein refers to an enzyme pathway to produce a product in which at least one of the enzymes is not endogenous to the host cell containing the biosynthetic pathway.

[0052] The term “extractant” as used herein refers to one or more organic solvents which can be used to extract alcohol (e.g., butanol) from a fermentation broth.

[0053] The term “effective isobutanol productivity” as used herein refers to the total amount in grams of isobutanol produced per gram of cells.

[0054] The term “effective titer” as used herein, refers to the total amount of a particular alcohol (e.g., butanol) produced by fermentation per liter of fermentation medium. The total amount of butanol includes: (i) the amount of butanol in the fermentation medium; (ii) the amount of butanol recovered from the organic extractant; and (iii) the amount of butanol recovered from the gas phase, if gas stripping is used.

[0055] The term “effective rate” as used herein, refers to the total amount of alcohol (e.g., butanol) produced by fermentation per liter of fermentation medium per hour of fermentation.

[0056] The term “effective yield” as used herein, refers to the amount of alcohol (e.g., butanol) produced per unit of fermentable carbon substrate consumed by the biocatalyst.

[0057] The term “separation” as used herein is synonymous with “recovery” and refers to removing a chemical compound from an initial mixture to obtain the compound in greater purity or at a higher concentration than the purity or concentration of the compound in the initial mixture.

[0058] The term “aqueous phase,” as used herein, refers to the aqueous phase of a biphasic mixture obtained by contacting a fermentation broth with a water-immiscible organic extractant. In an embodiment of a process described herein that includes fermentative extraction, the term “fermentation broth” then specifically refers to the aqueous phase in biphasic fermentative extraction.

[0059] The term “organic phase,” as used herein, refers to the non-aqueous phase of a biphasic mixture obtained by contacting a fermentation broth with a water-immiscible organic extractant.

[0060] The term “carbon substrate” or “fermentable carbon substrate” refers to a carbon source capable of being metabolized by host organisms of the present invention and particularly carbon sources selected from the group consisting of monosaccharides, oligosaccharides, polysaccharides, and one-carbon substrates or mixtures thereof. Non-limiting examples of carbon substrates are provided herein and include, but are not limited to, monosaccharides, disaccharides, oligosaccharides, polysaccharides, ethanol, lactate, succinate, glycerol, carbon dioxide, methanol, glucose, fructose, lactose, sucrose, xylose, arabinose, dextrose, cellulose, methane, amino acids, or mixtures thereof.

[0061] "Fermentation broth" as used herein means the mixture of water, sugars (fermentable carbon sources), dissolved solids (if present), microorganisms producing alcohol, product alcohol and all other constituents of the material in which product alcohol is being made by the reaction of sugars to alcohol, water and carbon dioxide (CO₂) by the microorganisms present. From time to time, as used herein the term "fermentation medium" and "fermented mixture" can be used synonymously with "fermentation broth."

[0062] As used herein a "fermentor" refers to any container, containers, or apparatus that are used to ferment a substrate. A fermentor can contain a fermentation medium and microorganism capable of fermentation. The term "fermentation vessel" refers to the vessel in which the fermentation reaction is carried out whereby alcohol such as butanol is made. "Fermentor" can be used herein interchangeably with "fermentation vessel."

[0063] The term "fermentation product" includes any desired product of interest, including, but not limited to 1-butanol, 2-butanol, isobutanol, etc.

[0064] "Biomass" as used herein refers to a natural product containing a hydrolysable starch that provides a fermentable sugar, including any cellulosic or lignocellulosic material and materials comprising cellulose, and optionally further comprising hemicellulose, lignin, starch, oligosaccharides, disaccharides, and/or monosaccharides. Biomass can also comprise additional components, such as protein and/or lipids. Biomass can be derived from a single source, or biomass can comprise a mixture derived from more than one source. For example, biomass can comprise a mixture of corn cobs and corn stover, or a mixture of grass and leaves. Biomass includes, but is not limited to, bioenergy crops, agricultural residues, municipal solid waste, industrial solid waste, sludge from paper manufacture, yard waste, wood, and forestry waste. Examples of biomass include, but are not limited to, corn grain, corn cobs, crop residues such as corn husks, corn stover, grasses, wheat, rye, wheat straw, barley, barley straw, hay, rice straw, switchgrass, waste paper, sugar cane bagasse, sorghum, soy, components obtained from milling of grains, trees, branches, roots, leaves, wood chips, sawdust, shrubs and bushes, vegetables, fruits, flowers, animal manure, and mixtures thereof.

[0065] "Feedstock" as used herein means a product containing a fermentable carbon source. Suitable feedstock include, but are not limited to, rye, wheat, corn, corn mash, cane, cane mash, sugar cane, barley, cellulosic material, lignocellulosic material, and mixtures thereof.

[0066] The term “biomass” as used herein, in some instances, refers to the mass of the culture, e.g., the amount of recombinant microorganisms, typically provided in units of grams per liter (g/l) dry cell weight (dcw).

[0067] The term “aerobic conditions” as used herein means growth conditions in the presence of oxygen.

[0068] The term “microaerobic conditions” as used herein means growth conditions with low levels of oxygen (i.e., below normal atmospheric oxygen levels).

[0069] The term “anaerobic conditions” as used herein means growth conditions in the absence of oxygen.

[0070] The term “polynucleotide” is intended to encompass a singular nucleic acid as well as plural nucleic acids, and refers to a nucleic acid molecule or construct, e.g., messenger RNA (mRNA) or plasmid DNA (pDNA). A polynucleotide can contain the nucleotide sequence of the full-length cDNA sequence, or a fragment thereof, including the untranslated 5' and 3' sequences and the coding sequences. The polynucleotide can be composed of any polyribonucleotide or polydeoxyribonucleotide, which can be unmodified RNA or DNA or modified RNA or DNA. For example, polynucleotides can be composed of single- and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that can be single-stranded or, more typically, double-stranded or a mixture of single- and double-stranded regions. “Polynucleotide” embraces chemically, enzymatically, or metabolically modified forms.

[0071] A polynucleotide sequence can be referred to as “isolated,” in which it has been removed from its native environment. For example, a heterologous polynucleotide encoding a polypeptide or polypeptide fragment having ALS activity contained in a vector is considered isolated for the purposes of the present invention. Further examples of an isolated polynucleotide include recombinant polynucleotides maintained in heterologous host cells or purified (partially or substantially) polynucleotides in solution. Isolated polynucleotides or nucleic acids according to the present invention further include such molecules produced synthetically. An isolated polynucleotide fragment in the form of a polymer of DNA can be comprised of one or more segments of cDNA, genomic DNA or synthetic DNA.”

[0072] As used herein, “reduced activity” refers to any measurable decrease in a known biological activity of a polypeptide when compared to the same biological activity of the

polypeptide prior to the change resulting in the reduced activity. Such a change can include a modification of a polypeptide or a polynucleotide encoding a polypeptide as described herein. A reduced activity of a polypeptide disclosed herein can be determined by methods well known in the art and disclosed herein. Reduced activity of an enzyme refers to down-regulation, whether partial or total, of the activity of the enzyme as compared to the activity of the wildtype enzyme. Down-regulation may occur when a native gene has a "disruption" or "modification," referring to an insertion, deletion, or targeted mutation within a portion of that gene, that results in *e.g.*, a complete gene knockout such that the gene is deleted from the genome and no protein is translated or a translated subunit protein having an insertion, deletion, amino acid substitution or other targeted mutation. The location of the modification in the protein may be, for example, within the N-terminal portion of the protein or within the C-terminal portion of the protein. The modified protein will have impaired activity with respect to the protein that was not disrupted, and can be non-functional. Reduced activity in an enzyme could also result via manipulating the upstream regulatory domains or by use of sense, antisense or RNAi technology, etc. Another mechanism of reducing activity of an enzyme is introduction of a mutation that alters kinetic properties of the enzyme (*e.g.*, reducing the affinity for a substrate, lowering the k_{cat} , etc.).

[0073] As used herein, "eliminated activity" refers to the complete abolishment of a known biological activity of a polypeptide when compared to the same biological activity of the polypeptide prior to the change resulting in the eliminated activity. Such a change can include a modification of a polypeptide or a polynucleotide encoding a polypeptide as described herein. An eliminated activity includes a biological activity of a polypeptide that is not measurable when compared to the same biological activity of the polypeptide prior to the change resulting in the eliminated activity. An eliminated activity of a polypeptide disclosed herein can be determined by methods well known in the art and disclosed herein.

[0074] The terms "PDC-," "PDC knockout," or "PDC-KO" as used herein refer to a cell that has a genetic modification to inactivate or reduce expression of a gene encoding pyruvate decarboxylase (PDC) so that the cell substantially or completely lacks pyruvate decarboxylase enzyme activity. If the cell has more than one expressed (active) PDC gene, then each of the active PDC genes can be inactivated or have minimal expression thereby producing a PDC- cell.

[0075] The term "specific activity" as used herein is defined as the units of activity in a given amount of protein. Thus, the specific activity is not directly measured but is calculated by dividing 1) the activity in units/ml of the enzyme sample by 2) the concentration of protein in

that sample, so the specific activity is expressed as units/mg, where an enzyme unit is defined as moles of product formed/minute. The specific activity of a sample of pure, fully active enzyme is a characteristic of that enzyme. The specific activity of a sample of a mixture of proteins is a measure of the relative fraction of protein in that sample that is composed of the active enzyme of interest.

[0076] The terms “ k_{cat} ” and “ K_M ” are known to those skilled in the art and are described in *Enzyme Structure and Mechanism*, 2nd ed. (Ferst; W.H. Freeman Press, NY, 1985; pp 98-120). K_M , the Michaelis constant, is the concentration of substrate that leads to half-maximal velocity. The term “ k_{cat} ”, often called the “turnover number”, is defined as the maximum number of substrate molecules converted to products per active site per unit time, or the number of times the enzyme turns over per unit time. $k_{\text{cat}} = V_{\text{max}}/[E]$, where $[E]$ is the enzyme concentration (Ferst, *supra*). The terms “total turnover” and “total turnover number” are used herein to refer to the amount of product formed by the reaction of an enzyme with substrate.

[0077] The term “catalytic efficiency” is defined as the k_{cat}/K_M of an enzyme. Catalytic efficiency is used to quantify the specificity of an enzyme for a substrate.

[0078] The term “isolated nucleic acid molecule”, “isolated nucleic acid fragment” and “genetic construct” are used interchangeably and mean a polymer of RNA or DNA that is single or double-stranded, optionally containing synthetic, non natural or altered nucleotide bases. An isolated nucleic acid fragment in the form of a polymer of DNA can be comprised of one or more segments of cDNA, genomic DNA or synthetic DNA.

[0079] The term “amino acid” refers to the basic chemical structural unit of a protein or polypeptide. The abbreviations in Table 1 are used herein to identify specific amino acids.

Table 1: Amino acids and abbreviations thereof.

<u>Amino Acid</u>	<u>Three-Letter Abbreviation</u>	<u>One-Letter Abbreviation</u>
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Glutamine	Gln	Q
Glutamic acid	Glu	E
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

[0080] The term “gene” refers to a nucleic acid fragment that is capable of being expressed as a specific protein, optionally including regulatory sequences preceding (5’ non-coding sequences) and following (3’ non-coding sequences) the coding sequence. “Native gene” refers to a gene as found in nature with its own regulatory sequences. “Chimeric gene” refers to any gene that is not a native gene, comprising regulatory and coding sequences that are not found together in nature. Accordingly, a chimeric gene can comprise regulatory sequences and coding sequences that are derived from different sources, or regulatory sequences and coding sequences derived from the same source, but arranged in a manner different than that found in nature. “Endogenous gene” refers to a native gene in its natural location in the genome of a microorganism. A “foreign” gene refers to a gene not normally found in the host microorganism,

but that is introduced into the host microorganism by gene transfer. Foreign genes can comprise native genes inserted into a non-native microorganism, or chimeric genes. A “transgene” is a gene that has been introduced into the genome by a transformation procedure.

[0081] As used herein, “native” refers to the form of a polynucleotide, gene, or polypeptide as found in nature with its own regulatory sequences, if present.

[0082] As used herein the term “coding sequence” or “coding region” refers to a DNA sequence that encodes for a specific amino acid sequence.

[0083] As used herein, “endogenous” refers to the native form of a polynucleotide, gene or polypeptide in its natural location in the organism or in the genome of an organism.

“Endogenous polynucleotide” includes a native polynucleotide in its natural location in the genome of an organism. “Endogenous gene” includes a native gene in its natural location in the genome of an organism. “Endogenous polypeptide” includes a native polypeptide in its natural location in the organism transcribed and translated from a native polynucleotide or gene in its natural location in the genome of an organism.

[0084] The term “heterologous” when used in reference to a polynucleotide, a gene, or a polypeptide refers to a polynucleotide, gene, or polypeptide not normally found in the host organism. “Heterologous” also includes a native coding region, or portion thereof, that is reintroduced into the source organism in a form that is different from the corresponding native gene, *e.g.*, not in its natural location in the organism’s genome. The heterologous polynucleotide or gene can be introduced into the host organism by, *e.g.*, gene transfer. A heterologous gene can include a native coding region with non-native regulatory regions that is reintroduced into the native host. For example, a heterologous gene can include a native coding region that is a portion of a chimeric gene including non-native regulatory regions that is reintroduced into the native host. “Heterologous polypeptide” includes a native polypeptide that is reintroduced into the source organism in a form that is different from the corresponding native polypeptide. A “heterologous” polypeptide or polynucleotide can also include an engineered polypeptide or polynucleotide that comprises a difference from the “native” polypeptide or polynucleotide, *e.g.*, a point mutation within the endogenous polynucleotide can result in the production of a “heterologous” polypeptide. As used herein a “chimeric gene,” a “foreign gene,” and a “transgene,” can all be examples of “heterologous” genes.

[0085] A “transgene” is a gene that has been introduced into the genome by a transformation procedure.

[0086] As used herein, the term “modification” refers to a change in a polynucleotide disclosed herein that results in reduced or eliminated activity of a polypeptide encoded by the polynucleotide, as well as a change in a polypeptide disclosed herein that results in reduced or eliminated activity of the polypeptide. Such changes can be made by methods well known in the art, including, but not limited to, deleting, mutating (*e.g.*, spontaneous mutagenesis, random mutagenesis, mutagenesis caused by mutator genes, or transposon mutagenesis), substituting, inserting, down-regulating, altering the cellular location, altering the state of the polynucleotide or polypeptide (*e.g.*, methylation, phosphorylation or ubiquitination), removing a cofactor, introduction of an antisense RNA/DNA, introduction of an interfering RNA/DNA, chemical modification, covalent modification, irradiation with UV or X-rays, homologous recombination, mitotic recombination, promoter replacement methods, and/or combinations thereof. Guidance in determining which nucleotides or amino acid residues can be modified can be found by comparing the sequence of the particular polynucleotide or polypeptide with that of homologous polynucleotides or polypeptides, *e.g.*, yeast or bacterial, and maximizing the number of modifications made in regions of high homology (conserved regions) or consensus sequences.

[0087] The term “recombinant genetic expression element” refers to a nucleic acid fragment that expresses one or more specific proteins, including regulatory sequences preceding (5' non-coding sequences) and following (3' termination sequences) coding sequences for the proteins. A chimeric gene is a recombinant genetic expression element. The coding regions of an operon can form a recombinant genetic expression element, along with an operably linked promoter and termination region.

[0088] “Regulatory sequences” refers to nucleotide sequences located upstream (5' non-coding sequences), within, or downstream (3' non-coding sequences) of a coding sequence, and which influence the transcription, RNA processing or stability, or translation of the associated coding sequence. Regulatory sequences can include promoters, enhancers, operators, repressors, transcription termination signals, translation leader sequences, introns, polyadenylation recognition sequences, RNA processing site, effector binding site and stem-loop structure.

[0089] The term “promoter” refers to a nucleic acid sequence capable of controlling the expression of a coding sequence or functional RNA. In general, a coding sequence is located 3' to a promoter sequence. Promoters can be derived in their entirety from a native gene, or be composed of different elements derived from different promoters found in nature, or even comprise synthetic nucleic acid segments. It is understood by those skilled in the art that

different promoters can direct the expression of a gene in different tissues or cell types, or at different stages of development, or in response to different environmental or physiological conditions. Promoters which cause a gene to be expressed in most cell types at most times are commonly referred to as “constitutive promoters”. “Inducible promoters,” on the other hand, cause a gene to be expressed when the promoter is induced or turned on by a promoter-specific signal or molecule. It is further recognized that since in most cases the exact boundaries of regulatory sequences have not been completely defined, DNA fragments of different lengths can have identical promoter activity. For example, it will be understood that “FBA1 promoter” can be used to refer to a fragment derived from the promoter region of the FBA1 gene.

[0090] The term “terminator” as used herein refers to DNA sequences located downstream of a coding sequence. This includes polyadenylation recognition sequences and other sequences encoding regulatory signals capable of affecting mRNA processing or gene expression. The polyadenylation signal is usually characterized by affecting the addition of polyadenylic acid tracts to the 3’ end of the mRNA precursor. The 3’ region can influence the transcription, RNA processing or stability, or translation of the associated coding sequence. It is recognized that since in most cases the exact boundaries of regulatory sequences have not been completely defined, DNA fragments of different lengths can have identical terminator activity. For example, it will be understood that “CYC1 terminator” can be used to refer to a fragment derived from the terminator region of the CYC1 gene.

[0091] The term “operably linked” refers to the association of nucleic acid sequences on a single nucleic acid fragment so that the function of one is affected by the other. For example, a promoter is operably linked with a coding sequence when it is capable of effecting the expression of that coding sequence (i.e., that the coding sequence is under the transcriptional control of the promoter). Coding sequences can be operably linked to regulatory sequences in sense or antisense orientation.

[0092] The term “expression”, as used herein, refers to the transcription and stable accumulation of sense (mRNA) or antisense RNA derived from the nucleic acid fragment of the invention. Expression can also refer to translation of mRNA into a polypeptide.

[0093] The term “overexpression,” as used herein, refers to expression that is higher than endogenous expression of the same or related gene. A heterologous gene is overexpressed if its expression is higher than that of a comparable endogenous gene.

[0094] The term overexpression refers to an increase in the level of nucleic acid or protein in a host cell. Thus, overexpression can result from increasing the level of transcription or translation of an endogenous sequence in a host cell or can result from the introduction of a heterologous sequence into a host cell. Overexpression can also result from increasing the stability of a nucleic acid or protein sequence.

[0095] As used herein the term “transformation” refers to the transfer of a nucleic acid fragment into the genome of a host microorganism, resulting in genetically stable inheritance. Host microorganisms containing the transformed nucleic acid fragments are referred to as “transgenic” or “recombinant” or “transformed” microorganisms.

[0096] The terms “plasmid”, “vector” and “cassette” refer to an extra chromosomal element often carrying genes which are not part of the central metabolism of the cell, and usually in the form of circular double-stranded DNA fragments. Such elements can be autonomously replicating sequences, genome integrating sequences, phage or nucleotide sequences, linear or circular, of a single- or double-stranded DNA or RNA, derived from any source, in which a number of nucleotide sequences have been joined or recombined into a unique construction which is capable of introducing a promoter fragment and DNA sequence for a selected gene product along with appropriate 3' untranslated sequence into a cell. “Transformation cassette” refers to a specific vector containing a foreign gene and having elements in addition to the foreign gene that facilitates transformation of a particular host cell. “Expression cassette” refers to a specific vector containing a foreign gene and having elements in addition to the foreign gene that allow for enhanced expression of that gene in a foreign host.

[0097] As used herein the term “codon degeneracy” refers to the nature in the genetic code permitting variation of the nucleotide sequence without affecting the amino acid sequence of an encoded polypeptide. The skilled artisan is well aware of the “codon-bias” exhibited by a specific host cell in usage of nucleotide codons to specify a given amino acid. Therefore, when synthesizing a gene for improved expression in a host cell, it is desirable to design the gene such that its frequency of codon usage approaches the frequency of preferred codon usage of the host cell.

[0098] The term “codon-optimized” as it refers to genes or coding regions of nucleic acid molecules for transformation of various hosts, refers to the alteration of codons in the gene or coding regions of the nucleic acid molecules to reflect the typical codon usage of the host organism without altering the polypeptide encoded by the DNA. Such optimization includes

replacing at least one, or more than one, or a significant number, of codons with one or more codons that are more frequently used in the genes of that organism.

[0099] Deviations in the nucleotide sequence that comprise the codons encoding the amino acids of any polypeptide chain allow for variations in the sequence coding for the gene. Since each codon consists of three nucleotides, and the nucleotides comprising DNA are restricted to four specific bases, there are 64 possible combinations of nucleotides, 61 of which encode amino acids (the remaining three codons encode signals ending translation). The “genetic code” which shows which codons encode which amino acids is reproduced herein as Table 2A. As a result, many amino acids are designated by more than one codon. For example, the amino acids alanine and proline are coded for by four triplets, serine and arginine by six, whereas tryptophan and methionine are coded by just one triplet. This degeneracy allows for DNA base composition to vary over a wide range without altering the amino acid sequence of the proteins encoded by the DNA.

Table 2A: The Standard Genetic Code

	T	C	A	G
T	TTT Phe (F) TTC “ TTA Leu (L) TTG “	TCT Ser (S) TCC “ TCA “ TCG “	TAT Tyr (Y) TAC “ TAA Stop TAG Stop	TGT Cys (C) TGC TGA Stop TGG Trp (W)
C	CTT Leu (L) CTC “ CTA “ CTG “	CCT Pro (P) CCC “ CCA “ CCG “	CAT His (H) CAC “ CAA Gln (Q) CAG “	CGT Arg (R) CGC “ CGA “ CGG “
A	ATT Ile (I) ATC “ ATA “ ATG Met (M)	ACT Thr (T) ACC “ ACA “ ACG “	AAT Asn (N) AAC “ AAA Lys (K) AAG “	AGT Ser (S) AGC “ AGA Arg (R) AGG “
G	GTT Val (V) GTC “ GTA “ GTG “	GCT Ala (A) GCC “ GCA “ GCG “	GAT Asp (D) GAC “ GAA Glu (E) GAG “	GGT Gly (G) GGC “ GGA “ GGG “

[00100] Many organisms display a bias for use of particular codons to code for insertion of a particular amino acid in a growing peptide chain. Codon preference, or codon bias, differences in codon usage between organisms, is afforded by degeneracy of the genetic code, and is well documented among many organisms. Codon bias often correlates with the efficiency of translation of messenger RNA (mRNA), which is in turn believed to be dependent on, *inter alia*,

the properties of the codons being translated and the availability of particular transfer RNA (tRNA) molecules. The predominance of selected tRNAs in a cell is generally a reflection of the codons used most frequently in peptide synthesis. Accordingly, genes can be tailored for optimal gene expression in a given organism based on codon optimization.

[00101] Given the large number of gene sequences available for a wide variety of animal, plant and microbial species, it is possible to calculate the relative frequencies of codon usage. Codon usage tables are readily available, for example, at the “Codon Usage Database” available at <http://www.kazusa.or.jp/codon/> (visited March 20, 2008), and these tables can be adapted in a number of ways. See Nakamura, Y., *et al. Nucl. Acids Res.* 28:292 (2000). Codon usage tables for yeast, calculated from GenBank Release 128.0 [15 February 2002], are reproduced below as Table 2B. This table uses mRNA nomenclature, and so instead of thymine (T) which is found in DNA, the tables use uracil (U) which is found in RNA. Table 2B has been adapted so that frequencies are calculated for each amino acid, rather than for all 64 codons.

Table 2B: Codon Usage Table for *Saccharomyces cerevisiae*.

Amino Acid	Codon	Number	Frequency per thousand
Phe	UUU	170666	26.1
Phe	UUC	120510	18.4
Leu	UUA	170884	26.2
Leu	UUG	177573	27.2
Leu	CUU	80076	12.3
Leu	CUC	35545	5.4
Leu	CUA	87619	13.4
Leu	CUG	68494	10.5
Ile	AUU	196893	30.1
Ile	AUC	112176	17.2
Ile	AUA	116254	17.8
Met	AUG	136805	20.9
Val	GUU	144243	22.1
Val	GUC	76947	11.8
Val	GUA	76927	11.8
Val	GUG	70337	10.8
Ser	UCU	153557	23.5
Ser	UCC	92923	14.2

Amino Acid	Codon	Number	Frequency per thousand
Ser	UCA	122028	18.7
Ser	UCG	55951	8.6
Ser	AGU	92466	14.2
Ser	AGC	63726	9.8
Pro	CCU	88263	13.5
Pro	CCC	44309	6.8
Pro	CCA	119641	18.3
Pro	CCG	34597	5.3
Thr	ACU	132522	20.3
Thr	ACC	83207	12.7
Thr	ACA	116084	17.8
Thr	ACG	52045	8.0
Ala	GCU	138358	21.2
Ala	GCC	82357	12.6
Ala	GCA	105910	16.2
Ala	GCG	40358	6.2
Tyr	UAU	122728	18.8
Tyr	UAC	96596	14.8
His	CAU	89007	13.6
His	CAC	50785	7.8
Gln	CAA	178251	27.3
Gln	CAG	79121	12.1
Asn	AAU	233124	35.7
Asn	AAC	162199	24.8
Lys	AAA	273618	41.9
Lys	AAG	201361	30.8
Asp	GAU	245641	37.6
Asp	GAC	132048	20.2
Glu	GAA	297944	45.6
Glu	GAG	125717	19.2
Cys	UGU	52903	8.1
Cys	UGC	31095	4.8

Amino Acid	Codon	Number	Frequency per thousand
Trp	UGG	67789	10.4
Arg	CGU	41791	6.4
Arg	CGC	16993	2.6
Arg	CGA	19562	3.0
Arg	CGG	11351	1.7
Arg	AGA	139081	21.3
Arg	AGG	60289	9.2
Gly	GGU	156109	23.9
Gly	GGC	63903	9.8
Gly	GGA	71216	10.9
Gly	GGG	39359	6.0
Stop	UAA	6913	1.1
Stop	UAG	3312	0.5
Stop	UGA	4447	0.7

[00102] By utilizing this or similar tables, one of ordinary skill in the art can apply the frequencies to any given polypeptide sequence, and produce a nucleic acid fragment of a codon-optimized coding region which encodes the polypeptide, but which uses codons optimal for a given species.

[00103] Randomly assigning codons at an optimized frequency to encode a given polypeptide sequence, can be done manually by calculating codon frequencies for each amino acid, and then assigning the codons to the polypeptide sequence randomly. Additionally, various algorithms and computer software programs are readily available to those of ordinary skill in the art. For example, the “EditSeq” function in the Lasergene Package, available from DNASTar, Inc., Madison, WI, the backtranslation function in the VectorNTI Suite, available from InforMax, Inc., Bethesda, MD, and the “backtranslate” function in the GCG-Wisconsin Package, available from Accelrys, Inc., San Diego, CA. In addition, various resources are publicly available to codon-optimize coding region sequences, *e.g.*, the “backtranslation” function at <http://www.entelechon.com/bioinformatics/backtranslation.php?lang=eng> (visited April 15, 2008) and the “backtranseq” function available at <http://bioinfo.pbi.nrc.ca:8090/EMBOSS/index.html> (visited July 9, 2002). Constructing a rudimentary algorithm to assign codons based on a given frequency can also easily be accomplished with basic mathematical functions by one of ordinary skill in the art.

[00104] Codon-optimized coding regions can be designed by various methods known to those skilled in the art including software packages such as “synthetic gene designer” (userpages.umbc.edu/~wug1/codon/sgd/, visited March 19, 2012).

[00105] A polynucleotide or nucleic acid fragment is “hybridizable” to another nucleic acid fragment, such as a cDNA, genomic DNA, or RNA molecule, when a single-stranded form of the nucleic acid fragment can anneal to the other nucleic acid fragment under the appropriate conditions of temperature and solution ionic strength. Hybridization and washing conditions are well known and exemplified in Sambrook, J., Fritsch, E. F. and Maniatis, T. *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory: Cold Spring Harbor, NY (1989), particularly Chapter 11 and Table 11.1 therein (entirely incorporated herein by reference). The conditions of temperature and ionic strength determine the “stringency” of the hybridization. Stringency conditions can be adjusted to screen for moderately similar fragments (such as homologous sequences from distantly related organisms), to highly similar fragments (such as genes that duplicate functional enzymes from closely related organisms). Post hybridization washes determine stringency conditions. One set of conditions uses a series of washes starting with 6X SSC, 0.5% SDS at room temperature for 15 min, then repeated with 2X SSC, 0.5% SDS at 45 °C for 30 min, and then repeated twice with 0.2X SSC, 0.5% SDS at 50 °C for 30 min. Another set of stringent conditions uses higher temperatures in which the washes are identical to those above except for the temperature of the final two 30 min washes in 0.2X SSC, 0.5% SDS was increased to 60 °C. Another set of highly stringent conditions uses two final washes in 0.1X SSC, 0.1% SDS at 65 °C. An additional set of stringent conditions include hybridization at 0.1X SSC, 0.1% SDS, 65 °C and washes with 2X SSC, 0.1% SDS followed by 0.1X SSC, 0.1% SDS, for example.

[00106] Hybridization requires that the two nucleic acids contain complementary sequences, although depending on the stringency of the hybridization, mismatches between bases are possible. The appropriate stringency for hybridizing nucleic acids depends on the length of the nucleic acids and the degree of complementation, variables well known in the art. The greater the degree of similarity or homology between two nucleotide sequences, the greater the value of T_m for hybrids of nucleic acids having those sequences. The relative stability (corresponding to higher T_m) of nucleic acid hybridizations decreases in the following order: RNA:RNA, DNA:RNA, DNA:DNA. For hybrids of greater than 100 nucleotides in length, equations for calculating T_m have been derived (*see Sambrook et al., supra*, 9.50-9.51). For

hybridizations with shorter nucleic acids, i.e., oligonucleotides, the position of mismatches becomes more important, and the length of the oligonucleotide determines its specificity (*see* Sambrook *et al.*, *supra*, 11.7 11.8). In one embodiment the length for a hybridizable nucleic acid is at least about 10 nucleotides. In one embodiment, a minimum length for a hybridizable nucleic acid is at least about 15 nucleotides; at least about 20 nucleotides; or the length is at least about 30 nucleotides. Furthermore, the skilled artisan will recognize that the temperature and wash solution salt concentration can be adjusted as necessary according to factors such as length of the probe.

[00107] As used herein, the term “polypeptide” is intended to encompass a singular “polypeptide” as well as plural “polypeptides,” and refers to a molecule composed of monomers (amino acids) linearly linked by amide bonds (also known as peptide bonds). The term “polypeptide” refers to any chain or chains of two or more amino acids, and does not refer to a specific length of the product. Thus, peptides, dipeptides, tripeptides, oligopeptides, “protein,” “amino acid chain,” or any other term used to refer to a chain or chains of two or more amino acids, are included within the definition of “polypeptide,” and the term “polypeptide” can be used instead of, or interchangeably with any of these terms. A polypeptide can be derived from a natural biological source or produced by recombinant technology, but is not necessarily translated from a designated nucleic acid sequence. It can be generated in any manner, including by chemical synthesis.

[00108] By an “isolated” polypeptide or a fragment, variant, or derivative thereof is intended a polypeptide that is not in its natural milieu. No particular level of purification is required. For example, an isolated polypeptide can be removed from its native or natural environment. Recombinantly produced polypeptides and proteins expressed in host cells are considered isolated for purposes of the invention, as are native or recombinant polypeptides which have been separated, fractionated, or partially or substantially purified by any suitable technique.

[00109] As used herein, the terms “variant” and “mutant” are synonymous and refer to a polypeptide differing from a specifically recited polypeptide by one or more amino acid insertions, deletions, mutations, and substitutions, created using, *e.g.*, recombinant DNA techniques, such as mutagenesis. Guidance in determining which amino acid residues can be replaced, added, or deleted without abolishing activities of interest, can be found by comparing the sequence of the particular polypeptide with that of homologous polypeptides, *e.g.*, yeast or

bacterial, and minimizing the number of amino acid sequence changes made in regions of high homology (conserved regions) or by replacing amino acids with consensus sequences.

[00110] “Engineered polypeptide” as used herein refers to a polypeptide that is synthetic, *i.e.*, differing in some manner from a polypeptide found in nature.

[00111] Alternatively, recombinant polynucleotide variants encoding these same or similar polypeptides can be synthesized or selected by making use of the “redundancy” in the genetic code. Various codon substitutions, such as silent changes which produce various restriction sites, can be introduced to optimize cloning into a plasmid or viral vector for expression. Mutations in the polynucleotide sequence can be reflected in the polypeptide or domains of other peptides added to the polypeptide to modify the properties of any part of the polypeptide. For example, mutations can be used to reduce or eliminate expression of a target protein and include, but are not limited to, deletion of the entire gene or a portion of the gene, inserting a DNA fragment into the gene (in either the promoter or coding region) so that the protein is not expressed or expressed at lower levels, introducing a mutation into the coding region which adds a stop codon or frame shift such that a functional protein is not expressed, and introducing one or more mutations into the coding region to alter amino acids so that a non-functional or a less enzymatically active protein is expressed.

[00112] Amino acid “substitutions” can be the result of replacing one amino acid with another amino acid having similar structural and/or chemical properties, *i.e.*, conservative amino acid replacements, or they can be the result of replacing one amino acid with an amino acid having different structural and/or chemical properties, *i.e.*, non-conservative amino acid replacements. “Conservative” amino acid substitutions can be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, or the amphipathic nature of the residues involved. For example, nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan, and methionine; polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine, and glutamine; positively charged (basic) amino acids include arginine, lysine, and histidine; and negatively charged (acidic) amino acids include aspartic acid and glutamic acid. Alternatively, “non-conservative” amino acid substitutions can be made by selecting the differences in polarity, charge, solubility, hydrophobicity, hydrophilicity, or the amphipathic nature of any of these amino acids. “Insertions” or “deletions” can be within the range of variation as structurally or functionally tolerated by the recombinant proteins. The variation allowed can be experimentally determined

by systematically making insertions, deletions, or substitutions of amino acids in a polypeptide molecule using recombinant DNA techniques and assaying the resulting recombinant variants for activity.

[00113] A “substantial portion” of an amino acid or nucleotide sequence is that portion comprising enough of the amino acid sequence of a polypeptide or the nucleotide sequence of a gene to putatively identify that polypeptide or gene, either by manual evaluation of the sequence by one skilled in the art, or by computer-automated sequence comparison and identification using algorithms such as BLAST (Altschul, S. F., *et al.*, *J. Mol. Biol.*, 215:403-410 (1993)). In general, a sequence of ten or more contiguous amino acids or thirty or more nucleotides is necessary in order to putatively identify a polypeptide or nucleic acid sequence as homologous to a known protein or gene. Moreover, with respect to nucleotide sequences, gene specific oligonucleotide probes comprising 20-30 contiguous nucleotides can be used in sequence-dependent methods of gene identification (*e.g.*, Southern hybridization) and isolation (*e.g.*, *in situ* hybridization of bacterial colonies or bacteriophage plaques). In addition, short oligonucleotides of 12-15 bases can be used as amplification primers in PCR in order to obtain a particular nucleic acid fragment comprising the primers. Accordingly, a “substantial portion” of a nucleotide sequence comprises enough of the sequence to specifically identify and/or isolate a nucleic acid fragment comprising the sequence. The instant specification teaches the complete amino acid and nucleotide sequence encoding particular proteins. The skilled artisan, having the benefit of the sequences as reported herein, can now use all or a substantial portion of the disclosed sequences for purposes known to those skilled in this art. Accordingly, the instant invention comprises the complete sequences as reported in the accompanying Sequence Listing, as well as substantial portions of those sequences as defined above.

[00114] The term “complementary” is used to describe the relationship between nucleotide bases that are capable of hybridizing to one another. For example, with respect to DNA, adenine is complementary to thymine and cytosine is complementary to guanine, and with respect to RNA, adenine is complementary to uracil and cytosine is complementary to guanine.

[00115] The term “percent identity”, as known in the art, is a relationship between two or more polypeptide sequences or two or more polynucleotide sequences, as determined by comparing the sequences. In the art, “identity” also means the degree of sequence relatedness between polypeptide or polynucleotide sequences, as the case may be, as determined by the match between strings of such sequences. “Identity” and “similarity” can be readily calculated

by known methods, including but not limited to those described in: 1.) Computational Molecular Biology (Lesk, A. M., Ed.) Oxford University: NY (1988); 2.) Biocomputing: Informatics and Genome Projects (Smith, D. W., Ed.) Academic: NY (1993); 3.) Computer Analysis of Sequence Data, Part I (Griffin, A. M., and Griffin, H. G., Eds.) Humana: NJ (1994); 4.) Sequence Analysis in Molecular Biology (von Heinje, G., Ed.) Academic (1987); and 5.) Sequence Analysis Primer (Gribskov, M. and Devereux, J., Eds.) Stockton: NY (1991).

[00116] Methods to determine identity are designed to give the best match between the sequences tested. Methods to determine identity and similarity are codified in publicly available computer programs. Sequence alignments and percent identity calculations can be performed using the MegAlign™ program of the LASERGENE bioinformatics computing suite (DNASTAR Inc., Madison, WI). Multiple alignments of the sequences are performed using the “Clustal method of alignment” which encompasses several varieties of the algorithm including the “Clustal V method of alignment” corresponding to the alignment method labeled Clustal V (described by Higgins and Sharp, *CABIOS*. 5:151-153 (1989); Higgins, D.G. *et al.*, *Comput. Appl. Biosci.*, 8:189-191 (1992)) and found in the MegAlign™ program of the LASERGENE bioinformatics computing suite (DNASTAR Inc.). For multiple alignments, the default values correspond to GAP PENALTY=10 and GAP LENGTH PENALTY=10. Default parameters for pairwise alignments and calculation of percent identity of protein sequences using the Clustal method are KTUPLE=1, GAP PENALTY=3, WINDOW=5 and DIAGONALS SAVED=5. For nucleic acids these parameters are KTUPLE=2, GAP PENALTY=5, WINDOW=4 and DIAGONALS SAVED=4. After alignment of the sequences using the Clustal V program, it is possible to obtain a “percent identity” by viewing the “sequence distances” table in the same program. Additionally the “Clustal W method of alignment” is available and corresponds to the alignment method labeled Clustal W (described by Higgins and Sharp, *CABIOS*. 5:151-153 (1989); Higgins, D.G. *et al.*, *Comput. Appl. Biosci.* 8:189-191(1992)) and found in the MegAlign™ v6.1 program of the LASERGENE bioinformatics computing suite (DNASTAR Inc.). Default parameters for multiple alignment (GAP PENALTY=10, GAP LENGTH PENALTY=0.2, Delay Divergen Seqs(%)=30, DNA Transition Weight=0.5, Protein Weight Matrix=Gonnet Series, DNA Weight Matrix=IUB). After alignment of the sequences using the Clustal W program, it is possible to obtain a “percent identity” by viewing the “sequence distances” table in the same program.

[00117] It is well understood by one skilled in the art that many levels of sequence identity are useful in identifying polypeptides, such as from other species, wherein such polypeptides have the same or similar function or activity, or in describing the corresponding polynucleotides. Useful examples of percent identities include, but are not limited to: 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95%, or any integer percentage from 55% to 100% can be useful in describing the present invention, such as 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99%. Suitable polynucleotide fragments not only have the above homologies but typically comprise a polynucleotide having at least 50 nucleotides, at least 100 nucleotides, at least 150 nucleotides, at least 200 nucleotides, or at least 250 nucleotides. Further, suitable polynucleotide fragments having the above homologies encode a polypeptide having at least 50 amino acids, at least 100 amino acids, at least 150 amino acids, at least 200 amino acids, or at least 250 amino acids.

[00118] The term “sequence analysis software” refers to any computer algorithm or software program that is useful for the analysis of nucleotide or amino acid sequences. “Sequence analysis software” can be commercially available or independently developed. Typical sequence analysis software will include, but is not limited to: 1.) the GCG suite of programs (Wisconsin Package Version 9.0, Genetics Computer Group (GCG), Madison, WI); 2.) BLASTP, BLASTN, BLASTX (Altschul *et al.*, *J. Mol. Biol.*, 215:403-410 (1990)); 3.) DNASTAR (DNASTAR, Inc. Madison, WI); 4.) Sequencher (Gene Codes Corporation, Ann Arbor, MI); and 5.) the FASTA program incorporating the Smith-Waterman algorithm (W. R. Pearson, *Comput. Methods Genome Res.*, [Proc. Int. Symp.] (1994), Meeting Date 1992, 111-20. Editor(s): Suhai, Sandor. Plenum: New York, NY). Within the context of this application it will be understood that where sequence analysis software is used for analysis, that the results of the analysis will be based on the “default values” of the program referenced, unless otherwise specified. As used herein “default values” will mean any set of values or parameters that originally load with the software when first initialized.

[00119] Standard recombinant DNA and molecular cloning techniques are well known in the art and are described by Sambrook, J., Fritsch, E. F. and Maniatis, T., *Molecular Cloning: A Laboratory Manual*, Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1989) (hereinafter “Maniatis”); and by Silhavy, T. J., Bennis, M. L. and Enquist, L. W.,

Experiments with Gene Fusions, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1984); and by Ausubel, F. M. *et al.*, *Current Protocols in Molecular Biology*, published by Greene Publishing Assoc. and Wiley-Interscience (1987). Additional methods used here are in *Methods in Enzymology*, Volume 194, *Guide to Yeast Genetics and Molecular and Cell Biology* (Part A, 2004, Christine Guthrie and Gerald R. Fink (Eds.), Elsevier Academic Press, San Diego, CA). Other molecular tools and techniques are known in the art and include splicing by overlapping extension polymerase chain reaction (PCR) (Yu, *et al.* (2004) *Fungal Genet. Biol.* 41:973–981), positive selection for mutations at the *URA3* locus of *Saccharomyces cerevisiae* (Boeke, J.D. *et al.* (1984) *Mol. Gen. Genet.* 197, 345-346; M A Romanos, *et al.* *Nucleic Acids Res.* 1991 January 11; 19(1): 187), the cre-lox site-specific recombination system as well as mutant lox sites and FLP substrate mutations (Sauer, B. (1987) *Mol Cell Biol* 7: 2087-2096; Senecoff, *et al.* (1988) *Journal of Molecular Biology*, Volume 201, Issue 2, Pages 405-421; Albert, *et al.* (1995) *The Plant Journal*. Volume 7, Issue 4, pages 649–659), “seamless” gene deletion (Akada, *et al.* (2006) *Yeast*;23(5):399-405), and gap repair methodology (Ma *et al.*, *Genetics* 58:201-216; 1981).

POLYPEPTIDES WITH GPD ACTIVITY

[00120] Endogenous NAD-dependent “glycerol-3-phosphate dehydrogenase” or “GPD” is a key enzyme in glycerol synthesis, converting dihydroxyacetone phosphate (DHAP) to glycerol-3-phosphate. The terms “glycerol-3-phosphate dehydrogenase” and “GPD” refer to any polypeptide (or polypeptides) having the biological function of GPD. Such polypeptides include polypeptides having an enzyme activity that catalyzes the conversion of dihydroxyacetone phosphate to glycerol-3-phosphate. GPDs are widespread in nature and can fall into three categories. In the first category, EC 1.1.1.8, a GPD is a soluble cytoplasmic enzyme where the redox cofactor is the NAD/NADH couple, GPDs in the EC 1.1.1.8 category are described as NADH specific, but this does not preclude that some of the GPDs may have measureable activity with NADPH. *Saccharomyces cerevisiae* GPD1 is an example of this type of GPD (Albertyn *et al.* Al, 1992, FEBS Lett 308: 130-132; Valadi, *et al.*, 2004, J. Biol Chem 279: 39677-39685). Another example is the human GPD1, for which there are multiple 3-dimensional structural studies (Ou *et al.*, 2005, J.Mol.Biol. 357: 858-869). Assays for enzymes in this category can utilize the spectrophotometric measurement of NADH oxidation in the presence of DHAP and the GPD enzyme (Niesel *et al.* 1982 *Methods Enzymol* 89: 296-301). The second category, EC

1.1.5.3, GPD enzymes are intrinsic membrane proteins of the mitochondrial inner membrane, and contain a flavin cofactor, and reducing equivalents are transferred to the quinone/quinol couple in the mitochondrion. There is a third minor category of GPDs, EC 1.1.1.94 which utilize either NADH or NADPH with substantially the same affinity. GPDs of the third minor category can also be feedback inhibited by glycerol-3-phosphate.

[00121] Recombinant microorganisms such as yeast can have one or more endogenous genes encoding glycerol-3-phosphate dehydrogenase (GPD). In some yeasts, such as *S. cerevisiae*, *S. pombe*, and *P. stipitis*, GPD1 and GPD2 are functional homologs. Any of the genes encoding GPD enzymes of yeast may be disrupted to reduce GPD activity in a yeast cell.

[00122] One of the key yield loss mechanisms in yeast butanol production is the loss of carbon and reducing equivalents that are diverted from glycolysis by the conversion of dihydroxyacetone phosphate to glycerol. Since GPD catalyzes the first step in this conversion of dihydroxyacetone phosphate to glycerol, the activity of GPD can contribute to the production of glycerol and the loss of butanol yield. As a result, some have considered eliminating the function of GPD (for example, by knocking out the gene encoding GPD protein) in butanol-producing yeast. However, glycerol is required for growth and is an important osmoprotectant. Thus, retaining the ability to make some glycerol offers certain advantages.

[00123] One way to retain the ability to make glycerol, but also improve the production of product alcohol is to alter the cofactor specificity of GPD. *Saccharomyces cerevisiae* GPD1 generally favors the cofactor nicotinamide adenine dinucleotide (“NADH”) in catalyzing the first step in the conversion of dihydroxyacetone phosphate to glycerol in a yeast cell. However, as demonstrated herein, GPD enzymes can also use the cofactor nicotinamide adenine dinucleotide phosphate (“NADPH”).

[00124] The use of GPD enzymes with preference for NADPH as compared to NADH can allow host cells to retain the ability to produce glycerol under different metabolic conditions when compared with enzymes with a preference for NADH. However, this glycerol production can advantageously be limited under anaerobic conditions when NADPH production is limited.

[00125] At the same time, decreasing the preference for NADH by GPD can increase the availability of NADH in a host cell. NADH is also used by other enzymes in a product alcohol production pathway, for example, in the isobutanol production pathway the available NADH can be utilized by KARI and alcohol dehydrogenase. Thus, decreasing the affinity of GPD for NADH can increase product alcohol (e.g., isobutanol) production. Accordingly, in some

embodiments, a heterologous and/or engineered GPD is expressed in a recombinant microorganism that also expresses an NADH-utilizing enzyme, for example, an NADH-utilizing enzyme that acts in the isobutanol production pathway such as KARI and alcohol dehydrogenase.

[00126] An additional way to improve the production of a product alcohol (e.g., butanol) is to alter the GPD to decrease the K_M for NADPH. Decreasing the K_M for NADPH by altering GPD can increase the rate of NADPH oxidation catalyzed by GPD, thus allowing an increase in the availability of NADH in the host cell. The available NADH can be used by other enzymes in the product alcohol production pathway, for example, in the isobutanol production pathway the available NADH can be utilized by KARI and alcohol dehydrogenase. Thus, increasing the affinity of GPD for NADPH can increase product alcohol (e.g., isobutanol) production. Accordingly, in some embodiments, a heterologous and/or engineered GPD is expressed in recombinant microorganism that also expresses other NADH-utilizing enzymes, for example, an NADH-utilizing enzyme that acts in the isobutanol production pathway such as KARI and alcohol dehydrogenase.

[00127] Another way to retain the ability to make some glycerol and also improve the production of product alcohol is to use heterologous GPD enzymes that can reduce the amount of glycerol produced as compared to the amount produced by the endogenous GPD enzymes. An example heterologous enzyme is *E. coli* gpsA. Two mechanistic features of *E. coli* gpsA that may contribute to its ability to produce less glycerol include (1) gpsA is product inhibited by glycerol-3-phosphate, and (2) gpsA utilizes the cofactors NADH and NADPH with substantially the same affinity (Edgar and Bell, JBC 255:3492-7 (1980)) and under certain conditions this can also allow for the production of glycerol using NADPH, thus allowing for the availability of NADH in the host cell. Product inhibition by glycerol-3-phosphate in *Saccharomyces* may result in reduced glycerol production, especially if the glycerol-3-phosphate phosphatase enzymatic reaction is slower than the GPD enzymatic reaction. The published Michaelis constants for the *Saccharomyces* phosphatases GPP1 and GPP2 are 3.1 and 3.9, respectively (Norbeck, JBC 271:13875-81 (1996), which is nearly 1000-fold higher than the inhibition constant (K_i) of glycerol-3-phosphate on *E. coli* gpsA (Edgar and Bell, JBC 253:6345-63 (1978)). Most conditions are conducive to product inhibition by glycerol-3-phosphate.

[00128] GPD enzymes that can utilize NADH or NADPH and/or are feedback inhibited by glycerol-3-phosphate can include both naturally occurring proteins and engineered proteins. For instance, NADH-utilizing or NADPH-utilizing GPD enzymes are described by EC 1.1.1.94 and

have been found in *Aspergillus oryzae*, *Candida versatilis*, *Escherichia coli*, and *Oryctolagus cuniculus*.

[00129] In some embodiments, the heterologous GPD used herein is a *Leishmania mexicana*, *Dunaliella viridis*, *Jaculus orientalis*, *Archeoglobus fulgidus*, *Rickettsia prowazekii*, *Beggiatoa alba*, *Kangiella koreensis* *Aspergillus oryzae*, *Candida versatilis*, *Escherichia coli*, or *Oryctolagus cuniculus* GPD.

[00130] In certain embodiments, the sequences of other GPD enzymes that can utilize either NADH or NADPH and/or are feedback inhibited by glycerol-3-phosphate can be identified in the literature and candidates can be identified in bioinformatics databases well known to the skilled person using sequences disclosed herein and available in the art. For example, such sequences can be identified through BLAST searching of publicly available databases with known GPD encoding polynucleotide or polypeptide sequences. In such a method, identities can be based on the Clustal W method of alignment using the default parameters of GAP PENALTY = 10, GAP LENGTH PENALTY = 0.1, and Gonnet 250 series of protein weight matrix.

[00131] Additionally, GPD polynucleotide or polypeptide sequences disclosed herein or known in the art can be used to identify other candidate GPD homologs in nature. For example, the GPD encoding nucleic acid sequences disclosed herein or known in the art can be used to isolate genes encoding homologous proteins. Isolation of homologous genes using sequence-dependent protocols include, but are not limited to (1) methods of nucleic acid hybridization; (2) methods of DNA and RNA amplification, as exemplified by various uses of nucleic acid amplification technologies (e.g., polymerase chain reaction (PCR), Mullis et al., U.S. Patent No. 4,683,202; ligase chain reaction (LCR), Tabor, S. et al., *PNAS USA* 82:1074 (1985); or strand displacement amplification (SDA), Walker et al., *PNAS USA* 89:392 (1992)]; and (3) methods of library construction and screening by complementation.

[00132] Another way to improve the production of a product alcohol is to alter the GPD to increase the K_M for NADH. Increasing the K_M for NADH by altering GPD can decrease the rate of NADH oxidation catalyzed by GPD, thus allowing an increase in the availability of NADH in a host cell. The available NADH can be used by other enzymes in the product alcohol production pathway, for example, in the isobutanol production pathway the available NADH can be utilized by KARI and alcohol dehydrogenase. Thus, decreasing the affinity of GPD for NADH can increase product alcohol (e.g., isobutanol) production. Accordingly, in some embodiments, a heterologous and/or engineered GPD is expressed in recombinant

microorganism that also expresses other NADH-utilizing enzymes, for example, an NADH-utilizing enzyme that acts in the isobutanol production pathway such as KARI and alcohol dehydrogenase.

[00133] GPD enzymes with an increased K_M for NADH can also be produced by means of protein engineering. In some embodiments, the GPD has at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90% or 95% identity to *Saccharomyces cerevisiae* GPD 1 (SEQ ID NO: 195), but is not 100% identical to SEQ ID NO:195. In some embodiments, the GPD comprises at least one substitution at a residue corresponding to position 42, 44, 45, 71, 73, 75, 95, 124, 126, 129, 151, 152, 183, 184, 185, 246, 310, 336, 337, or 339 of *Saccharomyces cerevisiae* GPD 1(SEQ ID NO: 195).

[00134] For example, in some embodiments, the GPD comprises a substitution of the residue corresponding to position 44 of SEQ ID NO:195 (Asn in SEQ ID NO:195) to an amino acid selected from the group consisting of A, C, G, I, L, M, S, and V.

[00135] In some embodiments, the GPD comprises a substitution of the residue corresponding to position 45 of SEQ ID NO:195 (Trp in SEQ ID NO:195) to an amino acid selected from the group consisting of A, C, G, H, I, K, L, M, N, Q, R, S, T, and V.

[00136] In some embodiments, the GPD comprises a substitution of the residue corresponding to position 73 of SEQ ID NO:195 (Phe in SEQ ID NO:195) to an amino acid selected from the group consisting of G, A, R, and K.

[00137] In some embodiments, the GPD comprises a substitution of the residue corresponding to position 129 of SEQ ID NO:195 (Phe in SEQ ID NO:195) to an amino acid selected from the group consisting of G, A, R, and K.

[00138] In some embodiments, the GPD comprises a substitution of the residue corresponding to position 337 of SEQ ID NO:195 (Ser in SEQ ID NO:195) to an amino acid selected from the group consisting of A, C, D, E, G, I, L, M, N, Q, and V.

[00139] In some embodiments, the GPD comprises a substitution of the residue corresponding to position 339 of SEQ ID NO:195 (Gln in SEQ ID NO:195) to an amino acid selected from the group consisting of A, C, G, I, L, M, S, and V.

[00140] In some embodiments, the GPD comprises a substitution of the residue corresponding to position 42, 71, 75, 95, 124, 126, 151, 152, 183, 184, 185, 246, 310, and/or 336 of SEQ ID NO:195 (Ser, Trp, Glu, Tyr, Gln, Pro, Leu, Lys, Asn, Ile, Ala, Asn, Arg, Gln of SEQ

ID NO: 195, respectively) to any other amino acid selected from the 19 naturally occurring amino acids.

[00141] In some embodiments, the GPD has a K_M for NADH that is about 0.01 mM to about 1mM. In some embodiments, the GPD has a K_M for NADH that is about 0.05 mM to about 1 mM. In some embodiments, the GPD has a K_M for NADH that is about 0.10 mM to about 1 mM. In some embodiments, the GPD has a K_M for NADH that is about 0.15 mM to about 1 mM. In some embodiments, the GPD has a K_M for NADH that is about 0.20 mM to about 1 mM. In some embodiments, the GPD has a K_M for NADH that is about 0.30 mM to about 1 mM. In some embodiments, the GPD has a K_M for NADH that is about 0.40 mM to about 1 mM. In some embodiments, the GPD has a K_M for NADH that is about 0.50 mM to about 1 mM. Assays for measuring the K_M for NADH of GPD are disclosed in Example 1 below and are known in the art, see, e.g., Niesel et al., *Methods Enzymol.* 89:296-301 (1982). Certain assays can be referred to as "NADH consumption assays," which refer to an enzyme assay for the determination of the specific activity of the GPD enzyme, involving measuring the disappearance of the GPD cofactor, NADH, from the enzyme reaction.

[00142] In some embodiments, the GPD has a K_M for NADPH that is about 0.01 mM to about 1mM. In some embodiments, the GPD has a K_M for NADPH that is about 0.05 mM to about 1 mM. In some embodiments, the GPD has a K_M for NADPH that is about 0.10 mM to about 1 mM. In some embodiments, the GPD has a K_M for NADPH that is about 0.15 mM to about 1 mM. In some embodiments, the GPD has a K_M for NADPH that is about 0.20 mM to about 1 mM. In some embodiments, the GPD has a K_M for NADPH that is about 0.30 mM to about 1 mM. In some embodiments, the GPD has a K_M for NADPH that is about 0.40 mM to about 1 mM. In some embodiments, the GPD has a K_M for NADPH that is about 0.50 mM to about 1 mM. The NADH assays disclosed below in Example 1 can be adapted to measure the K_M for NADPH of GPD by replacing NADH with NADPH. Additional assays for measuring the K_M for NADPH of GPD are known in the art, see, e.g., Niesel et al., *Methods Enzymol.* 89:296-301 (1982). Certain assays can be referred to as "NADPH consumption assays," which refer to an enzyme assay for the determination of the specific activity of the GPD enzyme, involving measuring the disappearance of the GPD cofactor, NADH, from the enzyme reaction.

[00143] In some embodiments, the heterologous and/or engineered GPD can increase the growth of a recombinant microorganism comprising the heterologous and/or engineered GPD as

compared to a recombinant microorganism that does not contain the heterologous and/or engineered GPD.

[00144] In some embodiments, the heterologous and/or engineered GPD can increase the product alcohol (*e.g.*, isobutanol) production of a recombinant microorganism comprising the heterologous and/or engineered GPD as compared to a recombinant microorganism that does not contain the heterologous and/or engineered GPD.

[00145] In some embodiments, the heterologous and/or engineered GPD can decrease the glycerol production of a recombinant microorganism comprising the GPD as compared to a recombinant microorganism that does not contain the heterologous and/or engineered GPD.

[00146] In some embodiments, the heterologous and/or engineered GPD can increase the ratio of product alcohol (*e.g.*, isobutanol) to glycerol produced by a recombinant microorganism comprising the heterologous and/or engineered GPD as compared to a recombinant microorganism that does not contain the heterologous and/or engineered GPD.

[00147] In some embodiments, the heterologous and/or engineered GPD can increase the yield (*e.g.*, gram of isobutanol produced per gram of substrate consumed) of a recombinant microorganism comprising the heterologous and/or engineered GPD as compared to a recombinant microorganism that does not contain the heterologous and/or engineered GPD.

[00148] Thus in a recombinant microorganism comprising a butanol biosynthetic pathway, a heterologous and/or engineered GPD that has a higher K_M for NADH than the microorganism's endogenous GPD, and a deletion or disruption of an endogenous gene encoding GPD, "improved production of butanol" can refer to increased production of butanol, a decreased production of glycerol, or both, as compared to a microorganism that lacks the heterologous and/or engineered GPD.

[00149] In a recombinant microorganisms comprising a heterologous and/or engineered GPD that has a higher K_M for NADH than the microorganism's endogenous GPD, and a deletion or disruption in an endogenous gene encoding GPD, "improved production of alcohol" can refer to an increased production of alcohol, a decreased production of glycerol, or both, as compared to a microorganism that lacks the heterologous and/or engineered GPD.

[00150] Thus, in a recombinant microorganism comprising a butanol biosynthetic pathway, a heterologous GPD that has substantially the same affinity for NADH and NADPH and/or is feedback inhibited by glycerol-3-phosphate, and a deletion or disruption of an endogenous gene encoding GPD, "improved production of butanol" can refer to increased

production of butanol, a decreased production of glycerol, or both, as compared to a microorganism that lacks the heterologous GPD.

[00151] In a recombinant microorganisms comprising a heterologous GPD that has substantially the same affinity for NADH and NADPH and/or is feedback inhibited by glycerol-3-phosphate, and a deletion or disruption in an endogenous gene encoding GPD, "improved production of alcohol" can refer to an increased production of alcohol, a decreased production of glycerol, or both, as compared to a microorganism that lacks the heterologous GPD.

RECOMBINANT MICROORGANISMS

[00152] While not wishing to be bound by theory, it is believed that the processes described herein are useful in conjunction with any alcohol producing microorganism, particularly recombinant microorganisms which produce alcohol.

[00153] Recombinant microorganisms which produce alcohol are also known in the art (e.g., Ohta *et al.*, *Appl. Environ. Microbiol.* 57:893-900 (1991); Underwood *et al.*, *Appl. Environ. Microbiol.* 68:1071-81 (2002); Shen and Liao, *Metab. Eng.* 10:312-20 (2008); Hahnai *et al.*, *Appl. Environ.* 73:7814-8 (2007); U.S. Patent No. 5,514,583; U.S. Patent No. 5,712,133; International Publication No. WO 1995/028476; Feldmann *et al.*, *Appl. Microbiol. Biotechnol.* 38:354-61 (1992); Zhang *et al.*, *Science* 267:240-3 (1995); U.S. Patent Publication No. 2007/0031918A1; U.S. Patent No. 7,223,575; U.S. Patent No. 7,741,119; U.S. Patent Publication No. 2009/0203099A1; U.S. Patent Publication No. 2009/0246846A1; and International Publication No. WO 2010/075241, which are herein incorporated by reference).

[00154] For example, the metabolic pathways of microorganisms may be genetically modified to produce butanol. These pathways may also be modified to reduce or eliminate undesired metabolites, and thereby improve yield of the product alcohol. The production of butanol by a microorganism is disclosed, for example, in U.S. Patent Nos. 7,851,188; 7,993,889; 8,178,328, 8,206,970; U.S. Patent Application Publication Nos. 2007/0292927; 2008/0182308; 2008/0274525; 2009/0305363; 2009/0305370; 2011/0250610; 2011/0313206; 2011/0111472; 2012/0258873; and 2013/0071898, the entire contents of each are herein incorporated by reference. In certain embodiments, the microorganism is genetically modified to comprise a butanol biosynthetic pathway or a biosynthetic pathway for a butanol isomer, such as 1-butanol, 2-butanol, or isobutanol. In certain embodiments, at least one, at least two, at least three, at least four, or at least five polypeptides catalyzing substrate to product conversions in the butanol

biosynthetic pathway are encoded by heterologous polynucleotides in the microorganism. In certain embodiments, all the polypeptides catalyzing substrate to product conversions of the butanol biosynthetic pathway are encoded by heterologous polynucleotides in the microorganism. It will be appreciated that microorganisms comprising a butanol biosynthetic pathway may further comprise one or more additional genetic modifications as disclosed in U.S. Patent Application Publication No. 2013/0071898, which is herein incorporated by reference in its entirety.

[00155] In some embodiments, the microorganism may be bacteria, cyanobacteria, filamentous fungi, or yeasts. Suitable microorganisms capable of producing product alcohol (e.g., butanol) via a biosynthetic pathway include a member of the genera *Clostridium*, *Zymomonas*, *Escherichia*, *Salmonella*, *Serratia*, *Erwinia*, *Klebsiella*, *Shigella*, *Rhodococcus*, *Pseudomonas*, *Bacillus*, *Lactobacillus*, *Enterococcus*, *Alcaligenes*, *Paenibacillus*, *Arthrobacter*, *Corynebacterium*, *Brevibacterium*, *Schizosaccharomyces*, *Kluveromyces*, *Yarrowia*, *Pichia*, *Zygosaccharomyces*, *Debaryomyces*, *Candida*, *Brettanomyces*, *Pachysolen*, *Hansenula*, *Issatchenkia*, *Trichosporon*, *Yamadazyma*, or *Saccharomyces*. In one embodiment, recombinant microorganisms may be selected from the group consisting of *Escherichia coli*, *Alcaligenes eutrophus*, *Bacillus licheniformis*, *Paenibacillus macerans*, *Rhodococcus erythropolis*, *Pseudomonas putida*, *Lactobacillus plantarum*, *Enterococcus faecium*, *Enterococcus gallinarum*, *Enterococcus faecalis*, *Bacillus subtilis*, *Candida sonorensis*, *Candida methanosorbosa*, *Kluyveromyces lactis*, *Kluyveromyces marxianus*, *Kluveromyces thermotolerans*, *Issatchenkia orientalis*, *Debaryomyces hansenii*, and *Saccharomyces cerevisiae*. In one embodiment, the genetically modified microorganism is yeast. In one embodiment, the genetically modified microorganism is a crabtree-positive yeast selected from *Saccharomyces*, *Zygosaccharomyces*, *Schizosaccharomyces*, *Dekkera*, *Torulopsis*, *Brettanomyces*, and some species of *Candida*. Species of crabtree-positive yeast include, but are not limited to, *Saccharomyces cerevisiae*, *Saccharomyces kluyveri*, *Schizosaccharomyces pombe*, *Saccharomyces bayanus*, *Saccharomyces mikitaе*, *Saccharomyces paradoxus*, *Saccharomyces uvarum*, *Saccharomyces castelli*, *Zygosaccharomyces rouxii*, *Zygosaccharomyces bailli*, and *Candida glabrata*.

[00156] In some embodiments, the host cell is *Saccharomyces cerevisiae*. *Saccharomyces cerevisiae* are known in the art and are available from a variety of sources including, but not limited to, American Type Culture Collection (Rockville, MD), Centraalbureau voor

Schimmelcultures (CBS) Fungal Biodiversity Centre, LeSaffre, Gert Strand AB, Ferm Solutions, North American Bioproducts, Martrex, and Lallemand. *S. cerevisiae* include, but are not limited to, BY4741, CEN.PK 113-7D, Ethanol Red® yeast, Ferm Pro™ yeast, Bio-Ferm® XR yeast, Gert Strand Prestige Batch Turbo alcohol yeast, Gert Strand Pot Distillers yeast, Gert Strand Distillers Turbo yeast, FerMax™ Green yeast, FerMax™ Gold yeast, Thermosacc® yeast, BG-1, PE-2, CAT-1, CBS7959, CBS7960, and CBS7961.

[00157] In some embodiments, the microorganism may be immobilized or encapsulated. For example, the microorganism may be immobilized or encapsulated using alginate, calcium alginate, or polyacrylamide gels, or through the induction of biofilm formation onto a variety of high surface area support matrices such as diatomite, celite, diatomaceous earth, silica gels, plastics, or resins. In some embodiments, ISPR may be used in combination with immobilized or encapsulated microorganisms. This combination may improve productivity such as specific volumetric productivity, metabolic rate, product alcohol yields, tolerance to product alcohol. In addition, immobilization and encapsulation may minimize the effects of the process conditions such as shearing on the microorganisms.

[00158] Biosynthetic pathways for the production of isobutanol that may be used include those as described by Donaldson *et al.* in U.S. Patent No. 7,851,188; U.S. Patent No. 7,993,388; and International Publication No. WO 2007/050671, which are incorporated herein by reference. In one embodiment, the isobutanol biosynthetic pathway comprises the following substrate to product conversions:

- a) pyruvate to acetolactate, which may be catalyzed, for example, by acetolactate synthase;
- b) the acetolactate from step a) to 2,3-dihydroxyisovalerate, which may be catalyzed, for example, by ketol-acid reductoisomerase;
- c) the 2,3-dihydroxyisovalerate from step b) to α -ketoisovalerate, which may be catalyzed, for example, by dihydroxyacid dehydratase;
- d) the α -ketoisovalerate from step c) to isobutyraldehyde, which may be catalyzed, for example, by a branched-chain α -keto acid decarboxylase; and,
- e) the isobutyraldehyde from step d) to isobutanol, which may be catalyzed, for example, by a branched-chain alcohol dehydrogenase.

[00159] In another embodiment, the isobutanol biosynthetic pathway comprises the following substrate to product conversions:

- a) pyruvate to acetolactate, which may be catalyzed, for example, by acetolactate synthase;
- b) the acetolactate from step a) to 2,3-dihydroxyisovalerate, which may be catalyzed, for example, by ketol-acid reductoisomerase;
- c) the 2,3-dihydroxyisovalerate from step b) to α -ketoisovalerate, which may be catalyzed, for example, by dihydroxyacid dehydratase;
- d) the α -ketoisovalerate from step c) to valine, which may be catalyzed, for example, by transaminase or valine dehydrogenase;
- e) the valine from step d) to isobutylamine, which may be catalyzed, for example, by valine decarboxylase;
- f) the isobutylamine from step e) to isobutyraldehyde, which may be catalyzed by, for example, omega transaminase; and,
- g) the isobutyraldehyde from step f) to isobutanol, which may be catalyzed, for example, by a branched-chain alcohol dehydrogenase.

[00160] In another embodiment, the isobutanol biosynthetic pathway comprises the following substrate to product conversions:

- a) pyruvate to acetolactate, which may be catalyzed, for example, by acetolactate synthase;
- b) the acetolactate from step a) to 2,3-dihydroxyisovalerate, which may be catalyzed, for example, by ketol-acid reductoisomerase;
- c) the 2,3-dihydroxyisovalerate from step b) to α -ketoisovalerate, which may be catalyzed, for example, by dihydroxyacid dehydratase;
- d) the α -ketoisovalerate from step c) to isobutyryl-CoA, which may be catalyzed, for example, by branched-chain keto acid dehydrogenase;
- e) the isobutyryl-CoA from step d) to isobutyraldehyde, which may be catalyzed, for example, by acylating aldehyde dehydrogenase; and,
- f) the isobutyraldehyde from step e) to isobutanol, which may be catalyzed, for example, by a branched-chain alcohol dehydrogenase.

[00161] Biosynthetic pathways for the production of 1-butanol that may be used include those described in U.S. Patent Application Publication No. 2008/0182308 and WO2007/041269, which are incorporated herein by reference. In one embodiment, the 1-butanol biosynthetic pathway comprises the following substrate to product conversions:

a) acetyl-CoA to acetoacetyl-CoA, which may be catalyzed, for example, by acetyl-CoA acetyltransferase;

b) the acetoacetyl-CoA from step a) to 3-hydroxybutyryl-CoA, which may be catalyzed, for example, by 3-hydroxybutyryl-CoA dehydrogenase;

c) the 3-hydroxybutyryl-CoA from step b) to crotonyl-CoA, which may be catalyzed, for example, by crotonase;

d) the crotonyl-CoA from step c) to butyryl-CoA, which may be catalyzed, for example, by butyryl-CoA dehydrogenase;

e) the butyryl-CoA from step d) to butyraldehyde, which may be catalyzed, for example, by butyraldehyde dehydrogenase; and,

f) the butyraldehyde from step e) to 1-butanol, which may be catalyzed, for example, by butanol dehydrogenase.

[00162] Biosynthetic pathways for the production of 2-butanol that may be used include those described by Donaldson *et al.* in U.S. Patent No. 8,206,970; U.S. Patent Application Publication Nos. 2007/0292927 and 2009/0155870; International Publication Nos. WO 2007/130518 and WO 2007/130521, all of which are incorporated herein by reference. In one embodiment, the 2-butanol biosynthetic pathway comprises the following substrate to product conversions:

a) pyruvate to alpha-acetolactate, which may be catalyzed, for example, by acetolactate synthase;

b) the alpha-acetolactate from step a) to acetoin, which may be catalyzed, for example, by acetolactate decarboxylase;

c) the acetoin from step b) to 3-amino-2-butanol, which may be catalyzed, for example, acetoin aminase;

d) the 3-amino-2-butanol from step c) to 3-amino-2-butanol phosphate, which may be catalyzed, for example, by aminobutanol kinase;

e) the 3-amino-2-butanol phosphate from step d) to 2-butanone, which may be catalyzed, for example, by aminobutanol phosphate phosphorylase; and,

f) the 2-butanone from step e) to 2-butanol, which may be catalyzed, for example, by butanol dehydrogenase.

[00163] In another embodiment, the 2-butanol biosynthetic pathway comprises the following substrate to product conversions:

- a) pyruvate to alpha-acetolactate, which may be catalyzed, for example, by acetolactate synthase;
- b) the alpha-acetolactate from step a) to acetoin, which may be catalyzed, for example, by acetolactate decarboxylase;
- c) the acetoin to 2,3-butanediol from step b), which may be catalyzed, for example, by butanediol dehydrogenase;
- d) the 2,3-butanediol from step c) to 2-butanone, which may be catalyzed, for example, by diol dehydratase; and,
- e) the 2-butanone from step d) to 2-butanol, which may be catalyzed, for example, by butanol dehydrogenase.

[00164] Biosynthetic pathways for the production of 2-butanone that may be used include those described in U.S. Patent No. 8,206,970 and U.S. Patent Application Publication Nos. 2007/0292927 and 2009/0155870, which are incorporated herein by reference. In one embodiment, the 2-butanone biosynthetic pathway comprises the following substrate to product conversions:

- a) pyruvate to alpha-acetolactate, which may be catalyzed, for example, by acetolactate synthase;
- b) the alpha-acetolactate from step a) to acetoin, which may be catalyzed, for example, by acetolactate decarboxylase;
- c) the acetoin from step b) to 3-amino-2-butanol, which may be catalyzed, for example, acetoin aminase;
- d) the 3-amino-2-butanol from step c) to 3-amino-2-butanol phosphate, which may be catalyzed, for example, by aminobutanol kinase; and,
- e) the 3-amino-2-butanol phosphate from step d) to 2-butanone, which may be catalyzed, for example, by aminobutanol phosphate phosphorylase.

[00165] In another embodiment, the 2-butanone biosynthetic pathway comprises the following substrate to product conversions:

- a) pyruvate to alpha-acetolactate, which may be catalyzed, for example, by acetolactate synthase;
- b) the alpha-acetolactate from step a) to acetoin which may be catalyzed, for example, by acetolactate decarboxylase;

c) the acetoin from step b) to 2,3-butanediol, which may be catalyzed, for example, by butanediol dehydrogenase;

d) the 2,3-butanediol from step c) to 2-butanone, which may be catalyzed, for example, by diol dehydratase.

[00166] The terms “acetohydroxyacid synthase,” “acetolactate synthase,” and “acetolactate synthetase” (abbreviated “ALS”) are used interchangeably herein to refer to an enzyme that catalyzes the conversion of pyruvate to acetolactate and CO₂. Example acetolactate synthases are known by the EC number 2.2.1.6 (Enzyme Nomenclature 1992, Academic Press, San Diego). These enzymes are available from a number of sources, including, but not limited to, *Bacillus subtilis* (GenBank Nos: CAB07802.1, Z99122, NCBI (National Center for Biotechnology Information) amino acid sequence, NCBI nucleotide sequence, respectively), CAB15618, *Klebsiella pneumoniae* (GenBank Nos: AAA25079, M73842), and *Lactococcus lactis* (GenBank Nos: AAA25161, L16975).

[00167] The term “ketol-acid reductoisomerase” (“KARI”), “acetohydroxy acid isomeroreductase,” and “acetohydroxy acid reductoisomerase” will be used interchangeably and refer to enzymes capable of catalyzing the reaction of (S)-acetolactate to 2,3-dihydroxyisovalerate. Example KARI enzymes may be classified as EC number EC 1.1.1.86 (Enzyme Nomenclature 1992, Academic Press, San Diego), and are available from a vast array of microorganisms, including, but not limited to, *Escherichia coli* (GenBank Nos: NP_418222, NC_000913), *Saccharomyces cerevisiae* (GenBank Nos: NP_013459, NC_001144), *Methanococcus maripaludis* (GenBank Nos: CAF30210, BX957220), and *Bacillus subtilis* (GenBank Nos: CAB14789, Z99118). KARIs include *Anaerostipes caccae* KARI variants “K9G9” (SEQ ID NO:85), “K9D3” (SEQ ID NO:86), and “K9JB4P” (SEQ ID NO:87). Ketol-acid reductoisomerase (KARI) enzymes are described in U.S. Patent Nos. 7,910,342 and 8,129,162; U.S. Patent Application Publication Nos. 2008/0261230, 2009/0163376, 2010/0197519, PCT Application Publication No. WO/2011/041415, PCT Application Publication No. WO2012/129555; and U.S. Patent Application No. 14/038,455, filed on September 26, 2013, all of which are incorporated herein by reference. Examples of KARIs disclosed therein are those from *Lactococcus lactis*, *Vibrio cholera*, *Pseudomonas aeruginosa* PAO1, and *Pseudomonas fluorescens* PF5 mutants. In some embodiments, the KARI utilizes NADH. In some embodiments, the KARI utilizes NADPH. In some embodiments, the KARI utilizes NADH or NADPH.

[00168] The term “acetohydroxy acid dehydratase” and “dihydroxyacid dehydratase” (“DHAD”) refers to an enzyme that catalyzes the conversion of 2,3-dihydroxyisovalerate to α -ketoisovalerate. Example acetohydroxy acid dehydratases are known by the EC number 4.2.1.9. Such enzymes are available from a vast array of microorganisms, including, but not limited to, *E. coli* (GenBank Nos: YP_026248, NC000913), *Saccharomyces cerevisiae* (GenBank Nos: NP_012550, NC_001142), *M. maripaludis* (GenBank Nos: CAF29874, BX957219), *B. subtilis* (GenBank Nos: CAB14105, Z99115), *L. lactis* (SEQ ID NO:88), and *N. crassa*. U.S. Patent Application Publication No. 2010/0081154, U.S. Patent No. 7,851,188, and U.S. Patent No. 8,241,878, which are incorporated herein by reference in their entireties, describe dihydroxyacid dehydratases (DHADs), including a DHAD from *Streptococcus mutans* (SEQ ID NO:89) and variants thereof.

[00169] The term “branched-chain α -keto acid decarboxylase,” “ α -ketoacid decarboxylase,” “ α -ketoisovalerate decarboxylase,” or “2-ketoisovalerate decarboxylase” (“KIVD”) refers to an enzyme that catalyzes the conversion of α -ketoisovalerate to isobutyraldehyde and CO₂. Example branched-chain α -keto acid decarboxylases are known by the EC number 4.1.1.72 and are available from a number of sources, including, but not limited to, *Lactococcus lactis* (GenBank Nos: AAS49166, AY548760; CAG34226, AJ746364), *Salmonella typhimurium* (GenBank Nos: NP_461346, NC_003197), *Clostridium acetobutylicum* (GenBank Nos: NP_149189, NC_001988), *M. caseolyticus*, and *L. grayi*. Suitable branched-chain α -keto acid decarboxylases can comprise SEQ ID NO:90 from *Lactococcus lactis* and SEQ ID NO:91 from *Listeria grayi*.

[00170] The term “branched-chain alcohol dehydrogenase” (“ADH”) refers to an enzyme that catalyzes the conversion of isobutyraldehyde to isobutanol. Example branched-chain alcohol dehydrogenases are known by the EC number 1.1.1.265, but may also be classified under other alcohol dehydrogenases (specifically, EC 1.1.1.1 or 1.1.1.2). Alcohol dehydrogenases may be NADPH dependent or NADH dependent. Such enzymes are available from a number of sources, including, but not limited to, *S. cerevisiae* (GenBank Nos: NP_010656, NC_001136, NP_014051, NC_001145), *E. coli* (GenBank Nos: NP_417484, NC_000913), *C. acetobutylicum* (GenBank Nos: NP_349892, NC_003030; NP_349891, NC_003030). U.S. Patent Application Publication No. 2009/0269823 describes SadB, an alcohol dehydrogenase (ADH) from *Achromobacter xylosoxidans* (SEQ ID NO:92). Alcohol dehydrogenases also include horse liver

ADH (SEQ ID NO:93) and *Beijerinckia indica* ADH (SEQ ID NO:94) (as described by U.S. Patent Application Publication No. 2011/0269199, which is incorporated herein by reference).

[00171] The term “butanol dehydrogenase” refers to a polypeptide (or polypeptides) having an enzyme activity that catalyzes the conversion of isobutyraldehyde to isobutanol or the conversion of 2-butanone and 2-butanol. Butanol dehydrogenases are a subset of a broad family of alcohol dehydrogenases. Butanol dehydrogenase may be NAD- or NADP-dependent. The NAD-dependent enzymes are known as EC 1.1.1.1 and are available, for example, from *Rhodococcus ruber* (GenBank Nos: CAD36475, AJ491307). The NADP dependent enzymes are known as EC 1.1.1.2 and are available, for example, from *Pyrococcus furiosus* (GenBank Nos: AAC25556, AF013169). Additionally, a butanol dehydrogenase is available from *Escherichia coli* (GenBank Nos: NP_417484, NC_000913) and a cyclohexanol dehydrogenase is available from *Acinetobacter sp.* (GenBank Nos: AAG10026, AF282240). The term “butanol dehydrogenase” also refers to an enzyme that catalyzes the conversion of butyraldehyde to 1-butanol, using either NADH or NADPH as cofactor. Butanol dehydrogenases are available from, for example, *C. acetobutylicum* (GenBank NOs: NP_149325, NC_001988; note: this enzyme possesses both aldehyde and alcohol dehydrogenase activity); NP_349891, NC_003030; and NP_349892, NC_003030) and *E. coli* (GenBank NOs: NP_417-484, NC_000913).

[00172] The term “branched-chain keto acid dehydrogenase” refers to an enzyme that catalyzes the conversion of α -ketoisovalerate to isobutyryl-CoA (isobutyryl-coenzyme A), typically using NAD⁺ (nicotinamide adenine dinucleotide) as an electron acceptor. Example branched-chain keto acid dehydrogenases are known by the EC number 1.2.4.4. Such branched-chain keto acid dehydrogenases are comprised of four subunits and sequences from all subunits are available from a vast array of microorganisms, including, but not limited to, *B. subtilis* (GenBank Nos: CAB14336, Z99116; CAB14335, Z99116; CAB14334, Z99116; and CAB14337, Z99116) and *Pseudomonas putida* (GenBank Nos: AAA65614, M57613; AAA65615, M57613; AAA65617, M57613; and AAA65618, M57613).

[00173] The term “acylating aldehyde dehydrogenase” refers to an enzyme that catalyzes the conversion of isobutyryl-CoA to isobutyraldehyde, typically using either NADH or NADPH as an electron donor. Example acylating aldehyde dehydrogenases are known by the EC numbers 1.2.1.10 and 1.2.1.57. Such enzymes are available from multiple sources, including, but not limited to, *Clostridium beijerinckii* (GenBank Nos: AAD31841, AF157306), *C. acetobutylicum* (GenBank Nos: NP_149325, NC_001988; NP_149199, NC_001988), *P. putida*

(GenBank Nos: AAA89106, U13232), and *Thermus thermophilus* (GenBank Nos: YP_145486, NC_006461).

[00174] The term “transaminase” refers to an enzyme that catalyzes the conversion of α -ketoisovalerate to L-valine, using either alanine or glutamate as an amine donor. Example transaminases are known by the EC numbers 2.6.1.42 and 2.6.1.66. Such enzymes are available from a number of sources. Examples of sources for alanine-dependent enzymes include, but are not limited to, *E. coli* (GenBank Nos: YP_026231, NC_000913) and *Bacillus licheniformis* (GenBank Nos: YP_093743, NC_006322). Examples of sources for glutamate-dependent enzymes include, but are not limited to, *E. coli* (GenBank Nos: YP_026247, NC_000913), *Saccharomyces cerevisiae* (GenBank Nos: NP_012682, NC_001142) and *Methanobacterium thermoautotrophicum* (GenBank Nos: NP_276546, NC_000916).

[00175] The term “valine dehydrogenase” refers to an enzyme that catalyzes the conversion of α -ketoisovalerate to L-valine, typically using NAD(P)H as an electron donor and ammonia as an amine donor. Example valine dehydrogenases are known by the EC numbers 1.4.1.8 and 1.4.1.9 and such enzymes are available from a number of sources, including, but not limited to, *Streptomyces coelicolor* (GenBank Nos: NP_628270, NC_003888) and *B. subtilis* (GenBank Nos: CAB14339, Z99116).

[00176] The term “valine decarboxylase” refers to an enzyme that catalyzes the conversion of L-valine to isobutylamine and CO₂. Example valine decarboxylases are known by the EC number 4.1.1.14. Such enzymes are found in *Streptomyces*, such as for example, *Streptomyces viridifaciens* (GenBank Nos: AAN10242, AY116644).

[00177] The term “omega transaminase” refers to an enzyme that catalyzes the conversion of isobutylamine to isobutyraldehyde using a suitable amino acid as an amine donor. Example omega transaminases are known by the EC number 2.6.1.18 and are available from a number of sources, including, but not limited to, *Alcaligenes denitrificans* (AAP92672, AY330220), *Ralstonia eutropha* (GenBank Nos: YP_294474, NC_007347), *Shewanella oneidensis* (GenBank Nos: NP_719046, NC_004347), and *P. putida* (GenBank Nos: AAN66223, AE016776).

[00178] The term “acetyl-CoA acetyltransferase” refers to an enzyme that catalyzes the conversion of two molecules of acetyl-CoA to acetoacetyl-CoA and coenzyme A (CoA). Example acetyl-CoA acetyltransferases are acetyl-CoA acetyltransferases with substrate preferences (reaction in the forward direction) for a short chain acyl-CoA and acetyl-CoA and are classified as E.C. 2.3.1.9 [Enzyme Nomenclature 1992, Academic Press, San Diego]; although,

enzymes with a broader substrate range (E.C. 2.3.1.16) will be functional as well. Acetyl-CoA acetyltransferases are available from a number of sources, for example, *Escherichia coli* (GenBank Nos: NP_416728, NC_000913; NCBI (National Center for Biotechnology Information) amino acid sequence, NCBI nucleotide sequence), *Clostridium acetobutylicum* (GenBank Nos: NP_349476.1, NC_003030; NP_149242, NC_001988, *Bacillus subtilis* (GenBank Nos: NP_390297, NC_000964), and *Saccharomyces cerevisiae* (GenBank Nos: NP_015297, NC_001148).

[00179] The term “3-hydroxybutyryl-CoA dehydrogenase” refers to an enzyme that catalyzes the conversion of acetoacetyl-CoA to 3-hydroxybutyryl-CoA. 3-Example hydroxybutyryl-CoA dehydrogenases may be reduced nicotinamide adenine dinucleotide (NADH)-dependent, with a substrate preference for (S)-3-hydroxybutyryl-CoA or (R)-3-hydroxybutyryl-CoA. Examples may be classified as E.C. 1.1.1.35 and E.C. 1.1.1.30, respectively. Additionally, 3-hydroxybutyryl-CoA dehydrogenases may be reduced nicotinamide adenine dinucleotide phosphate (NADPH)-dependent, with a substrate preference for (S)-3-hydroxybutyryl-CoA or (R)-3-hydroxybutyryl-CoA and are classified as E.C. 1.1.1.157 and E.C. 1.1.1.36, respectively. 3-Hydroxybutyryl-CoA dehydrogenases are available from a number of sources, for example, *C. acetobutylicum* (GenBank NOs: NP_349314, NC_003030), *B. subtilis* (GenBank NOs: AAB09614, U29084), *Ralstonia eutropha* (GenBank NOs: YP_294481, NC_007347), and *Alcaligenes eutrophus* (GenBank NOs: AAA21973, J04987).

[00180] The term “crotonase” refers to an enzyme that catalyzes the conversion of 3-hydroxybutyryl-CoA to crotonyl-CoA and H₂O. Example crotonases may have a substrate preference for (S)-3-hydroxybutyryl-CoA or (R)-3-hydroxybutyryl-CoA and may be classified as E.C. 4.2.1.17 and E.C. 4.2.1.55, respectively. Crotonases are available from a number of sources, for example, *E. coli* (GenBank NOs: NP_415911, NC_000913), *C. acetobutylicum* (GenBank NOs: NP_349318, NC_003030), *B. subtilis* (GenBank NOs: CAB13705, Z99113), and *Aeromonas caviae* (GenBank NOs: BAA21816, D88825).

[00181] The term “butyryl-CoA dehydrogenase” refers to an enzyme that catalyzes the conversion of crotonyl-CoA to butyryl-CoA. Example butyryl-CoA dehydrogenases may be NADH-dependent, NADPH-dependent, or flavin-dependent and may be classified as E.C. 1.3.1.44, E.C. 1.3.1.38, and E.C. 1.3.99.2, respectively. Butyryl-CoA dehydrogenases are available from a number of sources, for example, *C. acetobutylicum* (GenBank NOs: NP_347102, NC_003030), *Euglena gracilis* (GenBank NOs: Q5EU90, AY741582),

Streptomyces collinus (GenBank NOs: AAA92890, U37135), and *Streptomyces coelicolor* (GenBank NOs: CAA22721, AL939127).

[00182] The term “butyraldehyde dehydrogenase” refers to an enzyme that catalyzes the conversion of butyryl-CoA to butyraldehyde, using NADH or NADPH as cofactor.

Butyraldehyde dehydrogenases with a preference for NADH are known as E.C. 1.2.1.57 and are available from, for example, *Clostridium beijerinckii* (GenBank NOs: AAD31841, AF157306) and *C. acetobutylicum* (GenBank NOs: NP_149325, NC_001988).

[00183] The term “isobutyryl-CoA mutase” refers to an enzyme that catalyzes the conversion of butyryl-CoA to isobutyryl-CoA. This enzyme uses coenzyme B₁₂ as cofactor. Example isobutyryl-CoA mutases are known by the EC number 5.4.99.13. These enzymes are found in a number of *Streptomyces*, including, but not limited to, *Streptomyces cinnamonensis* (GenBank Nos: AAC08713, U67612; CAB59633, AJ246005), *S. coelicolor* (GenBank Nos: CAB70645, AL939123; CAB92663, AL939121), and *Streptomyces avermitilis* (GenBank Nos: NP_824008, NC_003155; NP_824637, NC_003155).

[00184] The term “acetolactate decarboxylase” refers to a polypeptide (or polypeptides) having an enzyme activity that catalyzes the conversion of alpha-acetolactate to acetoin. Example acetolactate decarboxylases are known as EC 4.1.1.5 and are available, for example, from *Bacillus subtilis* (GenBank Nos: AAA22223, L04470), *Klebsiella terrigena* (GenBank Nos: AAA25054, L04507) and *Klebsiella pneumoniae* (GenBank Nos: AAU43774, AY722056).

[00185] The term “acetoin aminase” or “acetoin transaminase” refers to a polypeptide (or polypeptides) having an enzyme activity that catalyzes the conversion of acetoin to 3-amino-2-butanol. Acetoin aminase may utilize the cofactor pyridoxal 5'-phosphate or NADH (reduced nicotinamide adenine dinucleotide) or NADPH (reduced nicotinamide adenine dinucleotide phosphate). The resulting product may have (R) or (S) stereochemistry at the 3-position. The pyridoxal phosphate-dependent enzyme may use an amino acid such as alanine or glutamate as the amino donor. The NADH- and NADPH-dependent enzymes may use ammonia as a second substrate. A suitable example of an NADH dependent acetoin aminase, also known as amino alcohol dehydrogenase, is described by Ito, *et al.* (U.S. Patent No. 6,432,688). An example of a pyridoxal-dependent acetoin aminase is the amine:pyruvate aminotransferase (also called amine:pyruvate transaminase) described by Shin and Kim (*J. Org. Chem.* 67:2848-2853, 2002).

[00186] The term “acetoin kinase” refers to a polypeptide (or polypeptides) having an enzyme activity that catalyzes the conversion of acetoin to phosphoacetoin. Acetoin kinase may

utilize ATP (adenosine triphosphate) or phosphoenolpyruvate as the phosphate donor in the reaction. Enzymes that catalyze the analogous reaction on the similar substrate dihydroxyacetone, for example, include enzymes known as EC 2.7.1.29 (Garcia-Alles, *et al.*, *Biochemistry* 43:13037-13046, 2004).

[00187] The term “acetoin phosphate aminase” refers to a polypeptide (or polypeptides) having an enzyme activity that catalyzes the conversion of phosphoacetoin to 3-amino-2-butanol O-phosphate. Acetoin phosphate aminase may use the cofactor pyridoxal 5'-phosphate, NADH or NADPH. The resulting product may have (R) or (S) stereochemistry at the 3-position. The pyridoxal phosphate-dependent enzyme may use an amino acid such as alanine or glutamate. The NADH and NADPH-dependent enzymes may use ammonia as a second substrate. Although there are no reports of enzymes catalyzing this reaction on phosphoacetoin, there is a pyridoxal phosphate-dependent enzyme that is proposed to carry out the analogous reaction on the similar substrate serinol phosphate (Yasuta, *et al.*, *Appl. Environ. Microbiol.* 67:4999-5009, 2001).

[00188] The term “aminobutanol phosphate phospholyase,” also called “amino alcohol O-phosphate lyase,” refers to a polypeptide (or polypeptides) having an enzyme activity that catalyzes the conversion of 3-amino-2-butanol O-phosphate to 2-butanone. Amino butanol phosphate phospho-lyase may utilize the cofactor pyridoxal 5'-phosphate. There are reports of enzymes that catalyze the analogous reaction on the similar substrate 1-amino-2-propanol phosphate (Jones, *et al.*, *Biochem J.* 134:167-182, 1973). U.S. Patent Application Publication No. 2007/0259410 describes an aminobutanol phosphate phospho-lyase from the organism *Erwinia carotovora*.

[00189] The term “aminobutanol kinase” refers to a polypeptide (or polypeptides) having an enzyme activity that catalyzes the conversion of 3-amino-2-butanol to 3-amino-2-butanol O-phosphate. Amino butanol kinase may utilize ATP as the phosphate donor. Although there are no reports of enzymes catalyzing this reaction on 3-amino-2-butanol, there are reports of enzymes that catalyze the analogous reaction on the similar substrates ethanolamine and 1-amino-2-propanol (Jones, *et al.*, *supra*). U.S. Patent Application Publication No. 2009/0155870 describes, in Example 14, an amino alcohol kinase of *Erwinia carotovora subsp. Atroseptica*.

[00190] The term “butanediol dehydrogenase” also known as “acetoin reductase” refers to a polypeptide (or polypeptides) having an enzyme activity that catalyzes the conversion of acetoin to 2,3-butanediol. Butanediol dehydrogenases are a subset of the broad family of alcohol dehydrogenases. Butanediol dehydrogenase enzymes may have specificity for production of (R)-

or (S)-stereochemistry in the alcohol product. (S)-specific butanediol dehydrogenases are known as EC 1.1.1.76 and are available, for example, from *Klebsiella pneumoniae* (GenBank Nos: BBA13085, D86412). (R)-specific butanediol dehydrogenases are known as EC 1.1.1.4 and are available, for example, from *Bacillus cereus* (GenBank Nos. NP 830481, NC_004722; AAP07682, AE017000), and *Lactococcus lactis* (GenBank Nos. AAK04995, AE006323).

[00191] The term “butanediol dehydratase,” also known as “dial dehydratase” or “propanediol dehydratase” refers to a polypeptide (or polypeptides) having an enzyme activity that catalyzes the conversion of 2,3-butanediol to 2-butanone. Butanediol dehydratase may utilize the cofactor adenosyl cobalamin (also known as coenzyme B₁₂ or vitamin B₁₂; although vitamin B₁₂ may refer also to other forms of cobalamin that are not coenzyme B₁₂). Adenosyl cobalamin-dependent enzymes are known as EC 4.2.1.28 and are available, for example, from *Klebsiella oxytoca* (GenBank Nos: AA08099 (alpha subunit), D45071; BAA08100 (beta subunit), D45071; and BBA08101 (gamma subunit), D45071 (Note all three subunits are required for activity), and *Klebsiella pneumonia* (GenBank Nos: AAC98384 (alpha subunit), AF102064; GenBank Nos: AAC98385 (beta subunit), AF102064, GenBank Nos: AAC98386 (gamma subunit), AF102064). Other suitable dial dehydratases include, but are not limited to, B₁₂-dependent dial dehydratases available from *Salmonella typhimurium* (GenBank Nos: AAB84102 (large subunit), AF026270; GenBank Nos: AAB84103 (medium subunit), AF026270; GenBank Nos: AAB84104 (small subunit), AF026270); and *Lactobacillus collinoides* (GenBank Nos: CAC82541 (large subunit), AJ297723 ; GenBank Nos: CAC82542 (medium subunit); AJ297723; GenBank Nos: CAD01091 (small subunit), AJ297723); and enzymes from *Lactobacillus brevis* (particularly strains CNRZ 734 and CNRZ 735, Speranza, *et al.*, *J. Agric. Food Chem.* 45:3476-3480, 1997), and nucleotide sequences that encode the corresponding enzymes. Methods of diol dehydratase gene isolation are well known in the art (*e.g.*, U.S. Patent No. 5,686,276).

[00192] The term “pyruvate decarboxylase” refers to an enzyme that catalyzes the decarboxylation of pyruvic acid to acetaldehyde and carbon dioxide. Pyruvate decarboxylases are known by the EC number 4.1.1.1. These enzymes are found in a number of yeast, including *Saccharomyces cerevisiae* (GenBank Nos: CAA97575, CAA97705, CAA97091).

[00193] It will be appreciated that host cells comprising an isobutanol biosynthetic pathway as provided herein may further comprise one or more additional modifications. U.S. Patent Application Publication No. 2009/0305363 (incorporated by reference) discloses increased

conversion of pyruvate to acetolactate by engineering yeast for expression of a cytosol-localized acetolactate synthase and substantial elimination of pyruvate decarboxylase activity. In some embodiments, the host cells comprise modifications to reduce glycerol-3-phosphate dehydrogenase activity and/or disruption in at least one gene encoding a polypeptide having pyruvate decarboxylase activity or a disruption in at least one gene encoding a regulatory element controlling pyruvate decarboxylase gene expression as described in U.S. Patent Application Publication No. 2009/0305363 (incorporated herein by reference), modifications to a host cell that provide for increased carbon flux through an Entner-Doudoroff Pathway or reducing equivalents balance as described in U.S. Patent Application Publication No. 2010/0120105 (incorporated herein by reference). Other modifications include integration of at least one polynucleotide encoding a polypeptide that catalyzes a step in a pyruvate-utilizing biosynthetic pathway.

[00194] Other modifications include at least one deletion, mutation, and/or substitution in an endogenous polynucleotide encoding a polypeptide having acetolactate reductase activity. As used herein, “acetolactate reductase activity” refers to the activity of any polypeptide having the ability to catalyze the conversion of acetolactate to DHMB. Such polypeptides can be determined by methods well known in the art and disclosed herein. As used herein, “DHMB” refers to 2,3-dihydroxy-2-methyl butyrate. DHMB includes “fast DHMB,” which has the 2S, 3S configuration, and “slow DHMB,” which has the 2S, 3R configuration. *See Kaneko et al., Phytochemistry 39: 115-120 (1995)*, which is herein incorporated by reference in its entirety and refers to fast DHMB as anglyceric acid and slow DHMB as tiglyceric acid. In embodiments, the polypeptide having acetolactate reductase activity is YMR226C of *Saccharomyces cerevisiae* or a homolog thereof.

[00195] Additional modifications include a deletion, mutation, and/or substitution in an endogenous polynucleotide encoding a polypeptide having aldehyde dehydrogenase and/or aldehyde oxidase activity. As used herein, “aldehyde dehydrogenase activity” refers to any polypeptide having a biological function of an aldehyde dehydrogenase. Such polypeptides include a polypeptide that catalyzes the oxidation (dehydrogenation) of aldehydes. Such polypeptides include a polypeptide that catalyzes the conversion of isobutyraldehyde to isobutyric acid. Such polypeptides also include a polypeptide that corresponds to Enzyme Commission Numbers EC 1.2.1.3, EC 1.2.1.4 or EC 1.2.1.5. Such polypeptides can be determined by methods well known in the art and disclosed herein. As used herein, “aldehyde

oxidase activity” refers to any polypeptide having a biological function of an aldehyde oxidase. Such polypeptides include a polypeptide that catalyzes production of carboxylic acids from aldehydes. Such polypeptides include a polypeptide that catalyzes the conversion of isobutyraldehyde to isobutyric acid. Such polypeptides also include a polypeptide that corresponds to Enzyme Commission Number EC 1.2.3.1. Such polypeptides can be determined by methods well known in the art and disclosed herein. In some embodiments, the polypeptide having aldehyde dehydrogenase activity is *ALD6* from *Saccharomyces cerevisiae* or a homolog thereof.

[00196] A genetic modification which has the effect of reducing glucose repression wherein the yeast production host cell is *pdc-* is described in U.S. Patent Application Publication No. 2011/0124060, incorporated herein by reference. In some embodiments, the pyruvate decarboxylase that is deleted or down-regulated is selected from the group consisting of: *PDC1*, *PDC5*, *PDC6*, and combinations thereof. In some embodiments, the pyruvate decarboxylase is selected from *PDC1* pyruvate decarboxylase from *Saccharomyces cerevisiae*, *PDC5* pyruvate decarboxylase from *Saccharomyces cerevisiae*, *PDC6* pyruvate decarboxylase from *Saccharomyces cerevisiae*, pyruvate decarboxylase from *Candida glabrata*, *PDC1* pyruvate decarboxylase from *Pichia stipites*, *PDC2* pyruvate decarboxylase from *Pichia stipites*, pyruvate decarboxylase from *Kluyveromyces lactis*, pyruvate decarboxylase from *Yarrowia lipolytica*, pyruvate decarboxylase from *Schizosaccharomyces pombe*, and pyruvate decarboxylase from *Zygosaccharomyces rouxii*. In some embodiments, host cells contain a deletion or down-regulation of a polynucleotide encoding a polypeptide that catalyzes the conversion of glyceraldehyde-3-phosphate to glycerate 1,3, bisphosphate. In some embodiments, the enzyme that catalyzes this reaction is glyceraldehyde-3-phosphate dehydrogenase.

[00197] WIPO publication number WO 2011/103300 discloses recombinant host cells comprising (a) at least one heterologous polynucleotide encoding a polypeptide having dihydroxy-acid dehydratase activity; and (b)(i) at least one deletion, mutation, and/or substitution in an endogenous gene encoding a polypeptide affecting Fe-S cluster biosynthesis; and/or (ii) at least one heterologous polynucleotide encoding a polypeptide affecting Fe-S cluster biosynthesis. In embodiments, the polypeptide affecting Fe-S cluster biosynthesis is encoded by *AFT1*, *AFT2*, *FRA2*, *GRX3*, or *CCCI*. In embodiments, the polypeptide affecting Fe-S cluster biosynthesis is constitutive mutant *AFT1* L99A, *AFT1* L102A, *AFT1* C291F, or *AFT1* C293F.

[00198] Additionally, host cells may comprise heterologous polynucleotides encoding a polypeptide with phosphoketolase activity and/or a heterologous polynucleotide encoding a polypeptide with phosphotransacetylase activity as described in U.S. Patent Application No. 2012/0156735, incorporated herein by reference.

[00199] In some embodiments, any particular nucleic acid molecule or polypeptide may be at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% identical to a nucleotide sequence or polypeptide sequence described herein. The term “percent identity” as known in the art, is a relationship between two or more polypeptide sequences or two or more polynucleotide sequences, as determined by comparing the sequences. In the art, “identity” also means the degree of sequence relatedness between polypeptide or polynucleotide sequences, as the case may be, as determined by the match between strings of such sequences. “Identity” and “similarity” can be readily calculated by known methods, including but not limited to those disclosed in: 1.) *Computational Molecular Biology* (Lesk, A. M., Ed.) Oxford University: NY (1988); 2.) *Biocomputing: Informatics and Genome Projects* (Smith, D. W., Ed.) Academic: NY (1993); 3.) *Computer Analysis of Sequence Data, Part I* (Griffin, A. M., and Griffin, H. G., Eds.) Humana: NJ (1994); 4.) *Sequence Analysis in Molecular Biology* (von Heinje, G., Ed.) Academic (1987); and 5.) *Sequence Analysis Primer* (Gribskov, M. and Devereux, J., Eds.) Stockton: NY (1991).

[00200] Standard recombinant DNA and molecular cloning techniques are well known in the art and are described by Sambrook, *et al.* (Sambrook, J., Fritsch, E. F. and Maniatis, T. (Molecular Cloning: A Laboratory Manual; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 1989, here in referred to as Maniatis) and by Ausubel, *et al.* (Ausubel, *et al.*, *Current Protocols in Molecular Biology*, pub. by Greene Publishing Assoc. and Wiley-Interscience, 1987). Examples of methods to construct microorganisms that comprise a butanol biosynthetic pathway are disclosed, for example, in U.S. Patent No. 7,851,188, and U.S. Patent Application Publication Nos. 2007/0092957; 2007/0259410; 2007/0292927; 2008/0182308; 2008/0274525; 2009/0155870; 2009/0305363; and 2009/0305370, the entire contents of each are herein incorporated by reference.

EXPRESSION OF A BUTANOL BIOSYNTHETIC PATHWAY IN SACCHAROMYCES CEREVISIAE

[00201] Methods for gene expression in *Saccharomyces cerevisiae* are known in the art (e.g., Methods in Enzymology, Volume 194, Guide to Yeast Genetics and Molecular and Cell Biology, Part A, 2004, Christine Guthrie and Gerald R. Fink, eds., Elsevier Academic Press, San Diego, CA). Expression of genes in yeast typically requires a promoter, followed by the gene of interest, and a transcriptional terminator. A number of yeast promoters, including those used in the Examples herein, can be used in constructing expression cassettes for genes encoding an isobutanol biosynthetic pathway, including, but not limited to constitutive promoters FBA, GPD, ADH1, and GPM, and the inducible promoters GAL1, GAL10, and CUP1. Suitable transcriptional terminators include, but are not limited to FBA_t, GPD_t, GPM_t, ERG10_t, GAL1_t, CYC1, and ADH1. For example, suitable promoters, transcriptional terminators, and the genes of an isobutanol biosynthetic pathway can be cloned into *E. coli*-yeast shuttle vectors and transformed into yeast cells as described in U.S. App. Pub. No. 2010/0129886. These vectors allow strain propagation in both *E. coli* and yeast strains. Typically the vector contains a selectable marker and sequences allowing autonomous replication or chromosomal integration in the desired host. Typically used plasmids in yeast are shuttle vectors pRS423, pRS424, pRS425, and pRS426 (American Type Culture Collection, Rockville, MD), which contain an *E. coli* replication origin (e.g., pMB1), a yeast 2 μ origin of replication, and a marker for nutritional selection. The selection markers for these four vectors are His3 (vector pRS423), Trp1 (vector pRS424), Leu2 (vector pRS425) and Ura3 (vector pRS426). Construction of expression vectors with genes encoding polypeptides of interest can be performed by either standard molecular cloning techniques in *E. coli* or by the gap repair recombination method in yeast.

[00202] The gap repair cloning approach takes advantage of the highly efficient homologous recombination in yeast. Typically, a yeast vector DNA is digested (e.g., in its multiple cloning site) to create a “gap” in its sequence. A number of insert DNAs of interest are generated that contain a ≥ 21 bp sequence at both the 5' and the 3' ends that sequentially overlap with each other, and with the 5' and 3' terminus of the vector DNA. For example, to construct a yeast expression vector for “Gene X”, a yeast promoter and a yeast terminator are selected for the expression cassette. The promoter and terminator are amplified from the yeast genomic DNA, and Gene X is either PCR amplified from its source organism or obtained from a cloning vector comprising Gene X sequence. There is at least a 21 bp overlapping sequence between the 5' end of the linearized vector and the promoter sequence, between the promoter and Gene X, between Gene X and the terminator sequence, and between the terminator and the 3' end of the linearized

vector. The “gapped” vector and the insert DNAs are then co-transformed into a yeast strain and plated on the medium containing the appropriate compound mixtures that allow complementation of the nutritional selection markers on the plasmids. The presence of correct insert combinations can be confirmed by PCR mapping using plasmid DNA prepared from the selected cells. The plasmid DNA isolated from yeast (usually low in concentration) can then be transformed into an *E. coli* strain, e.g., *TOP10*, followed by mini preps and restriction mapping to further verify the plasmid construct. Finally the construct can be verified by sequence analysis.

[00203] Like the gap repair technique, integration into the yeast genome also takes advantage of the homologous recombination system in yeast. Typically, a cassette containing a coding region plus control elements (promoter and terminator) and auxotrophic marker is PCR-amplified with a high-fidelity DNA polymerase using primers that hybridize to the cassette and contain 40-70 base pairs of sequence homology to the regions 5' and 3' of the genomic area where insertion is desired. The PCR product is then transformed into yeast and plated on medium containing the appropriate compound mixtures that allow selection for the integrated auxotrophic marker. For example, to integrate “Gene X” into chromosomal location “Y”, the promoter-coding regionX-terminator construct is PCR amplified from a plasmid DNA construct and joined to an autotrophic marker (such as *URA3*) by either SOE PCR or by common restriction digests and cloning. The full cassette, containing the promoter-coding regionX-terminator-*URA3* region, is PCR amplified with primer sequences that contain 40-70 bp of homology to the regions 5' and 3' of location “Y” on the yeast chromosome. The PCR product is transformed into yeast and selected on growth media lacking uracil. Transformants can be verified either by colony PCR or by direct sequencing of chromosomal DNA.

GROWTH FOR PRODUCTION

[00204] Recombinant host cells disclosed herein are contacted with suitable carbon substrates, typically in fermentation media. Additional carbon substrates may include, but are not limited to, monosaccharides such as fructose, oligosaccharides such as lactose, maltose, galactose, or sucrose, polysaccharides such as starch or cellulose or mixtures thereof and unpurified mixtures from renewable feedstocks such as cheese whey permeate, cornsteep liquor, sugar beet molasses, and barley malt. Other carbon substrates can include ethanol, lactate, succinate, or glycerol.

[00205] Additionally the carbon substrate may also be one-carbon substrates such as carbon dioxide, or methanol for which metabolic conversion into key biochemical intermediates has been demonstrated. In addition to one and two carbon substrates, methylotrophic organisms are also known to utilize a number of other carbon containing compounds such as methylamine, glucosamine and a variety of amino acids for metabolic activity. For example, methylotrophic yeasts are known to utilize the carbon from methylamine to form trehalose or glycerol (Bellion *et al.*, *Microb. Growth CI Compd.*, [Int. Symp.], 7th (1993), 415-32, Editors: Murrell, J. Collin, Kelly, Don P.; Publisher: Intercept, Andover, UK). Similarly, various species of *Candida* will metabolize alanine or oleic acid (Sulter *et al.*, *Arch. Microbiol.* 153:485-489 (1990)). Hence it is contemplated that the source of carbon utilized in the present invention may encompass a wide variety of carbon containing substrates and will only be limited by the choice of organism.

[00206] Although it is contemplated that all of the above mentioned carbon substrates and mixtures thereof are suitable in the present invention, in some embodiments, the carbon substrates are glucose, fructose, and sucrose, or mixtures of these with C5 sugars such as xylose and/or arabinose for yeasts cells modified to use C5 sugars. Sucrose may be derived from renewable sugar sources such as sugar cane, sugar beets, cassava, sweet sorghum, and mixtures thereof. Glucose and dextrose can be derived from renewable grain sources through saccharification of starch based feedstocks including grains such as corn, wheat, rye, barley, oats, and mixtures thereof. In addition, fermentable sugars can be derived from renewable cellulosic or lignocellulosic biomass through processes of pretreatment and saccharification, as described, for example, in U.S. Patent Application Publication No. 2007/0031918 A1, which is herein incorporated by reference. Biomass, when used in reference to carbon substrate, refers to any cellulosic or lignocellulosic material and includes materials comprising cellulose, and optionally further comprising hemicellulose, lignin, starch, oligosaccharides and/or monosaccharides. Biomass can also comprise additional components, such as protein and/or lipid. Biomass can be derived from a single source, or biomass can comprise a mixture derived from more than one source; for example, biomass may comprise a mixture of corn cobs and corn stover, or a mixture of grass and leaves. Biomass includes, but is not limited to, bioenergy crops, agricultural residues, municipal solid waste, industrial solid waste, sludge from paper manufacture, yard waste, wood and forestry waste. Examples of biomass include, but are not limited to, corn grain, corn cobs, crop residues such as corn husks, corn stover grasses, wheat, wheat straw, barley, barley straw, hay, rice straw, switchgrass, waste paper, sugar cane bagasse, sorghum, soy,

components obtained from milling of grains, trees, branches, roots, leaves, wood chips, sawdust, shrubs and bushes, vegetables, fruits, flowers, animal manure, and mixtures thereof.

[00207] In addition to an appropriate carbon source, fermentation media must contain suitable minerals, salts, cofactors, buffers and other components, known to those skilled in the art, suitable for the growth of the cultures and promotion of an enzymatic pathway described herein.

CULTURE CONDITIONS

[00208] Typically cells are grown at a temperature in the range of about 20 °C to about 40 °C in an appropriate medium. Suitable growth media in the present invention are common commercially prepared media such as Luria Bertani (LB) broth, Sabouraud Dextrose (SD) broth or Yeast Medium (YM) broth or broth that includes yeast nitrogen base, ammonium sulfate, and dextrose (as the carbon/energy source) or YPD Medium, a blend of peptone, yeast extract, and dextrose in optimal proportions for growing most *Saccharomyces cerevisiae* strains. Other defined or synthetic growth media can also be used, and the appropriate medium for growth of the particular microorganism will be known by one skilled in the art of microbiology or fermentation science. The use of agents known to modulate catabolite repression directly or indirectly, e.g., cyclic adenosine 2',3'-monophosphate (cAMP), can also be incorporated into the fermentation medium.

[00209] Suitable pH ranges for the fermentation are between pH 5.0 to pH 9.0, where pH 6.0 to pH 8.0 is preferred for the initial condition. Suitable pH ranges for the fermentation of yeast are typically between about pH 3.0 to about pH 9.0. In one embodiment, about pH 5.0 to about pH 8.0 is used for the initial condition. Suitable pH ranges for the fermentation of other microorganisms are between about pH 3.0 to about pH 7.5. In one embodiment, about pH 4.5 to about pH 6.5 is used for the initial condition.

[00210] Fermentations can be performed under aerobic or anaerobic conditions. In one embodiment, anaerobic or microaerobic conditions are used for fermentation.

INDUSTRIAL BATCH AND CONTINUOUS FERMENTATIONS

[00211] Butanol, or other products, can be produced using a batch method of fermentation. A classical batch fermentation is a closed system where the composition of the medium is set at the beginning of the fermentation and not subject to artificial alterations during the fermentation.

A variation on the standard batch system is the fed-batch system. Fed-batch fermentation processes are also suitable in the present invention and comprise a typical batch system with the exception that the substrate is added in increments at the fermentation progresses. Fed-batch systems are useful when catabolite repression is apt to inhibit the metabolism of the cells and where it is desirable to have limited amounts of substrate in the media. Batch and fed-batch fermentations are common and well known in the art and examples can be found in Thomas D. Brock in *Biotechnology: A Textbook of Industrial Microbiology*, Second Edition (1989) Sinauer Associates, Inc., Sunderland, MA., or Deshpande, Mukund V., *Appl. Biochem. Biotechnol.*, 36:227, (1992), herein incorporated by reference.

[00212] Butanol, or other products, may also be produced using continuous fermentation methods. Continuous fermentation is an open system where a defined fermentation medium is added continuously to a bioreactor and an equal amount of conditioned media is removed simultaneously for processing. Continuous fermentation generally maintains the cultures at a constant high density where cells are primarily in log phase growth. Continuous fermentation allows for the modulation of one factor or any number of factors that affect cell growth or end product concentration. Methods of modulating nutrients and growth factors for continuous fermentation processes as well as techniques for maximizing the rate of product formation are well known in the art of industrial microbiology and a variety of methods are detailed by Brock, *supra*.

[00213] It is contemplated that the production of butanol, or other products, can be practiced using batch, fed-batch or continuous processes and that any known mode of fermentation would be suitable. Additionally, it is contemplated that cells can be immobilized on a substrate as whole cell catalysts and subjected to fermentation conditions for butanol production.

METHODS FOR BUTANOL ISOLATION FROM THE FERMENTATION MEDIUM

[00214] Bioproduced butanol may be isolated from the fermentation medium using methods known in the art for ABE fermentations (*see, e.g., Durre, Appl. Microbiol. Biotechnol.* 49:639-648 (1998), Groot *et al.*, *Process. Biochem.* 27:61-75 (1992), and references therein). For example, solids may be removed from the fermentation medium by centrifugation, filtration, decantation, or the like. The butanol may be isolated from the fermentation medium using

methods such as distillation, azeotropic distillation, liquid-liquid extraction, adsorption, gas stripping, membrane evaporation, or pervaporation.

[00215] Because butanol forms a low boiling point, azeotropic mixture with water, distillation can be used to separate the mixture up to its azeotropic composition. Distillation may be used in combination with the processes described herein to obtain separation around the azeotrope. Methods that may be used in combination with distillation to isolate and purify butanol include, but are not limited to, decantation, liquid-liquid extraction, adsorption, and membrane-based techniques. Additionally, butanol may be isolated using azeotropic distillation using an entrainer (see, *e.g.*, Doherty and Malone, *Conceptual Design of Distillation Systems*, McGraw Hill, New York, 2001).

[00216] The butanol-water mixture forms a heterogeneous azeotrope so that distillation may be used in combination with decantation to isolate and purify the isobutanol. In this method, the butanol containing fermentation broth is distilled to near the azeotropic composition. Then, the azeotropic mixture is condensed, and the butanol is separated from the fermentation medium by decantation. The decanted aqueous phase may be returned to the first distillation column as reflux or to a separate stripping column. The butanol-rich decanted organic phase may be further purified by distillation in a second distillation column.

[00217] The butanol can also be isolated from the fermentation medium using liquid-liquid extraction in combination with distillation. In this method, the butanol is extracted from the fermentation broth using liquid-liquid extraction with a suitable solvent. The butanol-containing organic phase is then distilled to separate the butanol from the solvent.

[00218] Distillation in combination with adsorption can also be used to isolate butanol from the fermentation medium. In this method, the fermentation broth containing the butanol is distilled to near the azeotropic composition and then the remaining water is removed by use of an adsorbent, such as molecular sieves (Aden *et al.*, *Lignocellulosic Biomass to Ethanol Process Design and Economics Utilizing Co-Current Dilute Acid Prehydrolysis and Enzymatic Hydrolysis for Corn Stover*, Report NREL/TP-510-32438, National Renewable Energy Laboratory, June 2002).

[00219] Additionally, distillation in combination with pervaporation can be used to isolate and purify the butanol from the fermentation medium. In this method, the fermentation broth containing the butanol is distilled to near the azeotropic composition, and then the remaining

water is removed by pervaporation through a hydrophilic membrane (Guo *et al.*, *J. Membr. Sci.* 245, 199-210 (2004)).

[00220] *In situ* product removal (ISPR) (also referred to as extractive fermentation) can be used to remove butanol (or other fermentative alcohol) from the fermentation vessel as it is produced, thereby allowing the microorganism to produce butanol at high yields. One method for ISPR for removing fermentative alcohol that has been described in the art is liquid-liquid extraction. In general, with regard to butanol fermentation, for example, the fermentation medium, which includes the microorganism, is contacted with an organic extractant at a time before the butanol concentration reaches a toxic level. The organic extractant and the fermentation medium form a biphasic mixture. The butanol partitions into the organic extractant phase, decreasing the concentration in the aqueous phase containing the microorganism, thereby limiting the exposure of the microorganism to the inhibitory butanol.

[00221] Liquid-liquid extraction can be performed, for example, according to the processes described in U.S. Patent Appl. Pub. Nos. 2009/0305370 and 2011/0097773, the disclosures of which are hereby incorporated in their entirety. U.S. Patent Appl. Pub. Nos. 2009/0305370 and 2011/0097773 describe methods for producing and recovering butanol from a fermentation broth using liquid-liquid extraction, the methods comprising the step of contacting the fermentation broth with a water immiscible extractant to form a two-phase mixture comprising an aqueous phase and an organic phase. Typically, the extractant can be an organic extractant selected from the group consisting of saturated, mono-unsaturated, polyunsaturated (and mixtures thereof) C₁₂ to C₂₂ fatty alcohols, C₁₂ to C₂₂ fatty acids, esters of C₁₂ to C₂₂ fatty acids, C₁₂ to C₂₂ fatty aldehydes, and mixtures thereof. The extractant(s) for ISPR can be non-alcohol extractants. The ISPR extractant can be an exogenous organic extractant such as oleyl alcohol, behenyl alcohol, cetyl alcohol, lauryl alcohol, myristyl alcohol, stearyl alcohol, 1-undecanol, oleic acid, lauric acid, myristic acid, stearic acid, methyl myristate, methyl oleate, undecanal, lauric aldehyde, 20-methylundecanal, and mixtures thereof.

[00222] In some embodiments, an alcohol ester can be formed by contacting the alcohol in a fermentation medium with an organic acid (*e.g.*, fatty acids) and a catalyst capable of esterifying the alcohol with the organic acid. In such embodiments, the organic acid can serve as an ISPR extractant into which the alcohol esters partition. The organic acid can be supplied to the fermentation vessel and/or derived from the biomass supplying fermentable carbon fed to the fermentation vessel. Lipids present in the feedstock can be catalytically hydrolyzed to organic

acid, and the same catalyst (*e.g.*, enzymes) can esterify the organic acid with the alcohol. In certain embodiments, lipids present in the feedstock can be converted into fatty acids and glycerol utilizing the catalysts (*e.g.*, enzymes) described above. The glycerol can, for example, be provided to the fermentation vessel to supplement the microorganisms with reduced glycerol production described herein. Supplementing the microorganisms can, for example, improve biomass production and microorganism cell health. The glycerol will be provided in sufficient amounts beyond that produced by yeast under fermentation conditions. Carboxylic acids that are produced during the fermentation can additionally be esterified with the alcohol produced by the same or a different catalyst. The catalyst can be supplied to the feedstock prior to fermentation, or can be supplied to the fermentation vessel before or contemporaneously with the supplying of the feedstock. When the catalyst is supplied to the fermentation vessel, alcohol esters can be obtained by hydrolysis of the lipids into organic acid and substantially simultaneous esterification of the organic acid with butanol present in the fermentation vessel. Organic acid and/or native oil not derived from the feedstock can also be fed to the fermentation vessel, with the native oil being hydrolyzed into organic acid. Any organic acid not esterified with the alcohol can serve as part of the ISPR extractant. The extractant containing alcohol esters can be separated from the fermentation medium, and the alcohol can be recovered from the extractant. The extractant can be recycled to the fermentation vessel. Thus, in the case of butanol production, for example, the conversion of the butanol to an ester reduces the free butanol concentration in the fermentation medium, shielding the microorganism from the toxic effect of increasing butanol concentration. In addition, unfractionated grain can be used as feedstock without separation of lipids therein, since the lipids can be catalytically hydrolyzed to organic acid, thereby decreasing the rate of build-up of lipids in the ISPR extractant.

[00223] *In situ* product removal can be carried out in a batch mode or a continuous mode. In a continuous mode of *in situ* product removal, product is continually removed from the reactor. In a batchwise mode of *in situ* product removal, a volume of organic extractant is added to the fermentation vessel and the extractant is not removed during the process. For *in situ* product removal, the organic extractant can contact the fermentation medium at the start of the fermentation forming a biphasic fermentation medium. Alternatively, the organic extractant can contact the fermentation medium after the microorganism has achieved a desired amount of growth, which can be determined by measuring the optical density of the culture. Further, the organic extractant can contact the fermentation medium at a time at which the product alcohol

level in the fermentation medium reaches a preselected level. In the case of butanol production according to some embodiments of the present invention, the organic acid extractant can contact the fermentation medium at a time before the butanol concentration reaches a toxic level, so as to esterify the butanol with the organic acid to produce butanol esters and consequently reduce the concentration of butanol in the fermentation vessel. The ester-containing organic phase can then be removed from the fermentation vessel (and separated from the fermentation broth which constitutes the aqueous phase) after a desired effective titer of the butanol esters is achieved. In some embodiments, the ester-containing organic phase is separated from the aqueous phase after fermentation of the available fermentable sugar in the fermentation vessel is substantially complete.

CONFIRMATION OF ISOBUTANOL PRODUCTION

[00224] The presence and/or concentration of isobutanol in the culture medium can be determined by a number of methods known in the art (see, for example, U.S. Patent 7,851,188, incorporated by reference). For example, a specific high performance liquid chromatography (HPLC) method utilizes a Shodex SH-1011 column with a Shodex SHG guard column, both may be purchased from Waters Corporation (Milford, Mass.), with refractive index (RI) detection. Chromatographic separation is achieved using 0.01 M H₂SO₄ as the mobile phase with a flow rate of 0.5 mL/min and a column temperature of 50 °C. Isobutanol has a retention time of 46.6 min under the conditions used.

[00225] Alternatively, gas chromatography (GC) methods are available. For example, a specific GC method utilizes an HP-INNOWax column (30 m X 0.53 mm id, 1 µm film thickness, Agilent Technologies, Wilmington, DE), with a flame ionization detector (FID). The carrier gas is helium at a flow rate of 4.5 mL/min, measured at 150 °C with constant head pressure; injector split is 1:25 at 200 °C; oven temperature is 45 °C for 1 min, 45 to 220 °C at 10 °C/min, and 220 °C for 5 min; and FID detection is employed at 240 °C with 26 mL/min helium makeup gas. The retention time of isobutanol is 4.5 min.

[00226] While various embodiments of the present invention have been described herein, it should be understood that they have been presented by way of example only, and not limitation. It will be apparent to persons skilled in the relevant art that various changes in form and detail can be made therein without departing from the spirit and scope of the invention.

Thus, the breadth and scope of the present invention should not be limited by any of the above-described exemplary embodiments, but should be defined only in accordance with the claims and their equivalents.

[00227] All publications, patents, and patent applications mentioned in this specification are indicative of the level of those skilled in the art to which this invention pertains, and are herein incorporated by reference to the same extent as if each individual publication, patent, or patent application was specifically and individually indicated to be incorporated by reference.

EXAMPLES

[00228] The present invention is further defined in the following Examples. It should be understood that these Examples, while indicating embodiments of the invention, are given by way of illustration only. From the above discussion and these Examples, one skilled in the art can ascertain the essential characteristics of this invention, and without departing from the spirit and scope thereof, can make various changes and modifications of the invention to adapt it to various uses and conditions.

General Methods

[00229] Standard recombinant DNA, molecular cloning techniques and transformation protocols used in the Examples are well known in the art and are described by Sambrook et al. (Sambrook, J., Fritsch, E. F. and Maniatis, T. (Molecular Cloning: A Laboratory Manual; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 1989, here in referred to as Maniatis), by Ausubel et al. (Ausubel et al., Current Protocols in Molecular Biology, pub. by Greene Publishing Assoc. and Wiley-Interscience, 1987) and by Amberg et al (Amberg, D. C., Burke, D. J. and Strathern, J. N. (Methods in Yeast Genetics: A Cold Spring Harbor Laboratory Course Manual, Cold Spring Harbor Press, 2005). Materials and methods suitable for the maintenance and growth of bacterial cultures are well known in the art. Techniques suitable for use in the following examples may be found as set out in Manual of Methods for General Bacteriology (Phillipp et al., eds., American Society for Microbiology, Washington, DC., 1994) or by Thomas D. Brock in (Brock, Biotechnology: A Textbook of Industrial Microbiology, Second Edition, Sinauer Associates, Inc., Sunderland, MA (1989). All reagents, restriction enzymes and materials used for the growth and maintenance of bacterial cells were obtained from Sigma-Aldrich Chemicals (St. Louis, MO), BD Diagnostic Systems (Sparks, MD), Invitrogen (Carlsbad,

CA), HiMedia (Mumbai, India), SD Fine chemicals (India), or Takara Bio Inc. (Shiga,, Japan), unless otherwise specified.

[00230] The meaning of abbreviations is as follows: "sec" means second(s), "min" means minute(s), "h" means hour(s), "nm" means nanometers, "uL" means microliter(s), "mL" means milliliter(s), "mg/mL" means milligram per milliliter, "L" means liter(s), "nm" means nanometers, "mM" means millimolar, "M" means molar, "mmol" means millimole(s), "μmole" means micromole(s), "kg" means kilogram, "g" means gram(s), "μg" means microgram(s) and "ng" means nanogram(s), "PCR" means polymerase chain reaction, "OD" means optical density, "OD600" means the optical density measured at a wavelength of 600 nm, "kDa" means kilodaltons, "g" can also mean the gravitation constant, "bp" means base pair(s), "kbp" means kilobase pair(s), "kb" means kilobase, "%" means percent, "% w/v" means weight/volume percent, "% v/v" means volume/volume percent, "HPLC" means high performance liquid chromatography, "g/L" means gram per liter, "μg/L" means microgram per liter, "ng/μL" means nanogram per microliter, "pmol/μL" means picomol per microliter, "RPM" means rotation per minute, "μmol/min/mg" means micromole per minute per milligram, "w/v" means weight per volume, "v/v" means volume per volume.

Strain Construction

Construction of strain PNY2115

[00231] *Saccharomyces cerevisiae* strain PNY0827 is used as the host cell for further genetic manipulation for PNY2115. PNY0827 refers to a strain derived from *Saccharomyces cerevisiae* which has been deposited at the ATCC under the Budapest Treaty on September 22, 2011 at the American Type Culture Collection, Patent Depository 10801 University Boulevard, Manassas, VA 20110-2209 and has the patent deposit designation PTA-12105.

Deletion of URA3 and sporulation into haploids

[00232] In order to delete the endogenous URA3 coding region, a deletion cassette was PCR-amplified from pLA54 (SEQ ID NO: 95) which contains a P_{TEF1} -*kanMX4*-*TEF1*t cassette flanked by loxP sites to allow homologous recombination in vivo and subsequent removal of the KANMX4 marker. PCR was done by using Phusion High Fidelity PCR Master Mix (New England BioLabs; Ipswich, MA) and primers BK505 (SEQ ID NO: 96) and BK506 (SEQ ID NO: 97). The *URA3* portion of each primer was derived from the 5' region 180bp upstream of the

URA3 ATG and 3' region 78bp downstream of the coding region such that integration of the *kanMX4* cassette results in replacement of the *URA3* coding region. The PCR product was transformed into PNY0827 using standard genetic techniques (Methods in Yeast Genetics, 2005, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 201-202) and transformants were selected on YEP medium supplemented 2% glucose and 100 µg/ml Geneticin at 30°C. Transformants were screened by colony PCR with primers LA468 (SEQ ID NO: 98) and LA492 (SEQ ID NO: 99) to verify presence of the integration cassette. A heterozygous diploid was obtained: NYLA98, which has the genotype MATa/α *URA3/ura3::loxP-kanMX4-loxP*. To obtain haploids, NYLA98 was sporulated using standard methods (Codón AC, Gasent-Ramírez JM, Benítez T. Factors which affect the frequency of sporulation and tetrad formation in *Saccharomyces cerevisiae* baker's yeast. Appl Environ Microbiol. 1995 PMID: 7574601). Tetrads were dissected using a micromanipulator and grown on rich YPE medium supplemented with 2% glucose. Tetrads containing four viable spores were patched onto synthetic complete medium lacking uracil supplemented with 2% glucose, and the mating type was verified by multiplex colony PCR using primers AK109-1 (SEQ ID NO: 100), AK109-2 (SEQ ID NO: 101), and AK109-3 (SEQ ID NO: 102). The resulting identified haploid strain called NYLA103, which has the genotype: MATα *ura3Δ::loxP-kanMX4-loxP*, and NYLA106, which has the genotype: MATa *ura3Δ::loxP-kanMX4-loxP*.

Deletion of His3

[00233] To delete the endogenous *HIS3* coding region, a scarless deletion cassette was used. The four fragments for the PCR cassette for the scarless *HIS3* deletion were amplified using Phusion High Fidelity PCR Master Mix (New England BioLabs; Ipswich, MA) and CEN.PK 113-7D genomic DNA as template, prepared with a Gentra Puregene Yeast/Bact kit (Qiagen; Valencia, CA). *HIS3* Fragment A was amplified with primer oBP452 (SEQ ID NO: 103) and primer oBP453 (SEQ ID NO: 104), containing a 5' tail with homology to the 5' end of *HIS3* Fragment B. *HIS3* Fragment B was amplified with primer oBP454 (SEQ ID NO: 105), containing a 5' tail with homology to the 3' end of *HIS3* Fragment A, and primer oBP455 (SEQ ID NO: 106) containing a 5' tail with homology to the 5' end of *HIS3* Fragment U. *HIS3* Fragment U was amplified with primer oBP456 (SEQ ID NO: 107), containing a 5' tail with homology to the 3' end of *HIS3* Fragment B, and primer oBP457 (SEQ ID NO: 108), containing a 5' tail with homology to the 5' end of *HIS3* Fragment C. *HIS3* Fragment C was amplified with

primer oBP458 (SEQ ID NO: 109), containing a 5' tail with homology to the 3' end of *HIS3* Fragment U, and primer oBP459 (SEQ ID NO: 110). PCR products were purified with a PCR Purification kit (Qiagen). *HIS3* Fragment AB was created by overlapping PCR by mixing *HIS3* Fragment A and *HIS3* Fragment B and amplifying with primers oBP452 (SEQ ID NO: 103) and oBP455 (SEQ ID NO: 106). *HIS3* Fragment UC was created by overlapping PCR by mixing *HIS3* Fragment U and *HIS3* Fragment C and amplifying with primers oBP456 (SEQ ID NO: 107) and oBP459 (SEQ ID NO: 110). The resulting PCR products were purified on an agarose gel followed by a Gel Extraction kit (Qiagen). The *HIS3* ABUC cassette was created by overlapping PCR by mixing *HIS3* Fragment AB and *HIS3* Fragment UC and amplifying with primers oBP452 (SEQ ID NO: 103) and oBP459 (SEQ ID NO: 110). The PCR product was purified with a PCR Purification kit (Qiagen). Competent cells of NYLA106 were transformed with the *HIS3* ABUC PCR cassette and were plated on synthetic complete medium lacking uracil supplemented with 2% glucose at 30 °C. Transformants were screened to verify correct integration by replica plating onto synthetic complete medium lacking histidine and supplemented with 2% glucose at 30°C. Genomic DNA preps were made to verify the integration by PCR using primers oBP460 (SEQ ID NO: 111) and LA135 (SEQ ID NO: 112) for the 5' end and primers oBP461 (SEQ ID NO: 113) and LA92 (SEQ ID NO: 114) for the 3' end. The *URA3* marker was recycled by plating on synthetic complete medium supplemented with 2% glucose and 5-FOA at 30°C following standard protocols. Marker removal was confirmed by patching colonies from the 5-FOA plates onto SD –URA medium to verify the absence of growth. The resulting identified strain, called PNY2003 has the genotype: MATa *ura3Δ::loxP-kanMX4-loxP his3Δ*.

Deletion of PDC1

[00234] To delete the endogenous *PDC1* coding region, a deletion cassette was PCR-amplified from pLA59 (SEQ ID NO: 115), which contains a *URA3* marker flanked by degenerate loxP sites to allow homologous recombination in vivo and subsequent removal of the *URA3* marker. PCR was done by using Phusion High Fidelity PCR Master Mix (New England BioLabs; Ipswich, MA) and primers LA678 (SEQ ID NO: 116) and LA679 (SEQ ID NO: 117). The *PDC1* portion of each primer was derived from the 5' region 50bp downstream of the *PDC1* start codon and 3' region 50bp upstream of the stop codon such that integration of the *URA3* cassette results in replacement of the *PDC1* coding region but leaves the first 50bp and the last 50bp of the coding region. The PCR product was transformed into PNY2003 using standard

genetic techniques and transformants were selected on synthetic complete medium lacking uracil and supplemented with 2% glucose at 30°C. Transformants were screened to verify correct integration by colony PCR using primers LA337 (SEQ ID NO: 118), external to the 5' coding region and LA135 (SEQ ID NO: 112), an internal primer to *URA3*. Positive transformants were then screened by colony PCR using primers LA692 (SEQ ID NO: 119) and LA693 (SEQ ID NO: 120), internal to the *PDC1* coding region. The *URA3* marker was recycled by transforming with pLA34 (SEQ ID NO: 121) containing the CRE recombinase under the *GAL1* promoter and plated on synthetic complete medium lacking histidine and supplemented with 2% glucose at 30°C. Transformants were plated on rich medium supplemented with 0.5% galactose to induce the recombinase. Marker removal was confirmed by patching colonies to synthetic complete medium lacking uracil and supplemented with 2% glucose to verify absence of growth. The resulting identified strain, called PNY2008 has the genotype: MATa *ura3Δ::loxP-kanMX4-loxP his3Δ pdc1Δ::loxP71/66*.

Deletion of PDC5

[00235] To delete the endogenous *PDC5* coding region, a deletion cassette was PCR-amplified from pLA59 (SEQ ID NO: 115), which contains a *URA3* marker flanked by degenerate loxP sites to allow homologous recombination in vivo and subsequent removal of the *URA3* marker. PCR was done by using Phusion High Fidelity PCR Master Mix (New England BioLabs; Ipswich, MA) and primers LA722 (SEQ ID NO: 122) and LA733 (SEQ ID NO: 123). The *PDC5* portion of each primer was derived from the 5' region 50bp upstream of the *PDC5* start codon and 3' region 50bp downstream of the stop codon such that integration of the *URA3* cassette results in replacement of the entire *PDC5* coding region. The PCR product was transformed into PNY2008 using standard genetic techniques and transformants were selected on synthetic complete medium lacking uracil and supplemented with 1% ethanol at 30°C. Transformants were screened to verify correct integration by colony PCR using primers LA453 (SEQ ID NO: 124), external to the 5' coding region and LA135 (SEQ ID NO: 112), an internal primer to *URA3*. Positive transformants were then screened by colony PCR using primers LA694 (SEQ ID NO: 125) and LA695 (SEQ ID NO: 126), internal to the *PDC5* coding region. The *URA3* marker was recycled by transforming with pLA34 (SEQ ID NO: 121) containing the CRE recombinase under the *GAL1* promoter and plated on synthetic complete medium lacking histidine and supplemented with 1% ethanol at 30°C. Transformants were plated on rich YEP

medium supplemented with 1% ethanol and 0.5% galactose to induce the recombinase. Marker removal was confirmed by patching colonies to synthetic complete medium lacking uracil and supplemented with 1% ethanol to verify absence of growth. The resulting identified strain, called PNY2009 has the genotype: MATa *ura3Δ::loxP-kanMX4-loxP his3Δ pdc1Δ::loxP71/66 pdc5Δ::loxP71/66*.

Deletion of FRA2

[00236] The *FRA2* deletion was designed to delete 250 nucleotides from the 3' end of the coding sequence, leaving the first 113 nucleotides of the *FRA2* coding sequence intact. An in-frame stop codon was present 7 nucleotides downstream of the deletion. The four fragments for the PCR cassette for the scarless *FRA2* deletion were amplified using Phusion High Fidelity PCR Master Mix (New England BioLabs; Ipswich, MA) and CEN.PK 113-7D genomic DNA as template, prepared with a Gentra Puregene Yeast/Bact kit (Qiagen; Valencia, CA). *FRA2* Fragment A was amplified with primer oBP594 (SEQ ID NO: 127) and primer oBP595 (SEQ ID NO: 128), containing a 5' tail with homology to the 5' end of *FRA2* Fragment B. *FRA2* Fragment B was amplified with primer oBP596 (SEQ ID NO: 129), containing a 5' tail with homology to the 3' end of *FRA2* Fragment A, and primer oBP597 (SEQ ID NO: 130), containing a 5' tail with homology to the 5' end of *FRA2* Fragment U. *FRA2* Fragment U was amplified with primer oBP598 (SEQ ID NO: 131), containing a 5' tail with homology to the 3' end of *FRA2* Fragment B, and primer oBP599 (SEQ ID NO: 132), containing a 5' tail with homology to the 5' end of *FRA2* Fragment C. *FRA2* Fragment C was amplified with primer oBP600 (SEQ ID NO: 133), containing a 5' tail with homology to the 3' end of *FRA2* Fragment U, and primer oBP601 (SEQ ID NO: 134). PCR products were purified with a PCR Purification kit (Qiagen). *FRA2* Fragment AB was created by overlapping PCR by mixing *FRA2* Fragment A and *FRA2* Fragment B and amplifying with primers oBP594 (SEQ ID NO: 127) and oBP597 (SEQ ID NO: 130). *FRA2* Fragment UC was created by overlapping PCR by mixing *FRA2* Fragment U and *FRA2* Fragment C and amplifying with primers oBP598 (SEQ ID NO: 131) and oBP601 (SEQ ID NO: 134). The resulting PCR products were purified on an agarose gel followed by a Gel Extraction kit (Qiagen). The *FRA2* ABUC cassette was created by overlapping PCR by mixing *FRA2* Fragment AB and *FRA2* Fragment UC and amplifying with primers oBP594 (SEQ ID NO: 127) and oBP601 (SEQ ID NO: 134). The PCR product was purified with a PCR Purification kit (Qiagen).

[00237] To delete the endogenous *FRA2* coding region, the scarless deletion cassette obtained above was transformed into PNY2009 using standard techniques and plated on synthetic complete medium lacking uracil and supplemented with 1% ethanol. Genomic DNA preps were made to verify the integration by PCR using primers oBP602 (SEQ ID NO: 135) and LA135 (SEQ ID NO: 112) for the 5' end, and primers oBP602 (SEQ ID NO: 135) and oBP603 (SEQ ID NO: 136) to amplify the whole locus. The *URA3* marker was recycled by plating on synthetic complete medium supplemented with 1% ethanol and 5-FOA (5-Fluoroorotic Acid) at 30°C following standard protocols. Marker removal was confirmed by patching colonies from the 5-FOA plates onto synthetic complete medium lacking uracil and supplemented with 1% ethanol to verify the absence of growth. The resulting identified strain, PNY2037, has the genotype: MATa *ura3Δ::loxP-kanMX4-loxP his3Δ pdc1Δ::loxP71/66 pdc5Δ::loxP71/66 fra2Δ*.

Addition of native 2 micron plasmid

[00238] The loxP71-URA3-loxP66 marker was PCR-amplified using Phusion DNA polymerase (New England BioLabs; Ipswich, MA) from pLA59 (SEQ ID NO: 115), and transformed along with the LA811x817 (SEQ ID NOs: 137, 138) and LA812x818 (SEQ ID NOs: 139, 140) 2-micron plasmid fragments (amplified from the native 2-micron plasmid from CEN.PK 113-7D; Centraalbureau voor Schimmelcultures (CBS) Fungal Biodiversity Centre) into strain PNY2037 on SE –URA plates at 30°C. The resulting strain PNY2037 2μ::loxP71-URA3-loxP66 was transformed with pLA34 (pRS423::cre) (also called, pLA34) (SEQ ID NO: 121) and selected on SE –HIS –URA plates at 30°C. Transformants were patched onto YP-1% galactose plates and allowed to grow for 48 hrs at 30°C to induce Cre recombinase expression. Individual colonies were then patched onto SE –URA, SE –HIS, and YPE plates to confirm URA3 marker removal. The resulting identified strain, PNY2050, has the genotype: MATa *ura3Δ::loxP-kanMX4-loxP, his3Δ pdc1Δ::loxP71/66 pdc5Δ::loxP71/66 fra2Δ* 2-micron.

Construction of PNY2115 from PNY2050

[00239] Construction of PNY2115 [MATa *ura3Δ::loxP his3Δ pdc5Δ::loxP66/71 fra2Δ* 2-micron plasmid (CEN.PK2) *pdc1Δ::P[PDC1]-ALS|alsS_Bs-CYC1t-loxP71/66 pdc6Δ::(UAS)PGK1-P[FBA1]-KIVD|Lg(y)-TDH3t-loxP71/66 adh1Δ::P[ADH1]-ADH|Bi(y)-ADHt-loxP71/66 fra2Δ::P[ILV5]-ADH|Bi(y)-ADHt-loxP71/66 gpd2Δ::loxP71/66*] from PNY2050 was as follows.

Pdc1Δ::P[PDC1]-ALS|alsS_Bs-CYC1t-loxP71/66

[00240] To integrate *alsS* into the *pdc1Δ::loxP66/71* locus of PNY2050 using the endogenous PDC1 promoter, An integration cassette was PCR-amplified from pLA71 (SEQ ID NO: 146), which contains the gene acetolactate synthase from the species *Bacillus subtilis* with a FBA1 promoter and a CYC1 terminator, and a URA3 marker flanked by degenerate loxP sites to allow homologous recombination in vivo and subsequent removal of the URA3 marker. PCR was done by using KAPA HiFi and primers 895 (SEQ ID NO: 149) and 679 (SEQ ID NO: 150). The PDC1 portion of each primer was derived from 60bp of the upstream of the coding sequence and 50bp that are 53bp upstream of the stop codon. The PCR product was transformed into PNY2050 using standard genetic techniques and transformants were selected on synthetic complete media lacking uracil and supplemented with 1% ethanol at 30°C. Transformants were screened to verify correct integration by colony PCR using primers 681 (SEQ ID NO: 151), external to the 3' coding region and 92 (SEQ ID NO: 152), internal to the URA3 gene. Positive transformants were then prepped for genomic DNA and screened by PCR using primers N245 (SEQ ID NO: 153) and N246 (SEQ ID NO: 154). The *URA3* marker was recycled by transforming with pLA34 (SEQ ID NO: 121) containing the CRE recombinase under the GAL1 promoter and plated on synthetic complete media lacking histidine and supplemented with 1% ethanol at 30°C. Transformants were plated on rich media supplemented with 1% ethanol and 0.5% galactose to induce the recombinase. Marker removal was confirmed by patching colonies to synthetic complete media lacking uracil and supplemented with 1% ethanol to verify absence of growth. The resulting identified strain, called PNY2090 has the genotype MATa *ura3Δ::loxP*, *his3Δ*, *pdc1Δ::loxP71/66*, *pdc5Δ::loxP71/66* *fra2Δ* 2-micron *pdc1Δ::P[PDC1]-ALS|alsS_Bs-CYC1t-loxP71/66*.

Pdc6Δ::(UAS)PGK1-P[FBA1]-KIVD|Lg(y)-TDH3t-loxP71/66

[00241] To delete the endogenous PDC6 coding region, an integration cassette was PCR-amplified from pLA78 (SEQ ID NO: 147), which contains the *kivD* gene from the species *Listeria grayi* with a hybrid FBA1 promoter and a TDH3 terminator, and a URA3 marker flanked by degenerate loxP sites to allow homologous recombination in vivo and subsequent removal of the URA3 marker. PCR was done by using KAPA HiFi and primers 896 (SEQ ID NO: 155) and 897 (SEQ ID NO: 156). The PDC6 portion of each primer was derived from 60bp upstream of

the coding sequence and 59bp downstream of the coding region. The PCR product was transformed into PNY2090 using standard genetic techniques and transformants were selected on synthetic complete media lacking uracil and supplemented with 1% ethanol at 30°C. Transformants were screened to verify correct integration by colony PCR using primers 365 (SEQ ID NO: 157) and 366 (SEQ ID NO: 158), internal primers to the PDC6 gene. Transformants with an absence of product were then screened by colony PCR N638 (SEQ ID NO: 159), external to the 5' end of the gene, and 740 (SEQ ID NO: 160), internal to the FBA1 promoter. Positive transformants were then prepped for genomic DNA and screened by PCR with two external primers to the PDC6 coding sequence. Positive integrants would yield a 4720bp product, while PDC6 wild type transformants would yield a 2130bp product. The *URA3* marker was recycled by transforming with pLA34 containing the CRE recombinase under the GAL1 promoter and plated on synthetic complete media lacking histidine and supplemented with 1% ethanol at 30°C. Transformants were plated on rich media supplemented with 1% ethanol and 0.5% galactose to induce the recombinase. Marker removal was confirmed by patching colonies to synthetic complete media lacking uracil and supplemented with 1% ethanol to verify absence of growth. The resulting identified strain is called PNY2093 and has the genotype MATa ura3Δ::loxP his3Δ pdc5Δ::loxP71/66 fra2Δ 2-micron pdc1Δ::P[PDC1]-ALS|alsS_Bs-CYC1t-loxP71/66 pdc6Δ::(UAS)PGK1-P[FBA1]-KIVD|Lg(y)-TDH3t-loxP71/66.

Adh1Δ::P[ADH1]-ADH|Bi(y)-ADHt-loxP71/66

[00242] To delete the endogenous ADH1 coding region and integrate BiADH using the endogenous ADH1 promoter, an integration cassette was PCR-amplified from pLA65 (SEQ ID NO: 148), which contains the alcohol dehydrogenase from the species *Beijerinckii* with an ILV5 promoter and a ADH1 terminator, and a URA3 marker flanked by degenerate loxP sites to allow homologous recombination in vivo and subsequent removal of the URA3 marker. PCR was done by using KAPA HiFi and primers 856 (SEQ ID NO: 161) and 857 (SEQ ID NO: 162). The ADH1 portion of each primer was derived from the 5' region 50 bp upstream of the ADH1 start codon and the last 50 bp of the coding region. The PCR product was transformed into PNY2093 using standard genetic techniques and transformants were selected on synthetic complete media lacking uracil and supplemented with 1% ethanol at 30°C. Transformants were screened to verify correct integration by colony PCR using primers BK415 (SEQ ID NO: 163), external to the 5' coding region and N1092 (SEQ ID NO: 164), internal to the BiADH gene. Positive

transformants were then screened by colony PCR using primers 413 (SEQ ID NO: 169), external to the 3' coding region, and 92 (SEQ ID NO: 152), internal to the URA3 marker. The URA3 marker was recycled by transforming with pLA34 (SEQ ID NO: 121) containing the CRE recombinase under the GAL1 promoter and plated on synthetic complete media lacking histidine and supplemented with 1% ethanol at 30°C. Transformants were plated on rich media supplemented with 1% ethanol and 0.5% galactose to induce the recombinase. Marker removal was confirmed by patching colonies to synthetic complete media lacking uracil and supplemented with 1% ethanol to verify absence of growth. The resulting identified strain, called PNY2101 has the genotype MATa ura3Δ::loxP his3Δ pdc5Δ::loxP71/66 fra2Δ 2-micron pdc1Δ::P[PDC1]-ALS|alsS_Bs-CYC1t-loxP71/66 pdc6Δ::(UAS)PGK1-P[FBA1]-KIVD|Lg(y)-TDH3t-loxP71/66 adh1Δ::P[ADH1]-ADH|Bi(y)-ADHt-loxP71/66.

Fra2Δ::P[ILV5]-ADH|Bi(y)-ADHt-loxP71/66

[00243] To integrate BiADH into the fra2Δ locus of PNY2101, an integration cassette was PCR-amplified from pLA65 (SEQ ID NO: 148), which contains the alcohol dehydrogenase from the species *Beijerinckii indica* with an ILV5 promoter and an ADH1 terminator, and a URA3 marker flanked by degenerate loxP sites to allow homologous recombination *in vivo* and subsequent removal of the URA3 marker. PCR was done by using KAPA HiFi and primers 906 (SEQ ID NO: 165) and 907 (SEQ ID NO: 166). The FRA2 portion of each primer was derived from the first 60bp of the coding sequence starting at the ATG and 56bp downstream of the stop codon. The PCR product was transformed into PNY2101 using standard genetic techniques and transformants were selected on synthetic complete media lacking uracil and supplemented with 1% ethanol at 30°C. Transformants were screened to verify correct integration by colony PCR using primers 667 (SEQ ID NO: 167), external to the 5' coding region and 749 (SEQ ID NO: 168), internal to the ILV5 promoter. The URA3 marker was recycled by transforming with pLA34 (SEQ ID NO: 121) containing the CRE recombinase under the GAL1 promoter and plated on synthetic complete media lacking histidine and supplemented with 1% ethanol at 30°C. Transformants were plated on rich media supplemented with 1% ethanol and 0.5% galactose to induce the recombinase. Marker removal was confirmed by patching colonies to synthetic complete media lacking uracil and supplemented with 1% ethanol to verify absence of growth. The resulting identified strain, called PNY2110 has the genotype MATa ura3Δ::loxP his3Δ pdc5Δ::loxP66/71 2-micron pdc1Δ::P[PDC1]-ALS|alsS_Bs-CYC1t-loxP71/66

pdc6Δ::(UAS)PGK1-P[FBA1]-KIVD|Lg(y)-TDH3t-loxP71/66 adh1Δ::P[ADH1]-ADH|Bi(y)-ADHt-loxP71/66 fra2Δ::P[ILV5]-ADH|Bi(y)-ADHt-loxP71/66.

GPD2 deletion

[00244] To delete the endogenous *GPD2* coding region, a deletion cassette was PCR amplified from pLA59 (SEQ ID NO: 115), which contains a *URA3* marker flanked by degenerate loxP sites to allow homologous recombination in vivo and subsequent removal of the *URA3* marker. PCR was done by using KAPA HiFi and primers LA512 (SEQ ID NO: 141) and LA513 (SEQ ID NO: 142). The *GPD2* portion of each primer was derived from the 5' region 50bp upstream of the *GPD2* start codon and 3' region 50bp downstream of the stop codon such that integration of the *URA3* cassette results in replacement of the entire *GPD2* coding region. The PCR product was transformed into PNY2110 using standard genetic techniques and transformants were selected on synthetic complete medium lacking uracil and supplemented with 1% ethanol at 30°C. Transformants were screened to verify correct integration by colony PCR using primers LA516 (SEQ ID NO: 143) external to the 5' coding region and LA135 (SEQ ID NO: 112), internal to *URA3*. Positive transformants were then screened by colony PCR using primers LA514 (SEQ ID NO: 144) and LA515 (SEQ ID NO: 145), internal to the *GPD2* coding region. The *URA3* marker was recycled by transforming with pLA34 (SEQ ID NO: 121) containing the CRE recombinase under the *GALI* promoter and plated on synthetic complete medium lacking histidine and supplemented with 1% ethanol at 30°C. Transformants were plated on rich medium supplemented with 1% ethanol and 0.5% galactose to induce the recombinase. Marker removal was confirmed by patching colonies to synthetic complete medium lacking uracil and supplemented with 1% ethanol to verify absence of growth. The resulting identified strain, called PNY2115, has the genotype MATa *ura3Δ::loxP his3Δ pdc5Δ::loxP66/71 fra2Δ 2-micron pdc1Δ::P[PDC1]-ALS|alsS_Bs-CYC1t-loxP71/66 pdc6Δ::(UAS)PGK1-P[FBA1]-KIVD|Lg(y)-TDH3t-loxP71/66 adh1Δ::P[ADH1]-ADH|Bi(y)-ADHt-loxP71/66 fra2Δ::P[ILV5]-ADH|Bi(y)-ADHt-loxP71/66 gpd2Δ::loxP71/66.*

Creation of PNY2145 from PNY2115

[00245] PNY2145 was constructed from PNY2115 by the additional integration of a phosphoketolase gene cassette at the *pdc5Δ* locus and by replacing the native *AMN1* gene with a codon optimized version of the ortholog from *CEN.PK*. Integration constructs are further described below.

pdc5Δ::FBA(L8)-xpk1-CYC1t-loxP71/66

[00246] The *TEF(M4)-xpk1-CYC1t* gene from *pRS423::TEF(M4)-xpk1+ENO1-eutD* (SEQ ID NO: 170) was PCR amplified using primers N1341 and N1338 (SEQ ID NOS: 171 and 172), generating a 3.1 kb product. The *loxP*-flanked *URA3* gene cassette from *pLA59* (SEQ ID NO: 115) was amplified with primers N1033c and N1342 (SEQ ID NOS: 173 and 174), generating a 1.6 kb product. The *xpk1* and *URA3* PCR products were fused by combining them without primers for an additional 10 cycles of PCR using Phusion DNA polymerase. The resulting reaction mix was then used as a template for a PCR reaction with KAPA Hi Fi and primers N1342 and N1364 (SEQ ID NOS: 174 and 175). A 4.2 kb PCR product was recovered by purification from an electrophoresis agarose gel (Zymo kit). *FBA* promoter variant L8 (SEQ ID NO: 176) was amplified using primers N1366 and N1368 (SEQ ID NOS: 177 and 178). The *xpk1::URA3* PCR product was combined with the *FBA* promoter by additional rounds of PCR. The resulting product was phosphorylated with polynucleotide kinase and ligated into *pBR322* that had been digested with *EcoRV* and treated with calf intestinal phosphatase. The ligation reaction was transformed into *E. coli* cells (*Stbl3* competent cells from Invitrogen). The integration cassette was confirmed by sequencing. To prepare DNA for integration, the plasmid was used as a template in a PCR reaction with Kapa HiFi and primers N1371 and N1372 (SEQ ID NOS: 179 and 180). The PCR product was isolated by phenol-chloroform extraction and ethanol precipitation (using standard methods; eg. Maniatis, et al.). Five micrograms of DNA were used to transform strain PNY2115. Transformants were selected on medium lacking uracil (synthetic complete medium minus uracil with 1% ethanol as the carbon source). Colonies were screened for the integration event using PCR (JumpStart) with primers BK93 and N1114 (SEQ ID NOS: 181 and 182). Two clones were selected to carry forward. The *URA3* marker was recycled by transforming with *pJT254* (SEQ ID NO: 183) containing the *CRE* recombinase under the *GALI* promoter and plating on synthetic complete medium lacking histidine and supplemented with 1% ethanol at 30 °C. Transformants were grown in rich medium

supplemented with 1% ethanol to derepress the recombinase. Marker removal was confirmed for single colony isolates by patching to synthetic complete medium lacking uracil and supplemented with 1% ethanol to verify absence of growth. Loss of the recombinase plasmid, pJT254, was confirmed by patching the colonies to synthetic complete medium lacking histidine and supplemented with 1% ethanol. Proper marker removal was confirmed by PCR (primers N160SeqF5 (SEQ ID NO: 184) and BK380). One resulting clone was designated PNY2293.

amn1Δ::AMN1(y)-loxP71/66

[00247] To replace the endogenous copy of *AMN1* with a codon-optimized version of the *AMN1* gene from CEN.PK2, an integration cassette containing the CEN.PK *AMN1* promoter, *AMN1(y)* gene (nucleic acid SEQ ID NO: 185; amino acid SEQ ID NO: 186), and CEN.PK *AMN1* terminator was assembled by SOE PCR and subcloned into the shuttle vector pLA59. The *AMN1(y)* gene was ordered from DNA 2.0 with codon-optimization for *S. cerevisiae*. The completed pLA67 plasmid (SEQ ID NO: 187) contained: 1) pUC19 vector backbone sequence containing an *E. coli* replication origin and ampicillin resistance gene; 2) *URA3* selection marker flanked by loxP71 and loxP66 sites; and 3) P_{AMN1(CEN.PK)}-*AMN1(y)*-term_{AMN1(CEN.PK)} expression cassette. PCR amplification of the *AMN1(y)*-loxP71-*URA3*-loxP66 cassette was done by using KAPA HiFi from Kapa Biosystems, Woburn, MA and primers LA712 (SEQ ID NO: 188) and LA746 (SEQ ID NO: 189). The PCR product was transformed into PNY2293 using standard genetic techniques and transformants were selected on synthetic complete medium lacking uracil and supplemented with 1% ethanol at 30 °C. Transformants were observed under magnification for the absence of a clumping phenotype with respect to the control (PNY2293). The *URA3* marker was recycled using the pJT254 Cre recombinase plasmid as described above. After marker recycle, clones were again observed under magnification to confirm absence of the clumping phenotype. A resulting identified strain, PNY2145, has the genotype: MATa ura3Δ::loxP his3Δ pdc5Δ::P[FBA(L8)]-XPK|xpk1_Lp-CYCt-loxP66/71 fra2Δ 2-micron plasmid (CEN.PK2) pdc1Δ::P[PDC1]-ALS|alsS_Bs-CYC1t-loxP71/66 pdc6Δ::(UAS)PGK1-P[FBA1]-KIVD|Lg(y)-TDH3t-loxP71/66 adh1Δ::P[ADH1]-ADH|Bi(y)-ADHt-loxP71/66 fra2Δ::P[ILV5]-ADH|Bi(y)-ADHt-loxP71/66 gpd2Δ::loxP71/66 amn1Δ::AMN1(y).

Creation of PNY2310 from PNY2145

[00248] PNY2310 was generated by transforming strain PNY2145 with plasmids pLH804-L2V4 and pRS413::BiADH-kivD. Plasmid pLH804-L2V4 (SEQ ID NO: 12) is a yeast-*E. coli* shuttle vector based on pHR81 (ATCC#87541). It contains genes for the expression of KARI variant K9JB4P and DHAD variant L2V4. Plasmid pRS413::BiADH-kivD (SEQ ID NO: 13) is a yeast-*E. coli* shuttle vector based on pRS413 (ATCC#87518). It contains genes for the expression of BiADH and kivD. The positions of the relevant gene features in both plasmids are listed in the Tables 3 and 4. Plasmid transformants were selected by plating on synthetic complete medium lacking uracil and histidine with 1% (v/v) ethanol as the carbon source. Colonies were transferred to fresh plates by patching. After two days, cells from the patches were transferred to plates containing synthetic complete medium (minus uracil and histidine) with 2% (w/v) glucose as the carbon source. The resulting strain was designated PNY2310.

Table 3: Nucleotide positions of pathway gene features in plasmid pLH804-L2V4

Element	Description	Start	End	Strand
Promoter	ILV5p	427	1620	T
CDS	JB4P	1628	2659	T
Terminator	ILV5t	2685	3307	T
Terminator	FBA _t	3320	3632	B
CDS	ILVD-L2V4	3641	5356	B
Promoter	TEF1(M7)p	5366	5766	B

Table 4: Nucleotide positions of pathway gene features in plasmid pRS413::BiADH-kivD

Element	Description	Start	End	Strand
Promoter	FBA _{1p}	2293	2893	T
CDS	kivD Lg(y)	2902	4548	T
Terminator	TDH3 _t	4560	5139	T
Promoter	PDC1 _p	5983	6852	T
CDS	adhB _{iy}	6853	7896	T
Terminator	ADH1 _t	7905	8220	T

Creation of CPN97 from PNY2145

[00249] To replace the endogenous GPD1 of *Saccharomyces cerevisiae* with *E. coli gpaA*, primers were designed to amplify the *E. coli gpaA* open reading frame to insert in the endogenous GPD1 chromosomal location maintaining the region upstream of the ATG start codon and the endogenous *Saccharomyces cerevisiae* GPD1 stop codon. Overlapping PCR was used to obtain a PCR product containing 50 base pairs upstream of the *Saccharomyces cerevisiae* GPD1 for recombination, the *E. coli gpaA* gene, a loxP-URA3-loxP cassette, and 50 base pairs downstream

of the *Saccharomyces cerevisiae* GPD1 for recombination in PNY2145. The *E. coli* *gpsA* ORF (PCR product 1) was amplified utilizing *E. coli* BL21 chromosomal DNA as a template with primers F1 (SEQ ID NO: 234) and R1 (SEQ ID NO: 235). The loxP-URA3-loxP cassette (PCR product 2) was amplified utilizing pLA59 (SEQ ID NO:27) as a template using primers F2 (SEQ ID NO: 236) and R2 (SEQ ID NO: 237). The PCR product (PCR product 3) containing 50 base pairs upstream of the *Saccharomyces cerevisiae* GPD1, the *E. coli* *gpsA* gene, a loxP-URA3-loxP cassette, and 50 base pairs downstream of the *Saccharomyces cerevisiae* GPD1 was amplified utilizing the PCR products 1 and 2 as templates and primers F1 (SEQ ID NO: 234) and R2 (SEQ ID NO: 237). All PCR reactions were performed using the enzyme Herculase (Agilent; Santa Clara, CA) according to manufacturer's conditions.

[00250] PCR product 3 was recovered by purification and transformed into PNY2145 using a yeast transformation kit (Sigma-Aldrich; St. Louis, MO). Colonies were selected on yeast synthetic medium containing 1% ethanol but no uracil. Yeast synthetic medium: 6.7 g/L yeast nitrogen base without amino acids (Becton Dickinson; East Rutherford, NJ), 1.85 g/L Kaiser dropout His-Ura (Formedium; Norfolk, UK). Histidine or uracil were added at 76 mg/L when needed.

[00251] To recycle the URA3 marker, one colony was selected and transformed with plasmid pJT254 (SEQ ID NO: 183) containing CRE recombinase under the *GALI* promoter and was plated on yeast synthetic medium containing 1% ethanol and no histidine. One colony was selected and grown overnight in YPE medium (20 g/L peptone, 10 g/L yeast extract, 10 g/L ethanol) and restreaked on YPE plates (20 g/L peptone, 10 g/L yeast extract, 10 g/L ethanol, 15 g/L agar). Colonies were selected and patched on plates of yeast synthetic medium containing 1% ethanol and no uracil, 1% ethanol and no histidine, and YPE plates. A colony unable to grow on plates lacking uracil and histidine was selected and screened for marker removal and insertion of *E. coli* *gpsA* by PCR. The colony was designated CPN82.

[00252] To produce a strain with an isobutanol production pathway, CPN82 was transformed with pLH804::L2V4 (SEQ ID NO: 12) and pRS413::BiADH-kivD (SEQ ID NO: 13), described above. The transformation was plated on yeast synthetic medium lacking uracil and histidine and with 1% ethanol, and three colonies were selected and restreaked on yeast synthetic medium lacking histidine and uracil with 3 g/L glucose, 3 g/L ethanol. The colonies were tested for isobutanol production, and one colony was selected and designated CPN97.

Creation of Yeast Codon Optimized GPD1 M, M3, and M8 variant strains from PNY2145
Integration of Yeast Codon Optimized GPD1 variants

[00253] In order to test GPD1 mutants in the host *Saccharomyces cerevisiae* strain, native GPD1 was swapped with a codon optimized version of GPD1 synthesized by DNA 2.0 using *S. cerevisiae* codon usage.

Preparation of Integration Cassette

[00254] The gene swap cassette was prepared by cloning 2 fragments (upstream GPD1 upstream homology region and codon optimized yeast GPD1 fragment) in vector pBP3518 (SEQ ID NO:9) containing the URA3 marker gene along with the promoter and terminator and GPD downstream homology region cloned downstream of the URA3 marker gene.

[00255] Fragment 1 for the integration cassette was amplified using Phusion High Fidelity PCR Master Mix (New England Biolabs Inc.; Ipswich, MA), primers oBP1329 (SEQ ID NO:1) and oBP1333 (SEQ ID NO:2) and PNY2145 genomic DNA as template prepared using YeaStarTM Genomic DNA kit (Zymo Research). Fragment 2 was amplified using primers oBP1334 (SEQ ID NO:3) and oBP1335 (SEQ ID NO:4) and synthetic codon optimized Yeast GPD1 or appropriate GPD1 variants as the templates. Primer oBP1333 (SEQ ID NO:2) has a 5' tail with homology to the 5' region of Fragment 2 (synthetic codon optimized GPD1) and primer oBP1334 (SEQ ID NO:3) has a 5' tail with homology to the 3' end of Fragment 1 (GPD upstream region). The two fragments were combined using overlap PCR using primer oBP1329 (SEQ ID NO:1) and oBP1335 (SEQ ID NO:4). This combined fragment was cloned in AscI and PmeI sites in vector pBP3518 (SEQ ID NO:9) and the resulting vector referred as pBP3518GPD* (SEQ ID NO:10) was transformed into Agilent XL1Blue competent cells (Agilent Technologies, USA).

Transformation of Integration Cassette in PNY2145

[00256] Plasmid oBP3518GPD* (SEQ ID NO:10) was isolated using QIAprep Spin miniprep Kit (Qiagen, GmbH) and restricted using SacI and PacI restriction enzymes (New England Biolabs Inc. Ipswich, MA). The resulting 4.2 kb fragments (containing the entire integration cassette, GPD Upstream homology region, Codon Opt GPD, URA3 marker gene and Downstream GPD region) was transformed into PNY2145 using Frozen EZ Yeast

Transformation II Kit (Zymo Research). The transformation mix was plated on synthetic complete lacking Uracil with 0.5% ethanol at 30 °C for 48 hours. For confirmation of integration site, transformants were screened using two sets of primers oBP1342 (SEQ ID NO:6) and oBP1344 (SEQ ID NO:7) and oBP1341 (SEQ ID NO:5) and oBP1345 (SEQ ID NO:8) for confirmation of integration at both ends. The primers oBP1342 (SEQ ID NO:6) and oBP1345 (SEQ ID NO:8) were designed from a region outside the cassette to confirm integration at the right site.

Removal of URA3 Marker

[00257] The confirmed transformants (strain PNY2145 GPD1 Δ ::CO GPD1 URA3) were transformed with pRS423::P_{GAL1}-*cre* using a Frozen-EZ Yeast Transformation II™ kit (Zymo Research Corporation, Irvine, CA) and plating on synthetic complete medium lacking histidine and uracil supplemented with 0.5% ethanol and incubated at 30 °C for 48 hours. Transformants were grown in synthetic complete medium lacking histidine with 0.5% ethanol overnight and plated on synthetic complete medium with 0.5% ethanol and 0.1% 5FOA for URA3 marker. Marker removal was confirmed by colony PCR using primers oBP1341 (SEQ ID NO:5) and oBP1345 (SEQ ID NO:8).

Transformation of Pathway Plasmid

[00258] The strain PNY2145 GPD1 Δ ::CO GPD1 was then transformed with plasmid pLMH11-JM44 (SEQ ID NO:240) using Frozen-EZ Yeast Transformation II™ kit (Zymo Research Corporation, Irvine, CA) and plated on synthetic complete medium without uracil with 0.5% ethanol. The resulting strains were designated M, M3 (F73A), and M8 (F73G/F129G).

Creation of E. coli codon optimized yeast GPD1 E1, E3, and E8 strains from PNY2145

Integration of yeast GPD1 codon optimized for E. coli

[00259] In order to test GPD1 mutants in the host *Saccharomyces cerevisiae* strain, native GPD1 was swapped with an *E. coli* codon optimized version of GPD1 synthesized by DNA 2.0 using *E. coli* codon usage.

Preparation of Integration Cassette

[00260] The gene swap cassette was prepared by cloning 2 fragments (upstream GPD1 upstream homology region and codon optimized yeast GPD1 fragment) in vector pBP3518 (SEQ ID NO:9) containing the URA3 marker gene along with the promoter and terminator and GPD downstream homology region cloned downstream of the URA3 marker gene.

[00261] Fragment 1 for the integration cassette was amplified using Phusion High Fidelity PCR Master Mix (New England Biolabs Inc.; Ipswich, MA), primers oBP1329 (SEQ ID NO:1) and oBP1350 (SEQ ID NO:241) and PNY2145 genomic DNA as template prepared using YeaStar™ Genomic DNA kit (Zymo Research). Fragment 2 was amplified using primers oBP1351 (SEQ ID NO:242) and oBP1352 (SEQ ID NO:243) and synthetic *E. coli* codon Yeast optimized GPD1 gene and E3 and E8 variants as the templates. Primer oBP1350 (SEQ ID NO:241) has a 5' tail with homology to the 5' region of Fragment 2 (synthetic codon optimized GPD1) and primer oBP1351 (SEQ ID NO:242) has a 5' tail with homology to the 3' end of Fragment 1 (GPD upstream region). The two fragments were combined using overlap PCR using primer oBP1329 (SEQ ID NO:1) and oBP1352 (SEQ ID NO:243). This combined fragment was cloned in AscI and PmeI sites in vector pBP3518 (SEQ ID NO:9) and the resulting vector referred as pBP3518GPD1_EcOpt (SEQ ID NO:249) was transformed into Agilent XL1Blue competent cells (Agilent Technologies, USA).

Transformation of Integration Cassette in PNY2145

[00262] Plasmid pBP3518GPD1_EcOpt (SEQ ID NO:249) was isolated using QIAprep Spin miniprep Kit (Qiagen, GmbH) and restricted using SacI and PacI restriction enzymes (New England Biolabs Inc. Ipswich, MA). The resulting 4.2 kb fragments (containing the entire integration cassette, GPD Upstream homology region, *E. coli* codon optimized GPD1, URA3 marker gene and Downstream GPD region) was transformed into PNY2145 using Frozen EZ Yeast Transformation II Kit (Zymo Research). The transformation mix was plated on synthetic complete lacking Uracil with 0.5% ethanol at 30 °C for 48 hours. For confirmation of integration site, transformants were screened using two sets of primers oBP1342 (SEQ ID NO:6) and oBP1352 (SEQ ID NO:243) and oBP1357 (SEQ ID NO:248) and oBP1345 (SEQ ID NO:8) for confirmation of integration at both ends. The primers oBP1342 (SEQ ID NO:6) and oBP1345 (SEQ ID NO:8) were designed from a region outside the cassette to confirm integration at the right site.

Removal of URA3 Marker

[00263] The confirmed transformants (strain PNY2145 GPD1 Δ ::EC CO GPD1 URA3) were transformed with pRS423::P_{GALI}-*cre* using a Frozen-EZ Yeast Transformation II™ kit (Zymo Research Corporation, Irvine, CA) and plating on synthetic complete medium lacking histidine and uracil supplemented with 0.5% ethanol and incubated at 30 °C for 48 hours. Transformants were grown in synthetic complete medium lacking histidine with 0.5% ethanol overnight and plated on synthetic complete medium with 0.5% ethanol and 0.1% 5FOA for URA3 marker. Marker removal was confirmed by colony PCR using primers oBP1357 (SEQ ID NO:248) and oBP1345 (SEQ ID NO:8).

Transformation of Pathway Plasmid

[00264] The strain PNY2145 GPD1 Δ ::EC CO GPD1 was then transformed with plasmid pLMH11-JM44 (SEQ ID NO:240) using Frozen-EZ Yeast Transformation II™ kit (Zymo Research Corporation, Irvine, CA) and plated on synthetic complete medium without uracil with 0.5% ethanol. The resulting strains were designated EC1, E3 (F73A), and E8 (F73G/F129G).

*Creation of native yeast GPD1 N, N3, and N8 strains from PNY2145**Integration of yeast native GPD1 variants*

[00265] In order to test yeast native GPD1 mutants in the host *Saccharomyces cerevisiae* strain, yeast codon optimized GPD1 was swapped with yeast native version of GPD1 in strain PNY2145 GPD1 Δ ::CO GPD1.

Preparation of Integration Cassette

[00266] The gene swap cassette was prepared by cloning 2 fragments (upstream GPD1 upstream homology region and yeast native GPD1 fragment) in vector pBP3518 (SEQ ID NO:9) containing the URA3 marker gene along with the promoter and terminator and GPD downstream homology region cloned downstream of the URA3 marker gene.

[00267] Fragment 1 for the integration cassette was amplified using Phusion High Fidelity PCR Master Mix (New England Biolabs Inc.; Ipswich, MA), primers oBP1329 (SEQ ID NO:1) and oBP1353 (SEQ ID NO:244) and PNY2145 genomic DNA as template prepared using YeaStar™ Genomic DNA kit (Zymo Research). Fragment 2 was amplified using primers oBP1354 (SEQ ID NO:245) and oBP1355 (SEQ ID NO:246) and PNY2145 genomic DNA and

appropriate variants as the templates. Primer oBP1353 (SEQ ID NO:244) has a 5' tail with homology to the 5' region of Fragment 2 (yeast native GPD1) and primer oBP1354 (SEQ ID NO:245) has a 5' tail with homology to the 3' end of Fragment 1 (GPD upstream region). The two fragments were combined using overlap PCR using primer oBP1329 (SEQ ID NO: 1) and oBP1355 (SEQ ID NO:246). This combined fragment was cloned in AscI and PmeI sites in vector pBP3518 (SEQ ID NO:9) and the resulting vector referred as pBP3518GPD1_Native (SEQ ID NO:250) was transformed into Agilent XL1Blue competent cells (Agilent Technologies, USA).

Transformation of Integration Cassette in PNY2145

[00268] Plasmid pBP3518GPD1_Native (SEQ ID NO:250) was isolated using QIAprep Spin miniprep Kit (Qiagen, GmbH) and restricted using SacI and PacI restriction enzymes (New England Biolabs Inc. Ipswich, MA). The resulting 4.2 kb fragments (containing the entire integration cassette, GPD upstream homology region, yeast native GPD1, URA3 marker gene and downstream GPD region) was transformed into PNY2145 GPD1 Δ ::CO GPD1 using Frozen EZ Yeast Transformation II Kit (Zymo Research). The transformation mix was plated on synthetic complete lacking Uracil with 0.5% ethanol at 30 °C for 48 hours. For confirmation of integration site, transformants were screened using two sets of primers oBP1342 (SEQ ID NO:6) and oBP1355 (SEQ ID NO:246) and oBP1356 (SEQ ID NO:247) and oBP1345 (SEQ ID NO:8) for confirmation of integration at both ends.

Removal of URA3 Marker

[00269] The confirmed transformants (strain PNY2145 CO GPD1 Δ ::Native GPD1 URA3) were transformed with pRS423::P_{GAL1}-cre using a Frozen-EZ Yeast Transformation II™ kit (Zymo Research Corporation, Irvine, CA) and plating on synthetic complete medium lacking histidine and uracil supplemented with 0.5% ethanol and incubated at 30 °C for 48 hours. Transformants were grown in synthetic complete medium lacking histidine with 0.5% ethanol overnight and plated on synthetic complete medium with 0.5% ethanol and 0.1% 5FOA for URA3 marker. Marker removal was confirmed by colony PCR using primers oBP1356 (SEQ ID NO:247) and oBP1345 (SEQ ID NO:8).

Transformation of Pathway Plasmids

[00270] The strain PNY2145 CO GPD1 Δ ::Native GPD1 was then transformed with plasmid pLMH11-JM44 (SEQ ID NO:240) using Frozen-EZ Yeast Transformation IITM kit (Zymo Research Corporation, Irvine, CA) and plated on synthetic complete medium without uracil with 0.5% ethanol. The resulting strains were designated N, N3 (F73A), and N8 (F73G/F129G).

Integration of GPD1 Variants

[00271] Integration cassettes for GPD1 variants were prepared in the same way as described above for codon optimized GPD1 swap except that the template used for amplifying fragment 2 for each GPD1 variant was the corresponding coding sequence with the mutation listed in Table 5.

Table 5: GPD variants with corresponding mutations

GPD Variant	Mutation	Strain
GPD2	F129G	
GPD3	F73A	E3, M3, N3
GPD4	F73G	
GPD7	F73A129G	
GPD8	F73G129G	E8, M8, N8

E3 and E8 are variants of GPD1 using *E. coli* codon optimization; M3 and M8 are variants GPD1 using *S. cerevisiae* codon optimization; and N3 and N8 are variants GPD1 with native *S. cerevisiae* codon usage.

Integration of heterologous GPDs

[00272] Integration cassettes for heterologous GPDs can be prepared in the same way as described above for codon optimized GPD1 except one skilled in the art would redesign the primers based on the heterologous GPD1 to be inserted to get proper assembly of the integration cassette.

Example 1: Variant GPD1 enzymes

[00273] In this example, variant versions of *Saccharomyces* GPD1 were created and tested by expression in *E. coli* followed by enzymatic assays on the crude cell extract.

GPD1 mutagenesis

[00274] Mutagenesis of yeast GPD1 was directed by the desire to increase the K_M for NADH without having an impact on other kinetic parameters of the enzyme. The approach to mutagenesis was based on the high resolution crystal structures of human GPD1 (Ou et al, 2005, J.Mol.Biol. 357: 858-869), which allowed for the determination of the amino acids within contact distance of NAD in the cofactor binding pocket. By analyzing the amino acid residues and the type of contact made, it was possible to limit the number of amino acid changes to result in an increase in the NADH K_M value. Table 6 shows the results of this analysis, where the amino acid residues within 5 Angstroms of the bound NAD have been enumerated, and the role in NAD binding was interpreted. Because of their role in the pi-stacking stabilization of bound NAD, an initial focus was placed on mutagenesis of positions homologous to the phe41 and phe97 of truncated human GPD1 (SEQ ID NO:84) (phe73 and phe129 of yeast GPD1)

Table 6. Amino acid residues within 5 angstroms (Å) of bound NAD in the human crystal structure, and potential mutagenesis targets to increase NADH K_M value.

Residue (Number Human GPD1)	Yeast GPD Number (Residue if diff.)	No. of Contacts with NAD	Interaction Type	Possible Role	Equiv. Alt. Residues	Increase K_M with	Rationale
Ser(11)	42	3					
Asn(13)	44	18	Polar	H-bonds with NAD		A, C, G, I, L, M, S, V	A non-bulky non-polar residue can reduce attraction to NAD
Trp(14)	45	66	π -stacking	Possibly stabilizes pyrimidine ring for stable electron transfer	Phe, Tyr	A, C, G, H, I, K, L, M, N, Q, R, S, T, V	A non-ring residue which does not create too much of a Van der Waal clash and is non-negative can increase K_M for NAD but destabilize the electron transfer by removing the π -stacking fixing. May also alter K_{cat} .
Trp(39)	71	6					
Phe(41)	73	64	π -stacking	Stabilizes after binding	Tyr, Trp, possibly Arg, Lys due to cation- π	G, A, R, K	Small residues can increase adenine entropy. Large residues can make it difficult to get in. Long positive residues close gate with other ring due to cation- π .
Glu(43)	75	2					

Tyr(63)	95	39		Water stabilization			
Val(92)	124 (Gln)	6					
Pro(94)	126	25					
Phe(97)	129	64	π -stacking	Stabilizes after binding	Tyr, Trp, possibly Arg, Lys due to cation- π	G, A, R, K	Small residues can increase adenine entropy. Large residues can make it difficult to get in. Long positive residues close gate with other ring due to cation- π .
Ile(119)	151 (leu)	5					
Lys(120)	152	20	Polar	H-bonds with DHAP and stabilizes it			Reversing or removing polarity can increase K_M for NAD, but can alter K_{cat} or K_M for DHAP
Asn(151)	183	3					
Ile(152)	184	2					
Ala(153)	185	4					
Asn(205)	246	1		Possibly assists in electron transfer			Can alter K_{cat} or K_M for DHAP
Arg(269)	310	33		Stabilizes electron transfer. Likely can create attraction for diphosphate of NAD			Changing to a non-positive residue can make the diphosphate of NAD uneasy, but can alter K_{cat} or K_M for DHAP
Gln(295)	336	14					
Lys(296)	337 (Ser)	17		H-bonds with NAD		A, C, D, E, G, I, L, M, N, Q, S, V	Prevent h-bond formation, reverse polarity and reduce attraction for NAD
Gln(298)	339	14		H-bonds and stabilizes residues around NAD		A, C, G, I, L, M, S, V	A non-bulky non-polar residue can reduce attraction to NAD

Strain and Media

[00275] *Escherichia coli* TOP10 was obtained from Life Technologies Corp. (Cat. # C404003, Grand Island, NY). Expression plasmid pBAD was previously described (U.S. Patent No. 7,910,342). Synthetic yeast GPD1 gene, optimized for yeast expression, was obtained from DNA2.0 (Menlo Park, CA). Cells were grown at 37 °C in Miller's LB broth (Cat. # 46-050-CM, Mediatech, Inc., Herndon, VA) with 0.02% L-(+)-arabinose (Cat. # A3256, Sigma-Aldrich, Inc., St. Louis, MO) and 100 μ g/mL ampicillin (Cat. # A1066, Sigma-Aldrich, Inc., St. Louis, MO).

Cells were plated on LB agar plates with 100µg/mL ampicillin (Cat. # L1004, Teknova, Inc., Hollister, CA).

Construction of GPD1 Variants

[00276] Mutations were introduced at 4 different amino acid positions, according to Table 7. Mutagenesis was performed using a QuikChange Lightning Kit (Cat. #210519, Agilent Technologies, La Jolla, CA), according to manufacturer's directions. Mutagenesis primers were obtained from Sigma-Aldrich Co. LLC, St. Louis MO. Reactions were thermocycled in a Gene Amp 9700 (Perkin Elmer Applied Biosystems, Norwalk, CT). *Escherichia coli* TOP10 were transformed with 1 µl of QuikChange reaction product according to manufacturer's directions, and transformants were selected on LB agar plates with 100µg/mL ampicillin. DNA sequences were obtained for multiple isolates from each transformation in order to identify those with the desired mutations.

[00277] Double mutants were constructed in the same manner, except that the template in the mutagenesis reaction already contained one of the mutations, and the appropriate primers were used to introduce the second mutation.

Table 7: GPD1 Variants and Primers Used in Their Construction

Position	SEQ ID NO:	Forward Primer SEQ ID NO:	Reverse Primer SEQ ID NO:
Asn44		20	
N44A	212	21	54
N44C	213	22	55
N44G	214	23	56
N44I	215	24	57
N44L	216	25	58
N44M	217	26	59
N44S	218	27	60
N44V	219	28	61
Trp45		29	
W45A	220	30	62
W45C	221	31	63
W45G	222	32	64
W45H	223	33	65
W45I	224	34	66
W45K	225	35	67
W45L	226	36	68

W45M	227	37	69
W45N	228	38	70
W45Q	229	39	71
W45R	230	40	72
W45S	231	41	73
W45T	232	42	74
W45V	233	43	75
Phe73		44	
F73G	196	45	76
F73A	197	46	77
F73R	198	47	78
F73K	199	48	79
Phe129		49	
F129G	200	50	80
F129A	201	51	81
F129R	202	52	82
F129K	203	53	83

GPD1 Assay

[00278] Soluble fraction cell extracts were prepared from 5ml of culture by bead beating 2 x 10 seconds in 100 mM MOPS pH 6.8, 10 mM MgCl₂, 1mM EDTA in a Mini-Bead-beater (Cat. #3110BX, Biospec Products, Bartlesville, OK). Cell extract protein concentration was determined by Pierce BCA assay (Cat. #23224 and 23228, Thermo Fisher Scientific, Inc., Rockford, IL).

[00279] Assays were performed in a 100 µl volume containing 100 mM MOPS pH6.8, 1mM EDTA, 1mM glucose-6-phosphate, 3m U/µl glucose-6-phosphate dehydrogenase (Cat. #G8404, Sigma-Aldrich, Inc., St. Louis, MO), 1mM dihydroxyacetone phosphate (Cat. #D7137, Sigma-Aldrich, Inc., St. Louis, MO), varying concentrations of NADH, and varying concentrations of cell extract. Reactions were terminated by 4-fold dilution into 0.1% formic acid (Suprapur, #1167-1, EMD Chemicals, Gibbstown, NJ) in water (Omnisolv, #WX0001-1, EMD Chemicals, Gibbstown, NJ). Glycerol-3-phosphate production was measured by LC/MS.

LC/MS Method

[00280] 2 µL of each sample was injected on a Waters Acuity UPLC/SQD System, using a HSS T3 column (2.1×100 mm, 1.8 µm, #186003539, Waters, Milford, MA) at a temperature of 30 °C. UPLC mobile phases consisted of 0.1% formic acid in water (Mobile A) and 0.1% formic

acid in acetonitrile (Omnisolv, #AX0156-1, EMD Chemicals, Gibbstown, NJ) (Mobile B) with a constant flow rate of 0.3 mL/min. The gradient consisted of an initial 1 minute period at 99% A, followed by a 0.5 minute linear gradient ending at 75% B, and then a 0.5 minute linear gradient back down to 99% A, before injecting the next sample. MS analysis was performed by electrospray negative ionization at a cone voltage of 30V and $m/z = 171$. Retention time and peak intensities were determined using MassLynx4.1 software (Waters, Milford, MA). External standard (glycerol-3-phosphate, Cat. #G7886, Sigma-Aldrich, Inc., St. Louis, MO) was analyzed in the same manner was used for quantitation.

Analysis of GPD1 Variants

[00281] The activity of the GPD1 variants was initially screened by measuring the initial rate of formation of glycerol 3 phosphate (G3P) at two concentrations of NADH (30 and 300 μM). This serves as a measurement indicating the K_M of the variants for NADH: for a K_M much less than 30 μM , the ratio would be 1.0, for a K_M much higher than 300 μM the ratio would be 10. The results for the individual single mutants are shown in Table 8.

Table 8: Activity of GPD1 Variants Measured by Initial G3P Formation Rate

Variant	Activity Ratio
129R	78.0
129K	73.1
129G	67.8
129A	65.9
44G	2.4
44M	2.4
44S	2.3
45G	2.2
44V	2.2
73A	2.0
44A	1.8
73G	1.8
44C	1.8
44L	1.8
44I	1.7
45C	1.7
73R	1.5
45M	1.4
45A	1.4
45H	1.3
73K	1.2
45I	1.0

45S	1.0
45V	1.0
45T	0.9
45K	0.8
45L	0.7
45N	0.6
45Q	0.6
45R	0.6
GPD1 wt	1.3 +/- 0.3

[00282] This data can be interpreted to indicate that any variant with a ratio greater than 1.6 (average value of three control measurements plus standard deviation) has a higher NADH K_M than the wild-type GPD1. Double mutants of the individual high- K_M single mutants were created as described above. Full-scale analysis of the NADH K_M values for a selection of the double and single mutants is shown in Table 9.

Table 9: Michaelis constants (K_M) for NADH of GPD1 Variants

Variant	Vmax (U/ml)	K_M (μ M)
GPD1 WT_y	9	13
F129A	14	234
F129G	21	200
F129K	14	136
F129R	11	76
F73A	11	101
F73G	9	42
N44A	6	14
N44C	3	5
N44G	29	26
N44L	1	8
N44M	9	8
N44S	6	51
N44V	3	3
F129A	33	199
F129R	32	70
F73A129G (SEQ ID NO:204)	30	605
F73A129A (SEQ ID NO:205)	12	595
F73A129R (SEQ ID NO:206)	14	734
F73A129K (SEQ ID NO:207)	9	2989
F73G129G (SEQ ID NO:208)	29	515
F73G129A (SEQ ID NO:209)	15	554
F73G129R (SEQ ID NO:210)	7	1364
F73G129K (SEQ ID NO:211)	6	1671

[00283] These data indicate that mutations of amino acids corresponding to residues 44, 45, 73, and 129 (alone or in combination) of *S. cerevisiae* GPD can increase the K_M of GPD for NADH. As shown in Figure 2, amino acids 73 and 129 of *S. cerevisiae* GPD correspond to amino acids 41 and 97 of human GPD, respectively.

Example 2: Heterologous GPD enzymes with higher K_M for NADH than *S. cerevisiae* GPD

[00284] In this example, alternate glycerol-3-phosphate dehydrogenase enzymes with Michaelis constants (K_M) for NADH that are higher than yeast GPD1 were identified.

[00285] One strategy to identify higher NADH K_M enzymes is to use values published in literature for those enzymes that have been previously identified. Table 10 enumerates publications where the NADH K_M is higher than that reported for *Saccharomyces* GPD1.

Table 10: Published NADH K_M for GPDs

NADH K_M (mM)	Source	Reference
0.023	<i>Saccharomyces Cerevisiae</i> GPD1	Albertyn <i>et al.</i> 1992, <i>FEBS Lett.</i> 308:130
0.024	<i>Leishmania mexicana</i>	Marche <i>et al.</i> , 2000, <i>Mol Biochem Parasitol.</i> 106: 83-91
0.032	<i>Jaculus orientalis</i>	Berrada <i>et al.</i> , 2002 <i>Mol Cell Biochem.</i> 231: 117-27
0.0589	<i>Dunaliella viridis</i> GPDH1	He <i>et al.</i> , 2009, <i>Plant Mol Biol.</i> 71:193-205
0.0592	<i>Dunaliella viridis</i> GPDH2	He <i>et al.</i> , 2009, <i>Plant Mol Biol.</i> 71:193-205
0.078	<i>Drosophila melanogaster</i>	Niesel <i>et al.</i> , 1982, <i>Methods Enzymol.</i> 89, 296-301

[00286] The enzyme in the *Drosophila melanogaster* reference cited above identified “GPDH1” as coming from the *Drosophila* flight muscle. However, the reference predates sequence information about the gene encoding the enzyme. Subsequently, the sequence for this enzyme was identified (gi: 22945708) as glycerol 3 phosphate dehydrogenase, isoform C from the *Drosophila* genomic sequence (Carmon & MacIntyre, 2010, *Journal of Heredity* 101: 225-234). Using the techniques outlined in Example 1, this enzyme was expressed in *E. coli*, and the K_M value was measured in parallel with, and under the same conditions, used to measure *E. coli* expressed *Saccharomyces* GPD1. In one experiment, the NADH K_M was measured as 5 μ M for both enzymes, i.e., not significantly different.

[00287] An alternate strategy to identify naturally occurring GPDs with a high K_M for NADH is to evaluate members of the GPDs as defined by the enzyme commission nomenclature EC 1.1.1.94. While some of these enzymes use both NADH and NADPH equally well (*e.g.*, Edgar & Bell, 1980, *J Biol Chem* 255: 3492-34-97), others have been characterized as having a preference for NADPH (Frohlich *et al*, 2010, *J Bacteriol* 192: 4281-4288; Watanabe *et al*, 2008, *Yeast* 25:107-116; Sakasegawa *et al*, 2004, *Protein Science* 13: 1361-1371; Ruijter *et al*, 2004, *Microbiology* 150: 1095-1101). It is possible that this preference might be manifest as a high K_M for NADH (as compared to the K_M for NADPH of the same enzyme). Using the techniques outlined in Example 1, synthetic genes in the pBad expression vector for the following GPD enzymes were prepared: (a) *Archaeoglobus fulgidus* DSM 4304 (gi|11497621:c775889-774882) (SEQ ID NO:14); (b) *Candida versatilis* CvGPD1 gene for glycerol-3-phosphate dehydrogenase (gi|157060214|dbj|AB296385.1) (SEQ ID NO:15); and (c) *Rickettsia prowazekii* str. BuV67-CWPP chromosome (gi|383499256:539930-540880) (SEQ ID NO:16).

[00288] As in Example 1, these proteins were expressed in *E. coli*, and crude cell extracts were used to measure NADH K_M values. *Candida versatilis* GPD did not yield significant measureable activity. The *Archeoglobus fulgidus* enzyme had measureable activity with a $K_M = 7 \mu\text{M}$ for NADPH and $5 \mu\text{M}$ for NADH, and the *Rickettsia prowazekii* enzyme had measureable activity with a $K_M = 4 \mu\text{M}$ for NADPH and $664 \mu\text{M}$ for NADH.

[00289] The *Rickettsia prowazekii* enzyme K_M for NADH was higher than the K_M for *Saccharomyces*. In order to further evaluate what aspect of the enzyme might be contributing to this decreased affinity for NADH, the sequence of the *Rickettsia* enzyme was compared to the crystal structure of the human enzyme with NAD⁺ in the binding site. A notable feature of the human enzyme:NAD complex is the pi-stacking of phe41 and phe97 sandwiched around the adenine ring of NAD⁺ (pdb: 1X0X; Ou *et al*, 2006 *J Mol Biol* 357: 858-869). The pi-stacking is a very stable structure, and sequence alignment reveals that the homologous positions are conserved in yeast GPD1 sequence (phe73, phe129). However in *Rickettsia*, the homologous positions in an alignment are arg35 and ala85 (see Figure 2). These amino acids would be expected to destabilize NADH binding, thus leading to an increased NADH K_M . This was confirmed by mutagenesis at these positions in Example 1.

[00290] As further confirmation of the role of these amino acid positions in increasing NADH K_M , two of the most closely related GPD sequences outside of the *Rickettsia* genus were identified by BLAST search of the NCBI database. Two of the most closely related sequences are

from *Beggiatoa alba* (BLAST E value = $2e-51$, 37% sequence identity; SEQ ID NO:17) and *Kangiella koreensis* (BLAST E value = $2e-50$, 37% sequence identity; SEQ ID NO:18). These proteins were synthesized and tested as previously described.

[00291] Although the expression level in *E coli* for these enzymes was low, the values measured with the *Beggiatoa* enzyme were $K_M = 6 \mu\text{M}$ for NADPH and $101 \mu\text{M}$ for NADH, while the *Kangiella* enzyme values were $K_M = 1 \mu\text{M}$ for NADPH and $2018 \mu\text{M}$ for NADH. The amino acids in homologous positions to the two phenylalanines forming the pi-stacking in the human enzyme are lys85 and gly86 in the *Beggiatoa* enzyme, and arg35 and ala86 in the *Kangiella* enzyme (see Figure 2).

[00292] These results confirm that certain GPDs from other organisms have a higher K_M for NADH and also further support that the K_M for NADH can be raised by engineering GPD enzymes, for example, by modification of amino acids involved in the pi-stacking phenylalanine pair.

Example 3: Isobutanol and glycerol production for yeast strains comprising GPD1 enzyme variants.

[00293] In this example, yeast strains with variant GPD1 enzymes produced and described above were tested for isobutanol and glycerol production.

[00294] PNY2145 GPD and GPD variant strains with isobutanol pathway plasmid were plated on synthetic complete agar plates [1X yeast nitrogen base without amino acids (Difco), 1X amino acid drop-out without uracil (Clone-tech) containing 2% agar (Difco), 0.2% ethanol] and incubated for 72 hours at 300C incubator (New Brunswick)

[00295] Cells were patched on synthetic complete medium [1X yeast nitrogen base without amino acids (Difco), 1X amino acid drop-out without uracil (Clonotech) containing 2% agar (Difco), 1% glucose (sigma), 0.2% ethanol] and incubated for 72 hours at 30°C incubator (New Brunswick). Cells were adapted by repetitive plating every three days on same media for 30 days.

[00296] Patches of the adapted cells were inoculated in 10ml of synthetic complete liquid medium [1X yeast nitrogen base without amino acids (Difco), 1X amino acid drop-out without uracil (Clonotech) containing 2% agar (Difco), 1% glucose (Sigma), 0.2% ethanol] as primary cultures in 125 ml flasks (BD) and incubated at 30°C for 24 hours in an incubator shaker (New Brunswick) at 250 rpm. Secondary cultures were inoculated from the primary cultures in the

same medium with an initial OD of 0.5 and allowed to grow for another 24 hours. After 24 hours tertiary cultures were inoculated in the same medium from the secondary cultures with an initial O.D of 0.5 and allowed to grow for another 24 hours. These cells were then used for evaluation studies of GPD variants.

[00297] Cells were harvested by centrifugation at 3600 rpm for 5 minutes at room temperature in a centrifuge (Eppendorf) and suspended in production medium (1X yeast nitrogen base without amino acids (Difco), 1X amino acid drop-out without uracil (Clonetech), 35 g/L glucose (Sigma), 2 g/L ethanol, 100mM MES (Sigma) 1X peptone (Difco), 1X yeast extract (Difco), 1M HCL (Sigma) at pH 5.2) and initial OD of 2 in 15 ml falcon tubes. Cultures were then incubated at 30°C in an incubator shaker (New Brunswick) at 225 rpm for 20 hours. Samples were collected at 20 hours and analyzed by HPLC (Agilent Life Sciences) (Figure 3).

Sample preparation

[00298] Cultures were harvested at the end of production at 20 hours and each cell pellet was resuspended in 100 mM of MOPS pH 6.8 which contains 1X protease inhibitor (Roche). Lysis was achieved by subjecting cells to bead beating (Mini-Beadbeater-16, Biospec) for 5 cycles of 30 seconds each with an interval of 2 min between each cycle. Lysed sample was subjected to centrifugation at 13,000 rpm for 30 min in centrifuge at 40°C. Supernatant was carefully transferred to another tube. Protein estimation was done using Bradford reagent (Cat. #500-0205, Bio-Rad). GPD assay was done immediately on fresh samples without any freeze-thaw step.

Assay condition

[00299] Glycerol-3-phosphate (GPD) assays were performed using Cary 100 UV-Vis spectrophotometer in a 1 ml volume containing 100 mM MOPS (Sigma) pH 6.8, 1mM EDTA, 1mM glucose-6-phosphate, 3 mU/ μ l glucose-6-phosphate dehydrogenase (Cat. #G8404, Sigma-Aldrich, Inc., St. Louis, MO), 1mM dihydroxyacetone phosphate (Cat. #D51269, Sigma-Aldrich, Inc., St. Louis, MO), 0.3 mM NADH, and varying concentrations of cell extract. Rate of the reaction was calculated by taking slope of first 1 min for decrease in NADH concentration at 340 nm. Extinction coefficient of NADH was taken as 6.22 mM⁻¹cm⁻¹. For variants N8_1, E8_1 and M8_1 (i.e. those with the double mutant F73G/F129G) the NADH concentration was increased to 0.45 mM. This is non-saturating level of NADH for these variants.

Regression analysis of results

[00300] The data used in this section are provided in Table 11, and includes measurements of the metabolic products and the *in vitro* measurements of GPD as described above. To account for differences in the measured GPD activity that arises from measurement at subsaturating amounts of NADH, the total activity at V_{\max} was calculated by solving the single substrate Michaelis–Menton equation for V_{\max} , using GPD (U/mg) for rate and the K_M value as determined in Table 9.

Table 11: *In vitro* measurement of metabolic products and specific activity of GPD in various control and variant isobutanologen strains harboring different substitution in heterologous/native GPD sequence.

Strain	Glucose Cons. (g/L)	Gly (g/L)	EtOH (g/L)	iBuOH (g/L)	Yield (g/g)	iBuOH/Gly ratio	GPD (U/mg)	GPD K_M (μ M)	GPD V_{\max} (U/mg)
EC 1	21.52	1.55	0.4	4.14	0.19	2.67	0.0014	11	0.0015
E3 1	18.82	1.91	0.13	3.63	0.19	1.90	0.0018	101	0.0024
E8 1	12.16	0.75	0.13	3.11	0.26	4.15	0.0003	554	0.0007
M3 1	23.81	2.44	0.35	4.36	0.18	1.79	0.0053	101	0.0071
M8 1	22.91	2.89	0.2	4.64	0.20	1.61	0.0035	554	0.0078
N3 1	22.32	3.4	0.49	3.3	0.15	0.97	0.0097	101	0.0130
N8 1	13.95	2.1	0.19	2.16	0.15	1.03	0.0084	554	0.0187
2145	30.68	3.43	0.2	5.46	0.18	1.59	0.0025	11	0.0026

[00301] Initially, it was observed that integration of both unaltered and variant GPDs following the methods described above led to varying amounts of GPD activity, as analyzed in the yeast cell extract following isobutanol production. The specific activity of GPD could vary by as much as 10-fold. The isobutanol/glycerol ratio also exhibited an inverse correlation with the level of GPD activity detected as shown in Figure 4. As shown in the figure, the R^2 value for the linear regression of this relationship was 60.1%, but the cross-validated $R-Sq_{(pred)}$ value, which indicates the ability of the GPD activity data to predict unknown values of the isobutanol/glycerol ratio, was only 25.04%. This result indicated that while there was a correlation, the GPD activity alone was not a good predictor of this ratio.

[00302] This factor made it difficult to detect changes in the contribution of the high K_M GPD variants to the production of glycerol and isobutanol, as the beneficial effect of the increased NADH K_M was masked by the unpredictable activity levels. In order to further understand these effects, multiple linear regression analysis was applied to the data in Table 11 to determine if the contribution of the variant properties could be more clearly quantified. Modeled

parameters were subjected to multiple rounds of linear regression using Minitab (Minitab V16.2.1, Minitab Inc.; State College, PA), manually removing the contributing parameter with the greatest P value until the P value of the remaining coefficients were all below 0.05. This produced a regression equation with a maximum R-sq(pred) value, which indicated the ability of the model to predict the value of new observations. Two of the metabolic measurements in Table 11 yielded models with independent contributing parameters that could be interpreted physiologically.

[00303] Isobutanol titer was modeled using the following parameters from Table 11: glucose consumed, ethanol, glycerol, yield, GPD K_M , and GPD V_{max} . Eliminating the least significant parameters yielded a regression model with 3 parameters (predictors): glucose consumed, GPD K_M , and GPD V_{max} (shown in the Figure 5 and Table 12). The R-Sq value for this model is 98.6%. The R-Sq_(pred) value for this model is 94.32% indicating that this regression equation provides a high degree of predictive value. Interpreting the regression model physiologically, it suggests that isobutanol titer is predicted by positive contributions from the glucose consumed and GPD K_M values, and a negative contribution from the GPD V_{max} value. Thus, at any amount of glucose consumed, a higher GPD K_M will result in an increase in isobutanol titer. Similarly, increases in the level of the measured activity of the GPD enzyme (as V_{max} value) will result in a decrease in the isobutanol titer.

Table 12: Regression model for isobutanol titer (g/L).

Predictor	Coefficient	SE Coefficient	T	P	VIF
Constant	0.3316	0.3499	0.95	0.397	
Glucose Consumed (g/L)	0.17741	0.01382	12.83	0.000	1.804
GPD K_M (μ M)	0.0013789	0.0003309	4.17	0.014	1.977
GPD V_{max} (U/mg)	-75.81	10.10	-7.50	0.002	1.146

[00304] The regression model allowed an estimate of the magnitude of these effects. At a glucose consumption level of 30 g/L, and GPD V_{max} of 0.0026 U/mg, increasing the K_M from 11 μ M (wild type) to 550 μ M resulted in a 14% increase in isobutanol titer. However, at a glucose consumption level of 30 g/L, K_M at 11 μ M, and increase in the GPD level from 0.0015 U/mg to 0.019 U/mg (the maximum change in activity from Table 11) resulted in a 25% decrease in isobutanol titer. This regression model therefore demonstrated that decreasing GPD affinity for NADH (increased K_M), increased the isobutanol titer in the samples shown here.

[00305] Isobutanol yield (grams isobutanol/gram glucose consumed) was similarly modeled using the following parameters from Table 12: glucose consumed, ethanol, glycerol, isobutanol, GPD K_M , and GPD V_{max} . Elimination of the least significant parameters yielded a regression model using 2 parameters, GPD K_M , and GPD V_{max} (see Figure 6 and Table 13). This regression model had an R-Sq value of 93.9%. This regression model predicted yield with R-Sq_(pred) value of 76.3%, and notably is solely dependent on the activity level and K_M of the GPD enzyme. Similar to the regression model for isobutanol titer, the GPD V_{max} has a negative contribution to yield. In this regression model, increasing the K_M of GPD from 11 μ M (wild type) to 550 μ M, at the lowest observed GPD activity level of 0.0015 U/mg, resulted in a yield improvement of 28%. At the highest GPD activity level observed here (0.019 U/mg) the yield improvement was 78%.

Table 13: Regression model for yield (g/g).

Predictor	Coefficient	SE Coefficient	T	P	VIF
Constant	0.199176	0.005908	33.71	0.000	
GPD K_M (μ M)	0.00009410	0.00001597	5.89	0.002	1.144
GPD V_{max} (U/mg)	-5.2197	0.6404	-8.15	0.000	1.144

Example 4: Prophetic

[00306] In this example, heterologous GPD1 yeast integrants described above are tested for isobutanol and glycerol production.

Growth Media and Procedure

[00307] Two types of media are used during the growth procedure of yeast strains: an aerobic pre-culture media and an anaerobic culture media. All chemicals are obtained from Sigma unless otherwise noted (St. Louis, MO)

[00308] Aerobic pre-culture media (SE-Ura-His): 6.7 g/L yeast nitrogen base without amino acids (Difco, 291940, Sparks, MD), 1.4 g/L yeast synthetic drop-out medium supplement without histidine, leucine, tryptophan and uracil, 0.2% ethanol, 0.2% glucose, 0.01% w/v leucine and 0.002% w/v tryptophan.

[00309] Anaerobic culture media (SEG-Ura-His): 50 mM MES (pH 5.5, 6.7 g/L yeast nitrogen base without amino acids (Difco, 291940, Sparks, MD), 1.4 g/L yeast synthetic drop-out medium supplement without histidine, leucine, tryptophan and uracil, 0.1% ethanol, 3% glucose,

0.01% leucine, 0.002% tryptophan, 30 mg/L nicotinic acid, 30 mg/L thiamine and 10 mg/L ergosterol made up in 50/50 v/v Tween /ethanol solution.

[00310] The patched cells are inoculated into 25 mL SEG-Ura,His media with 0.2% glucose and 0.2% ethanol, and grown under progressively oxygen-limited conditions with lid closed for approximately 48 hours at 30°C with shaking, until a target OD₆₀₀ value of approximately 1.5 to 2 is achieved. OD₆₀₀ values are recorded. Cells are pelleted via centrifugation and the supernatant is discarded. Cell pellets are transferred into a Coy Anaerobic Bag (Grass Lake, MI) where pellets are resuspended in 1.0 mL anaerobic growth media (SEG-Ura-His). The resuspended cell pellets are used to inoculate 30 mL SEG-Ura-His media in 50 mL serum bottles (Wheaton, 223748, Millville, NJ) to a target initial OD₆₀₀ value of 0.2. All anaerobic media, serum vials, stoppers and crimps are allowed to degas in the anaerobic bag for at least 24 hours prior to inoculation. Serum bottles are stoppered, crimped and transferred out of the anaerobic bag and grown at 30°C with shaking at 240 rpm. Anaerobic cultures are grown for 24 to 72 hours with a target OD₆₀₀ value of at least 1.2. Additional anaerobic growth steps used the cells from the previous anaerobic culture step as inoculant. Three transformants were evaluated for each variant.

HPLC analysis of variant and heterologous yeast GPD1 strains

[00311] Samples are taken for HPLC analysis and to obtain OD₆₀₀ values at the end of the anaerobic growth period. HPLC analysis is performed using a Waters 2695 separations unit, 2996 photodiode array detector, and 2414 refractive index detector (Waters, Milford, MA) with a Shodex Sugar SH-G pre-column and Shodex Sugar SH1011 separations column (Shodex, JM Science, Grand Island, NY). Compounds are separated by isocratic elution at 0.01 N sulfuric acid with a flow rate of 0.5 mL/min. Chromatograms are analyzed using the Waters Empower Pro software.

Molar yields for glycerol, isobutanol and the isobutanol/glycerol ratio are determined. Mean and standard deviations are calculated from triplicate analyses for each variant and heterologous GPD. Student's t-test is employed to determine if the difference in the values are statistically significant from the codon-optimized GPD1 control values.

Example 5: Effect of *gpsA* on isobutanol production

[00312] Strains CPN97 and PNY2310 were grown on yeast synthetic medium containing 100mM MES (2-(N-morpholino)ethanesulfonic acid), 3 g/L glucose and 3 g/L ethanol and lacking histidine and uracil. A colony from each strain was selected and inoculated in 10 ml yeast synthetic medium containing 10 g/l glucose and 100 mM MES without histidine and uracil and incubated overnight at 30°C at 200 rpm. After overnight incubation, the cells were resuspended to an $OD_{600} = 0.4$ in 10 mL yeast synthetic medium containing 20 g/l glucose and 100 mM MES without histidine and uracil and incubated for 4 hours at 30°C and 200 rpm. The cells were then harvested and resuspended to an $OD_{600} = 0.2$ in 10 mL yeast synthetic medium containing 20 g/l glucose and 100 mM MES without histidine and uracil in a 20 ml serum vial (Wheaton; Millville, NJ), capped with a butylrubber stopper and sealed. Vials were placed in a 30°C incubator, rotated at 200 rpm, and incubated for 28 and 42.5 hours. Two vials were prepared for each strain tested.

[00313] After 28 and 42.5 hours, the cap of one of the vials was opened, OD_{600} was measured and the broth was analyzed by HPLC. HPLC analysis was performed on an Agilent 1100 series HPLC system containing a refractive index detector using a 300 mm x 7.8 mm BioRad-Aminex HPX-87H exclusion column (BioRad; Hercules, CA) incubated at 50°C and equipped with a BioRad-Microguard Cation H refill 30 mm x 4.6 mm. Samples were run at a flow rate of 0.6 ml/min in 0.01 N sulfuric acid running buffer. From the HPLC analysis, it was observed that isobutanol yield (Figure 7) and isobutanol/glycerol ratio (Figure 8) were increased in CPN97 as compared to PNY2310 and glucose consumption was decreased (Figure 9) in CPN97 as compared to PNY2310. While grown in aerobic conditions, the optical densities (ODs) as a function of time were similar.

Example 6: Generation of feedback resistant *gpsA*

Prophetic

[00314] The *gpsA* allele was amplified using *E. coli* MG1655 chromosomal DNA and primers P_{trc}-*gpsA* NcoI F (SEQ ID NO:238) and P_{trc}-*gpsA* PstI R (SEQ ID NO:239). The *gpsA* allele was cloned into NcoI/PstI-digested pTrcHis2B (Invitrogen; Carlsbad, CA) using the GeneArt seamless cloning and assembly kit (Life Technologies, Carlsbad, CA), to form pCPN124.

[00315] pCPN124 is submitted to error-prone mutagenesis using GeneMorphII Random Mutagenesis kit from Agilent (Santa Clara, CA). Plasmids that are obtained are transformed into

strain BB26-36 (Bell, J. Bact. 117:1065-1076 (1974)). Strain BB26-36 contains a mutation in the *plsB* gene. Additionally, the strain does not have the glycerol-3-phosphate auxotrophy of parent strain BB26 because of the loss of inhibition of glycerol kinase (GlpK) by fructose-1,6-diphosphate (fru-1,6-diP), so BB26-36 can produce glycerol-3-phosphate and grow on minimal media M9 plus glycerol 3 g/L and glucose 3 g/L (M9 contains 12.8 g sodium phosphate heptahydrate, 3 g potassium phosphate monobasic, 0.5 g sodium chloride, 1 g ammonium chloride, 0.24 g magnesium sulfate, and 11.1 mg calcium chloride, per liter).

[00316] The transformation reaction is plated on M9 medium containing 5 g/L glucose and 50 mg/L carbenicillin. Plasmids are extracted from the colonies growing on these plates and the *gpsA* gene is sequenced. GpsA activity of the mutated protein was determined and the K_i for glycerol-3-phosphate is measured. Proteins with increased K_i compared to wild-type protein are then used to replace GPD1 in the yeast chromosome and isobutanol and glycerol production are measured.

WHAT IS CLAIMED IS:

1. A recombinant microorganism comprising:
 - (a) a butanol biosynthetic pathway comprising at least one polypeptide that is heterologous to the recombinant microorganism;
 - (b) a heterologous glycerol-3-phosphate dehydrogenase (GPD), wherein the heterologous GPD has a higher K_M for NADH as compared to the K_M of the endogenous GPD of the microorganism; and
 - (c) a deletion or disruption in an endogenous gene encoding GPD;

wherein the recombinant microorganism has improved production of butanol as compared to a control recombinant microorganism that lacks the heterologous GPD.

2. The recombinant microorganism of claim 1, wherein the butanol biosynthetic pathway is selected from the group consisting of:

- (a) a 1-butanol biosynthetic pathway;
- (b) a 2-butanol biosynthetic pathway; and
- (c) an isobutanol biosynthetic pathway.

3. The recombinant microorganism of claim 2, wherein the 1-butanol biosynthetic pathway comprises at least one gene encoding a polypeptide that performs at least one of the following substrate to product conversions:

- (a) acetyl-CoA to acetoacetyl-CoA, as catalyzed by acetyl-CoA acetyltransferase;
- (b) acetoacetyl-CoA to 3-hydroxybutyryl-CoA, as catalyzed by 3-hydroxybutyryl-CoA dehydrogenase;
- (c) 3-hydroxybutyryl-CoA to crotonyl-CoA, as catalyzed by crotonase;
- (d) crotonyl-CoA to butyryl-CoA, as catalyzed by butyryl-CoA dehydrogenase;
- (e) butyryl-CoA to butyraldehyde, as catalyzed by butyraldehyde dehydrogenase; and
- (f) butyraldehyde to 1-butanol, as catalyzed by 1-butanol dehydrogenase.

4. The recombinant microorganism of claim 2, wherein the 2-butanol biosynthetic pathway comprises at least one gene encoding a polypeptide that performs at least one of the following substrate to product conversions:

- (a) pyruvate to alpha-acetolactate, as catalyzed by acetolactate synthase;

- (b) alpha-acetolactate to acetoin, as catalyzed by acetolactate decarboxylase;
 - (c) acetoin to 2,3-butanediol, as catalyzed by butanediol dehydrogenase;
 - (d) 2,3-butanediol to 2-butanone, as catalyzed by butanediol dehydratase; and
 - (e) 2-butanone to 2-butanol, as catalyzed by 2-butanol dehydrogenase.
5. The recombinant microorganism of claim 2, wherein the isobutanol biosynthetic pathway comprises at least one gene encoding a polypeptide that performs at least one of the following substrate to product conversions:
- (a) pyruvate to acetolactate, as catalyzed by acetolactate synthase;
 - (b) acetolactate to 2,3-dihydroxyisovalerate, as catalyzed by acetohydroxy acid isomeroreductase;
 - (c) 2,3-dihydroxyisovalerate to α -ketoisovalerate, as catalyzed by dihydroxyacid dehydratase;
 - (d) α -ketoisovalerate to isobutyraldehyde, as catalyzed by a branched chain keto acid decarboxylase; and
 - (e) isobutyraldehyde to isobutanol, as catalyzed by branched-chain alcohol dehydrogenase.
6. The recombinant microorganism of any one of claims 1-5, wherein the microorganism is from a genus selected from the group consisting of *Clostridium*, *Zymomonas*, *Escherichia*, *Salmonella*, *Serratia*, *Erwinia*, *Klebsiella*, *Shigella*, *Rhodococcus*, *Pseudomonas*, *Bacillus*, *Lactobacillus*, *Enterococcus*, *Alcaligenes*, *Paenibacillus*, *Arthrobacter*, *Corynebacterium*, *Brevibacterium*, *Schizosaccharomyces*, *Kluveromyces*, *Yarrowia*, *Pichia*, *Zygosaccharomyces*, *Debaryomyces*, *Candida*, *Brettanomyces*, *Pachysolen*, *Hansenula*, *Issatchenkia*, *Trichosporon*, *Yamadazyma*, and *Saccharomyces*.
7. The recombinant microorganism of claim 6, wherein the microorganism is from the genus *Saccharomyces*.
8. The recombinant microorganism of any one of claims 1-7, wherein the heterologous GPD is a naturally occurring GPD.

9. The recombinant microorganism of claim 8, wherein the heterologous GPD comprises the sequence of a GPD from an organism selected from the group consisting of *Leishmania mexicana*, *Dunaliella viridis*, *Jaculus orientalis*, *Archeoglobus fulgidus*, *Rickettsia prowazekii*, *Beggiatoa alba*, *Kangiella koreenis*, *Aspergillus oryzae*, *Candida versatilis*, *Escherichia coli*, and *Oryctolagus cuniculu*.
10. The recombinant microorganism of any one of claims 1-7, wherein the heterologous GPD is an engineered GPD.
11. The recombinant microorganism of claim 10, wherein the engineered GPD comprises at least one substitution at a residue corresponding to position 42, 44, 45, 71, 73, 75, 95, 124, 126, 129, 151, 152, 183, 184, 185, 246, 310, 336, 337, or 339 of SEQ ID NO: 195.
12. The recombinant microorganism of claim 11, wherein the engineered GPD comprises at least one substitution at a residue corresponding to position 73 of SEQ ID NO:195.
13. The recombinant microorganism of claim 11, wherein the engineered GPD comprises at least one substitution at a residue corresponding to position 129 of SEQ ID NO:195.
14. The recombinant microorganism of claim 11, wherein the engineered GPD comprises a substitution at a residue corresponding to position 73 of SEQ ID NO:195 and a substitution at a residue corresponding to position 129 of SEQ ID NO:195.
15. The recombinant microorganism of any one of claims 5-14, wherein the microorganism comprises a ketol-acid reductoisomerase (KARI) that utilizes NADH.
16. An engineered glycerol-3-phosphate dehydrogenase (GPD) enzyme having at least 85% identity to of SEQ ID NO: 195.
17. The engineered GPD enzyme of claim 15, wherein the enzyme comprises at least one substitution at a residue corresponding to position 42, 44, 45, 71, 73, 75, 95, 124, 126, 129, 151, 152, 183, 184, 185, 246, 310, 336, 337, or 339 of SEQ ID NO: 195.

18. The engineered GPD enzyme of claim 17, wherein the enzyme comprises at least one substitution corresponding to position 73 of SEQ ID NO: 195.
19. The engineered GPD enzyme of claim 17, wherein the enzyme comprises at least one substitution corresponding to position 129 of SEQ ID NO:195.
20. The engineered GPD enzyme of claim 17, wherein the enzyme comprises a substitution corresponding to position 73 of SEQ ID NO:195 and a substitution corresponding to position 129 of SEQ ID NO:195.
21. The engineered GPD enzyme of any one of claims 16-20, wherein the K_M for NADH is from about 0.01 mM to about 1mM.
22. A recombinant microorganism comprising the engineered GPD enzyme of any one of claims 16-21.
23. The recombinant microorganism of claim 22, wherein the microorganism comprises a butanol biosynthetic pathway that comprises at least one gene that is heterologous to the recombinant microorganism; wherein the microorganism comprises a deletion or disruption of an endogenous gene encoding GPD; and wherein the recombinant microorganism has improved production of butanol compared to a microorganism that lacks the engineered GPD enzyme.
24. The recombinant microorganism of claim 23, wherein the butanol biosynthetic pathway is selected from the group consisting of:
 - (a) a 1-butanol biosynthetic pathway;
 - (b) a 2-butanol biosynthetic pathway; and
 - (c) an isobutanol biosynthetic pathway.
25. The recombinant microorganism of claim 24, wherein the 1-butanol biosynthetic pathway comprises at least one gene encoding a polypeptide that performs at least one of the following substrate to product conversions:
 - (a) acetyl-CoA to acetoacetyl-CoA, as catalyzed by acetyl-CoA acetyltransferase;

- (b) acetoacetyl-CoA to 3-hydroxybutyryl-CoA, as catalyzed by 3-hydroxybutyryl-CoA dehydrogenase;
- (c) 3-hydroxybutyryl-CoA to crotonyl-CoA, as catalyzed by crotonase;
- (d) crotonyl-CoA to butyryl-CoA, as catalyzed by butyryl-CoA dehydrogenase;
- (e) butyryl-CoA to butyraldehyde, as catalyzed by butyraldehyde dehydrogenase; and
- (f) butyraldehyde to 1-butanol, as catalyzed by 1-butanol dehydrogenase.

26. The recombinant microorganism of claim 24, wherein the 2-butanol biosynthetic pathway comprises at least one gene encoding a polypeptide that performs at least one of the following substrate to product conversions:

- (a) pyruvate to alpha-acetolactate, as catalyzed by acetolactate synthase;
- (b) alpha-acetolactate to acetoin, as catalyzed by acetolactate decarboxylase;
- (c) acetoin to 2,3-butanediol, as catalyzed by butanediol dehydrogenase;
- (d) 2,3-butanediol to 2-butanone, as catalyzed by butanediol dehydratase; and
- (e) 2-butanone to 2-butanol, as catalyzed by 2-butanol dehydrogenase.

27. The recombinant microorganism of claim 24, wherein the isobutanol biosynthetic pathway comprises at least one gene encoding a polypeptide that performs at least one of the following substrate to product conversions:

- (a) pyruvate to acetolactate, as catalyzed by acetolactate synthase;
- (b) acetolactate to 2,3-dihydroxyisovalerate, as catalyzed by acetohydroxy acid isomeroreductase;
- (c) 2,3-dihydroxyisovalerate to α -ketoisovalerate, as catalyzed by dihydroxyacid dehydratase;
- (d) α -ketoisovalerate to isobutyraldehyde, as catalyzed by a branched chain keto acid decarboxylase; and
- (e) isobutyraldehyde to isobutanol, as catalyzed by branched-chain alcohol dehydrogenase.

28. The recombinant microorganism of any one of claims 22-27, wherein the microorganism is from a genus selected from the group consisting of *Clostridium*, *Zymomonas*, *Escherichia*, *Salmonella*, *Serratia*, *Erwinia*, *Klebsiella*, *Shigella*, *Rhodococcus*, *Pseudomonas*, *Bacillus*,

Lactobacillus, Enterococcus, Alcaligenes, Paenibacillus, Arthrobacter, Corynebacterium, Brevibacterium, Schizosaccharomyces, Kluyveromyces, Yarrowia, Pichia, Zygosaccharomyces, Debaryomyces, Candida, Brettanomyces, Pachysolen, Hansenula, Issatchenkia, Trichosporon, Yamadazyma, and Saccharomyces.

29. The recombinant microorganism of claim 28, wherein the microorganism is from the genus *Saccharomyces*.
30. The recombinant microorganism of any one of claims 27-29, wherein the microorganism comprises a ketol-acid reductoisomerase (KARI) that utilizes NADH.
31. A method for the production of isobutanol comprising:
- (a) providing a recombinant microorganism comprising
 - i. an engineered isobutanol biosynthetic pathway,
 - ii. a deletion or disruption in an endogenous gene encoding GPD1, and;
 - iii. at least one of
 - a. the engineered GPD enzyme of any one of claims 16-22; or
 - b. a heterologous glycerol-3-phosphate dehydrogenase (GPD), wherein the heterologous GPD has a higher K_M for NADH as compared to the K_M of the microorganism's endogenous GPD; and
 - (b) contacting the recombinant microorganism with at least one fermentable carbon substrate under conditions wherein isobutanol is produced.
32. The method of claim 31, wherein the recombinant microorganism is grown under anaerobic conditions.
33. The method of any one of claims 31-32, wherein the engineered GPD comprises at least one substitution at a residue corresponding to 42, 44, 45, 71, 73, 75, 95, 124, 126, 129, 151, 152, 183, 184, 185, 246, 310, 336, 337, or 339 of SEQ ID NO:195.
34. The method of any one of claims 31-33, wherein the K_M for NADH of the GPD is about 0.01 mM to about 1 mM.

35. A recombinant microorganism comprising:
- (a) a butanol biosynthetic pathway comprising at least one polypeptide that is heterologous to the recombinant microorganism;
 - (b) a heterologous glycerol-3-phosphate dehydrogenase (GPD), wherein the heterologous GPD has a higher K_m for NADH as compared to the K_m of the endogenous GPD of the microorganism; and
 - (c) a deletion or disruption in an endogenous gene encoding GPD;
- wherein the recombinant microorganism has decreased production of glycerol as compared to a control recombinant microorganism that lacks the heterologous GPD.
36. A recombinant microorganism comprising:
- (a) a butanol biosynthetic pathway comprising at least one polypeptide that is heterologous to the recombinant microorganism;
 - (b) a heterologous glycerol-3-phosphate dehydrogenase (GPD), wherein the heterologous GPD has a higher K_M for NADH as compared to the K_M of the endogenous GPD of the microorganism; and
 - (c) a deletion or disruption in an endogenous gene encoding GPD;
- wherein the recombinant microorganism has an increased butanol to glycerol molar ratio as compared to a control recombinant microorganism that lacks the heterologous GPD.
37. A recombinant microorganism comprising:
- (a) a heterologous glycerol-3-phosphate dehydrogenase (GPD), wherein the heterologous GPD has a higher K_M for NADH as compared to the K_M of the endogenous GPD of the microorganism; and
 - (b) a deletion or disruption in an endogenous gene encoding GPD;
- wherein the recombinant microorganism has decreased production of glycerol as compared to a control recombinant microorganism that lacks the heterologous GPD.
38. A recombinant microorganism comprising:
- (a) an engineered butanol biosynthetic pathway comprising at least one polypeptide that is heterologous to the recombinant microorganism;

(b) a heterologous glycerol-3-phosphate dehydrogenase (GPD), wherein the heterologous GPD has substantially the same affinity for NADH and NADPH; and

(c) a deletion or disruption in an endogenous gene encoding GPD;

wherein the recombinant microorganism has improved production of butanol as compared to a control recombinant microorganism that lacks the heterologous GPD.

39. A recombinant microorganism comprising:

(a) an engineered butanol biosynthetic pathway comprising at least one polypeptide that is heterologous to the recombinant microorganism;

(b) a heterologous glycerol-3-phosphate dehydrogenase (GPD), wherein the heterologous GPD has substantially the same affinity for NADH and NADPH; and

(c) a deletion or disruption in an endogenous gene encoding GPD;

wherein the recombinant microorganism has decreased production of glycerol as compared to a control recombinant microorganism that lacks the heterologous GPD.

40. A recombinant microorganism comprising:

(a) an engineered butanol biosynthetic pathway comprising at least one polypeptide that is heterologous to the recombinant microorganism;

(b) a heterologous glycerol-3-phosphate dehydrogenase (GPD), wherein the heterologous GPD has substantially the same affinity for NADH and NADPH; and

(c) a deletion or disruption in an endogenous gene encoding GPD;

wherein the recombinant microorganism has an increased butanol to glycerol molar ratio as compared to a control recombinant microorganism that lacks the heterologous GPD.

41. The recombinant microorganism of any one of claims 38-40, wherein the engineered butanol biosynthetic pathway is selected from the group consisting of:

(a) a 1-butanol biosynthetic pathway;

(b) a 2-butanol biosynthetic pathway; and

(c) an isobutanol biosynthetic pathway.

42. The recombinant microorganism of any one of claims 38-41, wherein the microorganism is from a genus selected from the group consisting of *Clostridium*, *Zymomonas*, *Escherichia*,

Salmonella, Serratia, Erwinia, Klebsiella, Shigella, Rhodococcus, Pseudomonas, Bacillus, Lactobacillus, Enterococcus, Alcaligenes, Paenibacillus, Arthrobacter, Corynebacterium, Brevibacterium, Schizosaccharomyces, Kluyveromyces, Yarrowia, Pichia, Zygosaccharomyces, Debaryomyces, Candida, Brettanomyces, Pachysolen, Hansenula, Issatchenkia, Trichosporon, Yamadazyma, and Saccharomyces.

43. The recombinant microorganism of claim 42, wherein the microorganism is from the genus *Saccharomyces*.
44. The recombinant microorganism of any one of claims 38-43, wherein the heterologous GPD is a GPD having the EC number 1.1.1.94.
45. The recombinant microorganism of claim 44, wherein the heterologous GPD is selected from an *E. coli* GPD, a *Candida versatilis* GPD, and an *Aspergillus oryzae* GPD.
46. A method for the production of butanol comprising:
(a) providing the recombinant microorganism of any one of claims 38-40; and
(b) contacting the recombinant microorganism with at least one fermentable carbon substrate under conditions wherein butanol is produced.
47. The method of claim 9, wherein the engineered butanol biosynthetic pathway is selected from the group consisting of:
(a) a 1-butanol biosynthetic pathway;
(b) a 2-butanol biosynthetic pathway; and
(c) an isobutanol biosynthetic pathway.
48. The method of claim 46 or 47, wherein the microorganism is from a genus selected from the group consisting of *Clostridium, Zymomonas, Escherichia, Salmonella, Serratia, Erwinia, Klebsiella, Shigella, Rhodococcus, Pseudomonas, Bacillus, Lactobacillus, Enterococcus, Alcaligenes, Paenibacillus, Arthrobacter, Corynebacterium, Brevibacterium, Schizosaccharomyces, Kluyveromyces, Yarrowia, Pichia, Zygosaccharomyces, Debaryomyces, Candida, Brettanomyces, Pachysolen, Hansenula, Issatchenkia, Trichosporon, Yamadazyma, and Saccharomyces.*

49. The method of claim 48, wherein the microorganism is from the genus *Saccharomyces*.
50. The method of any one of claims 46-49, wherein the heterologous GPD is a GPD having the EC number 1.1.1.94.
51. The method of claim 50, wherein the heterologous GPD is selected from an *E. coli* GPD, a *Candida versatilis* GPD, and an *Aspergillus oryzae* GPD.
52. A recombinant microorganism comprising:
(a) a heterologous glycerol-3-phosphate dehydrogenase (GPD), wherein the heterologous GPD has substantially the same affinity for NADH and NADPH; and
(b) a deletion or disruption in an endogenous gene encoding GPD;
wherein the recombinant microorganism has decreased production of glycerol as compared to a control recombinant microorganism that lacks the heterologous GPD.
53. A recombinant microorganism comprising a heterologous glycerol-3-phosphate dehydrogenase (GPD), wherein the heterologous GPD has substantially the same affinity for NADH or NADPH; and wherein the recombinant microorganism has decreased production of glycerol as compared to a control recombinant microorganism that lacks the heterologous GPD.
54. A recombinant microorganism comprising:
(a) an engineered butanol biosynthetic pathway comprising at least one polypeptide that is heterologous to the recombinant microorganism;
(b) a heterologous glycerol-3-phosphate dehydrogenase (GPD), wherein the heterologous GPD is inhibited by a glycerol-3-phosphate; and
(c) a deletion or disruption in an endogenous gene encoding GPD;
wherein the recombinant microorganism has improved production of butanol as compared to a control recombinant microorganism that lacks the heterologous GPD.
55. A recombinant microorganism comprising:
(a) an engineered butanol biosynthetic pathway comprising at least one polypeptide that is heterologous to the recombinant microorganism;

(b) a heterologous glycerol-3-phosphate dehydrogenase (GPD), wherein the heterologous GPD is inhibited by a glycerol-3-phosphate; and

(c) a deletion or disruption in an endogenous gene encoding GPD;

wherein the recombinant microorganism has decreased production of glycerol as compared to a control recombinant microorganism that lacks the heterologous GPD.

56. A recombinant microorganism comprising:

(a) an engineered butanol biosynthetic pathway comprising at least one polypeptide that is heterologous to the recombinant microorganism;

(b) a heterologous glycerol-3-phosphate dehydrogenase (GPD), wherein the heterologous GPD is inhibited by a glycerol-3-phosphate; and

(c) a deletion or disruption in an endogenous gene encoding GPD;

wherein the recombinant microorganism has an increased butanol to glycerol molar ratio as compared to a control recombinant microorganism that lacks the heterologous GPD.

57. The recombinant microorganism of any one of claims 54-56, wherein the engineered butanol biosynthetic pathway is selected from the group consisting of:

(a) a 1-butanol biosynthetic pathway;

(b) a 2-butanol biosynthetic pathway; and

(c) an isobutanol biosynthetic pathway.

58. The recombinant microorganism of any one of claims 54-57, wherein the microorganism is from a genus selected from the group consisting of *Clostridium*, *Zymomonas*, *Escherichia*, *Salmonella*, *Serratia*, *Erwinia*, *Klebsiella*, *Shigella*, *Rhodococcus*, *Pseudomonas*, *Bacillus*, *Lactobacillus*, *Enterococcus*, *Alcaligenes*, *Paenibacillus*, *Arthrobacter*, *Corynebacterium*, *Brevibacterium*, *Schizosaccharomyces*, *Kluveromyces*, *Yarrowia*, *Pichia*, *Zygosaccharomyces*, *Debaryomyces*, *Candida*, *Brettanomyces*, *Pachysolen*, *Hansenula*, *Issatchenkia*, *Trichosporon*, *Yamadazyma*, and *Saccharomyces*.

59. The recombinant microorganism of any one of claims 54-58, wherein the heterologous GPD is a GPD having the EC number 1.1.1.94.

60. The recombinant microorganism of claim 59, wherein the heterologous GPD is selected from an *E. coli* GPD, a *Candida versatilis* GPD, and an *Aspergillus oryzae* GPD.
61. A method for the production of butanol comprising:
- (a) providing the recombinant microorganism of any one of claims 54-60; and
 - (b) contacting the recombinant microorganism with at least one fermentable carbon substrate under conditions wherein butanol is produced.
62. A recombinant microorganism comprising:
- (a) a heterologous glycerol-3-phosphate dehydrogenase (GPD), wherein the heterologous GPD is inhibited by a glycerol-3-phosphate; and
 - (b) a deletion or disruption in an endogenous gene encoding GPD;
- wherein the recombinant microorganism has decreased production of glycerol as compared to a control recombinant microorganism that lacks the heterologous GPD.
63. A recombinant microorganism comprising a heterologous glycerol-3-phosphate dehydrogenase (GPD), wherein the heterologous GPD is inhibited by a glycerol-3-phosphate; and wherein the recombinant microorganism has decreased production of glycerol as compared to a control recombinant microorganism that lacks the heterologous GPD.

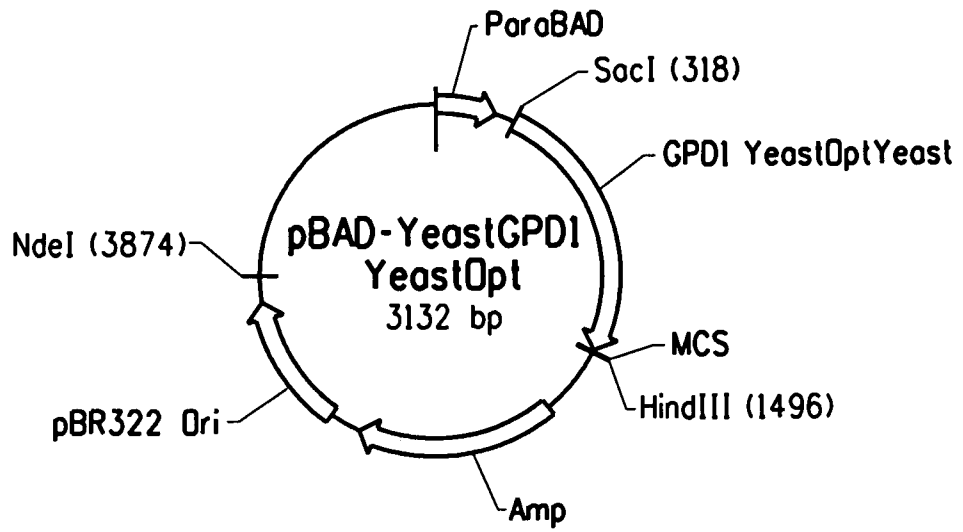


FIG. 1

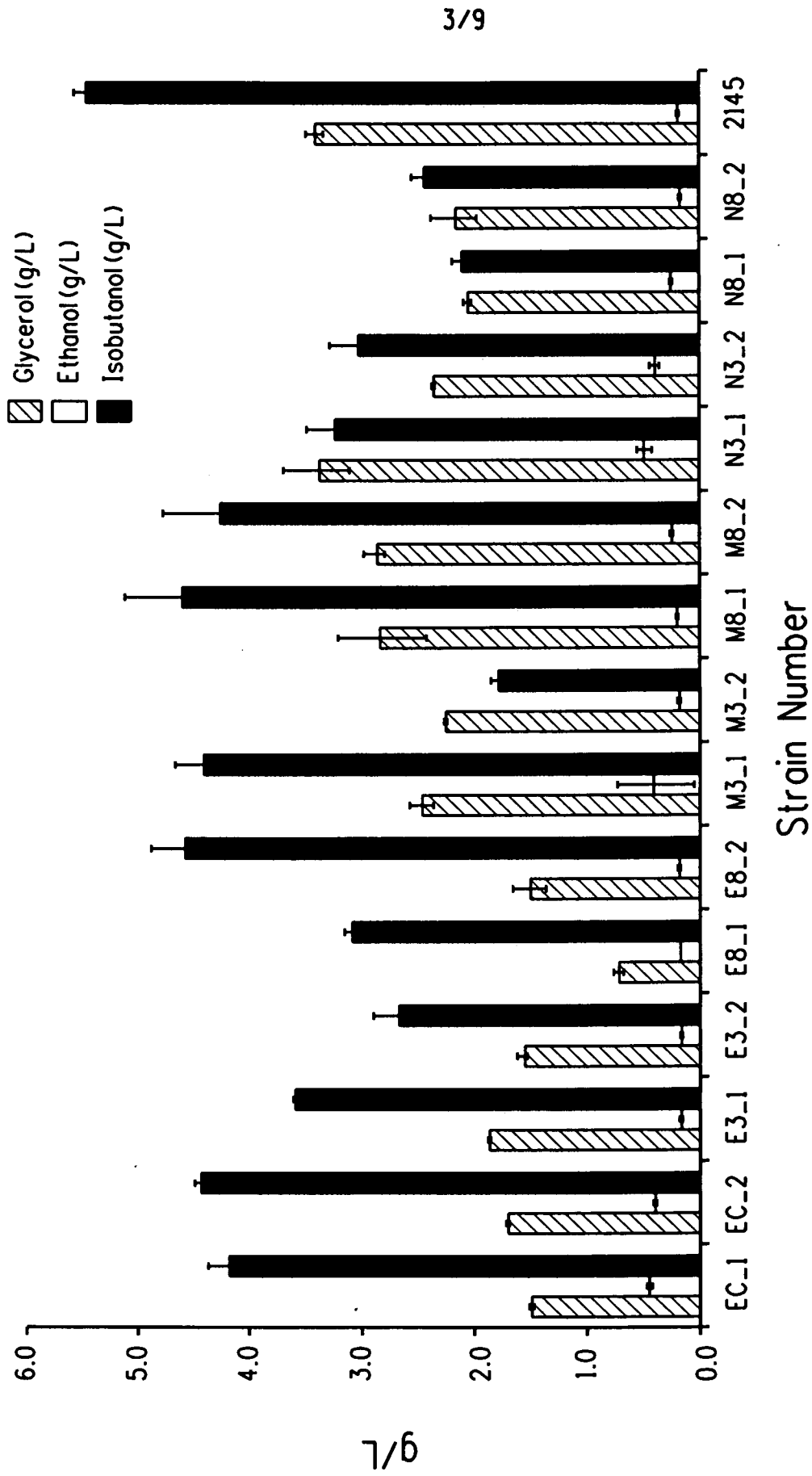


FIG. 3

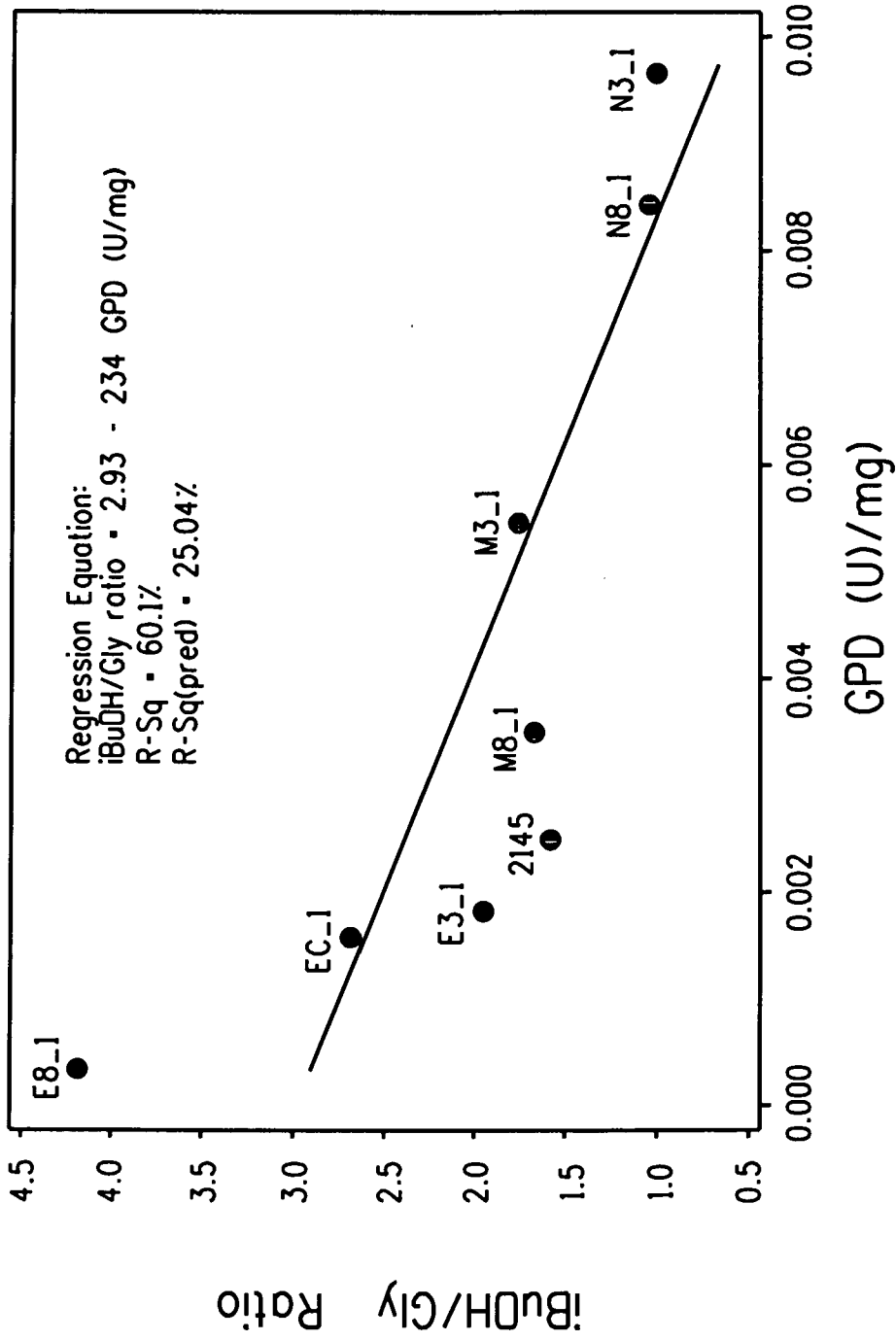
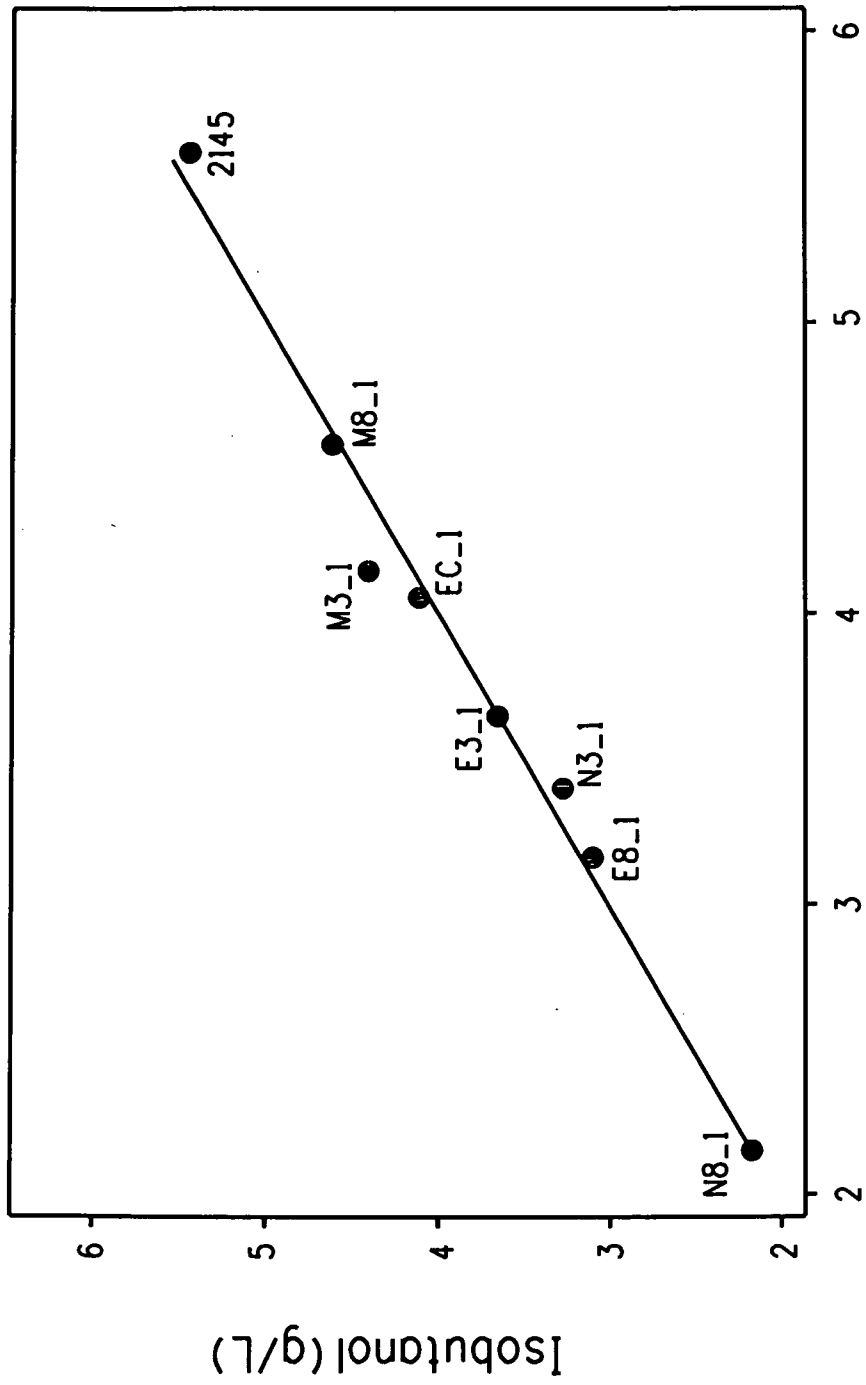
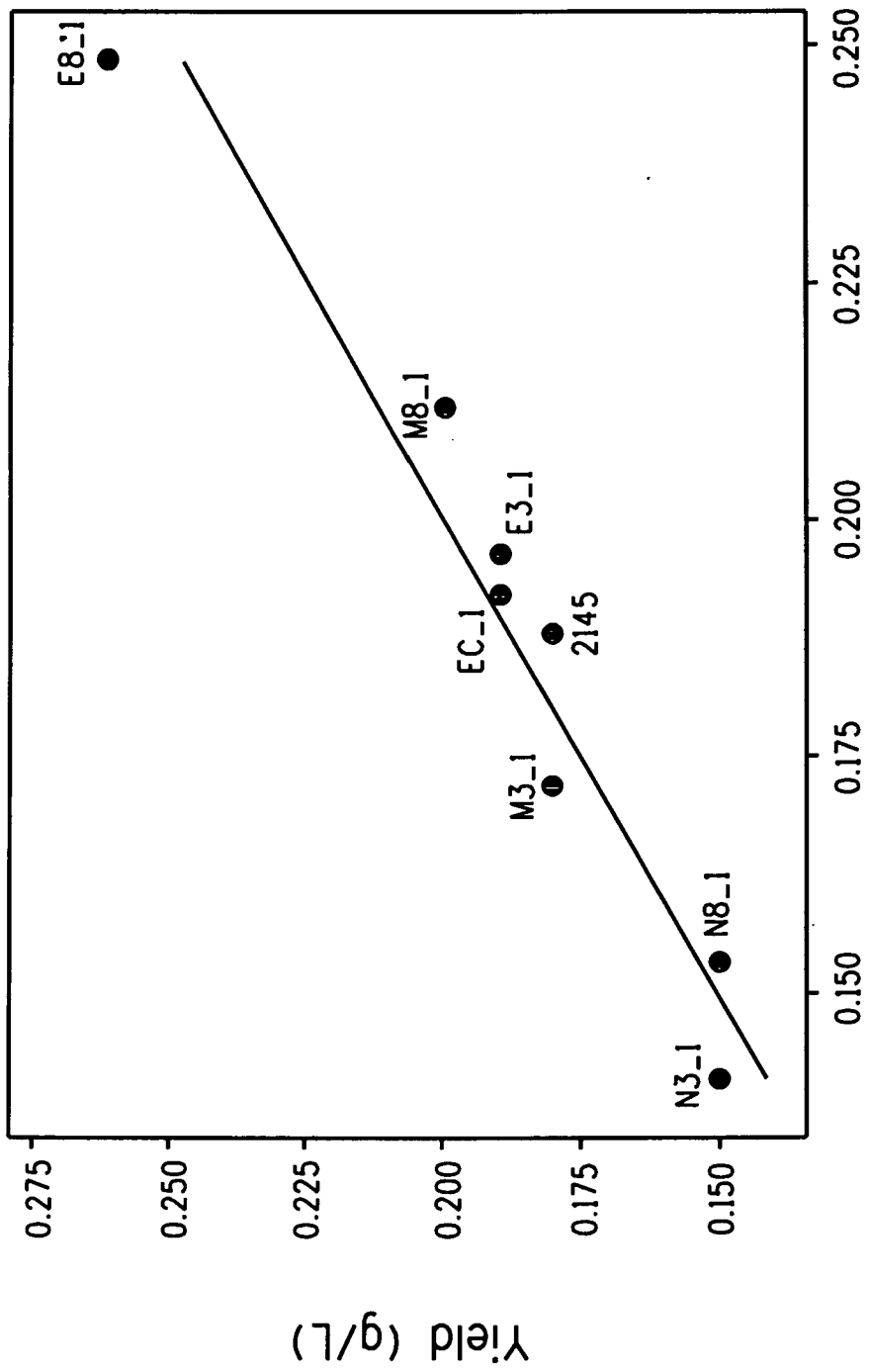


FIG. 4



FIT_2
FIG. 5



FIT_4

FIG. 6

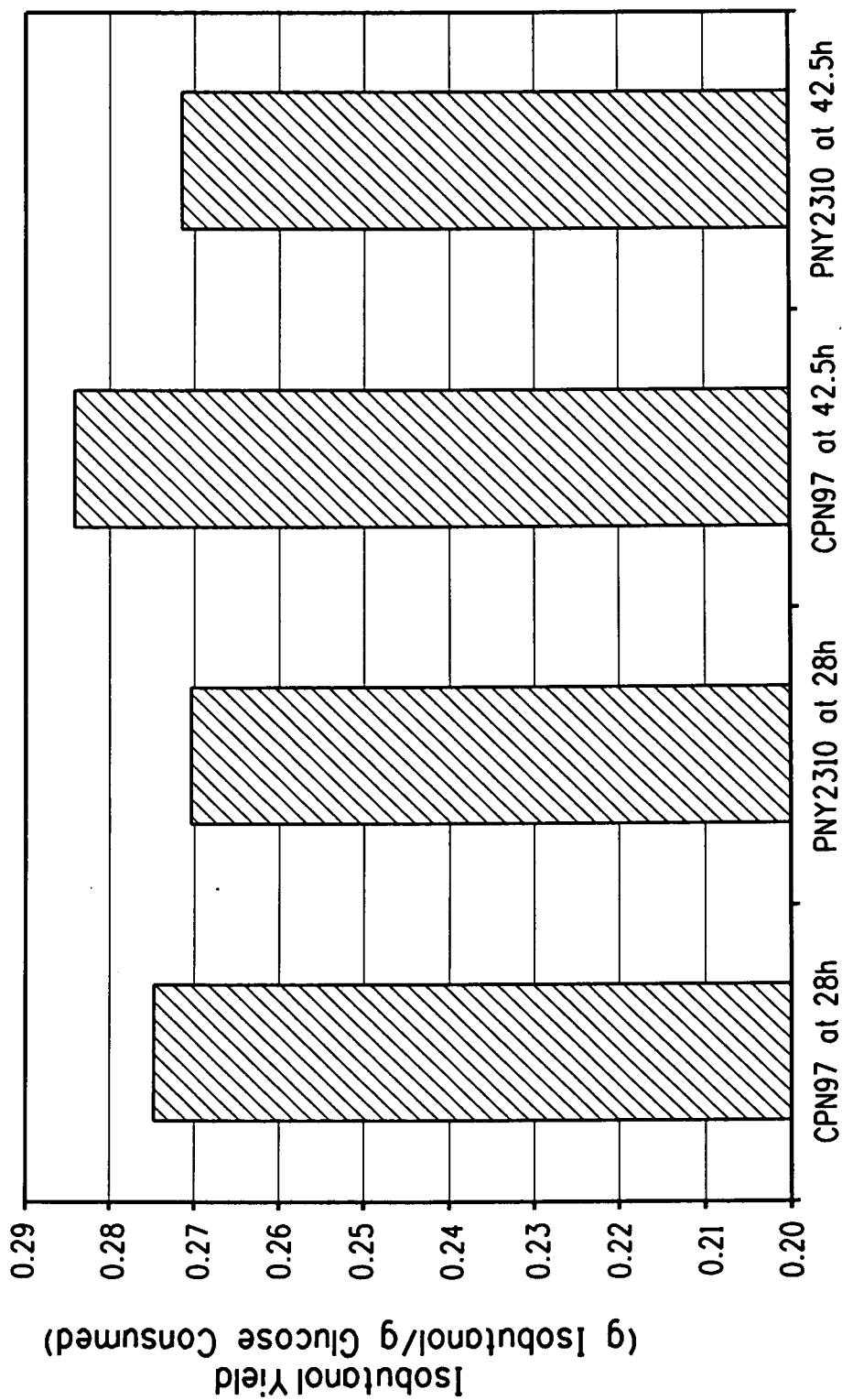


FIG. 7

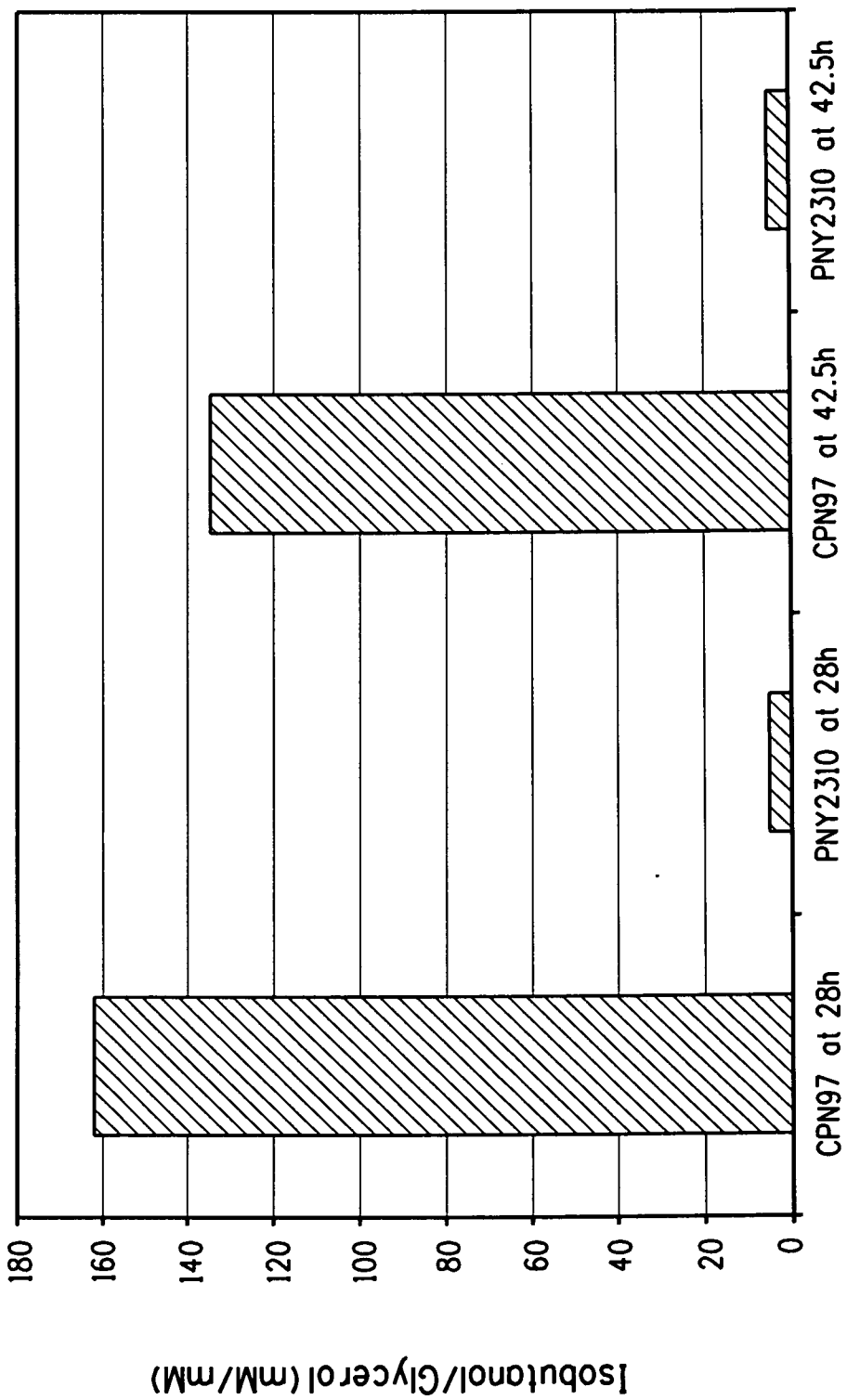


FIG. 8

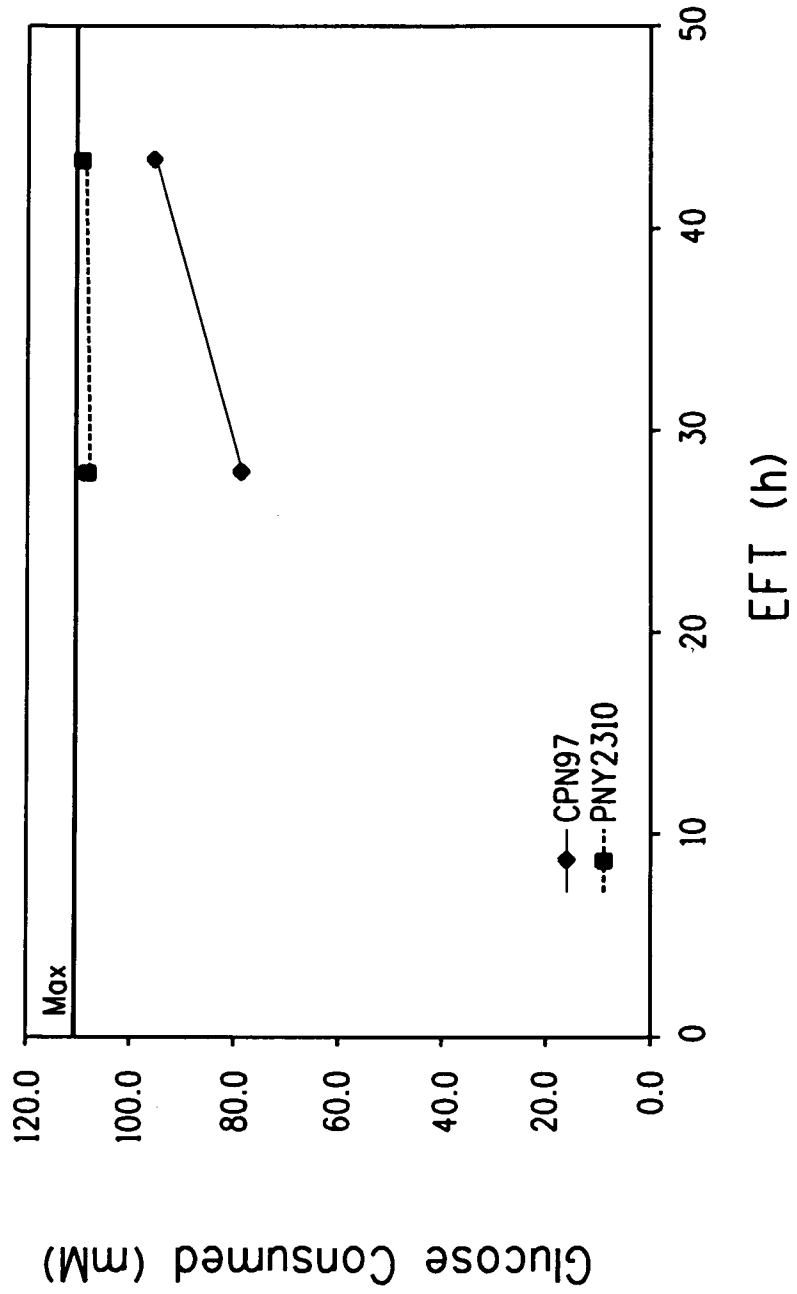


FIG. 9