



(86) Date de dépôt PCT/PCT Filing Date: 2010/08/20
(87) Date publication PCT/PCT Publication Date: 2011/02/24
(85) Entrée phase nationale/National Entry: 2012/02/21
(86) N° demande PCT/PCT Application No.: US 2010/046172
(87) N° publication PCT/PCT Publication No.: 2011/022651
(30) Priorités/Priorities: 2009/08/21 (US61/235,959);
2010/01/27 (US61/298,790)

(51) Cl.Int./Int.Cl. *C12P 7/00* (2006.01),
C12N 15/52 (2006.01), *C12P 7/08* (2006.01),
C12P 7/20 (2006.01)

(71) Demandeur/Applicant:
MASCOMA CORPORATION, US

(72) Inventeurs/Inventors:
MCBRIDE, JOHN, US;
RAJGARHIA, VINEET, US;
SHAW, ARTHUR J., IV, US;
TRIPATHI, SHITAL A., US;
BREVNOVA, ELENA, US;
CAIAZZA, NICKY, US;
...

(74) Agent: GOWLING LAFLEUR HENDERSON LLP

(54) Titre : PRODUCTION DE PROPANOLS, D'ALCOOLS, ET DE POLYOLS DANS DES ORGANISMES DE BIOTRAITEMENT CONSOLIDÉS

(54) Title: PRODUCTION OF PROPANOLS, ALCOHOLS, AND POLYOLS IN CONSOLIDATED BIOPROCESSING ORGANISMS

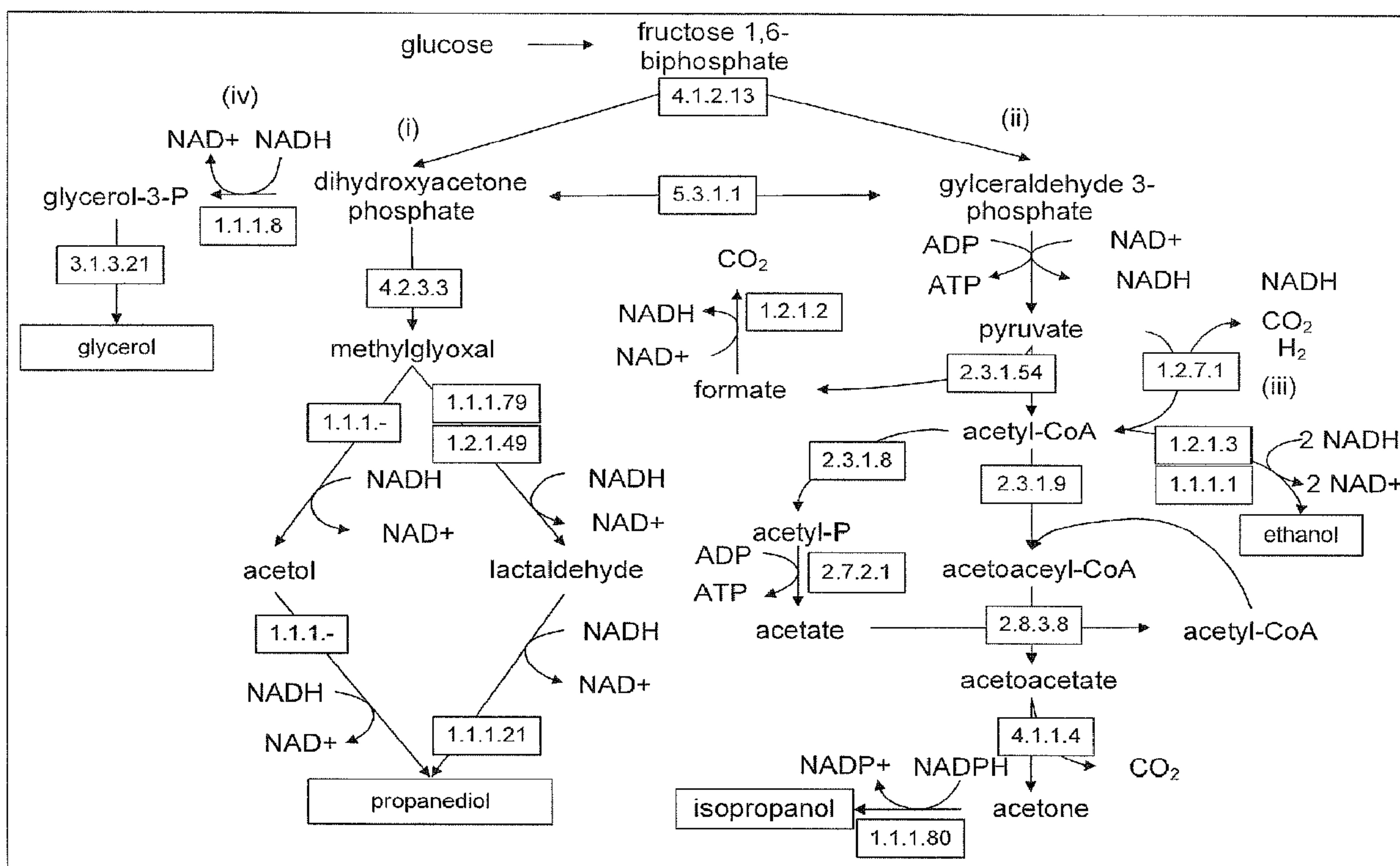


FIG. 1

(57) **Abrégé/Abstract:**

The present invention provides for novel metabolic pathways leading to propanol, alcohol or polyol formation in a consolidated bioprocessing system (CBP), where lignocellulosic biomass is efficiently converted to such products. More specifically, the

(72) **Inventeurs(suite)/Inventors(continued):** VAN DIJKEN, HANS, NL; FROEHLICH, ALLAN, US;
SILLERS, WILLIAM RYAN, US; FLATT, JIM, US

(57) **Abrégé(suite)/Abstract(continued):**

invention provides for a recombinant microorganism, where the microorganism expresses one or more native and/or heterologous enzymes; where the one or more enzymes function in one or more engineered metabolic pathways to achieve: (1) conversion of a carbohydrate source to 1,2-propanediol, isopropanol, ethanol and/or glycerol; (2) conversion of a carbohydrate source to n-propanol and isopropanol; (3) conversion of a carbohydrate source to isopropanol and methanol; or (4) conversion of a carbohydrate source to propanediol and acetone; wherein the one or more native and/or heterologous enzymes is activated, upregulated or downregulated.

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

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
24 February 2011 (24.02.2011)(10) International Publication Number
WO 2011/022651 A1

(51) International Patent Classification:

C12P 7/00 (2006.01) C12P 7/20 (2006.01)
C12P 7/08 (2006.01) C12N 15/52 (2006.01)

(21) International Application Number:

PCT/US2010/046172

(22) International Filing Date:

20 August 2010 (20.08.2010)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

61/235,959 21 August 2009 (21.08.2009) US
61/298,790 27 January 2010 (27.01.2010) US

(72) Inventors; and

(71) Applicants : **MCBRIDE, John** [US/US]; 67 Etna Road, Suite 300, Lebanon, New Jersey 03766 (US). **RAJ-GARHIA, Vineet** [IN/US]; 67 Etna Road, Suite 300, Lebanon, New Hampshire 03766 (US). **SHAW, Arthur J., IV** [US/US]; 67 Etna Road, Suite 300, Lebanon, New Hampshire 03766 (US). **TRIPATHI, Shital, A.** [IN/US]; 67 Etna Road, Suite 300, Lebanon, New Hampshire

03766 (US). **BREVENOVA, Elena** [RU/US]; 67 Etna Road, Suite 300, Lebanon, New Hampshire 03766 (US). **CAIAZZA, Nicky** [US/US]; 67 Etna Road, Suite 300, Lebanon, New Hampshire 03766 (US). **VAN DIJKEN, Hans** [NL/NL]; Westfrankelandsedijk 1, NL-3115 HG Schiedam (NL). **FROEHLICH, Allan** [US/US]; 67 Etna Road, Suite 300, Lebanon, New Jersey 03766 (US). **SILLERS, William, Ryan** [US/US]; 67 Etna Road, Suite 300, Lebanon, New Hampshire 03766 (US). **FLATT, Jim** [US/US]; 67 Etna Road, Suite 300, Lebanon, New Jersey 03766 (US).

(74) Agents: **JACKMAN, Peter, A.** et al.; 1100 New York Avenue N.W., Washington, District of Columbia 20005 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI,

[Continued on next page]

(54) Title: PRODUCTION OF PROPANOLS, ALCOHOLS, AND POLYOLS IN CONSOLIDATED BIOPROCESSING ORGANISMS

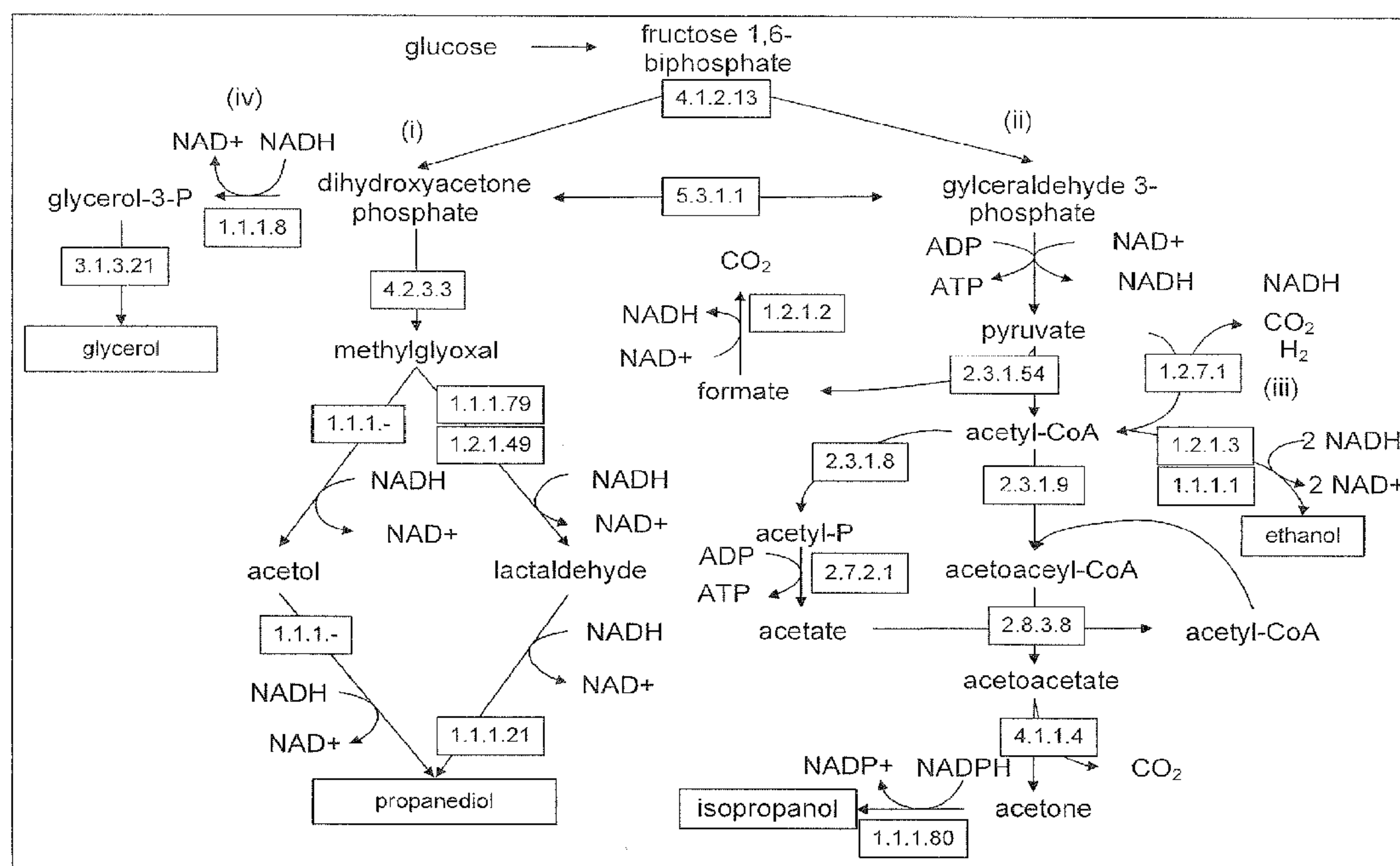


FIG. 1

[Continued on next page]

WO 2011/022651 A1

NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

(84) Designated States (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK,

- *with international search report (Art. 21(3))*
- *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*
- *with sequence listing part of description (Rule 5.2(a))*

(57) Abstract: The present invention provides for novel metabolic pathways leading to propanol, alcohol or polyol formation in a consolidated bioprocessing system (CBP), where lignocellulosic biomass is efficiently converted to such products. More specifically, the invention provides for a recombinant microorganism, where the microorganism expresses one or more native and/or heterologous enzymes; where the one or more enzymes function in one or more engineered metabolic pathways to achieve: (1) conversion of a carbohydrate source to 1,2-propanediol, isopropanol, ethanol and/or glycerol; (2) conversion of a carbohydrate source to n-propanol and isopropanol; (3) conversion of a carbohydrate source to isopropanol and methanol; or (4) conversion of a carbohydrate source to propanediol and acetone; wherein the one or more native and/or heterologous enzymes is activated, up-regulated or downregulated.

PRODUCTION OF PROPANOLS, ALCOHOLS, AND POLYOLS IN CONSOLIDATED BIOPROCESSING ORGANISMS

BACKGROUND OF THE INVENTION

Field of the Invention

[0001] Consolidated Bio-Processing (CBP) in essence describes a mode of operation where biocatalysts produce enzymes that can breakdown inexpensive cellulose into usable sugars and then simultaneously ferment them into value added products in a single vessel. CBP, which reduces the number of unit processes, significantly lowers operating and capital costs associated with cellulosic biofuel production. Furthermore, CBP processes reduce or eliminate the need for externally-added, expensive cellulases. *See Lynd et al. "Microbial cellulose utilization: Fundamentals and biotechnology," Microbiology and Molecular Biology Reviews 66(3):506-577 (2002); Lynd et al., "Consolidated bioprocessing of cellulosic biomass: An update," Current Opinion in Biotechnology 16(5):577-583 (2005); "Breaking the Biological Barriers to Cellulosic Ethanol: A Joint Research Agenda," December 2005, Rockville, Maryland Publication Date: June 2006; DOE/SC-0095. CBP is widely considered to be the "Ultimate low-cost configuration for cellulose hydrolysis and fermentation." DOE/USA Joint Research Agenda. See DOE/SC-0095 Joint Research Agenda. CBP on plant biomass, e.g., lignocellulosic biomass, also reduces the need to rely on petrochemical feedstocks to produce fermentable, value added products, such as propanols, alcohols, and polyols.*

[0002] Among forms of plant biomass, lignocellulosic biomass ("biomass") is particularly well-suited for producing fermentable, value added products because of its large-scale availability, low cost, and environmentally benign production. The primary obstacle impeding the processing of biomass feedstocks is the general absence of low-cost technology for overcoming the recalcitrance of these materials to conversion into useful products. Lignocellulosic biomass contains carbohydrate fractions (e.g., cellulose and hemicellulose) that can be converted into propanols, alcohols, and polyols. In order to convert these fractions, the cellulose and hemicellulose must ultimately be converted or

hydrolyzed into monosaccharides; it is the hydrolysis that has historically proven to be problematic.

[0003] Lignocellulosic feedstocks are recalcitrant to hydrolysis and subsequent release of sugars. Concentrated acid pre-treatment can release sugars with some associated loss of either pentose or hexose sugars. However, the larger issue with concentrated acid use is the additional capital cost associated with those pre-treatments. The capital cost implications involve using expensive materials of construction, handling corrosive chemicals and dealing with environmental implications. In fact, a group of eminent scholars in the area of lignocellulosic pretreatment have commented that although concentrated mineral acids are effective, they are too expensive to be practical when measured against the value of the resulting sugars. Mosier *et al.*, (2005), *Bioresource Technology* 96, 673 – 686.

[0004] More recently some companies have made technology claims where they have demonstrated concentrated acid recycle at laboratory scale as a means of reducing the cost associated with using concentrated acid pretreatments. A recent article on this recycling technology clarifies that they are only able to recycle 42% of the added acids and reiterates that this technology will only be tested in a pilot facility in the second half of 2010. Technology Review, Wednesday, June 10, 2009 (available at <http://www.technologyreview.com/energy/22774/>). Additionally, the article includes caution by industry experts against the use of concentrated HCl acids for pretreatment as the plant would require expensive materials of construction. CBP provides a viable alternative to the production of fermentable sugars from biomass.

[0005] CBP biomass processing schemes involving enzymatic or microbial hydrolysis commonly involve four biologically mediated transformations: (1) the production of saccharolytic enzymes (cellulases and hemicellulases); (2) the hydrolysis of carbohydrate components present in pretreated biomass to sugars; (3) the fermentation of hexose sugars (*e.g.*, glucose, mannose, and galactose); and (4) the fermentation of pentose sugars (*e.g.*, xylose and arabinose). These four transformations occur in a single step in CBP, which is distinguished from other less highly integrated configurations in that it does not involve a dedicated process step for cellulase and/or hemicellulase production.

[0006] Thus, CBP offers the potential for lower cost and higher efficiency than processes featuring dedicated cellulase production. The benefits result in part from avoided capital

costs, substrate and other raw materials, and utilities associated with cellulase production. In addition, several factors support the realization of higher rates of hydrolysis, and hence reduced reactor volume and capital investment using CBP, including enzyme-microbe synergy and the use of thermophilic organisms and/or complexed cellulase systems. Moreover, cellulose-adherent cellulolytic microorganisms are likely to compete successfully for products of cellulose hydrolysis with non-adhered microbes, *e.g.*, contaminants, which could increase the stability of industrial processes based on microbial cellulose utilization. Progress in developing CBP-enabling microorganisms is being made through two strategies: engineering naturally occurring cellulolytic microorganisms to improve product-related properties, such as yield and titer; and engineering non-cellulolytic organisms that exhibit high product yields and titers to express a heterologous cellulase and hemicellulase system enabling cellulose and hemicellulose utilization.

[0007] Many bacteria have the ability to ferment simple hexose sugars into a mixture of acidic and pH-neutral products via the process of glycolysis. The glycolytic pathway is abundant and comprises a series of enzymatic steps whereby a six carbon glucose molecule is broken down, via multiple intermediates, into two molecules of the three carbon compounds dihydroxyacetone phosphate and glyceraldehyde 3-phosphate. This process results in the net generation of ATP (biological energy supply) and the reduced cofactor NADH. From these three carbon compounds, a number of downstream value-added products can be made using the metabolic machinery of the CBP organisms, including, *e.g.*, propanols, alcohols, and polyols.

[0008] Industrial chemicals, such as propanols, alcohols, and polyols, are traditionally derived from petrochemical feedstocks. Production of such chemicals from petrochemical feedstocks, however, has its problems, not least of which is the use of a non-renewable resource that is subject to price fluctuations and heavy regulation. Thus, there is a need in the art for the production of propanols, alcohols, and polyols from resources that allow for large-scale availability, low cost, and environmentally benign production, all of which are advantages of CBP. In particular, there is a need for engineered organisms capable of converting biomass into propanols, alcohols, and polyols as part of a CBP system.

BRIEF SUMMARY OF THE INVENTION

- [0009] The present invention provides for novel metabolic pathways leading to propanol, alcohol or polyol formation in a consolidated bioprocessing system (CBP), where lignocellulosic biomass is efficiently converted to such products.
- [0010] The invention therefore provides for a recombinant microorganism, where the microorganism expresses one or more native and/or heterologous enzymes; where the one or more enzymes function in one or more engineered metabolic pathways to achieve: (1) conversion of a carbohydrate source to 1,2-propanediol, isopropanol, ethanol and/or glycerol; (2) conversion of a carbohydrate source to n-propanol and isopropanol; (3) conversion of a carbohydrate source to isopropanol and methanol; or (4) conversion of a carbohydrate source to propanediol and acetone.
- [0011] The engineered metabolic pathways of the invention are outlined in Figures 1-5 and 7. The enzymes that function at various steps along the pathways are identified in Tables 2-5. The engineered metabolic pathways of the invention are utilized to achieve high theoretical yields of products, particularly 1,2-propanediol, isopropanol, n-propanol, and methanol in bacteria and yeast.

BRIEF DESCRIPTION OF THE DRAWINGS/FIGURES

- [0012] Figure 1 shows a schematic of theoretical metabolic pathways for the production of mixed alcohols in bacterial and yeast CBP platforms. Yeast-specific branch pathways are depicted by EC numbers in dark gray boxes. Bacteria-specific branch pathways are depicted by EC numbers in light gray boxes.
- [0013] Figure 2 shows a schematic of theoretical metabolic pathways for the production of n-propanol and isopropanol in bacterial CBP platforms.
- [0014] Figure 3 shows a schematic of theoretical metabolic pathways for the production of isopropanol and methanol in bacterial CBP platforms.
- [0015] Figure 4 shows a schematic of theoretical metabolic pathways for the anaerobic production of propanediol and acetone in bacterial and yeast CBP platforms. A yeast-specific branch pathway is depicted by EC numbers in dark gray boxes. A bacteria-specific branch pathway is depicted by the EC number in a light gray box.

- [0016] Figure 5 shows a schematic of a theoretical metabolic pathway for the aerobic production of propanediol in yeast CBP platforms.
- [0017] Figure 6 shows a schematic of propanediol utilization (pdu) gene organization in *T. saccharolyticum*.
- [0018] Figure 7 shows a schematic of theoretical metabolic pathways for the production of n-propanol and isopropanol in bacterial and yeast CBP platforms.

DETAILED DESCRIPTION OF THE INVENTION

Metabolic Pathway Engineering

- [0019] Many bacteria have the ability to ferment simple hexose sugars into a mixture of acidic and pH-neutral products via the process of glycolysis. The glycolytic pathway is abundant and comprises a series of enzymatic steps whereby a six carbon glucose molecule is broken down, via multiple intermediates, into two molecules of the three carbon compound pyruvate. This process results in the net generation of ATP (biological energy supply) and the reduced cofactor NADH.
- [0020] Pyruvate is an important intermediary compound of metabolism. For example, under aerobic conditions pyruvate may be oxidized to acetyl coenzyme A (acetyl CoA), which then enters the tricarboxylic acid cycle (TCA), which in turn generates synthetic precursors, CO₂ and reduced cofactors. The cofactors are then oxidized by donating hydrogen equivalents, via a series of enzymatic steps, to oxygen resulting in the formation of water and ATP. This process of energy formation is known as oxidative phosphorylation.
- [0021] Under anaerobic conditions (no available oxygen), fermentation occurs in which the degradation products of organic compounds serve as hydrogen donors and acceptors. Excess NADH from glycolysis is oxidized in reactions involving the reduction of organic substrates to products, such as lactate and ethanol. In addition, ATP is regenerated from the production of organic acids, such as acetate, in a process known as substrate level phosphorylation. Therefore, the fermentation products of glycolysis and pyruvate metabolism include a variety of organic acids, alcohols and CO₂.
- [0022] Most facultative anaerobes metabolize pyruvate aerobically via pyruvate dehydrogenase (PDH) and the tricarboxylic acid cycle (TCA). Under anaerobic conditions, the main energy pathway for the metabolism of pyruvate is via pyruvate-

formate-lyase (PFL) pathway to give formate and acetyl-CoA. Acetyl-CoA is then converted to acetate, via phosphotransacetylase (PTA) and acetate kinase (ACK) with the co-production of ATP, or reduced to ethanol via acetaldehyde dehydrogenase (AcDH) and alcohol dehydrogenase (ADH). In order to maintain a balance of reducing equivalents, excess NADH produced from glycolysis is re-oxidized to NAD⁺ by lactate dehydrogenase (LDH) during the reduction of pyruvate to lactate. NADH can also be re-oxidized by AcDH and ADH during the reduction of acetyl-CoA to ethanol, but this is a minor reaction in cells with a functional LDH.

[0023] Ethanologenic organisms, including yeast (*e.g.*, *Saccharomyces cerevisiae*), are capable of a second type of anaerobic fermentation, commonly referred to as alcoholic fermentation, in which pyruvate is metabolized to acetaldehyde and CO₂ by pyruvate decarboxylase (PDC). Acetaldehyde is then reduced to ethanol by ADH regenerating NAD⁺. Alcoholic fermentation results in the metabolism of one molecule of glucose to two molecules of ethanol and two molecules of CO₂.

[0024] The present invention is directed to the modification of traditional glycolytic pathways in bacteria and yeast, as described above, to engineer novel metabolic pathways capable of generating or increasing the yield of certain products that could not otherwise be generated by the native organism. Such products include n-propanol or isopropanol along with alcohols, propanediol, ethanol, and glycerol.

[0025] In particular embodiments, the present invention is directed to the production of mixed alcohols in CBP yeast and bacterial platforms. In other embodiments, the present invention is directed to the production of n-propanol and isopropanol in a CBP bacterial platform. In additional embodiments, the present invention is directed to production of isopropanol and methanol in a CBP bacterial platform. In certain other embodiments, the present invention is directed to the production of propanediol in a CBP yeast or bacterial platform. In further embodiments, the propanediol could be directly utilized in industrial applications or condensed to propylene or converted via a chemical or microbial based biocatalysis to propanol.

[0026] The present invention is directed to the engineering of such alternative metabolic pathways in various microorganisms, including bacteria and yeast. The term "microorganism," as used herein, refers to an organism of microscopic or submicroscopic

size that can be seen only with the aid of a microscope and that typically consists of only a single cell. Microorganisms include bacteria, protozoans, and certain algae and fungi.

[0027] In certain embodiments, the bacterial microorganism is a species of the genera *Thermoanaerobacterium*, *Thermoanaerobacter*, *Clostridium*, *Geobacillus*, *Saccharococcus*, *Paenibacillus*, *Bacillus*, *Caldicellulosiruptor*, *Anaerocellum*, or *Anoxybacillus*. In certain embodiments, the microorganism is a bacterium selected from the group consisting of: *Thermoanaerobacterium thermosulfurigenes*, *Thermoanaerobacterium aotearoense*, *Thermoanaerobacterium polysaccharolyticum*, *Thermoanaerobacterium zeae*, *Thermoanaerobacterium xylanolyticum*, *Thermoanaerobacterium saccharolyticum*, *Thermoanaerobium brockii*, *Thermoanaerobacterium thermosaccharolyticum*, *Thermoanaerobacter thermohydrosulfuricus*, *Thermoanaerobacter ethanolicus*, *Thermoanaerobacter brocki*, *Clostridium thermocellum*, *Clostridium cellulolyticum*, *Clostridium phytofermentans*, *Clostridium straminosolvens*, *Geobacillus thermoglucosidasius*, *Geobacillus stearothermophilus*, *Saccharococcus caldoxylosilyticus*, *Saccharococcus thermophilus*, *Paenibacillus campinasensis*, *Bacillus flavothermus*, *Anoxybacillus kamchatkensis*, *Anoxybacillus gonensis*, *Caldicellulosiruptor acetigenus*, *Caldicellulosiruptor saccharolyticus*, *Caldicellulosiruptor kristjanssonii*, *Caldicellulosiruptor owensensis*, *Caldicellulosiruptor lactoaceticus*, and *Anaerocellum thermophilum*. In particular embodiments, the microorganism is *Clostridium thermocellum* or *Thermoanaerobacterium saccharolyticum*.

[0028] In certain other embodiments, the yeast microorganism is selected from the group consisting of *Saccharomyces cerevisiae*, *Kluyveromyces lactis*, *Kluyveromyces marxianus*, *Pichia pastoris*, *Yarrowia lipolytica*, *Hansenula polymorpha*, *Phaffia rhodozyma*, *Candida utliis*, *Arxula adeninivorans*, *Pichia stipitis*, *Debaryomyces hansenii*, *Debaryomyces polymorphus*, *Schizosaccharomyces pombe*, *Candida albicans*, and *Schwanniomyces occidentalis*. In particular embodiments, the yeast microorganism is *Saccharomyces cerevisiae*.

[0029] In certain instances, the microorganism of the invention is cellulolytic. The term "cellulolytic" means able to hydrolyze glycosidic linkages in oligohexoses and polyhexoses. Cellulolytic activity can also include the ability to depolymerize or debranch cellulose and hemicellulose.

- [0030] The term "ethanogenic" is intended to include the ability of a microorganism to produce ethanol from a carbohydrate as a fermentation product. The term is intended to include, but is not limited to, naturally occurring ethanogenic organisms, ethanogenic organisms with naturally occurring or induced mutations, and ethanogenic organisms which have been genetically modified.
- [0031] The terms "fermenting" and "fermentation" are intended to include the enzymatic process (*e.g.*, cellular or acellular, *e.g.*, a lysate or purified polypeptide mixture) by which ethanol is produced from a carbohydrate, in particular, as a product of fermentation.
- [0032] By "thermophilic" is meant an organism that thrives at a temperature of about 45°C or higher.
- [0033] By "mesophilic" is meant an organism that thrives at a temperature from about 20- about 45°C.
- [0034] The term "CBP organism" is intended to include microorganisms of the invention, *e.g.*, microorganisms that have properties suitable for CBP.
- [0035] In certain embodiments of the invention, one or more metabolic engineered pathways are utilized for the combined production of propanediol and isopropanol from glucose. The metabolic pathways and the various distinct enzymes (Table 2) required for the combined production of propanediol and isopropanol are shown in Figure 1 and described further below in Example 1. These metabolic pathways can be subdivided into the following distinct production routes: i) the conversion of dihydroxyacetone phosphate into propanediol; ii) the conversion of pyruvate into isopropanol; iii) the conversion of pyruvate into ethanol (bacterial platform only); and iv) the conversion of dihydroxyacetone phosphate into glycerol (yeast platform only).
- [0036] In certain other embodiments of the invention, one or more metabolic engineered pathways are utilized for the production of n-propanol and isopropanol. The metabolic pathways and the various distinct enzymes (Table 3) required for the production of n-propanol and isopropanol are shown in Figure 2 and described further below in Example 2. The metabolic pathways can be subdivided into two distinct production routes: i) the conversion of dihydroxyacetone phosphate into n-propanol; and ii) the conversion of pyruvate into isopropanol.
- [0037] In additional embodiments of the invention, one or more metabolic engineered pathways are utilized for the combined production of isopropanol and methanol from

carbohydrates. The metabolic pathways and the various distinct enzymes (Table 4) required for the production of isopropanol and methanol are shown in Figure 3 and described further below in Example 3. The metabolic pathways can be subdivided into distinct production routes: i) the conversion of pyruvate into isopropanol; and ii) the conversion of formate into CO₂ and methanol.

[0038] In other embodiments of the invention, one or more metabolic engineered pathways are utilized for the co-production of propanediol and acetone from hexose and pentose sugars in thermophilic clostridia and yeast, such as *S. cerevisiae*. The metabolic pathways and the various distinct enzymes (Table 5) required for the production of propanediol and acetone are shown in Figures 4 and 5 and described further below in Examples 4 and 5. The metabolic pathways can be subdivided as follows: i) the production of dihydroxyacetone phosphate and glyceraldehydes-3-phosphate from glucose; ii) the subsequent generation of propanediol from dihydroxyacetone phosphate; and iii) the generation of acetone from glyceraldehyde-3-phosphate.

[0039] A summary of the pathways of the present invention is provided in Table 1 as follows:

Table 1. Summary: Engineering of CBP biocatalysts for production of propanol

Pathway	Theoretical hexose/pentose	yield	Required Metabolic Engineering					Critical Step
			Gene EC#	Gene KO function	Gene EC#	Gene expression	Gene expression function	
Mixed Alcohol-Bacterial 2 glucose → 1,2-propanediol + isopropanol + ethanol + 4 CO ₂ + H ⁺ + 3 ATP	Hexose: 0.21 g/g propanediol 0.17 g/g propanol 0.13 g/g ethanol Pentose: 0.21 g/g propanediol 0.17 g/g propanol 0.13 g/g ethanol		1.1.1.27	<i>ldh</i>	4.2.3.3 1.1.1.- 2.3.1.9 2.8.1.8 4.1.1.4 1.1.1.80 1.1.1.202		<i>methylglyoxal synthase</i> <i>aldo-keto reductase</i> <i>thiolase</i> <i>acetyl-CoA transferase</i> <i>acetoacetate decarboxylase</i> <i>Isoprop dehyd oxidoreductase</i>	pathway flux control, methylglyoxal to 1,2 propanediol
			4.1.1.1	<i>pdh1</i> <i>pdh5</i> <i>pdh6</i>	4.2.3.3 1.1.1.6 2.3.1.54 2.7.2.1 2.3.1.8 2.8.3.8 4.1.1.4 1.1.1.79 1.1.1.80 1.2.1.49		<i>methylglyoxal synthase</i> <i>glycerol dehydrogenase</i> <i>pyruvate formate-lyase</i> <i>acetate kinase</i> <i>phosphate acetyltransferase</i> <i>acetate CoA-transferase</i> <i>acetoacetate decarboxylase</i> <i>glyoxylate reductase</i> <i>isopropanol dehydrogenase</i> <i>methylglyoxal dehydrogenase</i>	pathway flux control (without TPI deletion), methylglyoxal to propanol
Mixed Alcohol –Yeast 2 glucose → isopropanol + propanediol+ glycerol + 3 CO ₂ + 1 ATP	Hexose: 0.21 g/g propanediol 0.17 g/g isopropanol 0.26 g/g glycerol Pentose: (no <i>tpi</i> deletion) 0.21g/g propanediol 0.17 g/g isopropanol 0.26 g/g glycerol							

n-propanol/isopropanol in bacteria	Hexose: 0.44 g/g propanols Pentose: 0.44 g/g propanols	1.1.1.27 1.12.7.2 1.2.1.10	<i>ldh</i> <i>hyd</i> <i>acd</i>	4.2.3.3 1.1.1.- 2.3.1.9 2.8.1.8 4.1.1.4 1.1.1.80 1.1.1.202	<i>mgs</i> <i>aldo-keto reductase</i> <i>thiolase</i> <i>acetyl-CoA trans</i> <i>adc</i> <i>Isoprop dehyd</i> <i>oxidoreductase</i>	methylglyoxal to n-propanol
Isopropanol/Methanol –Bacterial 2 glucose → 2 isopropanol + 2 methanol + 4 CO ₂ + 4 ATP	Hexose: 0.33 g/g propanol 0.18 g/g methanol Pentose: 0.33 g/g propanol 0.18 g/g ethanol	1.1.1.27 1.2.1.10 4.2.3.3 1.2.7.1	<i>ldh</i> <i>Acdh</i> <i>mcs</i> <i>pfor</i>	1.2.1.2 1.2.1.46 1.1.1.- 2.8.3.8 4.1.1.4 2.3.1.9	<i>formate dehyd</i> <i>formaldehyde dehyd</i> <i>methanol dehyd</i> <i>(oxidoreductase)</i> <i>CoA transferase</i> <i>acetoacetate decarb</i> <i>thiolase</i>	formate to methanol
1,2-Propanediol/Acetone –Bacterial 2 glucose → 2 propanediol + acetone + 3 CO ₂ + 1 ATP	Hexose: 0.42 g/g propanediol 0.16 g/g acetone Pentose: 0.42 g/g propanediol 0.16 g/g acetone	1.1.1.1 1.1.1.27 1.12.7.2	<i>adh</i> <i>ldh</i> <i>hyd</i>	2.3.1.9 2.8.3.8 4.1.1.4	<i>thiolase</i> <i>acetate CoA-transferase</i> <i>acetoacetate decarboxylase</i>	methylglyoxal to propanediol
propanediol/Acetone –Yeast 2 glucose → 2 propanediol + acetone + 3 CO ₂ + 1 ATP	Hexose: 0.42 g/g propanediol 0.16 g/g acetone Pentose: (no <i>tpi</i> deletion) 0.42 g/g propanediol 0.16 g/g acetone	4.1.1.1 1.1.1.8 3.1.3.21	<i>pdh</i> <i>pdh</i> <i>pdh</i> <i>gpd</i> <i>gpd</i>	4.2.3.3 2.3.1.54 2.7.2.1 2.3.1.8 2.8.3.8 4.1.1.4 1.1.1.79 1.2.1.49	<i>methylglyoxal synthase</i> <i>pyruvate formate-lyase</i> <i>acetate kinase</i> <i>phosphate acetyltransferase</i> <i>acetate CoA-transferase</i> <i>acetoacetate decarboxylase</i> <i>glyoxylate reductase</i>	pathway flux control (without TPI deletion), methylglyoxal to propanediol

propanediol Aerobic-Yeast 17 glucose + 6 O ₂ → 24 propanediol+ 6 H ₂ O + 30 CO ₂ + 7 ATP	Hexose: 0.42 to 0.61 g/g propanediol Pentose: 100% xylose could not be converted via this pathway, but glucose/xylose mixtures could, with yield similar to glucose alone	4.1.1.1 1.1.1.8 3.1.3.21	<i>pdc1</i> <i>pdc5</i> <i>pdc6</i> <i>gpd2</i> <i>gpp1</i>	4.2.3.3 1.1.1.6 1.1.1.79 1.2.1.49	<i>methylglyoxal synthase</i> <i>glycerol dehydrogenase</i> <i>glyoxylate reductase</i> <i>methylglyoxal dehydrogenase</i>	<i>methylglyoxal dehydrogenase</i>	pathway flux control from glucose 6-phosphate to PPP and propanediol, methylglyoxal to propanediol		

Metabolic Enzymes

- [0100] As described above, the engineering of metabolic pathways in microorganisms requires certain enzymes to function at particular steps along the pathways, as shown in Figs. 1-5.
- [0101] The enzymes of the invention as described herein can be endogenous to the native strain of the microorganism, and can thus be understood to be referred to as "native" or "endogenous." An organism is in "a native state" if it has not been genetically engineered or otherwise manipulated by the hand of man in a manner that intentionally alters the genetic and/or phenotypic constitution of the organism. For example, wild-type organisms can be considered to be in a native state.
- [0102] For example, in certain embodiments, when the host cell is a particular *Thermoanaerobacter(ium)* strain, one or more metabolic enzymes can be an enzyme derived from that same *Thermoanaerobacter(ium)* strain. Source libraries with fragments of whole genomic DNA from such a *Thermoanaerobacter(ium)* strain can be host-modified with promoters, terminators, replication origins, or homologous recombination targeting. Screening of these libraries can identify DNA encoding for enzymes of interest that function in one or more metabolic engineered pathways of the invention.
- [0103] In other embodiments, the enzymes of the invention can be non-native or "heterologous" to the organism, and can be introduced into the organism on a vector by transformation or other methods known to one of ordinary skill in the art, as described further below.
- [0104] The terms "activity," "activities," "enzymatic activity," and "enzymatic activities" are used interchangeably and are intended to include any functional activity normally attributed to a selected polypeptide. Typically, the activity of a selected polypeptide encompasses the total enzymatic activity associated with the produced polypeptide. The polypeptide produced by a host cell and having enzymatic activity can be located in the intracellular space of the cell, cell-associated, secreted into the extracellular milieu, or a combination thereof.
- [0105] In certain embodiments, enzymes that function in the metabolic pathways of the invention are set forth below in Tables 2-5 and include the following: methylglyoxal synthase, aldo-keto reductase, glyoxylate reductase, methylglyoxal dehydrogenase,

aldehyde reductase, pyruvate formate lyase, thiolase, CoA transferase, acetoacetate decarboxylase, isopropanol, aldehyde dehydrogenase, alcohol dehydrogenase, diol-hydrolase, dehydrogenase, phosphotransacetylase, oxidoreductase, formate dehydrogenase, formaldehyde dehydrogenase and methanol dehydrogenase.

[0106] As used herein, the term "methylglyoxal synthase" or "mgs" refers to an enzyme that catalyzes the chemical reaction $\text{glycerone phosphate} \rightleftharpoons \text{methylglyoxal} + \text{phosphate}$

[0107] As used herein, the term "aldo-keto reductase" can refer to any number of related monomeric NADPH-dependent oxidoreductases, such as aldose reductase, prostaglandin F synthase, xylose reductase, and many others.

[0108] As used herein, the term "oxidoreductase" refers to an enzyme that catalyzes the transfer of electrons from one molecule (the reductant, also called the hydrogen or electron donor) to another (the oxidant, also called the hydrogen or electron acceptor).

[0109] As used herein, the term "glyoxylate reductase" refers to an enzyme that catalyzes the chemical reaction $\text{glycolate} + \text{NAD}^+ \rightleftharpoons \text{glyoxylate} + \text{NADH} + \text{H}^+$. This enzyme belongs to the family of oxidoreductases, specifically those acting on the CH-OH group of donor with NAD⁺ or NADP⁺ as acceptor.

[0110] As used herein, the term "methylglyoxal dehydrogenase" refers to an enzyme that oxidizes methylglyoxal to pyruvate.

[0111] As used herein, the term "CoA transferase" is an enzyme, for example, such as acetyl CoA transferase that catalyzes the chemical reaction $\text{acyl-CoA} + \text{acetate} \rightleftharpoons \text{a fatty acid anion} + \text{acetyl-CoA}$. The term "CoA transferase" also refers an enzyme that catalyzes the chemical reaction $\text{acetoacetyl-CoA} + \text{acetate} \rightleftharpoons \text{acetoacetate} + \text{acetyl-CoA}$.

[0112] As used herein, the term "acetoacetate decarboxylase" or "ADC" refers to an enzyme involved in both the ketone body production pathway in humans and other mammals, and solventogenesis in certain bacteria. Its reaction involves a decarboxylation of acetoacetate, forming acetone and carbon dioxide.

[0113] As used herein, the term "aldehyde dehydrogenase" refers to an enzyme that catalyzes the oxidation (dehydrogenation) of aldehydes.

[0114] As used herein, the term "dehydrogenase" refers to an enzyme that oxidizes a substrate by transferring one or more hydrides (H⁻) to an acceptor, usually NAD⁺/NADP⁺.

- [0115] As used herein, the term "formate dehydrogenase" is an enzyme that catalyzes the oxidation of formate to bicarbonate or carbon dioxide, donating the electrons to a second substrate, such as NAD^+ in formate: NAD^+ oxidoreductase.
- [0116] As used herein, the term "formaldehyde dehydrogenase" refers to an enzyme that catalyzes the chemical reaction $\text{formaldehyde} + \text{NAD}^+ + \text{H}_2\text{O} \rightleftharpoons \text{formate} + \text{NADH} + 2 \text{H}^+$. This enzyme belongs to the family of oxidoreductases, specifically those acting on the aldehyde or oxo group of donor with NAD^+ or NADP^+ as acceptor.
- [0117] As used herein, the term "methanol dehydrogenase" is an enzyme that catalyzes the chemical reaction $\text{methanol} + \text{NAD}^+ \rightleftharpoons \text{formaldehyde} + \text{NADH} + \text{H}^+$. This enzyme also belongs to the family of oxidoreductases, specifically those acting on the aldehyde or oxo group of donor with NAD^+ or NADP^+ as acceptor.
- [0118] As used herein, the term "pyruvate formate lyase" or "PFL" is intended to include the enzyme capable of converting pyruvate into Acetyl CoA and formate.
- [0119] As used herein the term "alcohol dehydrogenase" or "ADH" is intended to include the enzyme capable of converting aldehydes, such as acetaldehyde and propionaldehyde, and ketones, such as acetone, into an alcohol, such as ethanol, n-propanol, or isopropanol.
- [0120] As used herein, the term "phosphotransacetylase" or "PTA" is intended to include the enzyme capable of converting Acetyl CoA into acetyl phosphate.
- [0121] As used herein, the term "diol dehydratase" is intended to include the enzyme capable of converting propanediol to propanal.
- [0122] The term "upregulated" means increased in activity, *e.g.*, increase in enzymatic activity of the enzyme as compared to activity in a native host.
- [0123] The term "downregulated" means decreased in activity, *e.g.*, decrease in enzymatic activity of the enzyme as compared to activity in a native host.
- [0124] The term "activated" means expressed or metabolically functional.
- [0125] The polypeptide sequences corresponding to certain of the enzymes of the present invention are as follows:

***C. thermocellum* proteins**

- [0126] EC 2.3.1.54 (Cthe0505; SEQ ID NO: 1)

MDAWRGFNKGNWCQEIDVRDFIIRNYTPYEGDESFLVGPTDRTRKLWEKVSELLK
KERENGGVLDVDTHITITSHKPGYIDKELEVIVGLQTDEPLKRAIMPFGGIRMVIKGAE

AYGHSVDPQVVEIFTKYRKTHNQGVYDVYTPEMRKAKKAGIITGLPDAYGRGRIIGDYR
 RVALYGVDRLIAEKEKEMASLERDYIDYETVRDREEISEQIKSLKQLKEMALSYGFDISC
 PAKDAREAFQWLYFAYLAAVKEQNGAAMSIGRISTFLDIYIERDLKEGKLTEELAQELV
 DQLVIKLRIVRFLRTPEYEKLFSGDPTWVTESIGGMALDGRTLVTKSSFRFLHTLFLNLGH
 APEPNLTVLWSVNLPEGFKKYCAKVSIHSSSIQYESDDIMRKHWGDDYGIACCVSAMRI
 GKQMQLFFGARCNLAKALLYAINGGKDEMTGEQIAPMFAPVETELYDYEDVMKRFDMV
 LDWVARLYMNTLNIIHYMHDKYAYEALQMALHDKDVFRTMACGIAGLSVVADSLSAI
 KYAKVKPIRNENNLVVDYEVEGDYPKFGNNDERVDEIAVQVVKMFMNKLRLKQRAYRS
 ATPTLSILTITSNVVYGKKTGNTPDGRKAGEPLAPGANPMHGRDINGALAVLNSIAKLPY
 EYAQDGISYTFISIIPKALGRDEETRINNLSMLDGYFKQGGHHINVNVFEKETLLDAMEH
 PEKYPQLTIRVSGYAVNFIKLTREQQLDVINRTIHGKI

[0127] EC 2.3.1.8 (Cthe1029; SEQ ID NO: 2)

VIIISYKYYKYSFYDNSFGIMKGEEFMSFLEQIIERAKSDVKTIVLPESTDLRVIKA
 ASMIMKKGIKVVVLIGNEKEIKSLAGDIDLEGVMIEDSLNSEKLEDYANTLYELRKSKGM
 TIEAARETIKDPLYYGVMVMVKKGEADGMVAGAVNSTANTLRPALQILKTAPGTKLVSSF
 FVMVVPNCEYGHNGTFVYADCGLVENPDADQLSEIAISASKSFEMLVGAKPQVAMLSY
 SSYGSAKSELTEKVIKATQLAKEKAPHLAIDGELQVDAAIVPEVAKSKAKGSSVAGKAN
 VLIFPDLDAGNIAKLTQRLAKAEAYGPITQGLARPVNDLSRGCSAEDIVGVAAITAVQA
 QYVKA

[0128] EC 2.7.2.1 (Cthe1028; SEQ ID NO: 3)

MNILVINTGSSSLKYQLIDMTNESVLAKGVCDRIGLEHSFLKHTKTGGETVVIEKD
 LYNHKLAIQEVISALTDEKIGVIKSMSEISAVGHRIVHGGEKFKESAIIDEDVMKAIRDCV
 ELAPLHNPSNIIGIEACKQILPDVPMVAVFDTAHFQTMPRHAYIYALPYEIEKYKLRKY
 GFHGTSHKYVAHRAAQMLGKPIESLKLITCHLGNGASICAVKGGKSVDTSMGFTPLQGL
 CMGTRSGNVDPNAVITYLMEKEKMNINDINNFLNKKSGVLGISGVSSDFRDVQDAAEKG
 DDRAQLALDIFCYGVRKYIGKYIAVLNGVDVAVVFTAGIGENNAIYIRREVLKDMDFFGIKI
 DLDKNEVKGKEADISAPDAKVKTLVIPTNEELEIARETLRLVKNL

[0129] EC 1.1.1.80 (Cthe101; SEQ ID NO:4)

MINFVYKNPTKIIFGRGTELKVGEEVRQYSGKVLLHYGGGSIKKTGLYDRVVNSL
 KQAGVEVVVELGGVMPNPRLGLVNEGIIKICREKGIDFILAVGGGSAIDSAKAIAVGVPYDG
 DVWDFFCGKAEPKEALPVGVVLTIPAAGSEASPNSVITREDGLYKRGMYSELIRPVFAIM
 NPELTYTLPAYQTACGTADIMAHIMERYFTNETHTDLTDRLCEATLKTMIKNVPIALEEP
 DNYNARAEIMWAGTIAHNGLLGTGRIEDWASHNIEHEISAIYDVAHGAGLAVVFPAPWM
 KYVYKNNLDRFVQFAVRVWNVEMNFDEPERTALEGIERLKKFFKEIGLPVSLKEMNIGD
 DRLEEMASKCTNGGKATIGNFVKLNREDVY NILKLAV

[0130] Cthe0394 (SEQ ID NO: 5)

MKAFNYYAPTEIIFGCGRVQEIGSITAQYGKKALLVTVPFPEVKELYEKVKKSLR
 ENGVEVVHFDGVIPNPTTDVVTEGANMAKAAGVDVVIGLGGGSSIDTAKAIAVEATHPG
 TAWDYNCHTPGPTSATLPILAIAGTTAGTGSQCTQCAVITKTSEKDKSAIWHKNIFPKVAIV
 DPEVTVTMPKSVTAQTGFDAFAHNFEAYLSVKTSPLVEMMAIEAIKMIKEYLPKALENP
 NDIEARSKMSLADTLGGLTNSNAGVTLPHGLGMQVGGHAPHVSHGQALAIYPQFTRYT
 YAWAIEKFAKVGRIFNPALNELSDEEAAKEACVAIDDFLKKIGLWIGFKDVNVTKEQIRE
 IADDGQVLGDYLNPRVATIDEMYELLMNCYERKE

[0131] Cthe0423 (SEQ ID NO:6)

MTKIANKYEVIDNVEKLEKALKRLREAQSVYATYTQEQQVDKIFFEAAMAANKM
 RIPLAKMAVEETGMGVVEDKVIKNHYASEYIYNAYKNTKTCGVIEEDPAFGIKKIAEPLG
 VIAAVIPTTNPTSTAIFKTLIALKTRNAIIISPHPRAKNSTIEAAKIVLEAAVKAGAPEGIIGW
 IDVPSLELTNLVMREADVILATGGPGLVKAAYSSGKPAIGVGAGNTPAIIDDSADIVLAV
 NSIIHSKTFDNGMICASEQSVIVLDGVYKEVKKEFEKRGCYFLNEDETEKVRKTIINGAL
 NAKIVGQKAHTIANLAGFEVPETTKILIGEVTSDISEEFAHEKLCVPLAMYRAKDFDDA
 LDKAERLVADGGFGHTSSLYIDTVTQKEKLQKFSERMKTCRILVNTPSSQGGGIGDLYNFK
 LAPSLTLGCGSWGGSVSDNVGVKHLNLIKTVARRENMLWFRTPEKIYIKRGCLPVAL
 DELKNVMGKKKAFIVTDNFLYNNGYTKPITDKLDEMGIHVHKTFFDVSPDPSLASAKAGA
 AEMLAFQPDTHAVGGGSAMDAAKIMWVMYEHPEVDFMDMAMRFMDIRKRVYTFPK
 MGQKAYFIAIPTSAGTGSEVTPFAVITDEKTGIKYPLADYELLPDMAIVDADMMMNA PK
 GLTAASGIDALTHALEAYVSMLATDYTDSLALRAIKMIFEYLPRAYENGASDPVAREKM
 ANAATIAGMAFANAFLGVCHSMAHKLGAIFYHLPHGVANALMINEVIRFNSSEAPT KMG

TFPQYDHPRTLERYAEIADYIGLKGKNNEEKVENLIKAIDELKEKVGIRKTIKDYDIDEKE
FLDRLDEMVEQAFDDQCTGTNPRYPLMNEIRQMYLNAYYGGAKK

[0132] Cthe2445 (SEQ ID NO:7)

MKGKMKVCVLTGKEKLEWVERDIPQPGRGELQIKLKHVGVCGSDLHFYKEGRL
ANWELDGPLALGHEPGGIVSAIGEGVEGFEIGDKVALEPGVPCGECEDCRKGHYNLCKH
IKFMAIPHEKDGVF AEYCVHSASMCYKLPENVDTMEGGLMEPLSVALHATELSNAKIGE
TAIVLGSGCIGLCTVMALKARGVSEIYVTDVVDKRLEKALEVGATR VFNSQREDIVEFA
KTLPGGGADQVYECAGSRVTTLQTCKLIK RAGKVTLVGV SPEPVLELDIATLNAMEGTV
YSVYRYRNMYP IAAAVSSGVIPLKKIVSHVDFDKDCIEAIEYSTN HKDEVIKSVIKF

[0133] Cthe2579 (SEQ ID NO:8)

MNFKFKIGTKVFFGKECVKENKAVFKDFRKRALLVTGKNSAKASGAFSDVVEVL
EEYGIDYEIYDRVANNPSLENVKEGG EAARKFDADFIIGIGGGSPLDASKAVAVLATNDI
EPVDLYKNVFENKPLPIIAIPTTAGTGSEVTPYSILTRDDMKTKKSFGNEDTFPAVAFIDA
RYTESMSYETTVDTALDAFTHALEGYLGRRSTPVSDILAVEAIRIFGECLNLLNNKFDY
DVREKLLYMSMLGGMVISHTGTTIIHGMGYSLTYFKDIPHGRANGMLVREYLKYN YEA
AKEKTDNVLRLLLKVPSIDAFGEIIDRLIPQKPVLTK EEIELYASLAMKQNSTLSNARTVVK
EDMEEIFKNTFGKG

[0134] EC 4.2.2.3 (Cthe0095; SEQ ID NO:9)

MNIALIAHDKKKELMASFCIAYRSILKNHTLFATGTTGAIIVEATGLNVHRFLPGV
MGEQQISARAAYNELDLVIFFRDPISAKSDEPDIHSLRECDINNIPFATNLGTAEMLIKGL
ERGDLDWRELIKK

[0135] EC 1.1.1.- (Cthe0152; SEQ ID NO:10)

LKYCKLGNTGLEVSKLCFGGLIIGPLQANLPPETGAEIILKS FELGVNFIDTAELYG
TYSHIGKALKKTNKNIVVATKSYAYS AEGAKESLEKARKEMDIDVIDIFMLHEQESRLTL
KGHREALEY YISMKEKGIIKAVGVSTHNVEVVEACCEMPEVDVIHP IVNKAGIGIGDGTI
DDMLKAVEKAYS VGKGIYSMKPLGGGNLIKS YKEAMDFVLNIPYIHSIAVGMQSIEEVV
MNVCFEGKEVPQDVQKSLENKKRHLHIDWWCEGCGKCVERCKQKALKLVDGKAKVE
EEKCVLCSYCASVCPVFAIKVS

[0136] Cthe0236 (SEQ ID NO:11)

MQYRGLGKTGVKVSALGFGAMRLPQININGNTRVDEEKSIEMIHRAFELGVNYID
TAPGYCNGESEVVGKALKGWRDKIYLSTKNPIENASGDDWRKRENSLKKLDTDYID
FYHMGWGINWETYETKIDVKGGPLEAARKAKEEGLIRHISFSFHDKPENLIKLDITGNFET
VLCQYNLLDRSNEKAIAHAKRKGLGVIIMGPVGGGKLGEPSETIKKLLPKKTVSCAEIAL
RFVLANPNVDCALSGMSTIEMVEENVRVASNDTPLTKEELEMIRASMEENKRMEDLYC
TGCNYCMPCPVGVNIPLNFQLMNYHRVYKITDYARGQYSQIGKVEWYKGKPAHECIEC
GVCETKCPQKLEIRKQLKETARVLSVK

[0137] Cthe0283 (SEQ ID NO:12)

MKYRKMGRGTGLYISEISLGSWLTYGNSDKETAVKVIDTAYSLGINYFDTANVY
ANGRAEVIVGEALKKYPRESYILATKAFWPMGTGPNDKGLSRKHVFEQVHASLKRLNV
DYIDIFYCHRYDPETPLEETLRTIDDLLRQGKILYVGVSEWTAQAQMAQALHIADRYLLDR
IVVNQPQYNMFHRYIEKEIIPFGEKNGISQIVFSPLAQGVLTGKYKPGGNIPRDSRAADPN
SNMYIGQFLKEDKLLKVEKLKAVADEMGITLSQLAIAWVLRQPNVTSALIGASKPEQVE
ENVKASGINLSDEILNKIEAILQ

[0138] EC 5.3.1.1. (Cthe0139; SEQ ID NO:13)

MSRKVIAAGNWKMNKTPKEAVEFVQALKGRVADADTEVVVGVPFVCLPGVVE
AAKGSNIKVAAQNMHWEEKGAFTGEVSGPMLAELGVDYVIIGHSEERRQYFGETDETVN
KKVHAAFKYGLKPIICVGESLTQREQGVTAELVRYQVKIALGLLSAEQVKEAVIAYEPIW
AIGTGKTATNEQAEEVCGIIRECIKELYGQDVAEAIQYGGSVNAANAAELFNMPNIDG
GLVGGASLKLDDFEKIAKYNK

[0139] EC 1.2.7.1

[0140] Cthe2390 (SEQ ID NO:14)

MGKVVEIRWHGRGGQGAKTASLLLADAAFNTPGKYIQGFPEYGPMPGAPITAY
NRISDEKLTIHSNIYEPDYVVVVDDTLTSDVDTAGLKEDGAIIVNTPKTPDEIRPLLKGY
KGKVCTIDARKISIIETLGKYFPNTPMLGAVVKVSKIMDEEEFLKDMVESFKHKFANKPE
VVEGNIKALERSMQEVKGL

[0141] Cthe2391 (SEQ ID NO:15)

MSKELRDVKPDVTWKEITSGGVIDSPGNAHLFKTGDWRSMKPVWNEEKCKQCL
LCNPVCPDSSIMVSEEGKMTGIDYDHCKGCGICSKVCPFKAIDFVEEV

[0142] Cthe2392 (SEQ ID NO:16)

MGIRERLSGNEATAIAMRQINPDVVAAFPITPSTEIPQYFSSYVADGLVDTEFVAV
ESEHSAMSACIGAQAAGARAMTATSANGLAYMWEALYIAASMRLPIVLA AVNRALSGP
INIHNDHSDTMGARDSGWIQLYSENNQEAYDNMLMAHRIGEHPDVMLPVMVCQDGFIT
SHAIENIELVEDEKVKAFVGEYKPTHYLLDRENPISVGPLDLQMHYFEHKRQQAQAMEN
AKKVILEVAEEFYKLTGRKYGFFEEYKTDDADVAIVVMNSTAGTVKYVIDEYRAKGKK
VGLIKPRVFRPFPVDELAQALSKFKAVAVMDKADSFNAAGGPLFTEVTSALFTKGVFGP
KVINYKFGLGGRDVKVDDIEVVCEKLLEIASTGKVDSVYNLGVRE

[0143] Cthe2393 (SEQ ID NO:17)

MAYNLKEVAKKPERLTGGHRMCAGCGAPIVVRQVLKALKPEDHAVISAATGCL
EVSTFIYPYTAWKDSFIHS AFENTGATISGAEEAYKVLKKKGKIEGETKFIAFGGDGGTY
DIGLQALSGAMERGHDMVYVCYDNGAYMNTGIQRSSATPKYADTTTSPVGKKIPGKM
QPRKDLTEVLVNHRIPTYVAQTAPFGNMKDLYEKAEKAIYTPGPAFLNVLAPCPRGWRY
NTPDLMELSKLAVETCFWPLYEVIDGKYIINYKPKEKVPVKEFLKLQGRFKHLFKAGNE
YMLEEIQKEVDLRWERLLKLAGEA

[0144] EC 1.1.1.27 (Cthe1053; SEQ ID NO:18)

MNNNKVIKKVTTVVGAGFVGSTTAYTLMLSLGLISEIVLIDINAKKADGEVMDLNH
GMPFVRPVEIYRGDYKDCAGSDIVIITAGANQKEGETRIDLVKRNTEVFKNINEIVKYNN
DCILLVVTNPVDILTYVTYKLSGFPKNKVIGSGTVLDTARFRYLLSEHVKVDARNVHAYI
IGEHDTEVA AAWSLANIAGIPMDRYCDECHQCEEQISRNKIYESVKNAAYEIIRNKGATY
YAVALAVRRIVEAIVRNENSILTVSSLLEGQYGLSDVCLSVPTIVGVNGIEEILNVPFNDEE
IQLLRKSGNTLKEIIKTLDI

[0145] EC 1.12.7.2

[0146] Cthe425 (SEQ ID NO:19)

MKVSICIGSSCHLKGAKQIVEQLQSLVADYNLKEKVELGGAFCKMKNVCVNGVSVT
VDDKLFSVTPENVKSFFETEILKKLED

[0147] Cthe426 (SEQ ID NO:20)

MTECLQTKKSNCKNCYKCI RHCPVKSLKFTDGQAHIVRDECVLGCECYVVC PQN
AKQIRSDVEKAKQLVLKYDVYASIAPSFVAWFHNKSIHDMEQALIKLGFKGADE TAKG
AYIVKKQYEKMIEEKKSKIISSCCHTVNTLIQRHYTGAIQYLADV VSPMLAHAQMLKKE
HKGAKVVFIGPCISKKDEAEKYKGYVELVLT FDELDEWLKSENITIESNRGSSKEGRTRS
FPVSGGISSMDKDLGYHYMVVDGMENCINALENIERGEIDNCFIEMSACRGSCINGPPA
RRKSNNIVGAILAVNKNTGAKDFSVPMP EPEKLLKKEFRFEGVHKIMPGGT AIEEILKKMG
KTSIEHELNCGSCGYDTCRDKAVAVLNGKADLT MCLPYLKEKAESFSDAIKNTPN GVIV
LNEDLEIQQINNSAKRILNLSPTDLLGSPVSRILD PIDYILALREGKNCYYKRKYFAEYKK
YVDETHIYDKEYHVIIIMRDVTEEEKIKALKNKQSEAAIEIADKVVEKQMRVVQEIALLL
GETAAETKIALTKLKETMEDE

[0148] Cthe427 (SEQ ID NO:21)

MNDLCVDLGYKSLNKFGEQLCGDMIQVVKDDDTTILVLADGLGSGVKANILSTL
TSKIISTMIAAHMGIEECVNTIMSTLPVCKVRGIAYSTFTIITNNTYAEIIQYDNPLVILLR
NGKKYDYPTQTKIISGKKIVESKIRLNCDDVFVVM SDGAIYAGVGQTLNYGWQRENIIEF
IESHYDKSLSANALTSLIDTCNNLYANMPGDDTTIAAIKIRKRQVVNLMFGPPQNPEDV
HNMMSLFFAKQGRHIVCGGTTSTLAAKFLGKELETTIDYIDPRIPPIARIEGVDLVTEGVL
TISRVLEYAKDYIGKNILYNEWHSKNDGASIIARMLFEEATDINFYVGKAINPAHQNP NL
PIGFNIKMQLVEELSKILKQMGKTINLSYF

[0149] Cthe428 (SEQ ID NO:22)

MSVTMSEAFDYSMIDNILSEHGTSETAIIAILQSIQEEYHYIPKEVFPYLSKKLKVSE
ARIFSVATFYENFSLEPKGKYIHKVCDGTACHVRKSIPIIERLRKELGLSGTKPTTDDLMFT
VETVSCLGACGLAPVITVNDKVYAEMTPDKASELIKQLREGDADA

[0150] Cthe429 (SEQ ID NO:23)

MLKNREELRKAREMYSRYLKA EKRRVLVCAGTGCVSGGSMEIFERLSELVSKRG
MDCQVELKEEPHDNTIGMKKSGCHGFCMGPLVRIEPEGYLYTKVKLEDCEEIVDRTIV
AGEHIERLAYKQNGVVYKKQDEIPFYKKQTRLVLEHCGQIDSTSITELYLATGGYYALEK
ALFDMTGDEIINEITEANLRGRGGGGFPAGRKWAQVKRQNAKQKYVVCNGDEGDPGAF
MDRSIMEGDPHRMIEGMIIAGIACGASEGYIYVRAEYPLAVSRLKRAIEQAKEFGLLGENI
LGSNFSFNIHINRGAGAFVCGEGSALTASIEGKRGMPRVKPPRTVEQGLFDMPTVLNNVE
TFANVPLIIKNGADWYKSIGTEKSPGTKAFALTGNIENTGLIEIPMGTTLREVIFDIGGGMR
NGADFKAVQIGGPSGGCLSEKDLDLPLDFDSLKKAGAMIGSGGLVVMDSNTCMVEVAR
FFMNFTQNESCGKCVPCREGTKRMLEILERIVEGNGQDGDIELLLELADTISATALCGLG
KAAAFPVVSTIKNFREEYEAHIYDKRCPTGNCQKLKTITIDASLCKGCSKCARSCPVGAIT
GKVKEPFVIDQSKCIKCGACIETCAFHAILEG

[0151] Cthe430 (SEQ ID NO:24)

MDNREYMLIDGIPVEINGEKNLLELIRKAGIKLPTFCYHSELSVYGACRMCMVEN
EWGGLDAACSTPPRAGMSIKTNTERLQKYRKMILELLLANHCRDCTTCNNNGKCKLQD
LAMRYNISHIRFPNTASNPDVDDSSLCITRDRSKCILCGDCVRVCNEVQNVGAIDFAYRG
SKMTISTVFDKPIFESNCVGCQCACALACPTGAIVVKDDTQKVWKEIYDKNTRVSVQIAPA
VRVALGKELGLNDGENAIGKIVAALRRMGFDDIFDTSTGADLTVLEESAELLRRIREGKN
DMPLFTSCCPAWVNYCEKFYPELLPHVSTCRSPMQMFASIIKEEYSTSSKRLVHVAVMP
CTAKKFEAARKEFKVNGVPNVYVLTQTQELVRMIKESGIVFSELEPEAIDMPFGTYTGA
GVIFGVSGGVTEAVLRRVVSDKSPTSFRSLAYTGVRGMNGVKEASVMYGDRKLKVAV
VSGLKNAGDLIERIKAGEHYDLVEVMACPGGCINGGGQPFVQSEEREKRGKGLYSADKL
CNIKSSEENPLMMTL YKG ILKGRVHELLHVDYASKKEAK

[0152] Cthe431

[0153] SEQ ID NO:25:

MLEIKICVGSSCHLKGSYNVINEFQHLIEEKALHDKIDIKATFCMKQCQKNGVAV
EVNNEIFGVLPEAAEEFFKNVILPKV

[0154] EC 1.12.7.2

[0155] Hyd Cthe3019-24

[0156] Cthe3019 (SEQ ID NO: 26)

MSFFTMTKTLIKSIFHGPLYTVRYPLEKKEPFPASRGRIEINIQDCIFCGLCARRCPTG
AINVEKPESRWSINRLRCIQCGYCSEVCPKKCLKMNNMYPAPSFENIEDVYQNARVPDN
KENNRNIAGAC

[0157] Cthe3020 (SEQ ID NO:27)

MGKKTVIPFGPQHPVLPEPIHLDLVLEDETVEAIPSIGYIHRGLEKLVEKKDYQQ
FVYVAERICGICSMHGMGYCMSIENIMGVQIPERAFLRTIWAELSRIHSHMLWLGLLA
DALGFESLFMHSWRLREQILDIFEETTGGRVIFSVCDIGGVRRDIDSEMLKKINSILDGF
EFSEITKVFLNDSSVKLRTQGLGVLSREEAFELGAVGPMARASGIDIDMRKSGYAA
YGKLKIEPVVETAGDCYARTSVRIREFQSIDLRQCISLIPDGEIKVKIVGNPSGEYF
TRLEQPRGEVLYYVKANGTKFLERFRVRTPTFANIPALLHTLKGCLADVPVLILTIDPC
ISCTER

[0158] Cthe3021 (SEQ ID NO:28)

MAQQTINTISPNELLAYALRLKNANYRLVAISCTNAENGVEMSYSFDSGSDFTNL
RITVAPGDEIESISSIYSYSFLYENEIKELFGVNITGISPDYKDKLYRISVKTPFNMKE
GDKNG

[0159] Cthe3022 (SEQ ID NO:29)

MNFSKKSPWILHYDGSSCNGCDIEVLACLTPLYDIERFGVINTGNPKHADILLITGS
INEQNKSVMKQLYEQMADPKVVVAVGICAATGGIFSECYNVSGGVDDKIIPVDVYVPG
CAARPEAIIDGVVKALGILEERQKYARKKDK

[0160] Cthe3023 (SEQ ID NO:30)

MSQIIRLVLYIIAIIIVAPLLGGLLTGIDRVITARMQGRKGPSVLQPFYDVLKLFQKE
SIEVNTMHRFFVYISLIFVIFTTVIMLLGGDILLALFALTLSIFFVLGGYASNSPYSTIG
SERELLQMMAFEPMLLLAAIGLYYGDKSFFIKDIVTARIPSIVYLPGVFLGLLYVLT
FKLRKSPFDLSMSHHGHQEIVQGITTEYSGKDLAIQITHWYETIIALALVYLFFAFRSP
FSHVIAILACIIAFLLEIVVDNAFARAKWEFALKSTWIVTGVLASVNLIILSFFR

[0161] Cthe3024 (SEQ ID NO:31)

MNAILILILFPLLASVTVLSVRKDAIRNIIVRIFAFITGILTFLVVCRYFKDGLS
SIENRNIIDMTISLAEVLIAAYIIFTGIKNKKFIVSIFAAVQTALILWFEFTQKHGINVHSD
IVFDRL

SAVMVLIVGCIGSLILIYTVGYMKWYHIHHEGYKERKSFFFSVIFLFLFAMFGLIFSNNLI
 WMYFCWELTTLC SYLLIGYTRTPEAVNNSFHALAINLGGGLAFASAMVYIGTNFKTLEL
 SALTAMKLELAVLIPVFLLCIAALTKSAQMPFSSWLLGAMVAPTPSSALLHSATMVKAG
 VYLLIRLAPLLAGTTIGKVIALLGAVTFLASSIIAISKSDAKKILAYSTISNLGLIVTCAAIGT
 QESLWAAILLIFHSISKSLFLTGGGSVEHQIGSRNVEDMDILLQVSRRLSVYMI VGIAGM
 FLAPFGMLISKWVAMKAFIDSKNILT VIILGYGSATTIFYWTKWMGKLVANANRKDHIK
 HTFHIDEEIPIFIHAVLVVLSCFTFPLVSRYVLVPYLSGLFGPDVPIPIGTSDVNIMLIMLSM
 LLILPISFIPIYKSDRRRIVPIYMAGENTGDNESFYGAFDEKRRKVELHNWYMKNFFSVKKL
 TFWSNLLCAVVILVGVVLLIGGITK

[0162] Cthe342 (SEQ ID NO:32)

MQMVNVTIDNCKIQVPANYTVLEAAKQANIDIPTLCFLKDINEVGACRMCVVEV
 KGARSLQAACVYPVSEGLEVYTQTPAVREARKVTLELILSNHEKKCLTCVRSENCCELQR
 LAKDLNVKDIRFEGEMSNLPIDDLSPSVVRDPNKCVCRCVSMCKNVQTVGAIDVTER
 GFRTTVSTAFNKPLSEVPCVNCGQCINVCVPGALREKDDIDKVWEALANPELVVVQTA
 PAVRVALGEEFGMPIGSRVTGKMVAALSRLGFKKVFDTDTAADLTIMEEGTELINRIKN
 GGKLPLITSCSPGWIKFCEHNYPEFLDNLSSCKSPHEMFGAVLKSYAQAQKNGIDPSKV FV
 VSIMPCTAKKFEAQRPELSSTGYPDVDVLTRELARMIKETGIDFNSLPDKQFDDPMGE
 ASGAGVIFGATGGVMEAAIRTVGELLSGKPADKIEYTEVRGLDGIKEASIELDGFTL KAA
 VAHGLGNARKLLDKIKAGEADYHFIEIMACPGGCINGGGQPIQPSSVRNWKDIRCERAK
 AIYEEDESLPIRKSHENPKIKMLYEEFFGEPGSHKAHELLHTHYEKRENYPVK

[0163] Cthe430 (SEQ ID NO:33)

MDNREYMLIDGIPVEINGEKNLLELIRKAGIKLPTFCYHSEL SVYGACRMCMVEN
 EWGGLDAACSTPPRAGMSIKTNTERLQKYRKMILELLLANHCRDCTTCNNNGKCKLQD
 LAMRYNISHIRFPNTASNPDVDDSSLCITRDRSKCILCGDCVRVCNEVQNVGAIDFAYRG
 SKMTISTVFDKPIFESNCVGCGQCALACPTGAIVVKDDTQKVWKEIYDKNTRVSVQIAPA
 VRVALGKELGLNDGENAIGKIVAALRRMGFDDIFDTSTGADLTVLEESAELLRRIREGKN
 DMPLFTSCCPAWVNYCEKFYPELLPHVSTCRSPMQMFASIIKEEYSTSSKRLVHVAVMP
 CTAKKFEAARKEFKVNGVPNV DYVLTQTQELVRMIKESGIVFSELEPEAIDMPFGTYTGA
 GVIFGVSGGVTEAVLRRVVSDKSPTSFRSLAYTGVRGMNGVKEASV MYGDRKCLKVAV

VSGLKNAGDLIERIKAGEHYDLVEVMACPGGCINGGGQPFVQSEEREKRGKGLYSADKL
CNIKSSEENPLMMTL YKG ILKGRVHELLHVDYASKKEAK

[0164] Cthe3003 (SEQ ID NO:34)

MDSFLMKGYIKEANIDYSCSRGSMEDLPKWEFREIPKVPRAVMPSLSLEERKNNF
NEVELGLSEEVARKEARRCLKCGCSARFTCDLRKEASNHGIVYEEPIHDRPYIPKVDDHP
FIVRDHNCISCGRCAACAEIEGPGVLT FYMKNGRQLVGTKSGLPLRDTDCVSCGQCVT
ACPCAALDYRRERGGKVVRANDPKKTVVGVFVAPAVRSLISNTFGVSYEEASPFMAGLLK
KLGFDKVFDFTFAADLTIVEETTEFLSRIQNKGVMPQFTSCCPGWINFVEKRYPEIIPHLST
CKSPQMMMGA TVKNHYAKLMGINKEDLFVVSIVPCLAKKYEAARPEFIHDGIRDVDAV
LTTTEMLEMMELADIKPSEVVPQEFDEPYKQVSGAGILFGASGGVAEAALRMAVEKLTG
KVLTDHLEFEEIRGFEGVKESTIDVNGTKVRVAVVSGLKNAEPIIEKILNGVDVGYDLIEV
MACPGGCICGAGHPVPEKIDSLEKRQQVLVNIDKVSKYRKSQENPDILRLYNEFYGEPNS
PLAHELLHHTHYTPKHGDSTCSPERKKGTAAFDVQEFTICMCESCMEKGAENLYNDLSSK
IRLFKMDPFVQIKRIRLKETHPGKGVYIALNGKQIEEPMLSGNIPDESESE

[0165] Cthe3004 (SEQ ID NO:35)

MKTLENHNRIKVTVNGREIEVYDNL TILQALLQEDIHIPHLCYDIRLERSNGNCGL
CVVTLISPDGERDVKACQTPIKEGMVICTNTPKLENYRKIRLEQLLSHDNADC VAPCVMT
CPANIDIQSYLRHVGNNGDFEAAIRVIKERNPFPIVCGRVCPHTCESQCRRNLVDAPVAINY
VKRFAADWDMARPEPWTPEKKPPTGKKIAIVGAGPSGLSAA YYSAIKGHDVTVFERQPH
PGGMMRYGIPEYRLPKAILDKEIEMIKKLGVKIMTEKALGIHIRLEDLSKDFDAVYLAIGS
WQATPMHIEGEKLDGVWAGINYLEQVAKNVDIPLGDNVVVIGGGNTAIDCARTALRKG
AKSVKLVYRCTREEMPAAPYEVEEAIHEGVEMIFLMAPTKIIVKDGKKKKLVCIRMQLGE
PDRSGRRRPVPIEGSEVEIDADTIIGAIGQSTNTQFLYNDLPVKLNK WGDIEVNGKTLQTS
EYNIFAGGDCVTGPATVI

[0166] Cthe0349 (SEQ ID NO:36)

MPLVTSTEMFKKAYEGKYAIGAFNVNNMEIIQGITEAAKEVNAPLILQVSAGARK
YANHTYLVKLVEAAVEETGLPICLHLDHGDSFELCKSCIDGGFTSVMIDGSHLPFEENIKL
TKQVVDYAHSKGVVVEGELGRLAGIEDDVNVSEADAAFTDPDQAEEFVKRTGVDSLAI
AIGTSHGAYKFKGEAKLRFDILEEIEKRLPGFPIVLHGASSVIPEYVDMINKYGGDMPGA

KGVPEDMLRKAASMAVCKINIDSDLRLAMTATIRKYFAENPSHFDPRQYLGPARNAIKE
LVKHKIVNVLGCDGKA

[0167] Cthe1019 (SEQ ID NO:37)

MDIQLKKSGIGVKEKKSKNHLLYSIKQNLFAAYAMLIPTFVCMMCIHFIPMLQGIYL
SLLDLNQLTMTKFLNAPFIGLKNYYEILFDEKSLIRRGFWFALRNTAIYTVVVTFATFALG
IILAMLVNREFKGRGIVRTALLMPWVVPSYVVGMTWGFLWRQDSGLINIILCDILHILPE
KPYWLVGSNQIWAIIPTIWRGLPLSMILMLAGLQSPDYEEAADIDGANGWQKFWHIT
LPLLKPILAINVMFSLISNIYSFNIVSMMFGNGAGIPGEWGDLLMTYIQRNTFQMWRFGP
GAAALMIVMFFVLGIVALWYTLFKDDL VVK

[0168] Cthe0390 (SEQ ID NO:38)

VDKFTKLDLNSITSNNRMNIFNCILEAKEINRAVIAKKVGLSIPAVMSITDDLIQKG
IIVVIGKGKSSGGKRPELLAVVPDRFFFVGVDVGRTSVRVVVMNNCRDVVYKVSKPTES
VEPDELINQITEMTMESINESKFPLDRVVGIGVAMPGLIERGTGRVIFSPNFGWNNIALQD
ELKKHLPFNVLVENANRALVIGEIKNTQPNPTSCIVGVNLGYGIGSAIVLPNGLYYGVS
TSGEIGHIIVENHGSYCSCGNYGCIESIASGEAIAAREARIAIAANKIQSSVFEKCEGDLKKIDA
KMFVDAAKEGDHLAQSIVEKAADYIGKGLAITINMLDPEQIILCGGLTSLSGDFFIDMIKK
AVSKYQMRYAGGNVKIVVGKSGLYATAIGGAWIVANNIDFLSSN

[0169] Cthe2938 (SEQ ID NO:39)

MYIIGIDLGGTNIAVGLVNEEGKILHKDSVPTLRERPYQEIIKDMAMLTCLKVIKD
ADVSIQVKSIGVGSPGTPNCKDGILIYNNNLNFRNVPIRSEIQKYIDLVPVLDNDANCAA
LAESVAGAAKGANTSVTITLGTGIGGGVVIDGKIYSGFNYAGGELGHTVLMMDGEPCTC
GRKGCWEAYASATALIRQARKAAEANPDSLINLKVGGDLSKIDAKIPFDAAKQGDKTGE
MNVVQQYIRYIAEGLINMINIFMPEVLVIGGGVCKEYLLKPLRELKQGVYSKEDIPQTE
LRTAQMGNDAGIIGAAMLGKEC

[0170] Cthe0217 (SEQ ID NO:40)

MERIKFDYSKALPFVSEREVAYFENFVRSAMDMLHNKTGAGNDFVGWVDLPVN
YDREEFARIKAAAEKIKSDSDALVIGIGGSYLGARAAIEMLSHSFHNLMPSKSKRNAPEI
YFVGNNISSTYIADLLEVIEGKEISVNVISKSGTTTEPAIAFRIFKEYMENKYGKDGASKRI

YATTDKEKGALRKLATEEGYETFVVPDDIGGRFSVLTAVGLLPIAVAGIDIDSMMKGAA
DARELYSNPNLMENDCYKYAAVRNALYRKNKTIEIMVNYEPSLHYFTEWWKQLYGESE
GKDQKGIFPAGVDFTTDLHSMGQYIQDGLRNIFETVIRVEKPRKNIVIKEEKDNLDGLNFI
AGKDVDYVNKKAMEGTVLAHTDGGVPNLVVTVPELSAYYFGNMVYFFEKACGISGYL
LGVNPFDPQPGVEAYKKNMFALLGKPG YEEQRKKLEERL

[0171] Cthe1261 (SEQ ID NO:41)

MSSVRTIGVLTSGGDAPGMNAAIRSVVRTGLYYGFKVLGIRKGFNGLINGDIEEL
TARVGDIIHRGGTILQTARSPQFKTEEGLKKAMSMKVFIDALVVIGGDGSYRGARDI
SKLGLNVIGIPGTIDNDIGCTDYTIGFDTAMNTVQDAIDKIRDTAYSHERCSVLEVMGRH
AGYIAVNVSISGGAEAVVLPEKPFDMDTDVIKPIIEGRNRGKKHYLVIVAEGGEGKAIEIA
KEITEKTGIEARATILGHIQRGGSP TVYDRVMASQMGAKAVEVL MENKRN RVIVFKDNQ
IGDMDLEEALQVKKTISED LIQLSKILAL

T. saccharolyticum proteins

[0172] Reaction 6b, EC or0411 (SEQ ID NO:42)

MSYIPNENRYEKM IYRRCGRSGIMLPAISLGLWHNFGGYDVFENMREMVKKAFD
LGITHFDLANNYGPPPGSAEENFGKILRTDLRGYRDELLISTKAGYTMWPGPYGDWGSR
KYLSSLDQSLKRMGIDYVDIFYSHRRDPNTPLEETMSALAQA VRQGKALYVGISNYNA
EDTKKAAEILRQLGTPLLINQPSYSMFNRWIEDGLTDVLEEEGVGSIAFSPLAQGLLTDK
YLNGVPDDSR AVRKNTSLRGNLTEENINKVRELKKIADKRGQSIAQMALAWDLRKVTS
VIIGASRV SQIEENVKALDNLEFSHEELKQIDEILSK

[0173] EC4.2.3.4 or2316 (SEQ ID NO:43)

LNIALIAHDMKKSIMVDFAIAYKEILKKCNIYATGATGQLVEEATGIKVNKFLPGP
MGGDQQIGAMIAENNMDLVIFLRDPLTAQPHEPDILALLRVCDVHSIPLATNLATAEVL I
KGLDAGFLEWRDAVK

[0174] EC5.3.1.1 or2687 (SEQ ID NO:44)

LRRPIIAGNWKMYMTPSEAVNLVNELKPLVSGAEAEV VVIPPFVDLVDVKKAID
ASNIKLGAQNMHWEEKGAFTGEVSPIMLKEIGVEYVVIGH SERRQYFAETDETVNKKVK
SALSHGLKPIVCVGESLSQREAGEAFNVVREQTKKALDGIKSEDVLNVVIA YEPIWAIGT

GKTATSKDANDVIKVIRETIADIYSIDIANEVRIQYGGSVKPDNAKELMSESDIDGALVGG
ASLKAQDFAKIVNY

[0175] Reaction 7 (gldA) orl04 (SEQ ID NO:45)

MYMKTNFTYFMPTEIFGPGTLGKLATVKLPGKKALLVIGSGNSMRRHGYLDRVV
NYLKQNGVDYVVYDKILPNPIAEHVAEGAKVAKDNGCDFVIGLGGGSTIDSSKAIAVMA
KNPGDYWDYVSGGSGKGMEVKNGALPIVAIPTTAGTGTESDPWAVVTKTETNEKIGFG
CKYTYPTLSIVDPELMVSIPPKFTAYQGMDAFFHSVEGYLATVNQPGSDVLALQSSILITE
NLPKAVADGNNMEARTALAWASTAAGIVESLSSCISHHSLEHALSAYHPEIPHGAGLIML
SVSYFSFMASKAPERFVDIAKAMGEEIVGNTVEEQAMCFINGLKKLIRNIGMEDLSLSSF
GVTEDEATKLAKNAMDTMGGLFNVDPYKLSLDEVVSIYKNCF

[0176] EC2.3.1.54 (SEQ ID NO:46)

VDDKKVFDHLFILTDDTGMMQHSGSVDPDPKYGYTTDDNGRALIACAMMYEK
YKDDAYINLIKKYLSFLMYAQEDDGRFRNFMSFDRKFIDEDFSEDCFGRCMWALGYLIN
SNIDERVKLPAVKMIEKSLLLVDTLNYIRGKAYTLIGLYYIYNSFKNLDKDFVRKKMDKL
AHDIVEEYEKNSSSEDWQWFEDVVSVDNGVIPLSLLKYFSIAKDEEVLDIALKTIDFLDSV
CFKNGYFKA VGCKGWYRK GK DIAEYDEQPVEAYTMALMYIEAYKLTGDEKYKKRAID
CDKWFY GKN SKGLSLYDEDSGGCSDGITEDGVNSNEGAESLISIMISHCAIDQLK

[0177] EC2.3.1.8 (SEQ ID NO:47)

MKTSELLAMVVEKGASDLHITVGVPPVLRINGQLIKLNLPQLTPQDTEEITKDLLS
SDELKKLEDMGDIDLSYSVKGLGRFRINAYKQRGTYSLAIRSVALRIPTIDELGLPEVIKE
LALKTRGLIIVTGPTGSGKSTTLASMLINEERNCHILTLEDPIEYLHKHKKSIVNQREIG
HDAASYASALRAALREDPDVILVGEMRDLETIQIAITAAETGHLVLSTLHTIGSAKTIDRII
DVFPPHQQQQIKVQLSNVLEGIVSQQLLPKIDNSGRVVAVEVMIATPAIRNLIREGKSFQI
QSMVQTGNKFGMVTMDMWISQLLKRNLISMDDALTYCVDRENFSRLVV

[0178] EC 2.7.2.1 (SEQ ID NO:48)

MIKKKLGDLLVEVGLLDESQLNNAIKIQKKTGEKLGKILVKEGYLTEEQIIEALEF
QLGIPHIDMKKV FIDANVAKLIPESMAKRHV AIPK KENNSIFVAMADPLNIFAIDDIKLV
KLDVKPLIASEDGILKAIDRVFGKEEAERAVQDFKKELSHDSAEDDGNLLRDISSEDEINN

APAVRLVNSIIEQAVKNRASDVHIEPTENDLRIRFRIDGELHEAMRVFKSTQGPVITRIKIM
ANMNIAERRIPQDGKIEMNAGGKNIDIRVSSLPTIYGEKLVLRILDKSGYIITKDKLGLGN
DDLKLF DNLLKHPNGIILLTGPTGSGKTTTLYAMLNELNKPDKNIITVEDPVEYTLEGLN
QVQVNEKAGLTFASALRSILRQDPDIIMIGEIRDRETAIEAIRSSITGHLVLSTLHTNDSAG
AITRLIDMGIEPYLVSSSVVGVIQAQRLARKICDNCKIEYDASKREKIILGIDADESLKLYRS
KGCAVCNKTGYRGRVPPIYEIMMMTPKIKELTNEKAPADVILNEAVSNGMSTLKESAKKL
VLSGVTTVDEMLRLTYDDAY

[0179] EC 1.2.7.1 or0047 (SEQ ID NO:49)

MSKVMKTMDGNTAAAHVAYAFTEVAAIYPITPSSPMAEHVDEWSAHGRKNLFG
QEVKVIEMQSEAGAAGAVHGS LAAGALTTTFTASQGLLLMIPNMYKIAGELLPGVFHVS
ARALASHALSIFGDH QDVMACRQTGFALLASGSVQEVM DLGSAHLAAIKGRVPFLHFF
DGFRTSHEYQKIEVMDYEDLRKLLDMDAVREFKKRALNPEHPVTRGTAQNPDIYFQERE
ASNRYYNAVPEIVEEYMKEISKITGREYKLFNYYGAPDAERIVIAMG SVTETIETIDYLL
KKGEKVG VVKVHLYRPF SFKHFMDAIPKTVKKIAVLDR TKEAGAFGEPLYEDVRAAFY
DSEMKPIIVGGRYGLGSKDTTPAQIVAVFDNLKSDTPKNNFTIGIVDDVTYTS LPVGEEIE
TTAEGTISCKFWGFGSDGTVGANKSAIQIIGDNTDMYAQA YFSYDSKKSGGVTISHLRFG
KKPIRSTYLINNAD FVACHKQAYVYNYDVL AGLKKGGTFL LNCTWKPEELDEKL PASM
KRYIAKNNINFYIINA VDIAKELGLGARINMIMQSAFFKL ANIIPIDEAVKHLKDAIVKSYG
HKGEKIVNMNYAAVDRGIDALVKVDVPASWANA EDEAKVERNVPDFIKNIADVMNRQ
EGDKLPVSAFVG MEDGT FPMGTAA YEKRGIAVDVPEWQIDNCIQC NQCA YVC PHAAIR
PFLN EEEVKN APEGFTSKKAIGKGLEGLNFRIQVSVLDCTGCGVCANTCPSKEKSLIMK
PLETQLDQAKNWEYAMSLSYKENPLGTDTVKGSQFEKPLLEFSGACAGCGETPYARLV
TQLFGDRMLIANATGCSSIWGG SAPSTPYTVNKDGHGPAWANS LFEDNAEFGFGMALA
VKQQREKLADIVKEALELDTQDLKNALKLWLDNFNSSEITKKTANIIVSLIQDYKTDDS
KVKELLNEILDRKEYLVKKSQWIFGGDGWAYDIGFGGLDHV LASGEDVNVLVFDTEVY
SNTGGQSSKATPVGAIAQFAAAGKGIGKKDLGRIAMSYGYVYVAQIAMGANQAQTIKA
LKEAESYPGPSLIAYAPCINHGIKLG MGCSQIEEKKAVEAGYWHLYRYNPMLKAEGKN
PFILDSKAPTASYKEFIMGEVRYSSLAKTFPERAEALFEKAEELAKEKYETYKKLAEQN

[0180] EC 1.1.1.2

[0181] Or180 (SEQ ID NO:50)

MSKVAIIGSGFVGATSAFTLALSGTVTDIVLVDLNKDKAIGDALDISHGIPLIQPVN
VYAGDYKDVKGADVIVVTAGAAQKPGETRLDLVKKNTAIFKSMIPELLKYNDKAIYLIV
TNPVDILTYVITYKISGLPWGRVFGSGTVLDSSRFRYLLSKHCNIDPRNVHGRIIGEHGDTE
FAAWSITNISGISFNEYCSICGRVCNTNFRKEVEEEVVNAAYKIIDKKGATYYAVAVAVR
RIVECILRDENSILTVSSPLNGQYGVKDVSLSLPSIVGRNGVARILDLPLSDEEVEKFRHSA
SVMADVIKQLDI

[0182] EC 2.3.1.54 (SEQ ID NO:51)

MINEWRGFQEGKWQKTIDVQDFIQKNYTLYEGDDSFLEGPTEKTIKLWNKVLEL
MKEELKKGVLDIDTKTVSSITSHDAGYIDKDLEEIVGLQTDKPLKRAIMPYGGIRMVKKKA
CEAYGYKVDPKVEEIFTKYRKTHNDGVFDAYTPEIRAARHAGIITGLPDAYGRGRIIGDY
RRVALYGIDRLIEEKEKEKLELDYDEFDEATIRLREELTEQIKALNEMKEMALKYGYDIS
KPAKNAKEAVQWTYFAFLAAIKEQNGAAMSLGRVSTFLDIYIERDLKEGTLTEKQAQEL
MDHFVMKLRMVRFLRTPDYNELFSGDPVWVTESIGGVGVDGRPLVTKNSFRILNTLYN
LGPAPENLTVLWSKNLPENFKRFCAKVSIDTSSIQYENDDLMRPIYNDDYSIACCVSAM
KTGEQMQFFGARANLAKALLYAINGGIDERYKTQVAPKFNPITSEYLDYDEVMAAYDN
MLEWLAKVYVKAMNIIHYMHDKYAYERSLMALHDRDIVRTMAFGIAGLSVAADSLSAI
KYAKVKAIRDENGIAIDYEVEGDFPKFGNDDDRVDSIAVDIVERFMNKLKKHKTYRNSIP
TLSVLTITSNVVYGKKTGATPDGRKAGEPFAPGANPMHGRDTKGAIASMNSSKIPYDSSL
DGISYTFITVPNALGKDDDEDKINNLVGLLDGYAFNAGHHININVLNRDMLLDAMEHPEK
YPQLTIRVSGYAVNFNKLTREQQLEVISRTFHESM

[0183] EC1.2.7.1 (following four proteins)

[0184] Or1545 (SEQ ID NO:52)

MVITVCVGSSCHLKGSYDVINKLKEMIKNYGIEDKVELKADFCMGNCLRAVSVK
IDGGACLSIKPNSVERFFKEHVLGELK

[0185] Or1546 (SEQ ID NO:53)

MSVINFKEANCRNCYKCIKYCPVKAIKVNDEQAEIIEYRCIACGRCLNICPQNAKT
VRSDVERVQSFLNKGEKVAFTVAPSYPALVGHDGALNFLKALKSLGAEMIVETSVGAM
LISKEYERYYN DLKYDNLITTSCPSVNYLVEKYYPDLIKCLVPVVSPMVAVGRAIKNIHG
EGVKVVFIGPCLAKKAEMSDFSCEGAIDAVLTFEEVMNLFNTNKIGVECTKENLEDVDS

ESRFKLYPIEGKTMDCMDVDLNLRFKISVSSIENVKDILNDLRAGNLHGYWIEANACDG
GCINGPAFGKLESGLAKRKEEVISYSRMKERFSGDFSGITDFSLDLSRK FIDLSDRWKMP
SEMEIKEILSKIGKFSVEDELNCGACGYDTCREKAIAVFNGMAEPYMCLPYMRGRAETLS
NIIISSTPNAIIAVNNEYEIQDMNRAFEKMFLVNSAMVKGEDLSLIFDISDFVEVIENKKSIF
NKKVSFKNYGIIALESIIYYLEEKIAIGIFTDITKMEKQKESFSK LKRENYQLAQQVIDRQ
MKVAQEIASLLGETTAETKVILTKMKDMLLNQGDDE

[0186] or1547 (SEQ ID NO:54)

MSHYIDIAHASLNKYDEELCGDSVQIIRKKDYAMAVMADGLGSGVKANILSTLT
TRIVSKMLDMGSELRDVVETVAETLPICKERNIAYSTFTVVSIIYGDNAHLVEYDNPSVFY
FKNGVHKKVDRKCVEIGDKKIFESSFKLDLNDALIVVSDGVIHAGVGGILNLGWQWDN
VKQYLSKVLEVYSDASDICSQLITTCNNLYKNRPGDDTTAIVIKVNESKKVTVMVGPPIL
KNMDEWVVKKLMKSEGLKVVCGGTAAKIVSRILNKDVITSTEYIDPDIPPYAHIDGIDL
TEGVLT LRKTVEIFKEYMNDKDSNLLRFSKKDAATRLFKILNYATDVNFLVGQAVNSAH
QNPDPFSDLRIKVRIVEELISLLERLNKNVEVNYF

[0187] or1548 (SEQ ID NO:55)

LFKFNTDVQMLKYEVLVYNVAKLTLEDREDEYDEIPYEIIPGTKPRFRCCVYKER
AII EQRTKVAMGKNLKR TMKHAVDGEEPIIQVLDIACEECPIKRYRVTEACRGCITHRCT
EVC PKGAI TIINKKANIDYDKCIECGRCKDACPYNAISDNLRPCIRSCSAKAITMDEELKA
AINYEKCTSCGACTLACPFGAITDKSYIVDIIRAIKSGKKVYALVAPAIASQFKDVTVGQI
KSALKEFGFVDVIEVALGAD FVAMEEAK EFSHKIKDIK VMTSSCCPAFVAHIKKSYPELS
QNISTTVSPMTAISKYIKKHDPMAVTVFIPCTAKKSEVMRDDVKGITDFAMTFEEMVA
VLDAAKIDMKEQQDVEVDDATLFG RK FARSGGVLEAVVEAVKEIGADVEVNPVVCNG
LDECNKTLKIMKAGKLPNNFIEGMACIGGCIGGAGVINNNVNQAKLAVNKFGDSSYHKS
IKDRISQFDTDDVDFHVDSGEDESSETSFKEA

[0188] EC 1.2.1.43

[0189] or2328 (SEQ ID NO:56)

MDKVRITIDGIPAEVPANYTVLQAAKYAKIEIPTLCYLEEINEIGACRLCVVEIKGV
RNLQASCVPVSDGMEIYTNTPRVREARRSNLELILSAHDRSCLTCVRSGNCELQDLSRK
SGIDEIRFMGENIKYQKDESSPSIVRDPNKC VLCRR CVATC NNVQNVFAIGMVNRGFKTI

VAPSFGRGLNESPCISCGQCIEACPVGAIYEKDHTKIVYDALLDEKKYVVVQTAPAVRVA
 LGEEFGMPYGSIVTGKMVSALKRLGFDKVFDTDFAADLTIIIEGNELLKRLNEGGKLPMI
 TSCSPGWINYCERYYPEFIDNLSTCKSPHMMMGAIIKSIFYAEKEGIDPKDIFVVSIMPCTA
 KKYEIDRPQMIVDGMKDVDVAVLTRELARMIKQSGIDFVNLPDSEYDNPLGESSGAGVIF
 GATGGVMEAALRTVADIVEGKDIENFEYEEVRGLEGIKEAKIDIGGKEIKIAVANGTGNA
 KKLLDKIKNGEAEYHFIEVMGCPGGCIMGGGQPIHNPNEKDLVRKSRLKAIYEADKDLPI
 RKSHKNPMITKLYEEFLISPLGEKSHHLLHTTYSKKDLYPMND

[0190] EC 4.1.2.13

[0191] or0260 (SEQ ID NO:57)

LNDILVKARNNKYAIGGFNFNFYDDALGIISAAYELKSPIILMASEGCVKFLGVKH
 IVNFBVNQLKDEYNIPILHLDHKGKDIEIHKNCIDNKFDSIMYDGSLLNFEENIKNTKFIADLC
 HDKGMTIEGELGRISGAEEINIENSEDVFTDPDSVAEFTERSDVDSLAVAIGNAHGLYKGR
 PRLDFERLSKINKISKVPLVLHGGTGIPYEDIQKAIQLGISKVNVGTEIKIAYIKSIKKHLETI
 NDNDIRHLVSMVQNDIKELVKQYLDIFGTANKYSQLQSM

[0192] or0330 (SEQ ID NO:58)

MLVTGIELLKKANEEGYAVGAFNTSNLEITQAIVEAAEEMRSPAIQVSEGGLKY
 AGIETISAIVRTLATKASVPIALHLDHGTDFNNVMKCLRNGWTSVMMDASKLPLEKNIE
 VTKNVVTIAHGMGVSVEAEIGKIGGTEDNVTVDEREASMTDPDEAFKFAKETGVDYLAI
 SIGTAHGPYKGEPKLDFDRLVKIKEMLMPIVLHGASGVPEADIRKAVSLGVNKNINIDTDI
 RQAFAARLRELLKNDEEVYDPRKILGPCKEAMKEVIKNKM RMFGSEGRA

[0193] or0272 (SEQ ID NO:59)

MITGDQLLIKQINKSIVLNTIRKKGLISRADLANITGLNKSTVSSLVDELIKEGFVEE
 EGPGESKGGRPIMLMINSLAGCVIGVDLDVNYILVILTDILANILWQKRINLKLGESKED
 IISKMLELIDEAIKNSPNTVKGILGIGIGVPGITDYKRGVVLKAPNLNWENVELKKMVEER
 FNLKVYIDNEANTGAIGEKWFGGGRNAKNFVYVSAGIGIGTGIIINNELYRGSNGLAGEM
 GHMTIDINDHMCSCGNRGCWENYASEKSLFRYIKERLEAGQEDDFIDSENIDSLDINDIA
 GYAELGSKLAIDAINESKNLSVGIVNIVNTFNPDLVLIGHTLSAIGDMLIDAVKEYVREK
 CLVSRYNDIAIEISKLGMLEAIGAVTLVISEVFSYPGL

[0194] or1389 (SEQ ID NO:60)

MTNVLNFDYSNALNFVNEHEISYLEKQALLSLDMVLNKTAQGSDFLGWVDLPK
DYDKEEFARIKKA AEKIKSDSDALVVIGIGGSYLGARAAIEMLT HSFYNVLPQSVRK APEI
YFAGNSISSTYLQDLLEILEGKDV SINVISKSGTTTEPAIAFRVFRDFLEKKY GKEEAKSRI
YVTTDRQKGALKKLADEEGYET FVIPDDVGGRYSVLTAVGLLPIAAAGIDIDEMMKGA
YDASIVFKKPDIKENLSMQYAVLRNALYRK GKSVEILVNYEPRLHYFSEWWKQLYGESE
GKDHKGIYPASVDFSTDLHSMGQFIQDGS RIMFETVINVEKPLKEITINEDKDNVDGLNFL
TGKTVDLVNKKA FEGTVLAHNDGGVPNLIVNVPEISAYNFGYLVYFFEMACGISGYLNG
VNPFDQPGVEAYKKNMFALLGKPGYEKEKEELEKRLKR

[0195] or2875 (SEQ ID NO:61)

MYNIQLDSPNLGDKEKDYL VKCIESGYVSTVGPFVPEFERRFAEFLNVNHCVSVQ
SGTAALYMALYELGIKDGDEVIVPAITFVATVNPIVYCGATPVFVDVDKDTWNIDPKEIE
KAITPKTKAIIPVHLYGNPCDMDKIMEIAKENNIYVIEDATESLGALYKGRMTGTIGHIGC
FSFNGNKVITTGGGGMVASNEDWVSHIRFLVNQARDMTQGYFHTEIGFNYRMTNLEA
SLGIAQLERLAGFLEKKRMYFEIYKKIFNGIEEISLQTEYEGAKSSDWLSSVKIDCKKVGM
TIHQIQDELKRRGIPTRRIFNPIVDLPPYKKYKKGSYSNSYEIYENGLNLPSSTLNTYEDVK
YVAKTLLDILSIKKR

***T. saccharolyticum* pdu genes or228-or200**

[0196] or228

[0197] SEQ ID NO:62:

MLAIERRKRIMRLIQENQSVLVPELSKLFNVTEETIRRDLEKLEAEGLLKRTYGGA
VINENSSADIPLNIREITNIESKQAISMKVAEYIEDGDTLLLDSSSTVLQVAKQLKFKKKLT
VITNSEKIILELANAKDCKVISTGGVLKQNSMSLIGNFAEDMIKNFCVDKAISSKGFDMT
NGITESNEMEAEIKKAMANS AEKVFLLLDHNKFDKSSFVKMFDLDKIDYLF TDRKLSLE
WEEFLKKHNIDLIYC

[0198] SEQ ID NO:63:

ATGCTTGCGATAGAACGAAGGAAGAGGATAATGAGGCTTATACAGGAAAATC
AAAGCGTTTGGTGCCTGAGTTAAGTAAATTGTTTAATGTGACAGAGGAAACTATAAG
GAGAGATTTAGAGAAACTTGAAGCAGAAGGGCTTTTAAAGAGGACTTATGGTGGTG

CTGTTATAAATGAAAATTCAAGTGCTGATATCCCCTTAAATATAAGGGAAATAACGA
ATATAGAAAGCAAACAGGCCATAAGTATGAAGGTTGCCGAATACATTGAAGATGGT
GATACACTTTTGCTTGATTCAAGCTCTACAGTTCTTCAAGTAGCAAAGCAATTAAAA
TTCAAAAAGAAGCTTACAGTCATAACAAATTCGGAAAAGATAATATTAGAATTAGC
AAATGCGAAAGATTGCAAAGTCATTTCTACAGGAGGAGTATTGAAGCAAAATTCTAT
GTCGCTAATTGGAAATTTGCGCGGAAGATATGATAAAAAAATTTCTGTGTAGATAAAGC
CATAATATCATCAAAAAGGTTTTGACATGACAAATGGCATTACAGAGTCAAACGAAAT
GGAAGCTGAAATAAAAAAAGCCATGGCCAACCTCGGCAGAAAAAAGTGTTTTTACTTC
TTGATCACAACAAATTTGACAAGTCATCGTTCGTCAAGATGTTTGACTTAGATAAAAA
TCGATTATCTATTTACCGATAGAAAGCTGTCTTTAGAATGGGAAGAATTCTTGAAAA
AACACAATATTGATTTAATCTATTGTTAG

[0199] or227

[0200] SEQ ID NO:64:

VYSEYEVKKQICEIGKRIYMNGFVAANDGNITVRIGENEIITPTGVSKGFMTPTDM
LLNINLNGEVLKSSGDYKPSTEIKMHLRVYRERPDVKSVIHAHPPFGTGFAIVGIPLTKPI
MPEAVISLGCVPPIAEYGTPSTEELPDAVSKYLQNYDALLLENHGALTYGPDLISAYYKME
SLEFYAKLTFISTLLGGPKELSDSQVEKLYEIRRKFGLKGRHPGDLCSTLGCSTNSAKSND
DDISELVNVITKKVLEQLKYN

[0201] SEQ ID NO:65:

GTGTATTCTGAATATGAGGTAAAAAAACAGATCTGCGAAATAGGAAAGAGAA
TCTACATGAATGGGTTTGTGGCAGCGAATGACGGCAATATCACCGTTAGGATTGGTG
AAAATGAAATAATAACGACGCCTACCGGTGTCAGCAAAGGTTTCATGACTCCAGAC
ATGCTATTAAATATTAATTTAAACGGTGAAGTATTAAAATCTTCAGGCGACTACAAA
CCGTCCACAGAAATAAAGATGCATCTTAGAGTCTATAGAGAAAGGCCAGATGTCAA
ATCAGTCATACATGCACATCCACCATTTGGCACAGGTTTTGCTATTGTAGGGATCCC
GCTTACAAAGCCAATAATGCCAGAAGCAGTTATATCTTTAGGCTGTGTGCCGATAGC
CGAATACGGGACGCCTTCTACAGAAGAGCTGCCAGATGCCGTCTCTAAATATTTGCA
AAATTACGATGCGCTTTTATTAGAAAATCATGGTGCGTTGACATACGGTCCTGATTT
AATTAGCGCATACTACAAGATGGAATCACTTGAATTTTACGCAAAATTGACATTTAT
TTCTACACTTCTCGGAGGTCCAAAAGAATTATCAGATAGCCAAGTAGAAAAGCTTTA

TGAAATTAGGAGAAAATTCGGTTTAAAAGGAAGACATCCAGGCGATTTGTGCAGTA
CATTAGGATGCAGCACAAATTCTGCAAAATCGAATGATGATGACATTTCTGAACTTG
TGAATGTTATCACTAAGAAAGTATTAGAACAATTGAAATACAATTAA

[0202] or226

[0203] SEQ ID NO:66:

MKHSKRFEVLGKRPNVDGFINWPEKGFIAMCSPNDPKPSIKIENDKIVEMDGK
RREDFDFIDLFIADHAINIYQAEKSMKMNSLDIAKMLVDINVERKTIKVVSGLTPAKIME
VVNHLNVVEMMMAMQKMRARKIPANQSHITNLKDNPVQIAADAAECALRGFREEETT
VGVTKYAPFNIAIALLIGSQALKRGVLTQCAVEEATELELGMRGFTTYAETISVYGTESVF
IDGDDTPYSKAFLASAYASRGLKMRFTSGTGSEVLMGNAEGKSMLYLEIRCIMVTKGAG
VQGLQNGAISCIGITSSVPSGIRAVLAENLIASMLDLEVASGNDQTFTHSDIRRTARTMMQ
FLPGTDFIFSGYSGTPNYDNMFAGSNFDAEDFDDYNVLQRDLMVDGGLRPVKEEDVVE
VRRKAAKALQDVFRELNLGVVTDEEVEAAAYAHGSKDMPERDVLSDESIDEMMKRGI
TGIDIVKALYRSGHEDIAENILNMLKQRISGDYQLQTSAILDEDFNVISAINCPNDYLGPGT
GYRIDKDRWEEIKNIPYTINPDNL

[0204] SEQ ID NO:67:

ATGAAACATTCTAAGCGATTTGAGGTTCTCGGCAAAAGACCTGTAAATCAGG
ATGGATTTATAAATGAATGGCCAGAAAAAGGCTTCATAGCAATGTGTAGTCCCAATG
ATCCTAAGCCATCAATAAAGATTGAAAACGACAAGATCGTTGAGATGGATGGGAAG
AGAAGAGAAGACTTTGATTTTATAGATTTATTCATAGCTGATCACGCTATAAATATTT
ATCAGGCTGAGAAATCCATGAAAATGAACTCGCTTGATATAGCCAAAATGCTTGTAG
ATATAAATGTAGAGAGAGAAAGACTATAATAAAAGTAGTTTCGGGACTTACACCTGCC
AAAATAATGGAAGTTGTAAATCATCTTAATGTCGTTGAAATGATGATGGCTATGCAG
AAAATGCGAGCAAGAAAGATTCCGGCTAATCAATCACATATTACAAATCTTAAAGA
TAATCCTGTGCAGATTGCAGCGGATGCTGCCGAATGTGCTTTAAGAGGTTTTAGGGA
AGAAGAGACCACCGTAGGAGTGACAAAATATGCTCCGTTTAATGCAATAGCGTTATT
GATAGGGTCTCAGGCATTAAAAAGAGGGCGTGCTTACTCAATGTGCTGTTGAGGAGGC
GACGGAACCTGAATTAGGCATGAGGGGATTTACCACATACGCTGAGACTATATCTGT
TTATGGAACCTGAAAGTGTTTTTATAGATGGTGACGATACACCTTACTCCAAAGCATT
CCTTGCTTCTGCTTATGCGTCAAGAGGATTGAAAATGAGGTTTACGTCAGGTACAGG

TTCAGAAGTTCTTATGGGAAATGCAGAGGGTAAATCGATGTTGTACCTGGAAATCAG
GTGCATCATGGTTACAAAAGGTGCAGGAGTGCAGGGGCTTCAAAATGGTGCAATAA
GCTGTATAGGCATAACTAGCTCAGTTCCTTCAGGTATAAGGGCGGTGCTGGCTGAAA
ACCTTATAGCATCTATGCTTGATTTAGAGGTAGCATCAGGCAATGATCAGACTTTTA
CACATTCAGACATAAGAAGGACAGCAAGGACTATGATGCAGTTTTTACCCGGTACTG
ATTCATATTTTCAGGTTACAGTGGAACGCCTAATTATGACAATATGTTTGCAGGTTC
CAATTTTGATGCAGAAGATTTTGATGACTACAATGTACTGCAAAGGGATTTAATGGT
AGATGGAGGGTTAAGGCCTGTAAAAGAAGAAGATGTGGTAGAAGTGAGGCGAAAG
GCAGCTAAAGCTTTGCAGGATGTATTTAGAGAGTTAAATCTTGGAGTAGTTACAGAT
GAAGAAGTAGAAGCAGCAGCATATGCACACGGCAGCAAAGATATGCCTGAAAGAG
ATGTTTTGTCTGACCTTGAATCAATCGATGAGATGATGAAAAGAGGGATTACAGGCA
TTGACATCGTAAAGGCTTTATATAGATCTGGACATGAGGATATAGCGGAAAACATTT
TAAACATGTTAAACAGCGCATATCTGGAGACTATTTGCAGACATCAGCTATTCTTG
ATGAAGATTTTAATGTTATAAGCGCCATAAATTGTCCAAATGATTACTTAGGACCTG
GAACAGGATATAGGATTGATAAAGATAGATGGGAAGAGATAAAGAATATTCCTTAC
ACCATTAATCCTGACAATTTGTAA

[0205] or225

[0206] SEQ ID NO:68:

MYVDEELLKEITKRVI EELNNKHKTDNVPSYFIENG VAYK GKNI E EVVIGVGP AF
GKHIKKTINGLDHRDVIKEIIAGIEEEGMVHRIVRVLKTS DVAFIGKEAALLSGSGIGIGIQ
SKGTTVIHQKDL YPLSNLELFPQAPLLNLELYREIGKNAARYAKGMMVKPILIQNDY MV
RPKYQVKAAIMHIKETEKILKNAQSIQLTIDL

[0207] SEQ ID NO:69:

ATGTACGTAGATGAAGAACTGTTAAAAGAAATTACTAAACGTGTTATAGAAG
AATTAAATAATAAGCATAAA ACTGATAATGTGCCTTCGTATTTTATTGAAAATGGAG
TTGCCTATAAGGGTAAAAAATATAGAGGAAGTCGTCATTGGTGTTGGGCCTGCATTTG
GAAAGCATATAAAAAAAGACTATAAATGGCCTTGACCATAGAGATGTCATAAAAGAA
ATAATTGCAGGCATCGAAGAAGAAGGTATGGTTCATAGAATTGTAAGAGTTCTAAA
GACTTCTGATGTGGCGTTCATAGGCAAAGAAGCTGCTTTATTAAGCGGATCGGGAAT
AGGCATAGGCATACAATCAAAAGGTACTACAGTGATTCATCAAAAAGATTTATATCC

TTTAAGCAATTTAGAACTGTTTCCACAAGCTCCACTGCTAAATTTAGAAATTATACAG
GGAAATAGGCAAAAATGCGGCGAGATATGCTAAAGGCATGATGGTAAAGCCTATTT
TGATTCAAAATGATTACATGGTGAGACCTAAATACCAAGTGAAAGCTGCTATAATGC
ATATAAAAGAGACGGAAAAGATATTGAAAAATGCTCAATCAATCCAATTGACGATA
GACTTGTA

[0208] or224

[0209] SEQ ID NO:70:

MEEYPLSKSAFDKLVTKTGKHLNEINIENVMKGNV KPDDIKISKEVLLMQGQIAE
RYGRHQMKENFTRASELTDVPDEKILEIYESLRPFRSTKEELINLAYELRDKYNAINCANL
ILEAAEVYEKRNILKT

[0210] SEQ ID NO:71:

ATGGAAGAATATCCGCTATCAAAAAGTGCTTTTGATAAATTGGTGACAAAAA
CAGGCAAACATTTGAATGAAATAAATATTGAAAATGTAATGAAGGGAAACGTAAAA
CCCGATGATATCAAGATATCCAAAGAAGTGCTTTTAATGCAAGGGCAAATTGCAGA
AAGATACGGCAGGCATCAGATGAAGGAGAATTCACAAGAGCATCGGAGCTTACAG
ATGTTCCAGATGAAAAGATTTTGGAATATATGAGAGCTTAAGGCCGTTTAGATCTA
CAAAGGAAGAGCTTATAAATCTTGCCTATGAATTAAGAGATAAGTACAATGCCATTA
ACTGTGCAAACCTTGATACTTGAGGCTGCTGAAGTATATGAAAAAAGAAATATTTTGA
AAACTTAA

[0211] or223

[0212] SEQ ID NO:72:

MKLIAGVDIGNSTTEVCIAAIKDDNTLEFLSSSLTATTGVKGTVDNVTGVINGLTE
ALKKIGKNIRDLSLIRINEAAPVVCGAAMETITETVITGSTMIGHNPSTPGGVGLGVGEIIH
INDLADATKGKNYIVVIPKEIGYEEASIMINKSFENDIDVKAIVQSDEAVLINRLKKIPI
VDEVQRQIEKIPSGVVAAVEVAPEGKSISTLSNPYGIATIFDLTPEETKYVIPISKSLMGKKS
AVVIKTTPRGQVKERIIPAGNLLIMGPTMSSKVSVD SGAEAIMESVEEVGTIDDVEGEENT
NVGNMIKNLKNKMANITGQKVDKIKIKDIFAVDTPVPKVEGGLAGETSM EKAVVLAA
MVKTDTLPMIEIAEKLQRKLGVFVKIAGVEAVMATLGALTTPGTKLPLAILDIGGGSTDA
ALIDEKGIVKSIHMAGAGELVTMLIDSELGLNDRYLSEEIKRNPIGKVESLFHIRMENREI

KFFDKPLNPRYYGRIVILKENDMIPVFKEDLTMEKIIYVRRQAKDKVFKNAIRALKKIA
PENNLRRIPNVVLVGGSALDFEIPMILSELSKYKIIAGRGNIRKIEGPRNAVATGLVMSY
LG

[0213] SEQ ID NO:73:

ATGAAACTCATAGCAGGTGTTGATATTGGCAATTCTACAACAGAAGTGTGTAT
AGCCGCTATTAAAGATGACAATACATTAGAATTTTAAAGCAGTTCCTTGACAGCTAC
GACAGGTGTAAAAGGCACTGTGGATAATGTGACAGGGGTTATTAATGGATTGACTG
AGGCACTAAAAAAAATTGGCAAGAATATTAGGGATTTAAGCCTCATTAGAATCAAT
GAAGCCGCCCCAGTTGTCTGTGGTGCTGCTATGGAGACAATAACGGAAACTGTTATC
ACTGGTTCGACTATGATAGGTCATAATCCATCCACGCCGGGTGGTGTCGGACTTGGA
GTAGGCGAGATAATACATATAAATGATTTAGCTGATGCTACTAAAGGCAAAAATTAC
ATTGTGGTTATACCTAAGGAGATTGGCTATGAAGAAGCTTCAATAATGATAAACAAA
TCTTTTGAAAACGATATTGATGTAAAAGCTGCTATAGTTCAAAGCGATGAAGCAGTT
TTAATCAACAACAGGCTTAAAAAGATTATACCAATTGTTGACGAAGTAAGGCAGAT
AGAAAAGATTCCATCGGGTGTTGTAGCGGCTGTAGAGGTGGCACCAGAAGGCAAGT
CCATAAGCACGTTATCAAATCCTTATGGTATCGCAACAATATTTGACTTAACTCCAG
AAGAGACAAAGTATGTCATACCGATTTCGAAAAGTTTGATGGGGAAAAAGTCAGCA
GTTGTCATAAAAACACCGAGGGGACAAGTGAAAGAAAGAATAATTCCGGCTGGTAA
TCTCTTAATCATGGGGCCTACTATGTCATCAAAAGTAAGTGTTGATTCTGGTGCTGAA
GCTATAATGGAATCAGTTGAAGAAGTCGGCACAATTGATGACGTAGAAGGTGAAGA
AAATACAAATGTTGGGAATATGATAAAAAATCTAAAAAACAAGATGGCAAATATAA
CTGGGCAAAAAGTAGATAAGATAAAGATTAAAGATATCTTCGCTGTTGATACGACA
GTCCCTGTTAAAGTAGAGGGCGGACTTGCTGGTGAGACTTCAATGGAAAAAGCAGT
CGTGTTGGCGGCTATGGTAAAGACAGATACGCTTCGATGATAGAAATTGCAGAAAA
GCTTCAAAGAAAGTTGGGTGTATTTGTAAAAATAGCTGGAGTAGAAGCTGTGATGGC
TACATTAGGTGCGCTTACAACCTCCAGGCACAAAGTTGCCACTTGCAATACTGGATAT
CGGTGGGGGTTCTACAGATGCAGCTTTGATTGATGAAAAAGGCATTGTAAAATCTAT
ACACATGGCAGGTGCTGGAGAATTAGTCACAATGCTTATTGATTCAGAATTAGGGTT
AAATGATAGATATTTGTCTGAAGAAATAAAGAGAAATCCGATTGGAAAAGTTGAAA
GCCTATTTACATAAGAATGGAAAATAGGGAGATAAAGTTTTTTTGACAAACCTTTAA
ATCCTCGATATTACGGTAGGATCGTAATTTTAAAAGAAAATGACATGATCCCTGTAT

TTAAAGAAGATTTGACAATGGAAAAGATTATTTACGTGCGAAGACAAGCGAAGGAT
AAAGTTTTTCGTAAAAAATGCTATTAGAGCTTTGAAAAAAATTGCTCCGGAAAATAAT
TTAAGGCGAATACCAAATGTAGTCTTGGTTGGCGGTTCTGCTTTGGACTTTGAAATTC
CAGAGATGATTTTATCAGAGCTATCAAAATACAAAATCATAGCAGGCAGAGGGAAT
ATAAGAAAAATCGAAGGGCCAAGAAATGCTGTAGCGACAGGTCTTGTGATGTCTTA
TTAGGGTGA

[0214] or222

[0215] SEQ ID NO:74:

MEFIKPQIVIFANTENKYIINEVIAGIEEEGALYRLSYNECADVMKMAYDAAKAS
VLGIGIGISGDLVCLHSKNLEINTPLILSKTSENFDPRLVGCNAAKYVKGLPLKYLD

[0216] SEQ ID NO:75:

ATGGAATTTATAAAGCCTCAAATAGTGATTTTTTGCAAATACAGAAAACAAAT
ATATAATAAACGAGGTTATAGCTGGCATTGAAGAAGAAGGTGCATTATATAGATTAT
CTTACAATGAATGTGCTGATGTTATGAAAATGGCTTATGATGCAGCAAAAGCATCTG
TATTAGGTATCGGAATAGGCATATCTGGAGATTTAGTGTGTTTGCACCTCTAAAACT
TGGAAATCAATACACCTTTGATTCTTTCAAAGACAAGTGAAAACCTTTGATCCACGAC
TCGTTGGATGCAATGCTGCAAAATATGTAAAGGGTTTGCCACTTAAATACTTAGATT
AG

[0217] or221

[0218] SEQ ID NO:76:

MSVYTKTGDDGYTLNLLNGERIPKDDLRIETLGNLDELTSYLGFAKAQINDDSIKK
R

[0219] SEQ ID NO:77:

ATGAGTGTTTATACTAAAACTGGTGATGATGGTTACACGTTGCTATTAAATGG
AGAAAGAATTCCAAAGGACGATTTGAGAATAGAGACATTGGGAAATTTGGATGAAT
TGACAAGCTATTTAGGATTTGCAAAAGCTCAAATAAATGATGATTCCATAAAAAAGA
GATAG

[0220] or220

[0221] SEQ ID NO:78:

MVKIKNGFVIPGKNQISALLDIVRTITRKTERSLIKVDKKYPVNINSKVYINRLSDY
LFVLARYMEIRTEIEEKVKDVIRKHYGKNKGKGLNLDIAKNLMAKVEKKAESINLPVAI
AIVDMHGNNLIAAHFMDGTLLESMNLAINKAYTSVVLKMSTQELSKLAQPGQPLYGINTT
DNRIVVFGGGCPIKHQGEIVGGIGVSGGTVEQDIELSIYGADVFEVIS

[0222] SEQ ID NO:79:

ATGGTAAAGATTAAAAATGGTTTTGTAATACCTGGTAAAAACCAAATCTCAG
CATTATTAGATATTGTAAGGACTATAACGAGAAAAACTGAGAGAAGCTTAATCAAA
GTTGACAAGAAATATCCTGTAAATATTAATTCGAAAGTTTACATCAATAGATTGTCT
GATTATTTGTTTGTTTTAGCAAGGTATATGGAAATAAGAACGGAAATAGAAGAAAA
AGTAAAAGACGTGATAAGAAAGCATTATGGAAAGAACAAGGCGAAATAAAGCTA
AATTTAGATATAGCAAAAAATTTAATGGCTAAGGTAGAAAAGAAGGCAGAAAGCAT
TAATCTACCGGTTGCTATTGCAATAGTTGACATGCATGGCAATTTGATAGCGGCTCA
TTTTATGGATGGTACACTTCTTGAAAGCATGAATCTAGCTATAAATAAAGCTTATAC
ATCAGTGGTGCTTAAAATGTGCGACGCAAGAGTTATCAAACTTGCACAACCAGGGC
AGCCTCTTTACGGGATAAATACAACCTGATAATAGAATCGTAGTGTTTGGAGGTGGGT
GCCCTATAAAACATCAAGGTGAAATAGTTGGTGGAATTGGAGTTAGCGGTGGTACA
GTAGAACAAGATATAGAACTTTCTATTTATGGTGCAGATGTATTTGAGGAGGTTATA
TCATGA

[0223] or219

[0224] SEQ ID NO:80:

MKVKEEDIEAIVKKVLSEFNFEKNTKSFRDFGVFQDMNDAIRAAKDAQKKLRNM
SMESREKIIQNIRKKIMENKKILAEMGVSETGMGKVEHKIIKHELVALKTPGTEDIVTTA
WSGDKGLTLVEMGPFVIGTITPSTNPSETVLCNSIGMIAAGNSVVFNPHPGAVNVSNYA
VKLVNEAVMEAGGPENLVASVEKPTLETGNIMFKSPDVSLLVATGGPGVVTSVLSSGKR
AIGAGAGNPPVVVDETADIKKAAKDIVDGATFDNNLPCIAEKEVVSVDKITDELIYYMQ
QNGCYKIEGREIEKLIELVLDHKGGKITLNRKWVGKDAHLILKAIGIDADESVRCIIFAE
KDNPLVVEELMMPILGIVRAKNVDEAIMIATELEHGNNRHSAHMHSKNVDNLTGFGKIID
TAIFVKNAPSYAALGYGGEGYCTFTIASRTGEGLTSAFTFTKSRRCVLADGLSIR

[0225] SEQ ID NO: 81:

ATGAAAGTTAAAGAGGAAGATATTGAAGCGATCGTCAAAAAAGTCTTATCGG
AATTTAATTTTGAaaaaaaATACTAAAAGTTTCAGAGATTTTGGCGTATTTCAAGATA
TGAATGATGCTATTCGTGCTGCAAAAGATGCCCAGAAAAAATTGAGAAATATGTCCA
TGGAGTCGAGAGAAAAGATTATACAGAATATAAGAAAAAAGATTATGGAGAATAAA
AAAATACTTGCAGAGATGGGCGTCAGTGAAACTGGCATGGGGAAAGTAGAGCACAA
AATAATAAAACATGAGCTTGTAGCACTTAAGACACCTGGTACCGAAGATATAGTGA
CAACAGCATGGTCTGGCGATAAGGGACTGACATTGGTTGAAATGGGGCCATTTGGTG
TAATAGGTACGATTACTCCTTCGACAAATCCAAGTGAAACCGTCCTTTGCAATAGCA
TAGGTATGATAGCCGCAGGTAATTCAGTCGTATTTAATCCACATCCAGGTGCGGTAA
ATGTATCTAATTACGCTGTCAAGTTAGTAAATGAAGCGGTGATGGAAGCTGGCGGCC
CTGAGAATTTAGTCGCATCTGTTGAAAAACCTACACTTGAAACTGGAAATATTATGT
TCAAGAGTCCTGATGTTTCGCTATTAGTAGCGACAGGCGGACCTGGTGTAGTAACAT
CGGTTCTCTCATCTGGCAAAAGGGCAATAGGAGCAGGAGCAGGAAATCCACCAGTT
GTAGTTGATGAAACGGCAGATATAAAAAAAGCTGCGAAAGATATAGTCGATGGTGC
TACATTTGACAACAATTTGCCTTGTATTGCTGAAAAGGAAGTAGTTTCTGTAGATAA
AATAACAGATGAACTGATTTACTACATGCAACAGAATGGCTGCTACAAGATTGAGG
GGCGAGAAATTGAAAAGCTCATTGAACTTGTATTGGATCACAAAGGTGGCAAGATA
ACATTAAACAGGAAATGGGTTGGCAAAGATGCTCATTTAATACTAAAAGCTATAGG
CATAGATGCTGATGAAAGCGTAAGGTGCATAATTTTGGAGGCGGAAAAAGACAATC
CGTTAGTGGTAGAAGAGCTGATGATGCCTATTTTAGGAATAGTAAGAGCCAAAAAT
GTAGATGAAGCGATAATGATTGCGACAGAGTTAGAACATGGCAATAGGCATTCAGC
ACATATGCATTCTAAAAACGTTGATAATTTAACAAAGTTTGGAAAAATAATTGACAC
TGCTATATTTGTAAAAAATGCTCCATCGTATGCCGCGTTAGGATATGGTGGTGAAGG
TTATTGCACATTTACGATTGCAAGCAGAACAGGTGAAGGATTGACATCTGCAAGGAC
TTTTACTAAAAGTCGTAGATGTGTCTTGGCAGATGGATTATCAATAAGATAG

[0226] or218

[0227] SEQ ID NO:82:

MEVNQIDIEEIVKKILNDLRNEPKENIKESNSKIPSICRAAVLTDVKKIEVKEFNIPEI
NDDEMLVKVEGCGVCGTDVHEYKGDPFGLIPLVLGHEGTGEIVKLGKNVRRDSAGKEI
KEGDKIVTSVVPCGECDICLNHPDKTNLCENSKIYGLISDDNYHLNGWFSEYIVIRKGSTF

YKVNDINLNLRLLEPAAVVVHAVERAKSTGLMKFNSKVLVQGCPIGLLLLSVVKTL
GVENIIAVDGDENRLNMAKRLGATALINFTKYSNIDELVDAVKKASDGIGADFAFQCTG
VPSAASNIWKFVRRGGGLCEVGFFVNNGDCKINPHYDICNKEITAVGSWTYTPQDYLT
FDFLKRAKEIGLPIEELITHRFSLDKMNEAMEVNMKQEGIKVVYINDRF

[0228] SEQ ID NO:83:

ATGGAAGTCAATCAGATAGACATTGAGGAGATAGTTAAGAAAATATTAAATG
ATTTAAGAAATGAGCCTAAAGAAAACATTAAAGAGAGCAATTCAAAAATACCATCT
ATCTGCAGAGCTGCTGTACTTACAGATGTTAAAAAAATAGAAGTAAAAGAATTTAAT
ATTCCAGAAATAAATGATGATGAAATGCTTGTCAAGGTGGAAGGCTGTGGCGTTTGC
GGTACTGATGTTTCATGAATACAAAGGAGATCCTTTTGGACTTATACCATTTGGTTTTAG
GACACGAAGGTACAGGTGAGATAGTCAAGCTGGGGAAAAACGTGAGACGAGATTCT
GCTGGTAAAGAAATCAAAGAAGGCGATAAGATTGTTACATCTGTCGTTCCGTGCGGT
GAATGCGATATATGTTTGAATCATCCAGACAAGACAAATTTGTGTGAAAACCTCAAAG
ATTTACGGCTTAATATCCGATGATAATTACCATTTAAATGGTTGGTTCTCAGAGTACA
TCGTCATAAGGAAAGGCTCAACATTTTATAAGGTCAATGATATAAACCTTAATTTGA
GGCTTTTGGTAGAACCGGCTGCAGTAGTCGTACATGCAGTAGAGCGCGCAAAATCCA
CAGGTCTTATGAAATTCAACAGTAAAGTTCTCGTACAAGGCTGTGGCCCTATAGGAT
TACTGCTATTGTCGGTTGTAAAGACGCTTGGAGTAGAAAAATATCATAGCCGTCGACG
GCGATGAGAATAGACTCAACATGGCTAAAAGATTAGGTGCTACAGCACTCATTAATT
TTACTAAATACAGCAATATTGATGAGCTTGTTGATGCTGTTAAAAAAGCAAGCGATG
GAATTGGCGCAGATTTTGCATTTCAATGTACAGGCGTTCCTTCTGCAGCGTCTAATAT
TTGGAAGTTTGTAAAGGCGGGGAGGTGGTTTATGCGAAGTTGGATTTTTTTGTAAATAA
TGGTGATTGTAAGATAAACCCCCATTATGATATTTGCAATAAGGAGATAACAGCAGT
TGGCTCATGGACTTACACTCCTCAAGACTATTTGACAACCTTTTGATTTTCTCAAAGA
GCTAAAGAAATAGGACTTCCAATTGAAGAGCTGATAACACATAGATTTTCACTTGAT
AAAATGAATGAAGCTATGGAAGTTAATATGAAGCAGGAAGGGATAAAAGTAGTGTA
TATAAATGACAGATTTTAG

[0229] or217

[0230] SEQ ID NO:84:

MQAVGLIEVYGLVAAFVAADAACKKANVVIESFDNNKPLNAEALPVPLIIVVKL
RGDLEDVKIAVDAAVDAANKISGVVATNIIAKPEEDTEKLLKLNCLK

[0231] SEQ ID NO:85:

ATGCAGGCTGTTGGATTGATTGAAGTTTATGGATTAGTAGCGGCATTTGTGGC
AGCAGATGCTGCATGCAAAAAAGCGAATGTCGTAATAGAGTCTTTTGACAACAATA
AGCCATTAAATGCTGAAGCATTGCCAGTTCCATTGATAATAGTCGTTAAGCTCAGAG
GAGATCTTGAGGATGTAAAAATAGCGGTAGATGCTGCAGTTGATGCAGCTAATAAA
ATATCTGGTGTAGTTGCTACAAATATAATAGCAAAACCAGAAGAAGATACTGAAAA
GCTATTAAAGCTAAATTGTCTTAAATAA

[0232] or216

[0233] SEQ ID NO:86:

MVQEALGMVETRGLVAAIEAADAMVKAADVTLIGTEKIGSGLVTVMVRGDVG
AVKAATEVGASAASKLGLVAVHVIPRPHTDVEKILPTIK

[0234] SEQ ID NO:87:

ATGGTACAAGAAGCATTGGGAATGGTAGAAACGAGAGGATTGGTAGCAGCA
ATAGAAGCAGCAGATGCTATGGTAAAGGCTGCGGATGTCACCTTTGATAGGAACTGA
AAAAATAGGTTTCAGGACTTGTAACAGTCATGGTAAGAGGAGATGTCGGTGCAGTAA
AAGCAGCGACAGAAGTTGGCGCAAGTGCAGCTTCAAAATTGGGAGAGTTAGTGGCT
GTTACGTAATACCAAGGCCTCATACTGATGTTGAAAAGATACTGCCGACAATTAAA
TAA

[0235] or215

[0236] SEQ ID NO:88:

MYAIGLIEVNGFVTAVETLDAMLKTANVEFVTWEKKLGGRLVTIIKGDVSAVEE
AILTGKIEADKITRTVAYAVIPNPHPETIKMVNISAGKLFKADGGEINEF

[0237] SEQ ID NO:89:

ATGTATGCAATTGGACTTATTGAAGTAAATGGGTTTGTACACAGCGGTTGAAAC
ACTGGATGCAATGTTGAAAACAGCCAATGTAGAGTTTGTAAACATGGGAGAAAAAAC

TTGGAGGCAGACTTGTGACAATCATTATTAAAGGAGATGTTTCAGCAGTTGAAGAAG
CAATTTTAACTGGAAAGATTGAAGCTGACAAGATTACACGGACAGTAGCATACGCA
GTTATTCCAAATCCACATCCAGAACTATAAAGATGGTAAATATTAGTGCAGGAAAG
CTATTTAAAGCAGATGGTGGTGAAATAAATGAGTTCTGA

[0238] or214

[0239] SEQ ID NO:90:

MSSEEKDTNAKDVKVEKQKNNLTKTSNKEFKEELIMEQQALGMVETRGLVAAIE
AADAMVKAANVTLIGTEKIGSGLVTVMVRGDVGAVKAATETGANAAKKLGELVAVH
VIPRPHADVEKILPTIK

[0240] SEQ ID NO:91:

ATGAGTTCTGAAGAAAAGGATACGAATGCAAAAGATGTTAAAGTCGAAAAG
CAGAAAAATAATTTAACGAAAACATCAAATAAAGAATTTAAGGAGGAATTGATTAT
GGAACAACAAGCATTAGGAATGGTAGAGACGAGAGGATTGGTAGCAGCGATAGAA
GCTGCTGATGCAATGGTAAAGGCTGCTAATGTCACGTTAATAGGAACTGAAAAAAT
AGG TTCAGGACTTGTAACAGTCATGGTAAGAGGAGATGTTGGTGCAGTAAAAGCAG
CGACAGAGACTGGAGCAAATGCAGCTAAAAAGTTAGGGGAGTTAGTAGCTGTTCAC
GTAATACCAAGACCTCATGCAGATGTAGAGAAAATACTGCCTACGATAAAGTAG

[0241] or213

[0242] SEQ ID NO:92:

VITVNEKLIEIISKTIADTISERNSLKIPVGV SARHVHLTKEHLDILFGKDYILKKKK
ELMGGQFAAECEVTIIGFKLNAIEKVRVLGPLRDKTQVEISKTD AISLGLNPPIRESGDIKG
SSPITIVGPRGAISLKEGCIIAKRHHMSPEDSKRFNVKDDDIISVKINGQRGGILENVQIRV
DEKYTLEMHIDTDEANCMGLKSGDFVEIVRDNRS

[0243] SEQ ID NO:93:

GTGATAACAGTGAACGAAAAATTGATAGAGATTATATCAAAA ACTATAGCGG
ATACGATTAGTGAAAGGAATTCGCTTAAGATACCAGTAGGCGTATCAGCCCGACATG
TACATCTGACTAAAGAACATTTGGATATATTATTTGGAAAAGATTATATCCTTAAAA
AGAAAAAGGAATTGATGGGTGGACAGTTCGCAGCAGAGGAATGTGTGACAATTATC

GGATTTAAATTAAATGCTATTGAGAAAGTGAGAGTTTTGGGTCCTTTAAGAGATAAA
ACGCAGGTAGAAATATCGAAGACCGATGCAATAAGTTTAGGGTTAAACCCTCCTATA
CGGGAATCAGGTGATATAAAAGGTTTCATCGCCAATTACAATTGTAGGGCCGAGAGG
AGCAATATCATTTAAAAGAAGGATGTATAATAGCAAAACGACATATTCACATGTCAC
CGGAAGATTCCAAAAGATTCAATGTTAAAGACGACGATATAATATCAGTAAAAATA
AATGGTCAGCGAGGCGGAATTTTAGAAAATGTACAGATTAGAGTTGACGAAAAGTA
TACACTTGAGATGCATATTGACACAGATGAAGCTAATTGCATGGGACTAAAAAGCG
GCGATTTTGTTGAAATAGTAAGAGATAATAGGAGTTGA

[0244] or212

[0245] SEQ ID NO:94:

LIIAKVVGTVISTRKNQNLIGNKFLIVEPVSEMNYDSKNRVVAIDNVGAGVGEIVL
VTFGSSARIGCGMPDSPVDAAIVGIVDSIKDIIIDD

[0246] SEQ ID NO:95:

TTGATAATAGCTAAAGTTGTTGGTACTGTTATTTCTACCCGCAAGAATCAAAA
TTTAATAGGCAATAAATTTTAAATAGTAGAACCAGTAAGTGAAATGAATTATGACAG
TAAAAATAGGGTTGTTGCAATAGATAATGTAGGTGCAGGTGTAGGAGAGATAGTAT
TAGTTACCTTTGGAAGTTCAGCAAGAATCGGTTGTGGTATGCCAGATTCGCCTGTAG
ATGCGGCAATTGTCGGAATTGTTGATAGCATAAAAGATATTATCATTGATGATTAG

[0247] or211

[0248] SEQ ID NO:96:

MMNIDELKNIVFENGIVGAGGAGFPTHAKLTTGIDTIILNGAECEPLL RVDRQLLA
IYTDEILMTLSFIVDTLGAKRGIVAISAYKTAISSVKNLIGNYKNLELKVLPDVYPAGDE
VVLIYETTGRIVPEGSIPISVGTLVMNVETVLNVYNAILKHPVTEKYVTVTGDKYPSTF
KAKVGTSVARLIEKAGGCLEKDCEVIMGGPMTGKIVDVKTPITKTTKAIIVLPKDHPVIT
KRKTNIRIGLKRAMSVCSQCQMCTDLCPRNLLGHSIKPHKVMNAVANSIIDDTAA YTMT
MLCSECGLCEMYSCHQSLSPRKIISQIKIKLRQNGVKNPHNKR PETANVMRDERLVPME
RLISRLSLKKYDVDAPMNFDTVIPSHHVVMQLSQHVGA KAIPVVKVGDIVKEGDLIGDV
PNNKLGAKLHASIDGIIIDVTDDSIKPRGDFDGQSDRIG

[0249] SEQ ID NO:97:

ATGATGAATATTGATGAACTTAAAAATATCGTATTTGAAAATGGAATAGTCG
GTGCAGGCGGAGCTGGATTTCCTACACATGCAAACTTACTACAGGTATAGATACAA
TCATATTAAATGGCGCTGAATGTGAACCGCTTTTAAGAGTAGATAGGCAGCTACTTG
CAATATATACTGATGAAATATTGATGACTTTATCATTTCATAGTTGATACTTTAGGAGC
CAAACGTGGCATTGTAGCAATAAAAATCAGCATACAAAACCTGCCATCAGCTCAGTTAA
GAATTTGATTGGTAATTATAAAAACCTTGGAGTTAAAGGTATTGCCAGACGTTTATCC
TGCTGGTGATGAAGTTGTATTAATATATGAAACGACTGGAAGAATTGTGCCAGAAGG
TTCTATACCTATTTCTGTTGGCACGTTGGTAATGAATGTGGAAACTGTGCTTAATGTT
TATAATGCTATTTATTTAAAACATCCAGTCACAGAAAAGTATGTAAACAGTAACGGGA
GATGTCAAATATCCCAGCACATTTAAAGCAAAAGTAGGAACATCTGTAGCTCGTCTT
ATTGAAAAAGCAGGAGGATGCTTAGAAAAAGATTGTGAAGTGATAATGGGTGGTCC
TATGACTGGGAAAATAGTTGATGTAAAGACTCCAATAACAAAAACTACAAAAGCTA
TTATCGTTCTCCCAAAAAGACCACCCTGTGATAACAAAGAGAAAGACAAACATAAGG
ATAGGGTTAAAACGAGCAATGTCTGTTTGCTCTCAATGCCAAATGTGCACAGATCTA
TGTCCTAGAAATTTATTAGGTCATTCCATCAAACCTCATAAAGTCATGAATGCAGTT
GCAAATAGTATTATTGATGATACCGCTGCATATACGATGACAATGTTATGTTCTGAA
TGTGGATTGTGCGAGATGTATTCATGTCATCAAAGTTTGTCGCCGAGAAAGATAATA
AGCCAGATAAAGATAAAATTAAGGCAAAATGGTGTAAAAAAATCCACACAACAAAAG
ACCAGAAACAGCAAATGTCATGCGAGATGAGAGATTAGTGCCGATGGAAAGGCTTA
TTTCAAGACTTTCGCTCAAAAAAATACGATGTAGATGCTCCGATGAATTTTGATACTGT
TATTCCTTCACATCACGTTGTCATGCAACTAAGTCAGCATGTTGGTGCCAAAGCGAT
ACCTGTAGTAAAGGTAGGAGATATTGTGAAAGAAGGAGATCTGATAGGCGATGTGC
CTAATAATAAGCTGGGTGCTAAATTGCATGCCAGTATTGACGGCATTATAATAGATG
TAACTGATGACAGTATTGTTATCAAACCAAGAGGTGATTTTGATGGACAAAGCGATA
GGATTGGTTGA

[0250] or210

[0251] SEQ ID NO:98:

MDKAIGLVEYKSVATGITAADDMAKTADVEIIEAYTVCPGKYIVLLAGKLSAVN
SAIEKGINQYSENVIDSFILGNPHETIYKAMSGTSVIEDVEALGHIETFSAASIILAADTAAK

AAKVNLVEIRIARGMCGKSYLLLTGELA AVEASINAGCKALERTGMLLNKSIIPNPDRAI
WDKII

[0252] SEQ ID NO:99:

ATGGACAAAGCGATAGGATTGGTTGAATACAAATCAGTTGCTACAGGTATAA
CTGCTGCTGATGACATGGCTAAACTGCTGATGTGGAAATAATAGAAGCATATACAG
TATGTCCGGGGGAAATACATTGTTCTGTTAGCTGGGAAATTAAGTGCAGTTAATTCGG
CGATAGAAAAGGGCATAAATCAGTATTCGGAAAATGTCATTGATAGCTTTATATTGG
GAAATCCGCATGAAACAATATATAAAGCTATGAGTGGCACGTCTGTAATTGAAGAT
GTAGAAGCACTTGGTATCATAGAGACATTTTCTGCAGCATCAATAATACTTGCAGCA
GATACGGCTGCAAAAGCTGCAAAAGTGAATCTGGTAGAGATAAGAATAGCCAGAGG
TATGTGCGGCAAGTCATATCTACTGCTTACAGGAGAACTTGCTGCTGTTGAAGCATC
TATAAATGCAGGATGCAAAGCTTTGGAGAGAACGGGTATGCTTTTAAATAAGTCTAT
AATACCCAATCCAGATA GAGCTATTTGGGATAAGATAATTAA

[0253] or209

[0254] SEQ ID NO:100:

MYEAEKDKILNDYYNAKEIYAKFDIDIDKVLDKMKKIRISLHCWQGDDVTGF EK
SANGLSGGGILATGNWPGRARNGEELRQDIEKALSLIPGKHKINLHAIYAETDGEFVDRD
EINVEHFRKWIYWAKENGLGLDFNPTFFSHPKANDGYTLSSKDENIRKFWIQHGKRCREI
ANEIGRELKTQCVNNVWIPDGSKDLPANRIEHRKILKESLDEIFSVKYDKSNIVDSVESKL
FGIGSESYVVGSHFYMNYASRNDVMLCLDMGHFHPTENIADKISSILTFNDNLLIHVSR
GVRWDS DHVVILNEDLLSLAKEIRRC DAYDKVYIALDFFDASINRIMAWVIGARATLKAI
LISLLEPVHLLMEEENKGNFGARLALMEEFKTLPFYSVWNKYCMDENVPIGTSWIDDVK
EYEKEIVKNRA

[0255] SEQ ID NO:101:

ATGTATGAAGCAGAAAAAGATAAAATTTTAAATGATTATTATAATGCAAAAG
AGATTTATGCAAAGTTTGACATAGATATTGATAAAGTATTAGATAAAATGAAGAAG
ATTCGTATTTCACTTCACTGCTGGCAAGGCGATGATGTA ACTGGATT CGAAAAAAGT
GCCAATGGATTAAGCGGTGGAGGTATTTTGGCGACAGGAAACTGGCCTGGTAGAGC
AAGAAATGGTGAAGAATTAAGGCAAGACATTGAAAAAGCCTTAAGCCTTATACCAG

GCAAACACAAAATCAATTTACATGCCATTTACGCAGAAACGGATGGTGAATTTGTAG
ACAGAGATGAAATAAACGTGGAGCATTTCAGGAAATGGATTTACTGGGCAAAAGAA
AATGGCCTTGGCCTTGACTTCAATCCTACGTTTTTTTTTCGCATCCTAAAGCAAATGATG
GCTATACGCTTTCAAGCAAAGATGAAAACATAAGAAAATTTTGGATCCAACATGGTA
AAAGATGCCGTGAAATCGCAAATGAAATAGGAAGAGAGCTAAAAA ACTCAATGTGTG
AATAATGTTTGGATTCCTGATGGTTCAAAAGATTTGCCTGCTAATAGGATTGAACAC
AGAAAAATACTTAAAGAATCTTTAGATGAGATATTTTCAGTAAAATATGACAAATCA
AATATCGTTGATTCTGTTGAAAGCAAATTATTTGGCATTGGATCTGAAAGCTATGTG
GTTGGTTCACATGAGTTTTATATGAACTATGCGTCGAGAAATGATGTAATGCTGTGC
CTTGATATGGGACATTTTCATCCTACTGAGAATATTGCTGATAAGATATCATCAATAC
TTACATTCAATGACAATTTGTTGATTCATGTAAGCCGTGGTGTCCGGTGGGATAGCG
ACCATGTAGTCATTTTAAATGAAGATTTGCTTTCATTAGCAAAAGAAATAAGAAGAT
GTGATGCTTATGACAAAGTGTATATTGCATTAGATTTCTTTGATGCAAGCATAAATA
GGATAATGGCATGGGTAATAGGTGCAAGAGCGACGCTAAAAGCCATATTAATATCA
CTATTAGAGCCTGTGCATCTACTTATGGAAGAGGAGAATAAAGGAAATTTTGGTGCA
AGACTTGCTTTGATGGAGGAATTCAAAACATTGCCATTTTACTCTGTTTGGAACAAA
TACTGCATGGACGAAAATGTGCCTATTGGTACATCGTGGATTGATGATGTTAAAGAA
TATGAAAAAGAAATTGTAAAAAATAGGGCTTAA

[0256] or208

[0257] SEQ ID NO:102:

MKDIVYNLAFDFGASSGRLMLSAFDGEKITIEEIYRFPNEPVKLGQSFYWDFLRLF
HELKNGLKIASKRKIKISGIGIDTWGVVDYGLLDKNDQLISNPFHYRDKRTDGIKDFENM
ALLEEIYNVTGIQFMEFNTIFQLYCDYKKRPELLDNAKTLLFIPDLNFYLTNEKYNEYTV
ASTSQMLDANKKDWANDLIEKLNLP EGIFQKILMPGNTIGYLTKEIQEETGLSEVPVISVG
SHDTASAVAGTPIENGSSAYLICGTWSLLGVESEKPIINENTKKYNFTNEG GVEGLIRLLK
NINGLWIIQQLKQSWNSNGIKIGFPEISQMASKAEHEEFIINPDDKLFIAPDDMAE AIRQYC
TKTGQGLPQNIGDIARAA YNGIVEQYKNCLNNLEDIVGQEIDNIH MVGGGIQDKFLCKLT
ADVTGKKVITGPVEASIYGNVIVQLMALGYIKDLREGRKIIKNSIENDEEMFAK

[0258] SEQ ID NO:103:

ATGAAAGATATTGTGTATAATCTGGCTTTTGATTTTGGAGCTTCAAGTGGCCG
TCTTATGCTATCCGCGTTTGATGGCGAAAAAATCACAATTGAAGAGATTTATAGATT
TCCAAATGAGCCAGTCAAGCTGGGACAATCATTTTATTGGGATTTTTTAAGGCTTTTT
CACGAATTAAAAAACGGATTAAAAATAGCATCAAAGAGGAAAAATCAAAATATCCGG
CATTGGTATAGACACTTGGGGTGTCGATTATGGATTGCTTGATAAAAAATGATCAATT
GATTTCAAATCCTTTTCATTACAGAGATAAAAGAACGGATGGCATAATAAAAGATTT
TGAAAATATGGCGTTACTGGAGGAAATCTACAACGTAACCTGGTATACAGTTTATGGA
ATTTAATACAATATTCCAATTGTATTGCGATTATAAAAAGCGTCCAGAATTATTGGA
TAATGCAAAGACATTGTTGTTTATTCCAGATTTATTTAACTTTTATTTGACAAATGAG
AAATACAATGAATATACTGTTGCATCCACATCGCAAATGTTGGATGCTAACAAGAAA
GATTGGGCAAATGATCTTATAGAAAAGTTAAATTTGCCAGAAGGTATTTTTCAAAG
ATACTGATGCCAGGAAATACAATTGGTTATCTAACAAAAGAAATTCAAGAAGAAAC
AGGATTGTCTGAAGTTCCCGTGATTTCTGTTGGCAGCCATGATACGGCATCAGCAGT
TGCAGGTACACCTATTGAAAACGGTTCAAGTGCTTATTTGATTTGTGGTACTTGGTCA
TTATTAGGTGTTGAAAGTGAAAAACCTATAATAAATGAAAATACAAAGAAGTACAA
TTTTACAAATGAAGGCGGTGTCGAAGGCCTTATAAGGCTACTTAAAAATATTAATGG
TCTGTGGATAATTCAGCAATTAAAACAAAGTTGGAATTCAAATGGCATTAAAATAGG
ATTTCCAGAAATCAGCCAGATGGCATCTAAAGCAGAGCACGAAGAATTTATCATAA
ATCCTGATGACAAATTGTTTATAGCTCCAGATGATATGGCTGAGGCGATAAGGCAAT
ATTGTACAAAAACAGGACAGGGTTTGCCGCAGAATATTGGCGACATAGCAAGAGCC
GCTTACAATGGTATAGTTGAACAATACAAAAATTGCTTAAACAATTTAGAAGATATT
GTAGGGCAAGAAATAGATAATATTCACATGGTTGGTGGTGGGATACAGGATAAGTT
CCTGTGCAAGCTGACTGCAGATGTTACAGGGAAAAAAAGTCATAACAGGCCCTGTAG
AAGCTTCAATCTATGGCAATGTGATAGTCCAGCTTATGGCATTGGGATATATAAAAG
ACTTGAGAGAAGGAAGAAAGATAATAAAGAATTCTATAGAGAATGATGAAGAGATG
TTTGCTAAATAG

[0259] or207

[0260] SEQ ID NO:104:

VSNIYTLVVVEDEYEIRTGLVNCFPWNKMGFVVAEEFENGGECEFEYLCKNKVDT
ILCDIKMPVMSGIELAKKIFESNISTKIVIISGYTDFEYARQALRYGVKDYIVKPTKYNEIID
VFSRIKKELDNENTKEILNNSCNNEIDQYSSIISIIIEKYVDEHYRDVTLEDVAKVVYMNPY

YLSKYFKQKTGMNFSYITEVRMKKAVEFLKNPLYKTYEISYMIGYKNPKNFTRAFKKY
YKKSPREFVNSAINFKE

[0261] SEQ ID NO:105:

GTGTCTAATATTTATACGCTTGTAGTAGTAGAAGATGAATATGAGATAAGAA
CAGGATTAGTTAACTGCTTTCCATGGAACAAAATGGGTTTTTGTGTTGCAGAAGAAT
TTGAAAATGGAGGAGAATGTTTTGAGTATTTGTGTAAAAATAAGGTTGATACAATTT
TATGTGATATAAAAAATGCCAGTTATGTCTGGTATAGAGTTGGCAAAGAAAATTTTTG
AAAGTAATATAAGCACTAAAATAGTTATAATCAGTGGTTATACTGATTTTGAATATG
CCAGACAGGCGTTAAGATATGGTGTTAAAGATTATATAGTAAAACCTACTAAATATA
ATGAAATAATTGATGTTTTTCAGCAGAATAAAAAAAGAATTAGACAATGAAAATACA
AAGGAAATATTGAATAACTCATGTAACAATGAAATTGATCAGTACAGCAGCATAATT
TCAATCATAGAAAAATATGTTGATGAACATTACAGAGATGTGACATTGGAAGATGTA
GCTAAAGTAGTTTATATGAATCCGTATTATTTAAGCAAATATTTTAAACAAAAAACC
GGTATGAATTTTTCTGATTATATAACTGAGGTCAGAATGAAAAAAGCTGTAGAGTTT
CTAAAAAATCCTTTGTATAAAACTTATGAAATAAGTTATATGATTGGATATAAAAAAT
CCAAAAAATTTTACTAGAGCATTATAAAAAATATTATAAAAAATCCCCAAGAGAATTT
GTAAATTCAGCAATAAATTTTAAGGAATGA

[0262] or206

[0263] SEQ ID NO:106:

MRELNNKFFYKNLFLVLAALPLILIVIVLGSFSILITERYVRDEIYKNSREILKQSSNDL
SILFNDINKIYLTFGTNKDVTLYLERILNTNKYSLDDMWHLSMIESLFDSTSFSEPYIQSIY
LYFNNPNKNFLVTGNGINSVTNYIDNKWYDSFLNAPKDEISWIEVRNLKMYSFDKKGKIK
VLSIYKKIANFNGDKIDGVLVLNIYLDYIENLLNTSTIFPDQKILILDAHDNLCQNINGNFT
GKIDLDNYSKANIITKLESPNYNIKYVSIVPKKYL YEVPKLLKMTLVLLLT SIFFVILITFRI
TKRNYENVNKILKIIIEAEKTNEIFPEIPVESRDEYSYIINYNIINSYIEKSQLKMELAEKKYKM
KAMELLALQSQISPHFLSNALEIIYLRALS YTN GPNDVT KMIENLSQILKYLLSNPNETVT
VKEEIEN TKAYIQILKVR YRDKFKVNLIYDESILSCLMMKLMLQH LIENSIKHGLKKKNY
EGSIKIKIKAVDKKKIKISVIDNGIGMSKERLNYVKRILDSDFDFYEHIGLMNTNERLKL
YGKDCEILIRSKL NIGTAV YIIFPYQLKNQNND DYNK

[0264] SEQ ID NO:107:

ATGAGAGAATTAAACAATAAATTTTTTTTATAAAAATCTTTTTGTTTTGGCATT
GCCATTAATTTTAATTGTTATTGTATTAGGTTTCATTTTCAATATTAATAACAGAAAGA
TATGTTAGAGATGAAATATACAAAAATAGTAGAGAAATATTAAAGCAAAGCAGTAA
TGATTTGTCAATTTTATTTAATGATATAAATAAAAATTTATTTAACATTTGGAACAAAC
AAAGATGTGACATTGTATTTGGAAAGGATCTTAAATACAAATAAATATTCTTTAGAT
GATATGTGGCATCTTAGCATGATAGAAAGTTTATTTGATTCTACGTCGTTTTTCAGAAC
CTTATATACAATCAATTTATTTGTATTTTAACAATCCTAATAAAAATTTTTTTAGTGAC
AGGAAATGGTATTAATTCTGTAACAAATTATATTGATAATAAATGGTATGACAGCTT
TTTAAATGCACCAAAAGATGAGATTTCTTGGATAGAGGTTAGAAATTTAAAAATGTA
TAGTTTCGATAAAAAGGGGATAAAAAGTCCTAAGTATATACAAAAAAATTGCAAAC
TTAACGGGGATAAAAATTGATGGTGTGCTTGTACTAAATATATATTTGGACTATATTG
AAAATTTGCTAAATACTTCAACAATATTTCTGACCAAAAAAATTCTTATATTAGATGC
CCACGACAATTTAATATGTCAAAATATTAATGGGAATTTCACTGGGAAGATAGACTT
AGATAATTATAGCAAAGCAAACATCATAACAAAATTAGAATCTCCAAATTATAATAT
AAAATATGTATCTATTGTTCTAAAAAATACCTTTATGAAGTTCCTATAAAGCTTTTA
AAGATGACTTTAGTTTTACTTTTGACGTCAATTTTTTTTTGTGATATTGATAACATTAG
AATCACTAAACGAAATTACGAAAATGTAAATAAAAATATTAAAGATTATAGAGGCAG
AAAAGACAAATGAGATATTTCCAGAAATTCCAGTAGAAAGTAGAGATGAGTACAGC
TATATAATTTACAACATTATTAATAGTTATATTGAAAAAAGTCAATTGAAAATGGAA
TTAGCAGAAAAGAAGTATAAAAATGAAAGCAATGGAGTTATTAGCACTGCAATCGCA
AATTAGTCCTCATTTTTTTGTCTAATGCGTTGGAGATTATTTATCTTAGGGCATTGTCA
TACACAAACGGTCCTAATGATGTCACAAAAATGATTGAAAATTTGTCACAGATTTTA
AAGTATTTGTTAAGTAATCCAAATGAAACAGTAACTGTAAAAGAAGAAATTGAAAA
TACAAAGGCATATATACAAATATTGAAGGTCAGGTATAGAGATAAATTTAAAGTAA
ATCTAATTTATGATGAAAGTATTTTATCATGTCTCATGATGAAACTGATGCTGCAACA
TTTAATAGAAAATTCTATAAAACATGGGCTTAAGAAGAAAAAATTATGAAGGATCAA
TAAAAATCAAAATAAAAGCAGTTGATAAAAAGAAAATAAAAATTTTCAGTAATCGAT
AATGGCATAGGAATGTCCAAAGAGAGGCTAAATTATGTAAAAAGAATTCTTGACTCT
GACTTCGATTTTTTATGAACATATTGGACTAATGAATACAAATGAACGGTTAAACTT
CTCTATGGGAAAGATTGTGAAATATTAATAAGAAGTAAATTGAATATTGGTACTGCC

GTATATATAATTTTCCATATCAATTAAAAAATCAGAATAATGATGATTATAATAAG
TGA

[0265] or205

[0266] SEQ ID NO:108:

MGINRYDLVKRHNVILEKADIENPLSVGNGEIAFTADITGMQTFIDYKSIPLCTM
SQWGFHTTPAQNDKGYTLEDNLKYYDAFDRKVG YV TSAENQENVFNWLR SNPHRI
NLGNIGLNIILDDGT KAE LKDIFEIHQVLDLWNGILISDFKVEKVPVHVET FCHPYEDMIN
FSVESELLKQNKIYIEVKFPYGAANISGSDWDRNDRHDTNVVDYGRDFVELLRTVDEDV
YFVKIEYSKGVYLN RIGENHFALKQKEYNGRIEF SCLFSKQKPLKCLHSFSESKRMCKEY
WNSFWRGGAIDFSKCEDKRAFELERRVILSQYLTAIQCSGSMPPQETGLTCNSWYGKF
HLEMHW WHAVHFALWGRMPLLSRSIWWYRSIFNVSRDIARKQGYKGVRWPKMVGPD
GRDSPSPIGPLLVWQQPHLIYYSELFRENPT EETLDMFKDIVINTADFIASFVAYDRKND
RYILAPPLIPAQENHDPNVT LNPVFELEYFSFALEIAVKWIERLGLNVNQEWNEIRFKLAN
LPSKDGVYISHEKCINTYEKFNFDHPSMLAALGMLPGRKVDKETMRRTLHRVLKEWKF
EEMWGWDFPMMAMTATRLGEPETA INILLMDSPKNTY MVNGHNNQIPNKELPVYLP
NGGLLAAM ALMTAGWDGNSQSTPGFPKNGMWNVEWEGLKAMI

[0267] SEQ ID NO:109

ATGGGAATTAACAGATATGATCTTGTA AAAAGGCATAATGTAATTTTGGA
AAGCAGATATCGAAAATCCATTGTCAGTAGGTAATGGAGAAATTGCTTTTACAGCTG
ATATAACGGGAATGCAA ACTTTTATTGATGACTATAAGAGCATT CCTTTATGTACCA
TGTCACAGTGGGGGTTTCATACTACGCCGGCACAGAATGATAAGGGCTATTATACTT
TGGAAGATTTGAACCTCAAGTATTACGATGCATTTGACCGAAAGGTTGGATATGTAA
CATCAGCAGAAAATCAAGAGAATGTATTTAATTGGTTGAGGAGTAATCCTCATAGAA
TTAATTTAGGTAATATAGGATTAAATATAATTCTTGATGATGGCACAAAAGCAGAAT
TGAAAGATATTTTCGAAATACACCAAGTATTAGATTTGTGGAACGGAATATTGATAA
GTGACTTTAAAGTCGAAAAAGTCCCTGTTCACGTTGAGACTTTTGGCCATCCATATGA
AGATATGATAAATTTTTCTGTTGAATCAGAACTGCTAAAACAAAATAAAATTTATAT
TGAAGTAAAATTTCCATATGGTGCGGCAATATATCAGGCTCCGATTGGGATAGAAA
TGATAGACATGATACAAATGTGGTTGATTATGGCAGAGATTTTGTGGAATTATTGAG
AACTGTCGATGAAGATGTTTATTTTGTA AAAATAGAGTACTCAAAGGCGTTTATTT

AAATAGAATCGGGGAAAATCATTTTGCATTAAAGCAAAAAGAGTATAATGGGAGAA
TAGAATTTTCGTGCTTGTTTTTCGAAGCAAAAACCTCTTAAGTGCTTGCATTCATTTAG
TGAAAGCAAAAAGGATGTGTAAAGAATATTGGAATAGCTTTTGGAGAGGAGGTGGTG
CAATAGATTTTTCAAAGTGTGAGGATAAAAGAGCTTTTGAATTGGAGAGAAGGGTA
ATACTTTTCGCAATATCTTACAGCTATTCAATGTTTCGGGTTCTATGCCGCCGCAAGAAA
CAGGGCTCACCTGTAATAGCTGGTATGGTAAATTTCAATTTGGAAATGCATTGGTGGC
ATGCTGTACATTTTGCTTTATGGGGTAGAATGCCTTTGCTGAGTAGAAGTATATGGTG
GTACAGGAGCATTTTCAATGTATCACGTGACATTGCGAGAAAGCAAGGATACAAAG
GTGTACGCTGGCCTAAAATGGTTGGACCAGATGGAAGGGATAGCCCTTCTCCGATAG
GACCATTGCTTGTTTGGCAGCAGCCTCATCTTATATATTACAGTGAAGTGTTTTTTAG
AGAAAATCCTACGGAAGAAACATTAGATATGTTTAAAGACATAGTAATTAATACTGC
TGATTTTATTGCATCATTTGTTGCATATGATAGAAAAAATGATAGATATATACTTGCG
CCACCTTTGATTCCAGCACAAGAAAATCATGATCCTAACGTTACATTAAATCCGGTA
TTTGAATTGGAGTATTTTTCGTTTGCGCTGGAAATAGCAGTTAAATGGATTGAAAGG
TTAGGACTAAATGTGAACCAAGAGTGGAATGAAATACGTTTTAAATTAGCTAATTTA
CCTTCAAAAGACGGTGTATATATATCGCATGAAAAATGTATTAACACTTATGAGAAA
TTTAATTTTGACCATCCATCTATGCTTGCAGCATTGGGGATGCTACCAGGCCGCAAG
GTTGATAAAGAACTATGAGAAGGACTTTACATAGAGTATTAAAAGAGTGGAATT
TGAGGAAATGTGGGGTTGGGATTTTCCGATGATGGCTATGACTGCAACAAGATTAGG
CGAACCGGAGACAGCAATAAATATTCTTTTGATGGATTCACCAAAAAATACTTATAT
GGTAAATGGCCATAATAACCAAATACCGAATAAAGAACTACCAGTATATTTGCCTGG
AAATGGTGGACTATTGGCGGCAATGGCCCTCATGACAGCTGGTTGGGATGGGAATA
GCCAAAGCACACCTGGATTTCCTAAAAATGGGATGTGGAATGTTGAATGGGAAGGG
TTAAAAGCGATGATATGA

[0268] or204

[0269] SEQ ID NO:110:

MIKRKDLYIRDPFVVPVPNEKIYYMFGTTDINCWNDEKATGFDYYKSSDLENFEG
PFIAFRPDKNFIWDKNFWAPEVHKYNDMYYMFATFFADGRNRGTQILVSEKISGPYRPW
SIEPVTPKDWMCLDGTIFYVDENGEPWMIFCHEWVQIYDGEICAVRLSKDLKTTIGNPITL
FKASSANWTRSIKKIKDHECYVTDGPFIYRSEEGKLYMLWSSFIENNIYAVGISLSRTGKI
TGPWVHSENPFIAGDGGHGMIFKTFEGNLTLAVHTPNKRKEERPLFITLEKSVLNDTL

[0270] SEQ ID NO:111:

ATGATAAAACGAAAGGATCTTTATATACGTGATCCATTTGTAGTTCCAGTACC
GAATGAAAAAATATATTATATGTTTGGAACTACTGATATAAATTGCTGGAATGATGA
GAAAGCAACTGGATTTGATTACTATAAATCATCTGATTTAGAAAATTTTGAAGGACC
TTTTATTGCATTTAGACCAGATAAAAACTTTATTTGGGATAAAAAATTTTGGGCTCCA
GAAGTGCACAAATACAATGACATGTATTATATGTTTGTCTACATTTTTCGCTGATGGC
AGAAATAGAGGAACGCAAATTTTAGTATCTGAAAAAATAAGTGGGCCATATAGACC
ATGGAGTATTGAACCGGTGACGCCGAAGGATTGGATGTGTTTAGATGGGACTTTTTA
TGATAGATGAGAATGGGGAACCCTGGATGATATTTTGCCATGAATGGGTACAAATATA
TGATGGGGAAATTTGTGCTGTAAGATTGTTCGAAAGATTTAAAAACAACGATAGGAA
ATCCTATTACACTTTTTAAAGCTTCCAGTGCTAATTGGACAAGAAGTATTAAAAAGA
TTAAAGATCATGAATGCTACGTTACGGATGGCCCTTTTATTTATAGGTCTGAAGAGG
GAAAGCTTTATATGTTGTGGTCCAGTTTTATTGAAAACAATATATACGCTGTTGGTAT
ATCATTATCGAGAACAGGCAAAATAACCGGCCCGTGGGTACACAGTGAAAATCCAA
TTTTCGCAGGTGATGGTGGGCATGGTATGATATTTAAGACCTTTGAAGGGAATCTAA
CATTGGCAGTACACACACCTAATAAAAGGAAAGAAGAACGGCCCCTTTTTATAACTT
TAGAAAAATCTGTGCTTAATGATACCTTATAA

[0271] or203

[0272] SEQ ID NO:112:

MFKKITSLLISLLLIISLVTGCSSSSNSSSSSKNSSENNTSPKTVTLRFMWWGGDAR
HKATLDAISLYEKEHPNVKINAEYGGVTDYLQKLITQLSSGTAPDLIQIDVTWLQQLFsq
GDFFADLSKLKDINVNAFDQNFLKNYCYVNNKLIGLPTGINNSAMYINKDFFNKFGIDD
KTVWTDNLLQTAKMVHEKDKNAYLLDADSTICDYILVTYVGQKTGNQWVKDDYTL
GFDKQTLTEAFKYLNDLFEVGAIEPFSQSAPYEGKPDQNPMWLNGQTGMLWNWSSIYA
GVKANIKNLSLALPPIDPNAKQTGIVVRPSQLIAINKDSKNIDEAAKFLNWFFTNTDAIKT
LKDVRGVPATADARKILSENLLDSTLTDNANQAMEKMAPPENGISGNQELEKINTDIIQ
ELAYKKITPEQAADELINTYKQKLPELKSQQ

[0273] SEQ ID NO:113:

ATGTTTAAAAAAATTACATCTCTGTTAATATCGCTTCTTTTGATAATTTTCATTA
GTTACAGGATGTAGCAGTTCTTCGAATTCTTCGAGTTCATCGAAAAATAGTTCTGAA
AATAATACCAGCCCCAAAAACCGTAACATTAAGATTTATGTGGTGGGGTGGAGATGC
CAGACATAAAGCAACACTTGATGCCATAAGTCTTTATGAAAAAGAACATCCCAATGT
AAAGATTAATGCTGAATATGGCGGCGTTACTGACTATCTCCAAAAGCTGATAACTCA
ATTAAGCAGTGGTACAGCACCTGATCTTATACAAATAGATGTAAACATGGTTGCAGCA
ACTTTTTAGCCAAGGTGATTTTTTTTGCAGATTAAAGTAAGTTAAAAGATATCAATGTG
AATGCATTTGATCAAAATTTTCTTAAAAATTATTGCTATGTCAACAATAAGTTGATAG
GTTTGCCTACAGGAATAAACAATTCGGCAATGTATATTAACAAAGACTTTTTTAATA
AATTTGGCATAGACGATAAGACGGTTTGGACATGGGATAATCTCTTGCAAACCGCTA
AGATGGTGCATGAAAAGGATAAAAAATGCTTATCTTTTAGATGCTGATTCTACTATTT
GTGATTATATATTGGTCACATACGTGGGGCAAAAAACTGGAAATCAGTGGGTGAAA
GATGATTACACTTTAGGTTTTGATAAACAACATTGACAGAGGCATTCAAATATTTA
AACGATTTGTTCGAAGTAGGCGCTATAGAGCCATTTTCTCAAAGTGCTCCATACGAA
GGAAAACCTGATCAAAATCCTATGTGGCTTAATGGTCAAACGGGTATGCTTTGGAAC
TGGTCATCTATATATGCTGGTGTAAGCAAACATAAAGAACCTGTCATTGGCATTG
CCACCTATTGACCCTAATGCAAAACAGACAGGCATAGTTGTAAGACCATCACAGCTT
ATTGCTATTAACAAGGATTCTAAAAATATCGATGAAGCAGCAAAATTTTAAATTGG
TTCTTTACGAATACAGATGCTATAAAAACACTTAAAGATGTCAGAGGAGTTCCAGCT
ACCGCAGATGCACGCAAAATTTTATCAGAAAATAATTTGTTGGATTGACTTTAACT
GATAATGCAAATCAAGCTATGGAAAAGATGGCACCTCCTGAAAACGGTATAAGTGG
TAATCAAGAGTTAGAAAAGATAAATACTGATATCATACAAGAACTGGCTTATAAAA
AGATAACGCCAGAGCAGGCTGCTGATGAATTGATAAATACTTATAAACAGAACTT
CCAGAATTAAAAAGCCAGCAATAA

[0274] or202

[0275] SEQ ID NO:114:

MSYNKKRNLMGYLYISPWIIGFLIFTLYPFAMTFIYSFCNYSITKSPVFIGLGNYIT
MFTKDMYFWPSLINTIKYVLMTVPLKLCFALFVAMILNIDIKGVNVFRTTYYLPSIFGGS
VALSVIWKFLFMDNGIMNKFLSYFHIHGPSWLGNPHISLFTISLLSVWFEFGSSMVIFLAAL
KQVPNELYEASMLDGASKIRRFSSITLPMISPVLLFNLVMQTINAFQEFTGPYVITGGGPM
NSTYVYSMLIYDNAFRYFRMGYSSALSWILFLLILIVTVIIFKSSNTWVYYENGGR

[0276] SEQ ID NO:115:

ATGAGTTATAATAAAAAGAGAAATTTGATGGGGTATTTATATATTAGTCCATG
GATTATAGGCTTTTTTAATATTTACTCTGTATCCATTTGCTATGACTTTTATCTATTCAT
TTTGTAACACTACAGTATTACAAAATCACCTGTATTTATTGGATTAGGCAATTATATAAC
TATGTTTACTAAAGATATGTATTTTTTGGCCATCTTTAATTAATACTATAAAAATATGTA
TTAATGACAGTTCCTTTAAAATTATGTTTTTGCACCTTTTTGTTGCAATGATCTTAAATAT
TGATATTAAAGGAGTTAATGTGTTTAGAACAACCTATTATCTGCCTTCTATTTTTTGGGA
GGAAGTGTTGCTTTATCTGTTATATGGAAATTTTTTATTCATGGATAATGGTATTATGA
ATAAATTTCTTTCATACTTTCATATACACGGGCCAAGTTGGCTTGGAAACCCACACAT
ATCATTATTTACTATAAGTTTATTGTCAGTGTGGGAATTTGGGTCTTCTATGGTAATA
TTTTTGGCAGCCCTAAACAGGTCCCGAATGAGTTGTATGAAGCATCTATGTTAGAT
GGTGCAAGCAAAATAAGAAGGTTTTTCTCAATAACTTTACCTATGATATCGCCTGTG
CTATTATTTAATTTGGTTATGCAGACTATAAATGCTTTTTCAGGAATTTACAGGTCCAT
ACGTGATAACTGGTGGAGGACCGATGAACTCTACTTATGTGTACAGTATGTTGATTT
ATGATAATGCGTTTAGGTATTTTAGGATGGGTATTTCATCTGCCTTGTCTTGGATTTT
ATTTTTGTTAATATTGATTGTTACAGTTATAAATTTTAAATCTTCAAATACATGGGTG
TATTACGAAAATGGAGGTAGATGA

[0277] or201

[0278] SEQ ID NO:116:

MKAKNSQNNDIIRKVFIYVFLVAFGIFMIYPLLWVFASSFKSNDEIFKSISLIPKHIV
TNSYFEGWKGTGQYSFGTFILNSITLVVPVVVFTAISSTIVAYGFARFEFPLKTILFTLMIST
MMLPGTAVLIPRYILFNWLGWINTYKPFIVPALFGTTPFFIFMMVQFLRGLPKELEESATI
DGCNSFQILMKILIPLCKPAIISMCIQFIWTWNDFNPLIYINSVEKYTVSLGLNMTIDGTS
VVNWNQIMAMTHISMIPSIHFFSAQKYFVEGIATTGLKN

[0279] SEQ ID NO:117:

ATGAAAGCAAAGAATAGTCAAAATAACGATATAATCAGAAAAGTATTTATAT
ATGTTTTCTTGGTGGCTTTTGGTATTTTCATGATATATCCTTTACTTTGGGTTTTTTGCA
TCATCATTTAAATCAAATGATGAAATCTTTAAATCGATAAGCCTTATACCAAACAC
ATTGTGACAAATTCATATTTTGAAGGATGGAAAGGTACGGGACAATACTCTTTTGGT

ACATTTATTTTAAACAGCATTACGCTTGTTGTACCTGTTGTTGTATTTACTGCTATATC
ATCAACAATTGTAGCCTATGGATTTGCAAGATTTGAGTTTCCGCTTAAACTATTTTG
TTTACTTTGATGATATCTACTATGATGTTGCCGGGCACTGCAGTTTTGATACCAAGAT
ATATATTGTTTAATTGGTTAGGCTGGATAAACACTTATAAACCATTTATTGTTCCCGC
TTTGTTTCGGAACAACGCCTTTTTTTCATTTTTTATGATGGTTCAATTTTTTGAGAGGTCTTC
CTAAAGAATTAGAAGAATCGGCTACAATTGATGGTTGCAATTCATTTCAAATACTTA
TGAAGATTTTAATACCATTGTGTAAACCTGCAATTATTTCTATGTGTATATTTTCAGTT
CATTTGGACTTGGAATGACTTTTTTAATCCATTGATATATATCAACAGTGTAGAAAA
ATATACAGTTTCTCTCGGGCTTAATATGACAATTGATGGGACTTCAGTTGTAAATTGG
AACCAAATAATGGCAATGACAATTATTTCAATGATACCGAGCATCATAATATTTTTT
TCAGCGCAAAAATACTTCGTTGAAGGTATTGCAACAACCTGGATTAAAGAACTAA

[0280] or200

[0281] SEQ ID NO:118:

MRYTDGKVHDITIA YIGGGSRGWAWNLMTDLAKEESISGTVKLYDIDYDAAHD
NEIIGNALSMRQDVKGKWL YKACETLEESLKGAD FVIISILPGTFDEMESDVHAPEKYGI
YQSVGDTVGP GGIVRALRTIPMFVDIANAIKEHCPDAWVINYTNPMTLCVRTLYEIFPQI
KAFGCCHEVFGTQKLLSRALQDIEGIENVPREEIKINVLGINHFTWIDNARYKDIDLMYV
YKQFVNKYYESGFVSDANN NWMNNSFVSAERVKFDFLFLRYGVIAAAGDRHLAEFVPG
YWYLKDPETVREWMFGLTTVSWRKEDLKRRLERSKRLKTGEEKFELKETGEEGVRQIK
ALLGLGDLVTNVNMPNHGQIEGIPYGAVVETNALFSGNKLKPVLSGKLDPDNVNSLVLRQ
VYNQETTLKAALKRDFDLAFSAFVNDPLVTISLKDAKKLFKEMLENTKKYLDGWKIKA

[0282] SEQ ID NO:119:

ATGAGATATACAGATGGAAAGGTT CATGACATTACTATTGCTTATATCGGTGG
TGGTTCAAGAGGATGGGCGTGGAATTTAATGACTGACTTAGCAAAAGAGGAAAGTA
TTTCTGGTACAGTAAAGTTATACGACATAGATTACGATGCGGCACATGACAATGAGA
TAATAGGCAATGCTTTATCAATGAGACAGGATGTTAAAGGCAAATGGCTTTATAAAG
CTTGTGAGACGTTAGAAGAGTCACTAAAAGGTGCTGATTTTGT CATAATATCTATTTT
GCCAGGTACGTTTCGACGAGATGGAATCTGATGTT CATGCACCAGAAAAGTATGGCAT
TTATCAGTCAGTAGGTGATACAGTAGGACCTGGTGGAATAGTCAGAGCTTTAAGGAC
GATTCCGATGTTTGTGGACATTGCCAATGCGATTAAAGAGCATTGTCCAGATGCATG

GGTCATAAATTATACAAATCCTATGACACTTTGTGTAAGGACATTGTATGAAATTTTC
 CCTCAAATTAAAGCATTGTTGGATGCTGCCATGAAGTTTTTGGCACACAGAAGCTATTA
 TCTCGTGCTCTGCAGGATATAGAAGGCATTGAAAATGTTCCGAGGGAAGAGATAAA
 GATAAATGTTTTAGGTATAAATCATTGTTACGTGGATCGACAATGCAAGGTACAAAGA
 CATAGATTTAATGTATGTTTATAAACAATTTGTGAATAAGTACTATGAAAGCGGATT
 TGTCAGCGATGCTAACAATAATTGGATGAACAATTCATTGTTATCTGCAGAGAGAGT
 AAAGTTTGATCTGTTTTTGGAGGTATGGAGTAATAGCTGCAGCGGGAGATAGACATCT
 GGCGGAATTTGTGCCGGGATATTGGTATTTAAAAGATCCAGAGACAGTCAGAGAAT
 GGATGTTTGGCTTAACGACTGTAAGTTGGAGAAAAGAAGACTTAAAACGCAGGCTT
 GAAAGAAGTAAAAGGCTTAAGACAGGTGAGGAAAAATTTGAGTTAAAGGAAACAG
 GCGAAGAAGGTGTTAGGCAAATTAAAGCACTATTAGGCTTAGGCGATTTAGTGACTA
 ATGTCAACATGCCGAACCATGGACAGATTGAAGGAATACCATACGGTGCGGTAGTT
 GAAACAAACGCTTTATTTTCAGGTAATAAACTAAAGCCTGTATTATCAGGAAAATTG
 CCTGACAATGTAAACAGCCTCGTGTTAAGGCAAGTATACAACCAAGAAACGACGTT
 GAAAGCTGCTTTAAAGAGAGATTTTGATTGTTGGCTTTTAGTGCTTTTGTAATGATCCA
 CTTGTTACAATATCTTTAAAAGATGCAAAAAAATTATTTAAGGAAATGCTTGAAAAT
 ACGAAGAAATATCTAGATGGATGGAAAATAAAAGCTTGA

Non-Native proteins

[0283] EC 2.3.1.9

[0284] *C. acetobutylicum* ThlA (SEQ ID NO:120)

MKEVVIASAVRTAIGSYGKSLKDVPVAVDLGATAIKEAVKKAGIKPEDVNEVILGN
 VLQAGLGQNPARQASFKAGLPVEIPAMTINKVCGSGLRTVSLAAQIIKAGDADVIIAGGM
 ENMSRAPYLANNARWGYRMGNAKFVDEMITDGLWDAFNDYHMGITAENIAERWNISR
 EEQDEFALASQKKAEEAIKSGQFKDEIVPVVIKGRKGETVVDTEHPRFGSTIEGLAKLK
 PAFKKDGTVTAGNASGLNDCAAVLVIMSAEKAKELGVKPLAKIVSYGSAGVDPAIMGY
 GPFYATKAAIEKAGWTVDELDLIESNEAFAAQSLAVAKDLKFDMNKVNVNGGAIALGH
 PIGASGARILVTLVHAMQKRDAKKGLATLCIGGGQGTAILLEKC

[0285] EC 2.8.3.8

[0286] *C. acetobutylicum* CtfAB

[0287] CtfA (SEQ ID NO:121)

MNSKIIRFENLRSFFKDGMTIMIGGFLNCGTPTKLIDFLVNLNIKNLTIISNDTCYPN
 TGIGKLISNNQVKKLIASYIGSNPDTGKKLFNNELEVELSPQGTLVERIRAGGSGLGGVLT
 KTGLGTLIEKGKKKISINGTEYLLELPLTADVALIKGSIVDEAGNTFYKGTTKNFNPYMA
 MAAKTVIVEAENLVSCEKLEKEKAMTPGVLINYIVKEPA

[0288] CtfB (SEQ ID NO:122)

MINDKNLAKEIIAKRVARELKNGQLVNLGVGLPTMVADYIPKNFKITFQSENGIV
 GMGASPKINEADKDVVNAGGDYTTVLPDGTFFDSSVSFSLIRGGHVDVTVL GALQVDE
 KGNIANWIVPGKMLSGMGAMD LVNGAKKVIIAMRHTNKGQPKILKKCTLPLTAKSQA
 NLIVTELGVIEVINDGLLLTEINKNTTIDEIRSLTAADLLISNELRPM AV

[0289] EC 4.1.1.4

[0290] *C. acetobutylicum* Adc, Aad

[0291] Adc (SEQ ID NO:123)

MLKDEVIKQISTPLTSPA FPRGPYKFHNREYFNIVYRTDMDALRKVVPEPLEIDEP
 LVRFEIMAMHDTSGLG CYTESGQAIPVSFNGVKGDY LHMMYLDNEPAIAVGRELSAYP
 KKLGYPKLFVDS DTLVGTL DYGKL RVATATMGYKHKALDANEAKDQICRPNYMLKIIP
 NYDGSPRICELINAKITDVT VHEAWTGPTRLQLFDHAMAPLNDLPVKEIVSSSHILADIILP
 RAEVIYDYLK

[0292] Aad (SEQ ID NO:124)

MLKDEVIKQISTPLTSPA FPRGPYKFHNREYFNIVYRTDMDALRKVVPEPLEIDEP
 LVRFEIMAMHDTSGLG CYTESGQAIPVSFNGVKGDY LHMMYLDNEPAIAVGRELSAYP
 KKLGYPKLFVDS DTLVGTL DYGKL RVATATMGYKHKALDANEAKDQICRPNYMLKIIP
 NYDGSPRICELINAKITDVT VHEAWTGPTRLQLFDHAMAPLNDLPVKEIVSSSHILADIILP
 RAEVIYDYLK

[0293] EC 1.2.1.43 Formate dehydrogenase (*M. thermoacetica*) Moth_2312
 (SEQ ID NO:125)

MVNLTIDGQRVTAPEGMTILEVARENGIHIPTLCHHPKLRPLGYCRLCLVDIEGAA
 KPMTACNTPVAEGMVIRTSTPVIEEMRKGIEMLLSLHPEDCLTCEKAGNCQLQDCAYT
 YGVKHGELPVKREELPVLKENPFIVRDYNKCI VCGRCVRACQEVQVQRVVDLVGKGSA

ARVGATKAGAEVSLEEGGCVF CGNCVQVCPVGALTEKAGLGQGREWEFKKVRSICSYC
 GVGCNLTLYVKDGKVVKVRGYENPEVNNGWLCVKGRFGFDYIHNPDRITRPLIREGDR
 EKGYFREASWEEALALVSQKLTQIKGSYGSEALGFLCSAKCTNEENYLLQKLARGVLGT
 NNVDHCARLHSSTVAGLATTFGSGAMTNSIADIASADCIFVIGSNTTENHPVIALKVKEA
 VRRGARLIVADPRRIELVNFSYLWLRQKPGTDLALLNGLLHVIIKEELYDKEFIAQRTEGF
 EALKLAVEEYTPAKVSEVTGVPAGDII EAARTYARGPSSTILYAMGITQHITGTANVMAL
 ANLAMACGQVGKEGSGVNPLRGQSNVQGACDMGGLPNVLPGYQPVTDPGVRHKFSEA
 WGVDPDLPGEPLTLMEMMAAAQEGKLGMYILGENPVLTDPDVSHVKEALKNLEFLV
 VQDIFLTETARMADVVLPGASFAEKEGTFTSTERRVQLLHKAIEPPGEARPDWLILNDLL
 LLMGYPRKYSSPGEIMQEIAGLTPSYAGITYERLEDKGLQWPVLSLEHPGTPVLHREKFS
 RGYGQFQVVHYRPPAEPEDEEYPFLFTTGRNLYHYHTVISRKS RGLEEMCPAPVVEINDN
 DAARLGIREGEMIEIVSRRGKVRVKALVTDRIPRGQVFMNFHFHEAAANLLTIAALDPVA
 KIPEYKTCAVAIVKK

Proteins sequences for *Saccharomyces cerevisiae* engineering

[0294] EC 4.2.3.3

[0295] *Oryza sativa* - mgs (SEQ ID NO:126)

MELTTRTIAERKHIALVAHDHRKQALLEWVESHKTI LAQHQLYATGTTGNLIQR
 ASGIPVTSMLSGPMGGDQQVGALIAEGKIDMLIFFWDPLNAVPHDPDVKALLRLATVW
 NIPVATNRSTADFLIDSPLFKSEVAIAIPDYQRYLQDRLK

[0296] EC 2.3.1.8

[0297] *T. saccharolyticum* – or1741 (SEQ ID NO:127)

MKTSELLAMVVEKGASDLHITVGVPVLRINGQLIKLNLPQLTPQDTEEITKDLLS
 SDELKKLEDMGDIDLSYSVKGLGRFRINAYKQRGTYSLAIRSVALRIPTIDELGLPEVIKE
 LALKTRGLIIVTGPTGSGKSTTLAS MIDLINEERNCHILTLEDPIEYLHKHKK SIVNQREIG
 HDAASYASALRAALREDPDVILVGEMRDLETIQIAITAAETGHLVLSTLHTIGSAKTIDRII
 DVFPPHQQQQIKVQLSNVLEGIVSQQLLPKIDNSGRVVAVEVMIATPAIRNLIREGKSFQI
 QSMVQTGNKFGMVTMDMWISQLLKRNLISMDDALTYCVDRENFSRLVV

[0298] EC 1.1.1.6

[0299] *Pseudomonas putida* gldA (SEQ ID NO:128)

MDRAIQSPGKYVQGADALQRLGDYKPLADSWLVIADKFVLGFAEDTIRQSLSK
AGLAMDIVAFNGECSQGEVDRLCQLATQNGRSAIVGIGGGKTLDTAKAVAFFQKVPVA
VAPTIASTDAPCSALSVLYTDEGEFDRYLMLPTNPALVVVDTAIVARAPARLLAAGIGDA
LATWFEARAASRSSAATMAGGPATQTALNLARFCYDTLLEEGERKAMLA VQAQVVTPA
LERIVEANTYLSGVGFESGGVAAAHA VHNGLTAVAETHHFYHGEKVAFGVLVQLALEN
ASNAEMQEVMSLCHAVGLPITLAQLDITEDIPTKMRAVAELACAPGETIHNMPGGVTVE
QVYGALLVADQLGQHFLEF

[0300] EC 2.7.2.1

[0301] *T. saccharolyticum* or1742 (SEQ ID NO:129)

MIKKKLGDLLVEVGLLDESQNLNNAIKIQKKTGEKLGKILVKEGYLTEEQIIEALEF
QLGIPHIDMKKVFIDANVAKLIPESMAKRHVAIPIKKENNSIFVAMADPLNIFAIDDIKLV
KLDVKPLIASD GILKAIDRVFGKEEAERAVQDFKKELSHDSAEDDGNLLRDISSEDEINN
APAVRLVNSIIEQAVKNRASDVHIEPTENDLRIRFRIDGELHEAMRVFKSTQGPVITRIKIM
ANMNIAERRIPQDGKIE MNAGGKNIDIRVSSLPTIYGEKLVLRILDKSGYIITKDKLGLGN
DDLKLF DNLLKHPNGIILLTGPTGSGKTTTLYAMLNELNKPDKNIITVEDPVEYTLEGLN
QVQVNEKAGLTFASALRSILRQDPDIIMIGEIRDRETA EIAIRSSITGHLVLSTLHTNDSAG
AITRLIDMGIEPYLVSSSVGVIAQRLARKICDNCKIEYDASKREKIILGIDADESLKLYRS
KGC AVCNKTGYRGRVP IYEIMMMTPKIKELTNEKAPADVILNEAVSNGMSTLKESAKKL
VLSGVTTVDEMLRLTYDDAY

[0302] EC 2.8.3.8

[0303] *C. acetobutylicum* CtfAB

[0304] CtfA (SEQ ID NO:130)

MNSKIIRFENLRSFFKDGMTIMIGGFLNCGTPTKLIDFLVNLNIKNLTIISNDTCYPN
TGIGKLISNNQVKKLIASYIGSNPDTGKKLFNNELEVELSPQGTLVERIRAGGSGLGGLVLT
KTGLGLTIEKGKKKISINGTEYLLLELPLTADVALIKGSIVDEAGNTFYKGTTKNFNPYMA
MAAKTVIVEAENLVSCEKLEKEKAMTPGVLINYIVKEPA

[0305] CtfB (SEQ ID NO:131)

MINDKNLAKEIIAKRVARELKNGQLVNLGVGLPTMVADYIPKNFKITFQSENGIV
 GMGASPKINEADKDVVNAGGDYTTVLPDGTFFDSSVSFSLIRGGHVDVTVL GALQVDE
 KGNIANWIVPGKMLSGMGAMD LVNGAKKVIIAMRHTNKGQPKILKKCTLPLTAKSQA
 NLIVTELGVIEVINDGLLLTEINKNTTIDEIRSLTAADLLISNELRPM AV

[0306] EC 4.1.1.4

[0307] *C. acetobutylicum* - Adc (SEQ ID NO:132)

MLKDEVIKQISTPLTSPA FPRGPYKFHNREYFNIVYRTDMDALRKVVPEPLEIDEP
 LVRFEIMAMHDTSGLG CYTESGQAIPVSFNGVKGDY LHMMYLDNEPAIAV GRELSAYP
 KKLGYPKLFVDS DTLVGTLDYGKLRVATATMGYKHKALDANEAKDQICRPNYMLKIIP
 NYDGSPRICELINAKITDVT VHEAWTGPTRLQLFDHAMAPLNDLPVKEIVSSSHILADIILP
 RAEVIYDYLK

[0308] EC 2.3.1.54

[0309] *Escherichia coli* – pflA (SEQ ID NO:133)

MSVIGRIHSFESCGTV DGP GIRFITFFQGCLMRCLYCHNRDTWDTHGGKEVTVED
 LMKEVV TYRHF MNASGGGVTASGGEAILQAEFVRDWFRACKKEGIHTCLDTNGFVRRY
 DPVIDELLEVTDLV MLDLKQMNDEIHQNLVGVS NHRTLEFAKYLANKNVK V WIRYVVV
 PGWSDDDDSAHRLGEFTRDMGNVEKIELLPYHEL GKHKWVAMGEEYKLDGVKPPKKE
 TMERVKGILEQYGHKVMF

[0310] EC 2.3.1.54

[0311] *Escherichia coli* – pflB (SEQ ID NO:134)

MSELNEKLATAWEGFTKGDWQNEVNVRDFIQKNYTPYEGDESFLAGATEATTT
 LWDKVM EGVKLENRTHAPVDFDTAVASTITSHDAGYINKQLEKIVGLQTEAPLKRALIP
 FGGIKMIEGSCKAYNRELDPMIKKIFTEYRKTHNQGVFDVYTPDILRCRKSGVLTGLPDA
 YGRGRIIGDYRRVALY GIDYLMKDKLAQFTSLQADLENGVNLEQTIRLREEIAEQHRAL
 GQMKEMAAKYGYDISGPATNAQEAIQW TYFGYLA AVKSQNGAAMSFGRTSTFLDVYIE
 RDLKAGKITEQEAQEMVDHLVMKLRMVRFLRTPEYDELFSGDPIWATESIGGMGLDGR
 TLVTKNSFRFLNTLY TMGPSPEPNMTILWSEKLPLNFKKFAAKV SIDTSSLQYENDDLMR
 PDFNNDYAIACCVSPMIVGKQM QFFGARANLAKTMLY AINGGVDEKLKMQVGPKSEP
 IKGDVLNYDEVMERMDHFMDWLAKQYITALNIIHYMHDKYSYEASLMALHDRDVIRT

MACGIAGLSVAADSLSAIKYAKVKPIRDEDGLAIDFEIEGEYPQFGNNDPRVDDLAVDLV
ERFMKKIQKLHTYRDAIPTQSVLTITSNVVYGKKTGNTPDGRRAGAPFGPGANPMHGRD
QKGAVASLTSAKLFPAYAKDGISYTFSIVPNALGKDDEVKTNLAGLMDGYFHHEASI
EGGQHLNVNVMNREMLLDAMENPEKYPQLTIRVSGYAVRFNSLTKEQQQDVITRTFTQ
SM

[0312] EC 2.3.1.9

[0313] *Saccharomyces cerevisiae* ERG10 (SEQ ID NO:135)

MSQNVYIVSTARTPIGSFQGSLSKTAVELGAVALKGALAKVPELDASKDFDEIIF
GNVLSANLGQAPARQVALAAGLSNHIVASTVNKVCASAMKAILGAQSIKCGNADV
VAGGCESMTNAPYYMPAARAGAKFGQTVLVDGVERDGLNDAYDGLAMGVHAEKCARD
WDITREQQDNFAIESYQKSQKSQKEGKFDNEIVPVTIKGFRGKPDQTQVTKDEEPARLHVE
KLRSARTVFQKENGTVTAANASPINDGAAVILVSEKVLKEKNLKLPLAIKGWGEAAHQ
PADFTWAPSLAVPKALKHAGIEDINSVDYFEFNEAFSVVGLVNTKILKLDPSKVN
VYGG
AVALGHPLGCSGARVVVTLLSILQQEGGKIGVAAICNGGGGASSIVIEKI

[0314] EC 1.1.1.1

[0315] *Saccharomyces cerevisiae* ADH1 (SEQ ID NO:136)

MSIPETQKGVIFYESHGKLEYKDIPVPKPKANELLINVKYSGVCHTDLHAWHGD
WPLPVKLPLVGGHEGAGVVVGMGENVKGWKIGDYAGIKWLNGSCMACEYCELG
NES
NCPHADLSGYTHDGSFQQYATADAVQAAHIPQGTDLAQVAPILCAGITVYKALK
SANL
MAGHWVAISGAAGGLGSLAVQYAKAMGYRVLGIDGGEGKEELFRSIGGEVFIDFT
KEK
DIVGAVLKATDGGAHGVINVSVSEAAIEASTRYVRANGTTVLVGMPAGAKCCSDV
FNQ
VVKSISIVGSYVGNRADTREALDFFARGLVKSPIKVVGLSTLPEIYEKMEKGQIV
GRYVV
DTSK

[0316] EC 1.1.1.1

[0317] *Saccharomyces cerevisiae* ADH2 (SEQ ID NO:137)

MSIPETQKAIIFYESNGKLEHKDIPVPKPKPNELLINVKYSGVCHTDLHAWHGDW
PLPTKLPLVGGHEGAGVVVGMGENVKGWKIGDYAGIKWLNGSCMACEYCELG
NESNC
PHADLSGYTHDGSFQYATADAVQAAHIPQGTDLAQVAPILCAGITVYKALK
SANL
RAG
HWAASGAAGGLGSLAVQYAKAMGYRVLGIDGGPGKEELFTSLGGEVFIDFTKEK
DIVS

AVVKATNGGAHGIINVSVSEAAIEASTRYCRANGTVVLVGLPAGAKCSSDVFNHVVKSI
SIVGSYVGNRADTREALDFFARGLVKSPIKVVGLSSLPEIYEKMEKGQIAGRYVVDTSK

[0318] EC 1.1.1.1

[0319] *Saccharomyces cerevisiae* ADH3 (SEQ ID NO:138)

MLRTSTLFTRRVQPSLFSRNILRLQSTAAIPKTQKGVIFYENKGKLHYKDIPVPEPK
PNEILINVKYSGVCHTDLHAWHGDWPLPVKLPLVGGHEGAGVVVKLGSNVKGWKVGD
LAGIKWLNGSCMTCEFCESGHESNCPDADLSGYTHDGSFQQFATADAIQAAKIQQGTDL
AEVAPILCAGVTVYKALKEADLKAGDWVAISGAAGGLGSLAVQYATAMGYRVLGIDA
GEEKEKLFKKLGGEVFIDFTKTKNMVSDIQEATKGGPHGVINVSVSEAAISLSTEYVRPC
GTVVLVGLPANAYVKSEVFSHVVK SINIKGSYVGNRADTREALDFFSRGLIKSPIKIVGLS
ELPKVYDLMEKGKILGRYVVDTSK

[0320] EC 1.1.1.1

[0321] *Saccharomyces cerevisiae* ADH4 (SEQ ID NO:139)

MSSVTGFYIPPISFFGEGALEETADYIKNKDYKKALIVTDPGIAAIGLSGRVQKML
EERDLNVAIYDKTQPNPNIANVTAGLKVLKEQNSEIVVSIGGGSAHDNAKAIALLATNG
GEIGDYEGVNQSKKAALPLFAINTTAGTASEMTRFTIISNEEKKIKMAIIDNNVTPAVAVN
DPSTMFGGLPPALTAATGLDALTHCIEAYVSTASNPITDACALKGIDLINESLVAAYKD GK
DKKARTDMCYAEYLAGMAFNNASLG YVHALAHQLGGFYHLPHGVCNAVLLPHVQEA
NMQCPKAKKRLGEIALHFGASQEDPEETIKALHVLNRTMNIPRNLKELGVKTEDFEILAE
HAMHDACHLTNPVQFTKEQVVAIIKKAYEY

[0322] EC 1.1.1.1

[0323] *Saccharomyces cerevisiae* ADH5 (SEQ ID NO:140)

MPSQVIPEKQKAIVFYETDGKLEYKDVTVPPEPKPNEILVHVKYSGVCHSDLHAW
HGDWPFQLKFPLIGGHEGAGVVVKLGSNVKGWKVGDFAGIKWLNGTCMSCEYCEVGN
ESQCPYLDGTGFTHDGTFQEYATADAVQAAHIPPNNVLAEVAPILCAGITVYKALKRAN
VIPGQWVTISGACGGLGSLAIQYALAMGYRVIGIDGGNAKRKLFEQLGGEIFIDFTEEKDI
VGAIKATNGGSHGVINVSVSEAAIEASTRYCRPNGTVVLVGMPAHAYCNSDVFNQVVK
SISIVGSCVGNRADTREALDFFARGLIKSPIHLAGLSDVPEIFAKMEKGEIVGRYVVVETSK

[0324] EC 1.1.1.1

[0325] *Saccharomyces cerevisiae* ADH6 (SEQ ID NO:141)

MSYPEKFEGIAIQSHEDWKNPKKTKYDPKPFYDHDIDIKIEACGVCGSDIHCAAG
HWGNMKMPLVVGHEIVGKVVKLGPKSNSGLKVGQRVGVGAQVFSCLECDRCKNDNEP
YCTKFVTTYSQPYEDGYVSQGGYANYVRVHEHFVVPDENIPSHLAAPLLCGGLTVYSPL
VRNGCGPGKKVGIVGLGGIGSMGTLISKAMGAETYVISRSSRKREDAMKMGADHYIAT
LEEGDWGEKYFDTFDLIVVCASSLTDIDFNIMPKAMKVGGRIVSISIPEQHEMLSLKPYGL
KAVSISYSALGSIKELNQLLKLVSSEKDIKWVETLPVGEAGVHEAFERMEKGDVRYRFTL
VG YDKEFSD

[0326] EC 1.1.1.1

[0327] *Saccharomyces cerevisiae* ADH7 (SEQ ID NO:142)

MLYPEKFQGGIGISNAKDWKHPKLVSFDPKPFGDHDVDVEIEACGICGSDFHIAVG
NWGPVPENQILGHEIIGRVVKVGSKCHTGVKIGDRVGVGAQALACFECERCKSDNEQYC
TNDHVLTMWTPYKDGYSQGGFASHVRLHEHFVVPDENIPSPLAAPLLCGGITVFSPLLR
NGCGPGKRVGIVGIGGIGHMGILLAKAMGAEVYAFSRGHSKREDSMKLGADHYIAMLE
DKGWTEQYSNALDLLVVCSSSLSKVNFDIVKIMKIGGSIVSIAAPEVNEKLVKPLGLM
GVSISSSAIGSRKEIEQLLKLVSSEKNVKIWVEKLPISSEGVSHAFTRMESGDVKYRFTLVD
YDKKFHK

[0328] EC 1.1.1.1

[0329] *Saccharomyces cerevisiae* BDH2 (SEQ ID NO:143)

MRALAYFGKGNIRFTNHLKEPHIVAPDELVIDIEWCGICGTDLHEYTDGPIFFPED
GHTHEISHNPLPQAMGHEMAGTVLEVGPVKNLKVGDKVVVEPTGTCRDRYRWPLSP
NVDKEWCAACKKGYYNICSYLGLCGAGVQSGGFAERVVMNESH CYKVPDFVPLDVAA
LIQPLAVCWHAIRVCEFKAGSTALIIGAGPIGLGTILALNAAGCKDIVVSEPAKVRRELA
KMGARVYDPTAHAAKESIDYLRSIADGGDGFDTFDCSGLEVTLNAAIQCLTFRGTAVN
LAMWGH HKIQFSPMDITLHERKYTGSMCYTHHDFEAVIEALEEGRIDIDRARHMITGRV
NIEDGLDGAIMKLINEKESTIKIILTPNNHGELNREADNEKKEISELSSRKDQERLRESINE
AKLRHT

[0330] EC 1.1.1.1

[0331] *Saccharomyces cerevisiae* SFA1 (SEQ ID NO:144)

MSAATVGKPIKCIAAVAYDAKKPLSVEEITVDAPKAHEVRIKIEYTAVCHTDAYT
LSGSDPEGLFPCVLGHEGAGIVESVGDDVITVKPGDHVIALYTAECGKCKFCTSGKTNLC
GAVRATQGGKGVMPDGTTRFHNAKGEDIYHFMGCSTFSEYTVVADVSVVAIDPKAPLDA
ACLLGCGVTTGFGAALKTANVQKGDTVAVFGCGTVGLSVIQGAKLRGASKIIAIDINNK
KKQYCSQFGATDFVNPKEDLAKDQTIVEKLIEMTDGGLDFTFDCTGNTKIMRDALEACH
KGWGQSIIGVAAAGEEISTRPFQLVTGRVWKGSAFGGIKGRSEMGGLIKDYQKGALKV
EEFITHRRPFKEINQAFEDLHNGDCLRTVLKSDEIK

[0332] EC 1.1.1.1

[0333] *Saccharomyces cerevisiae* YPL088W (SEQ ID NO:145)

MVLVKQVRLGNSGLKISPIVIGCMSYGSKKWADWVIEDKTQIFKIMKHCYDKGL
RTFDTADFYSNGLSERIIKEFLEYYSIKRETVVIMTKIYFPVDETLDLHHNFTLNEFEELDL
SNQRGLSRKHIIAGVENSVKRLGTYIDLLQIHRLDHETPMKEIMKALNDVVEAGHVRYIG
ASSMLATEFAELQFTADKYGWFQFISSQSYYNLLYREDERELIPFAKRHNIGLLPWSPNA
RGMLTRPLNQSTDRIKSDPTFKSLHLDNLEEEQKEIINRVEKVSKDKKVSMAMLSIAWVL
HKGCHPIVGLNTTARVDEAIAALQVTLTEEEIKYLEEPYKPKRQRC*

[0334] EC 4.1.2.13

[0335] *Saccharomyces cerevisiae* FBA1 (SEQ ID NO:146)

MGVEQILKRKTGVIVGEDVHNLFTYAKEHKFAIPAINVTSSSTAVA ALEAARDSK
SPIILQTSNGGAAYFAGKGISNEGQNASIKGAIAAAHYIRSIAPAYGIPVVLHSDHCAKKL
LPWFDGMLEADEAYFKEHGEPLFSSHMLDLSEETDEENISTCVKYFKRMAAMDQWLEM
EIGITGGEEDGVNNENADKEDLYTKPEQVYNVYKALHPISP NFSIAAAFGNCHGLYAGDI
ALRPEILAEHQKYTREQVGCKEEKPLFLVFHGGSGSTVQEFHTGIDNGVVKNLDTDCQ
YAYLTGIRDYVLNKKDYIMSPVGNPEGPEKPNKKFFDPRVWVREGEKTMGAKITKSLET
FRTTNL

[0336] EC 5.3.1.1

[0337] *Saccharomyces cerevisiae* TPI1 (SEQ ID NO:147)

MARTFFVGGNFKLNGSKQSIKEIVERLNTASIPENVEVVICPPATYLDYSVSLVKK
PQVTVGAQNAYLKASGAFTGENSVDQIKDVGAKWVILGHSERRSYFHEDDKFIADKTK

FALGQGVGVILCIGETLEEKKAGKTLDVVERQLNAVLEEVKDWTNVVVAAYEPVWAIGT
GLAATPEDAQDIHASIRKFLASKLGDKAASELRILYGGSSANGSNAVTFKDKADVDGFLV
GGASLKPEFVDIINSRN

[0338] EC 1.2.1.2

[0339] *Saccharomyces cerevisiae* FDH1 (SEQ ID NO:148)

MSKGGKVLLVLYEGGKHAESEQEKL LGCIENELGIRNFIEEQGYELVTTIDKDPEPTS
TVDRELKDAEIVITTPFFPAYISRNRIAEAPNLKLCVTAGVGS DHVDLEAANERKITVTEV
TGSNVVSVAEHVMATILVLIRNYNGGHQQAINGEWDIAGVAKNEYDLEDKIISTVGAGR
IGYRVLERLVAFNPKKLLYYDYQELPAEAINRLNEASKLFNGRGDIVQRVEKLEDMVAQ
SDVVTINCPLHKDSRGLFNKKLISHMKDGAYLVNTARGAICVAEDVAEAVKSGKLAGY
GGDVWDKQPAPKDHPWRTMDNKHVGNAMTVHISGTS LDAQKRYAQGVKNILNSYF
SKKFDYRPQDIIVQNGSYATRAYGQKK

[0340] EC 1.1.1.21

[0341] *Saccharomyces cerevisiae* GRE3 (SEQ ID NO:149)

MSSLVTLNNG LKMPLVGLGCWKIDKKVCANQIYEAIKLG YRLFDGACDYGNEK
EVGEGIRKAISEGLVSRKDIFVVSKLWNNFHHPDHVKLALKKTLSDMGLDYLDLYYIHF
PIAFKYVPFEEKYPPGFYTGADDEKKGHITEAHVPIIDTYRALEECVDEGLIKSIGVSNFQ
GSLIQDLLRGCRIPVALQIEHHPYLTQEHLVEFCKLHDIQVVAYSSFGPQSFIEMDLQLA
KTTPTLFENDVIKKVSQNHPGSTTSQVLLRWATQRGIAVIPKSSKKERLLGNLEIEKKFTL
TEQELKDISALNANIRFNDPWTWLDGKFPTFA

[0342] EC 1.1.1.79

[0343] *Saccharomyces cerevisiae* GOR1 (SEQ ID NO:150)

MSKKPIVLKLGKDAFGDQAWGELEKIADVITIPESTTREQFLREV KDPQNKLSQV
QVITRTARSVKNTGRFDEELALALPSSVVA VCHTGAGYDQIDVEPFKKRHIQVANVPDL
VSNATADTHVFLLLGALRNFGIGNRRLIEGNWPEAGPACGSPFGYDPEGKTVGILGLGRI
GRCILERLKPFGFENFIYHNRHQLPSEEEHGC EYVGFEFLKRSDIVSVNVPLNHNTHHLI
NAETIEKMKDGVVIVNTARGAVIDEQAMTDALRSGKIRSAGLDVFEYEPKISKELLSMSQ
VLGLPHMGTHSVETRKKMEELVVENAKNVILT GKVLTIVPELQ NEDWPNESKPLV

[0344] EC 1.1.1.79

[0345] *Saccharomyces cerevisiae* YPL113C (SEQ ID NO:151)

MITSIDIADVITYSAKPRILVPYKTQWEVASHLPEYRKLAERVEFYKYEMSTKDDF
VKFLETHRINGFWLTEEFFTVLGNPSSYIEFFPASLKVILVPWVGCDFIDGKLLRSKGITLC
NIGPHAADHVTCLAIFLAISCFRMTSFWCYCFKYVENGNVEQCKKYISSDSYEIVTDSYH
GQEMKFPSRTDKCKPNKDRKVHHLAEKYTVGGKKMESPMNKKVLILGFGSIGQNIGSN
LHKVFNMSIEYYKRTGPVQKSLLDYNKYHSDLDLDPNTWKNADLIILALPSTASTNNIIN
RKSLAWCKDGVRIVNVGRGTCIDEDVLLDALESGKVASCGLDVFKNEETRVKQELLRR
WDVTALPHIGSTVADMVIKQTLITLENVQDIFVEGGDGKYVLN

[0346] EC 1.2.1.49

[0347] *Saccharomyces cerevisiae* GCY1 (SEQ ID NO:152)

MPATLHDSTKILSLNTGAQIPQIGLGTWQSKENDAYKAVLTALKDGYRHIDTAAL
YRNEDQVGQAIKDSGVPREEIFVTTKLWCTQHHEPEVALDQSLKRLGLDYVDLYLMHW
PARLDPAYIKNEDILSVPTKKDGSRAVDITNWNFIKTWELMQELPKTGKTKAVGVSNFSL
NNLKDLLASQGNKLTAAANQVEIHPLLPQDELINFCKSKGIVVEAYSPLGSTDAPLLKEP
VILEIAKKNNVQPGHVVISWHVQRGYVVLPKSVNPDRIKTNRKIFTLSTEDFEAINNISKE
KGEKRVVHPNWSPEVFK

[0348] EC 1.2.1.49

[0349] *Saccharomyces cerevisiae* ALD2 (SEQ ID NO:153)

MPTLYTDIEIPQLKISLKQPLGLFINNEFCPSSDGKTIETVNPATGEPITSFQAANEK
DVDKAVKAARAAFDNVWSKTSSEQRGIYLSNLLKLIEEEQDTLAALETLDAGKPYSNAK
GDLAQILQLTRYFAGSADKFDKGATIPLTFNKFAYTLKVPFGVVAQIVPWNYPPLAMAC
WKLQGALAAGNTVHKAENTSLSLLYFATLIKKAGFPPGVVNIVPGYGSLSVGQALASH
MDIDKISFTGSTKVGGFVLEASGQSNLKDVTLECGGKSPALVFEDADLDKAIDWIAAGIF
YNSGQNCTANSRVYVQSSIYDKFVEKFKETAKKEWDVAGKFDPFDEKCIVGPVISSTQY
DRIKSYIERGKREEKLDMFQTSEFPIGGAKGYFIPPTIFTDVPQTSKLLQDEIFGPVVVVS
FTNYDDALKLANDTCYGLASAVFTKDVKKAHMFARDIKAGTVWINSSNDEDVTVPFGG
FKMSGIGRELQSGVDTYLQTKAVHINLSLDN

[0350] EC 1.2.1.49

[0351] *Saccharomyces cerevisiae* ALD3 (SEQ ID NO:154)

MPTLYTDIEIPQLKISLKQPLGLFINNEFCPSSDGKTIETVNPATGEPITSFQAANEK
DVDKAVKAARAAFDNVWSKTSSEQRGIYLSNLLKLIEEEQDTLAALETLDAGKPFHSNA
KQDLAQIIELTRY YAGAVDKFNMGETIPLTFNKFAYTLKVPFGVVAQIVPWNYP LAMAC
RKMQGALAAGNTV I I KPAENTSLSLLYFATLIKKAGFPPGVVNVIPGYG SVVGKALGTH
MDIDKISFTGSTKVGGSVLEASGQSNLKDITLECGGKSPALVFEDADLDKAIEWVANGIF
FNSGQICTANSRVYVQSSYDKFVEKFKETAKKEWDVAGKFDPFDEKCIVGPVISSTQYD
RIKSYIERGKKEEKLDMFQTSEFPIGGAKGYFIPPTIFTDVPETSKLLRDEIFGPVVVVSKFT
NYDDALKLANDTCYGLASAVFTKDVKKAHMFARDIKAGTVWINQTNQEEAKVPFGGF
KMSGIGRESGDTGVDNYLQIKSVHVDLSLDK

[0352] EC 1.2.1.49

[0353] *Saccharomyces cerevisiae* ALD4 (SEQ ID NO:155)

MFSRSTLCLKTSASSIGRLQLRYFSHLPMTVPIKLPNGLEYEQPTGLFINNK FVPSK
QNKTFEVINPSTEEEI CH IYEGREDDVEEAVQAADRAFSNGSWNGIDPIDRGKALYRLAE
LIEQDKDVIAS IETLDNGKAISSSRGDV DL VINYLKSSAGFADKIDGRMIDTGRTHFSYTK
RQPLGVCGQIIPWNFPLLMWAWKIAPALVTGNTVVLKTAESTPLSALYVSKYIPQAGIPP
GVINIVSGFGKIVGEAITNHPKIKKVAFTGSTATGRHIYQSAAAGLKKVTLELG GKSPNIV
FADAELKKAVQNIILGIYYNSGEVCCAGSRVYVEESIYDKFIEEFKAASESIKVGDPFDES
TFQGAQTSQMQLNKILKYVDIGKNEGATLITGGERLGSKGYFIKPTVFGDVKEDMRIVK
EEIFGPVVTVT KFKSADEVINMANDSEYGLAAGIHTSNINTALKVADRVNAGTVWINTY
NDFHHA VPFGGFNASGLGREMSVDALQNYLQVKAVRAKLDE

[0354] EC 1.2.1.49

[0355] *Saccharomyces cerevisiae* ALD5 (SEQ ID NO:156)

MLSRTRAAAPNSRIFTRSLRLYSQAPLRVPITLPNGFTYEQPTGLFINGEFVASKQ
KKTFDVINPSNEEKITTVYKAMEDDVDEAVAAAKKAFETKWSIVEPEVRAKALFNLADL
VEKHQETLAAIESMDNGKSLFCARGDVALVSKYLRSCGGWADKIYGNVIDTGKNHFTY
SIKEPLGVCGQIIPWNFPLLMWSWKIGPALATGNTVVLKPAETTPLSALFASQLCQEAGIP
AGVVNILPGSGRVVGERLSAHPDVKKIAFTGSTATGRHIMKVAADTVKKVTLELG GKSP
NIVFADADLDKAVKNIAFGIFYNSGEVCCAGSRIYIQDTVYEEVLQKLKDYTESLKVGDP
FDEEVFQGAQTS DKQLHKILDYVDVAKSEGARLVTGGARHGSKGYFVKPTVFADVKG D

MRIVKEEVFGPIVTVSKFSTVDEVIAMANDSQYGLAAGIHTNDINKAVDVSKRVKAGTV
WINTYNNFHQNVFPGGFGQSGIGREMGEAALSNYTQTKSVRIAIDKPIR

[0356] EC 1.2.1.49

[0357] *Saccharomyces cerevisiae* ALD6 (SEQ ID NO:157)

MTKLHFDTAEPVKITLPNGLTYEQPTGLFINNKFMKAQDGKTYPVEDPSTENTVC
EVSSATTEDVEYAIECADRAFDTEWATQDPRERGRLLSKLADELESQIDLVSSIEALDN
GKTLALARGDVTIAINCLRDAAAYADKVNGRTINTGDGYMNFTTLEPIGVCGQIIPWNFP
IMMLAWKIPALAMGNVCILKPAAVTPLNALYFASLCKKVGIPAGVVNIVPGPGRTVGA
ALTNDPRIRKLAFTGSTEVGKSVAVDSSESNLKKITLELGGKSAHLVFDDANIKKTLPNL
VNGIFKNAGQICSSGSRIYVQEGIYDELLAAFKAYLETEIKVGNPFDKANFQGAI TNRQQF
DTIMNYIDIGKKEGAKILTGGEEKVGDKGYFIRPTVFYDVNEDMRIVKEEIFGPVVTVAKF
KTLEEGVEMANSSEFGLGSGIETESLSTGLKVAKMLKAGTVWINTYNDFDSRVFPGGVK
QSGYGREMGEENVYHAYTEVKAVRIKL

[0358] EC 1.2.1.49

[0359] *Saccharomyces cerevisiae* HFD1 (SEQ ID NO:158)

MSNDGSKILNYTPVSKIDEIVEISRNFFFEKQLKLSHENNPRKKDLEFRQLQLKKL
YYAVKDHEEELIDAMYKDFHRNKIESVLNETTKLMNDILHLIEILPKLIKPRRVSDSSPPF
MFGKTIVEKISRGSVLIHAPFNFPLLLAFAPLAAALAAGNTIVLKPSELTPHTAVVMENLLT
TAGFPDGLIQVVQGAIDETTRLLDCGKFDLIFYTGSPRVGSIVA EKA AAKSLTPCVLELG GK
SPTFITENFKASNIKIALKRIFFGAFGNSGQICVSPDYLLVHKSIYPKVIKECESVLNEFYPS
FDEQTDFTMRIHEPAYKKAVASINSTNGSKIVPSKISINSDTEDLCLVPPTIVYNIGWDDPL
MKQENFAPVLPPIEYEDLDETINKIIEEHDTPLVQYIFSDSQTEINRILTRLRSGDCVVGDTV
IHVGITDAPFGGIGTSGYGNYG GYYGFNTFSHER TIFKQPYWNDFTLFMRYP PNSAQKEK
LVRFAMERKPWFDRNGNNKWGLRQYFSLSA AVILISTIYAHCSS

[0360] EC 2.7.1.2

[0361] *Saccharomyces cerevisiae* GLK1 (SEQ ID NO:159)

MSFDDLHKATERAVIQAVDQICDDFEVTPEKLDELTA YFIEQMEKGLAPPKEGHT
LASDKGLPMIPAFVTGSPNGTERGVLLAADLGGTNFRICSVNLHGDHTFSMEQM KSKIP
DDLDDENVTSDDLFGFLARRTLAFMKKYHPDELA KGKDAKPMKLGFTFSYPVDQ TSL

NSGTLIRWTKGFRIADTVGKDVVQLYQEQLSAQGMPMIKVVALTNDTVGTYLSHCYTS
 DNTDSMTSGEISEPVIGCIFGTGTNGCYMEEINKITKLPQELRDKLIKEGKTHMIINVEWG
 SFDNELKHLPTTKYDVVIDQKLSTNPGFHLFEKRVSGMFLGEVLRNILVDLHSQGLLLQQ
 YRSKEQLPRHLTTPFQLSSEVLSHIEIDDSTGLRETELSLLQSLRLPTTPTERVQIQKLVRAI
 SRRSAYLA AVPLAAILIKTNALNKRYHGEVEIGCDGSGVVEYYPGFRSMLRHALALSPLG
 AGERKVHLKIAKDGS GVG AALCALVA

[0362] EC 5.3.1.9

[0363] *Saccharomyces cerevisiae* PGI1 (SEQ ID NO:160)

MSNNSFTNFKLATELPAWSKLQKIYESQGKTL SVKQEFQKDAKRFEKLNKTFTN
 YDGSKILFDYSKNLVNDEIIAALIELAKEANVTGLRDAMFKGEHINSTEDRAVYHVALRN
 RANKPMYVDGVNVAPEVDSVLKHMKEFSEQVRSGEWKGYTGKKITDVVNIGIGGSDLG
 PVMVTEALKHYAGVLDVHFVSNIDGTHIAETLK VVDPETTLFLIASKTFTTAETITNANT
 AKNWFLSKTGNDPSHIAKHFAALSTNETEVAKFGIDTKNMFGFESWVGGRYSVWSAIGL
 SVALYIGYDNFEAFLKGAEAVDNHFTQTPLEDNIPLLGGLLSVWYNNFFGAQTHLVAPF
 DQYLHRFPAYLQQLSMESNGKSVTRGNVFTDYSTGSILFGEPATNAQHSFFQLVHQGTK
 LIPSDFILAAQSHNPIENKLHQKMLASNFFAQAEALMVGKDEEQVKAEGATGGLVPHKV
 FSGNRPTTSILAQKITPATLGALIAYYEHVTFTEGAIWNINSFDQWGVELGKVLAKVIGKE
 LDNSSTISTHDASTNGLINQFKEWM

[0364] EC 2.7.1.11

[0365] *Saccharomyces cerevisiae* PFK1 (SEQ ID NO:161)

MQSQDSCYGVAFRSIITNDEALFKKTIHFYHTLGFATVKDFNKFKHGENSLLSSGT
 SQDSLRE VWLESFKLSEVDASGFRIPQQEATNKAQSQGALLKIRLVMSAPIDETFDTNET
 ATITYFSTD LNKIVEKFPKQAEKLSDTLVFLKDPMGNNITFSGLANATDSAPTSKDAFLEA
 TSEDEIISRASSDASDLLRQTLGSSQKKKKIAVMTSGGDSPGMNAAVRAVVRTGIHFGCD
 VFAVYEGYEGLLRGGKYLKKMAWEDVRGWLSEGGTLIGTARSMEFRKREGRRQAAGN
 LISQGIDALVVC GGDGSLTGADLFRHEWPSLVDELVAEGRFTKEEVAPYKNLSIVGLVGS
 IDNDMSGTDSTIGAYSALERICEMVDYIDATAKSHSRAFVVEVMGRHCGWLALMAGIA
 TGADYIFIPERAVPHGKWQDELKEVCQRHRSKGRRNNTIIVAEGALDDQLNPVTANDVK
 DALIELGLDTKVTILGHVQRGGTAVAHDRWLATLQGVD AVKAVLEFTPETPSPLIGILEN
 KIIRMPLVESVKLTKSVATAIENKDFDKAISLRDTEFIELYENFLSTTVKDDGSELLPVSDR

LNIGIVHVGAPSAALNAATRAATLYCLSHGHKPYAIMNGFSGLIQTGEVKELSWIDVEN
 WHNLGGSEIGTNRSVASEDLGTIAYYFQKNKLDGLIILGGFEGFRSLKQLRDGRTQHPIF
 NIPMCLIPATVSNNVPGTEYSLGVDTCLNALVNYTDDIKQSASATRRRVFVCEVQGGHS
 GYIASFTGLITGAVSVYTPEKKIDLASIREEDITLLKENFRHDKGENRNGKLLVRNEQASSV
 YSTQLLADIISEASKGKFGVRTAIPGHVQQGGVPSSKDRVTASRFAVKCIKFIEQWNKKN
 EASPNTDAKVLRFKFDTHGEKVPTVEHEDDSAASICVNGSHVSFKPIANLWENETNVEL
 RKGFEVHWAIEYNKIGDILSGRLKLRAEVAALAAENK

[0366] EC 2.7.1.11

[0367] *Saccharomyces cerevisiae* PFK2 (SEQ ID NO:162)

MTVTTPFVNGTSYCTVTAYSVQSYKAAIDFYTKFSLLENRSSPDENSTLLSNDSSIS
 LKILLRPDEKINKNVEAHLKELNSITKTQDWRSHATQSLVFNTSDILAVKDTLNAMEAPL
 QGYPTLFPMLQYTLDPNGVVGVTSTKNAVSTKPTPPAPEASAESGLSSKVHSYTDLA
 YRMKTTDTYPSLPKPLNRPQKAIAVMTSGGDAPGMNSNVRAIVRSAIFKGCRAFFVME
 GYEGLVRGGPEYIKEFWEDVRGWSAEGGTNIGTARCMFEKKREGRLLAGAHLIEAGV
 DALIVCGGDGSLTGADLFRSEWPSLIEELLKTNRISNEQYERMKHLNICGTVGSIDNDMS
 TTDATIGAYSALDRICKAIDYVEATANSHSRAFFVEVMGRNCGWLALLAGIATSADYIFI
 PEKPATSSEWQDQMCDIVSKHRSRGKRTTIVVVAEGAIAADLTPISPSDVHKVLVDRLGL
 DTRITTLGHVQRGGTAVAYDRILATLQGLEAVNAVLESTPDTPSPLIAVNENKIVRKPLM
 ESVKLTKAFAEAIQAKDFKRAMSLRDTEFIEHLNNFMANSADHNEPKLPKDKRLKIAIV
 NVGAPAGGINSAYVSMATYCMSQGHHPYAIYNGWSGLARHESVRSLNWKDMLGWQS
 RGGSEIGTNRVTPEEADLGMIAYYFQKYEFDGLIIVGGFEAFESLHQLERARESYPAFRIP
 MVLIPATLSNNVPGTEYSLGSDTALNALMEYCDVVKQSASSTRGRAFFVDCQGGNSGY
 LATYASLAVGAQVSYPVEEGISLEQLSEDIEYLAQSFEKAEGRGRFGKLILKSTNASKALS
 ATKLAEVITAEADGRFDAKPAYPGHVQQGGGLPSPIDRTRATRMAIKAVGFIKDNQAAIA
 EARAAEENFNADDKTISDTAAVVGVKGSHVVYNSIRQLYDYETEVSMRMPKVIHWQAT
 RLIADHLVGRKRVD

[0368] EC 4.1.1.1

[0369] *Saccharomyces cerevisiae* PDC1 (SEQ ID NO:163)

MSEITLGKYLFERLKQVNVNTVFGLPGDFNLSLLDKIYEVEGMRWAGNANELNA
 AYAADGYARIKGMSCIITTFGVGELSALNGIAGSYAEHVGVLHVVGVPSSISAQAKQLLL

HHTLGNGDFTVFHRMSANISSETTAMITDIATAPAEIDRCIRTTYVTQRPVYLGLPANLVD
 LNVPAKLLQTPIDMSLKPNDASEKEVIDTILALVKDAKNPVILADACCSRHVDVKAETKK
 LIDLTQFPFAFVTPMGKGSIDEQHPRYGGVYVGTLISKPEVKEAVESADLILSVGALLSDFN
 TGSFSYSYKTKNIVEFHSDHMKIRNATFPGVQMKFVLQKLLTTIADAAGYKPVAVPAR
 TPANAAVPASTPLKQEWMWNQLGNFLQEGDVVIAETGTSAFGINQTTFPNNTYGISQVL
 WGSIGFTTGATLGAAFAAEEIDPKKRVILFIGDGSLQLTVQEISTMIRWGLKPYLFLVNNND
 GYTIEKLIHGPKAQYNEIQGWDHLSLLPTFGAKDYETHRVATTGEWDKLTQDKSFNDNS
 KIRMIEIMLPVFDAPQNLVEQAKLTAATNAKQ

[0370] EC 4.1.1.1

[0371] *Saccharomyces cerevisiae* PDC5 (SEQ ID NO:164)

MSEITLGKYLFERLSQVNCNTVFGLPGDFNLSLLDKLYEVKGMRWAGNANELN
 AAYAADGYARIKGMSCIITTFGVGELSALNGIAGSYAEHVGVLHVVGVPSSISQAKQLLL
 HHTLGNGDFTVFHRMSANISSETTAMITDIANAPAEIDRCIRTTYTTQRPVYLGLPANLVD
 LNVPAKLLLETPIDLSLKPNDAAEAEAEVVRTVVELIKDAKNPVILADACASRHVDVKAETK
 KLMDLTQFPVYVTPMGKGAIDEQHPRYGGVYVGTLISRPEVKKAVESADLILSIGALLSD
 FNTGSFSYSYKTKNIVEFHSDHIKIRNATFPGVQMKFALQKLLDAIPEVVKDYPVAVPA
 RVPITKSTPANTPMKQEWMWNHLGNFLREGDIVIAETGTSAFGINQTTFPTDVYAIVQVL
 WGSIGFTVGALLGATMAAEELDPKKRVILFIGDGSLQLTVQEISTMIRWGLKPYIFVLNN
 NGYTIEKLIHGPHAEYNEIQGWDHLALLPTFGARNYETHRVATTGEWEKLTQDKDFQD
 NSKIRMIEVMPLVFDAPQNLVKQAQLTAATNAKQ

[0372] EC 4.1.1.1

[0373] *Saccharomyces cerevisiae* PDC6 (SEQ ID NO:165)

MSEITLGKYLFERLKQVNVNTIFGLPGDFNLSLLDKIYEVDGLRWAGNANELNA
 AYAADGYARIKGLSVLVTTFGVGELSALNGIAGSYAEHVGVLHVVGVPSSISAQAKQLLL
 HHTLGNGDFTVFHRMSANISSETTSMITDIATAPSEIDRLIRTTFITQRPSYLGLPANLVDLK
 VPGSLLEKPIDLSLKPNDPEAEKEVIDTVLELIQNSKNPVILSDACASRHNKKETQKLID
 LTQFPFAFVTPLGKGSIDEQHPRYGGVYVGTLISKQDVKKAVESADLILSVGALLSDFNTGS
 FSYSYKTKNVVEFHSDYVKVKNATFLGVQMKFALQNLLKVIPDVVKGYKSVPVPTKTP
 ANKGVPASTPLKQEWLWNELSKFLQEGDVISETGTSAFGINQTTIFPKDAYGISQVLWGSIG
 FTGATLGAAFAAEEIDPNKRVILFIGDGSLQLTVQEISTMIRWGLKPYLFLVNNNDGYTI

EKLIHGPHAEYNEIQTWDHLALLPAFGAKKYENHKIATTGEWDALTTDSEFQKNSVIRLI
ELKLPVFDAPESLIKQAQLTAATNAKQ

[0374] EC 1.1.1.8

[0375] *Saccharomyces cerevisiae* GPD2 (SEQ ID NO:166)

MLAVRRLTRYTFLKRTHPVLYTRRAYKILPSRSTFLRRSLLQTQLHSKMTAHTNI
KQHKHCHEDHPIRRSDSAVSIVHLKRAPFKVTVIGSGNWTIAKVIAENTELHSHIFEPE
VRMWVFDEKIGDENLTDIINTRHQNVKYLPNIDLPHNLVADPDLLHSIKGADILVFNIPH
QFLPNIVKQLQGHVAPHVRAISCLKGFELGSKGVQLLSSYVTDELGIQCGALSGANLAPE
VAKEHWSETTVAYQLPKDYQGDGKDVDHKILKLLFHRPYFHVNVIDDVAGISIAGALK
NVVALACGFVEGMGWGNNASAAIQRLGLGEIIFGRMFFPESKVETYYQESAGVADLIT
TCSGGRNVKVATYMAKTGKSALEAEKELLNGQSAQGIITCREVHEWLQTCELTQEFPLF
EAVYQIVYNNVRMEDLPEMIEELDIDDE

[0376] EC 3.1.3.21

[0377] *Saccharomyces cerevisiae* GPP1 (SEQ ID NO:167)

MPLTTKPLSLKINAALFDVDGTIISQPAIAAFWRDFGKDKPYFDAEHVIHISHGW
RTYDAIAKFAPDFADEEYVKNLEGEIPEKYGEHSIEVPGAVKLCNALNALPKEKWAVAT
SGTRDMAKKWFDILKIKRPEYFITANDVKQGKPHPEPYLKGRNGLGFPINEQDPSKSKVV
VFEDAPAGIAAGKAAGCKIVGIATTFDLDFLKEKGCDIIVKNHESIRVGEYNAETDEVELI
FDDYLYAKDDLLKW

[0378] In certain embodiments, an enzyme of the present invention includes any enzyme that is at least about 70%, 80%, 90%, 95%, 99% identical, or sharing at least about 60%, 70%, 80%, 90%, 95% sequence identity to any of the enzymes of the metabolic engineered pathways as described above. These enzymes sharing the requisite sequence identity or similarity can be wild-type enzymes from a different organism, or can be artificial, *i.e.*, recombinant, enzymes.

[0379] In certain embodiments, any genes encoding for enzymes with the same activity as any of the enzymes of the metabolically engineered pathways as described above may be used in place of the enzymes. These enzymes may be wild-type enzymes from a different organism, or may be artificial, recombinant or engineered enzymes.

[0380] Additionally, due to the inherent degeneracy of the genetic code, other nucleic acid sequences which encode substantially the same or a functionally equivalent amino acid sequence can also be used to express the polynucleotide encoding such enzymes. As will be understood by those of skill in the art, it can be advantageous to modify a coding sequence to enhance its expression in a particular host. The codons that are utilized most often in a species are called "optimal codons", and those not utilized very often are classified as "rare or low-usage codons". Codons can be substituted to reflect the preferred codon usage of the host, a process sometimes called "codon optimization" or "controlling for species codon bias." Methodology for optimizing a nucleotide sequence for expression in, *e.g.* *Saccharomyces cerevisiae*, are known to one of ordinary skill in the art.

Modified Strains

[0381] The present invention further provides for knockout strains in which the metabolic engineered pathways of the invention are carried out. Such a genetically modified microorganism would have an increased ability to produce lactate or acetate as a fermentation product. "Knock out" of the genes means partial, substantial, or complete deletion, silencing, inactivation, or down-regulation.

[0382] Thus, certain embodiments of the present invention provide for the "inactivation" or "deletion" of certain genes or particular polynucleotide sequences within thermophilic or mesophilic microorganisms, which "inactivation" or "deletion" of genes or particular polynucleotide sequences can be understood to encompass "genetic modification(s)" or "transformation(s)" such that the resulting strains of said thermophilic or mesophilic microorganisms can be understood to be "genetically modified" or "transformed." In certain embodiments, strains can be of bacterial, fungal, or yeast origin.

[0383] A genetically modified strain that is a knockout strain can have the advantage of eliminating the production of certain organic acids or products that interfere with the ability of the strain to generate a high yield of an alternative product, such as isopropanol or propanediol.

[0384] For example, if the conversion of pyruvate to lactate (the salt form of lactic acid) by the action of LDH was not available in the early stages of the glycolytic pathway, then

the pyruvate could be more efficiently converted to acetyl CoA by the action of pyruvate dehydrogenase or pyruvate-ferredoxin oxidoreductase.

[0385] Genes to be targeted for knockout for the present invention include lactate dehydrogenase (*ldh*), hydrogenase (*hyd*), acetaldehyde dehydrogenase (*acdh*), acetate kinase (*ack*), pyruvate-ferredoxin oxidoreductase (*por*) or pyruvate decarboxylase (*pdh*).

[0386] As used herein, the term "lactate dehydrogenase" or "LDH" is intended to include the enzyme capable of converting pyruvate into lactate. It is understood that LDH can also catalyze the oxidation of hydroxybutyrate.

[0387] As used herein, the term "acetate kinase" or "ACK" is intended to include the enzyme capable of converting acetyl phosphate into acetate.

[0388] As used herein, the term "pyruvate-ferredoxin oxidoreductase" or "POR" is intended to include the enzyme capable of converting pyruvate into acetyl CoA, carbon dioxide, and reduced ferredoxin.

[0389] The term "pyruvate decarboxylase activity" is intended to include the ability of a polypeptide to enzymatically convert pyruvate into acetaldehyde (*e.g.*, "pyruvate decarboxylase" or "PDC"). Typically, the activity of a selected polypeptide encompasses the total enzymatic activity associated with the produced polypeptide, comprising, *e.g.*, the superior substrate affinity of the enzyme, thermostability, stability at different pHs, or a combination of these attributes.

[0390] Certain embodiments of the present invention, alternatively, provide for the "insertion," (*e.g.*, the addition, integration, incorporation, or introduction) of certain genes or particular polynucleotide sequences within thermophilic or mesophilic microorganisms, which insertion of genes or particular polynucleotide sequences can be understood to encompass "genetic modification(s)" or "transformation(s)" such that the resulting strains of said thermophilic or mesophilic microorganisms can be understood to be "genetically modified" or "transformed." In certain embodiments, strains can be of bacterial, fungal, or yeast origin.

[0391] In one aspect of the invention, the genes or particular polynucleotide sequences are inserted to activate the activity for which they encode, such as the expression of an enzyme. In certain embodiments, genes encoding enzymes in the metabolic production of ethanol, *e.g.*, enzymes that metabolize pentose and/or hexose sugars, can be added to a mesophilic or thermophilic organism. In certain embodiments of the invention, the

enzyme can confer the ability to metabolize a pentose sugar and be involved, for example, in the D-xylose pathway and/or L-arabinose pathway.

[0392] In one aspect of the invention, the genes or particular polynucleotide sequences are partially, substantially, or completely deleted, silenced, inactivated, or down-regulated in order to inactivate the activity for which they encode, such as the expression of an enzyme. Deletions provide maximum stability because there is no opportunity for a reverse mutation to restore function. Alternatively, genes can be partially, substantially, or completely deleted, silenced, inactivated, or down-regulated by insertion of nucleic acid sequences that disrupt the function and/or expression of the gene (*e.g.*, P1 transduction or other methods known in the art). The terms "eliminate," "elimination," and "knockout" are used interchangeably with the terms "deletion," "partial deletion," "substantial deletion," or "complete deletion." In certain embodiments, strains of thermophilic or mesophilic microorganisms of interest can be engineered by site directed homologous recombination to knockout the production of organic acids. In still other embodiments, RNAi or antisense DNA (asDNA) can be used to partially, substantially, or completely silence, inactivate, or down-regulate a particular gene of interest.

Vectors and Host Cells

[0393] The present invention also relates to vectors which include genes encoding for enzymes of the present invention, as described above, as well as host cells which are genetically engineered with vectors of the invention and the production of polypeptides of the invention by recombinant techniques.

[0394] Host cells are genetically engineered (transduced or transformed or transfected) with the vectors of this invention which can be, for example, a cloning vector or an expression vector. The vector can be, for example, in the form of a plasmid, a viral particle, a phage, etc. The engineered host cells can be cultured in conventional nutrient media modified as appropriate for activating promoters, selecting transformants or amplifying the genes of the present invention. The culture conditions, such as temperature, pH and the like, are those previously used with the host cell selected for expression, and will be apparent to the ordinarily skilled artisan.

[0395] The DNA sequence in the expression vector is operatively associated with an appropriate expression control sequence(s) (promoter) to direct mRNA synthesis. Any

suitable promoter to drive gene expression in the host cells of the invention can be used. Additionally, promoters known to control expression of genes in prokaryotic or lower eukaryotic cells can be used. The expression vector also contains a ribosome binding site for translation initiation and a transcription terminator. The vector can also include appropriate sequences for amplifying expression, or can include additional regulatory regions.

[0396] The vector containing the appropriate selectable marker sequence as used herein, as well as an appropriate promoter or control sequence, can be employed to transform an appropriate thermophilic host to permit the host to express the protein.

[0397] The terms "promoter" or "surrogate promoter" is intended to include a polynucleotide segment that can transcriptionally control a gene-of-interest that it does not transcriptionally control in nature. In certain embodiments, the transcriptional control of a surrogate promoter results in an increase in expression of the gene-of-interest. In certain embodiments, a surrogate promoter is placed 5' to the gene-of-interest. A surrogate promoter can be used to replace the natural promoter, or can be used in addition to the natural promoter. A surrogate promoter can be endogenous with regard to the host cell in which it is used, or it can be a heterologous polynucleotide sequence introduced into the host cell, e.g., exogenous with regard to the host cell in which it is used.

[0398] The terms "gene(s)" or "polynucleotide segment" or "polynucleotide sequence(s)" are intended to include nucleic acid molecules, e.g., polynucleotides which include an open reading frame encoding a polypeptide, and can further include non-coding regulatory sequences, and introns. In addition, the terms are intended to include one or more genes that map to a functional locus. In addition, the terms are intended to include a specific gene for a selected purpose. The gene can be endogenous to the host cell or can be recombinantly introduced into the host cell, e.g., as a plasmid maintained episomally or a plasmid (or fragment thereof) that is stably integrated into the genome. In addition to the plasmid form, a gene can, for example, be in the form of linear DNA. In certain embodiments, the gene encodes a polypeptide, such as an enzyme of the present invention. The term gene is also intended to cover all copies of a particular gene, e.g., all of the DNA sequences in a cell encoding a particular gene product.

[0399] The term "transcriptional control" is intended to include the ability to modulate gene expression at the level of transcription. In certain embodiments, transcription, and

thus gene expression, is modulated by replacing or adding a surrogate promoter near the 5' end of the coding region of a gene-of-interest, thereby resulting in altered gene expression. In certain embodiments, the transcriptional control of one or more gene is engineered to result in the optimal expression of such genes, *e.g.*, in a desired ratio. The term also includes inducible transcriptional control as recognized in the art.

[0400] The term "expression" is intended to include the expression of a gene at least at the level of mRNA production.

[0401] The term "expression product" is intended to include the resultant product, *e.g.*, a polypeptide, of an expressed gene.

[0402] The term "increased expression" is intended to include an alteration in gene expression at least at the level of increased mRNA production and, preferably, at the level of polypeptide expression. The term "increased production" is intended to include an increase in the amount of a polypeptide expressed, in the level of the enzymatic activity of the polypeptide, or a combination thereof.

[0403] In certain aspects, the present invention relates to host cells containing the above-described constructs. The host cell can be an anaerobic thermophilic bacterial cell, including an anaerobic xylanolytic and/or cellulolytic host cell. The selection of an appropriate host is deemed to be within the scope of those skilled in the art from the teachings herein.

[0404] The present invention also includes recombinant constructs comprising one or more of the selectable marker sequences as broadly described above. The constructs comprise a vector, such as a plasmid or viral vector, into which a sequence of the invention has been inserted, in a forward or reverse orientation. In one aspect of this embodiment, the construct further comprises regulatory sequences, including, for example, a promoter, operably associated to the sequence. Large numbers of suitable vectors and promoters are known to those of skill in the art, and are commercially available. The following vectors are provided by way of example only.

[0405] The term "derived from" is intended to include the isolation (in whole or in part) of a polynucleotide segment from an indicated source or the purification of a polypeptide from an indicated source. The term is intended to include, for example, direct cloning, PCR amplification, or artificial synthesis from or based on a sequence associated with the indicated polynucleotide source.

[0406] Introduction of the construct in host cells can be done using methods known in the art. Introduction can also be effected by electroporation methods as described in U.S. Prov. Appl. No. 61/109,642, filed October 30, 2008, the contents of which are herein incorporated by reference.

[0407] Furthermore, the use of positive and/or negative selection markers, genetic tools, and homologous recombination-based genome integration adapted for use in, *e.g.*, thermophilic organisms, that can be used to efficiently select modified strains, including modified strains of *C. thermocellum* and *T. saccharolyticum* can be done using methods as described in U.S. Prov. Appl. No. 61/232,648, filed August 10, 2009, the contents of which are herein incorporated by reference. Methods for the expression of foreign genes, knockout and overexpression of native genes, and creation of clean industrial strains that do not contain antibiotic markers or other extraneous DNA can be performed, as described in U.S. Prov. Appl. No. 61/232,648.

Biomass

[0408] The terms "lignocellulosic material," "lignocellulosic substrate," and "cellulosic biomass" mean any type of biomass comprising cellulose, hemicellulose, lignin, or combinations thereof, such as but not limited to woody biomass, forage grasses, herbaceous energy crops, non-woody-plant biomass, agricultural wastes and/or agricultural residues, forestry residues and/or forestry wastes, paper-production sludge and/or waste paper sludge, waste-water-treatment sludge, municipal solid waste, corn fiber from wet and dry mill corn ethanol plants, and sugar-processing residues.

[0409] In a non-limiting example, the lignocellulosic material can include, but is not limited to, woody biomass, such as recycled wood pulp fiber, sawdust, hardwood, softwood, and combinations thereof; grasses, such as switch grass, cord grass, rye grass, reed canary grass, miscanthus, or a combination thereof; sugar-processing residues, such as but not limited to sugar cane bagasse; agricultural wastes, such as but not limited to rice straw, rice hulls, barley straw, corn cobs, cereal straw, wheat straw, canola straw, oat straw, oat hulls, and corn fiber; stover, such as but not limited to soybean stover, corn stover; succulent plants, such as but not limited to agave; and forestry wastes, such as but not limited to recycled wood pulp fiber, sawdust, hardwood (*e.g.*, poplar, oak, maple, birch, willow), softwood, or any combination thereof. Lignocellulosic material can comprise one species of fiber; alternatively, lignocellulosic material can comprise a

mixture of fibers that originate from different lignocellulosic materials. Particularly advantageous lignocellulosic materials are agricultural wastes, such as cereal straws, including wheat straw, barley straw, canola straw and oat straw; corn fiber; stovers, such as corn stover and soybean stover; grasses, such as switch grass, reed canary grass, cord grass, and miscanthus; or combinations thereof.

[0410] Paper sludge is also a viable feedstock for lactate or acetate production. Paper sludge is solid residue arising from pulping and paper-making, and is typically removed from process wastewater in a primary clarifier. At a disposal cost of \$30/wet ton, the cost of sludge disposal equates to \$5/ton of paper that is produced for sale. The cost of disposing of wet sludge is a significant incentive to convert the material for other uses, such as conversion to ethanol. Processes provided by the present invention are widely applicable. Moreover, the saccharification and/or fermentation products can be used to produce ethanol or higher value added chemicals, such as organic acids, aromatics, esters, acetone and polymer intermediates. During glycolysis, cells convert simple sugars, such as glucose, into pyruvic acid, with a net production of ATP and NADH. In the absence of a functioning electron transport system for oxidative phosphorylation, at least 95% of the pyruvic acid is consumed in short pathways which regenerate NAD^+ , an obligate requirement for continued glycolysis and ATP production. The waste products of these NAD^+ regeneration systems are commonly referred to as fermentation products.

EXEMPLIFICATION

EXAMPLE 1

1.1 Production of Mixed Alcohols in Bacterial and Yeast CBP Platforms

[0411] Production of mixed alcohols in bacteria and yeast makes use of bacterial and yeast CBP platforms, and their available toolboxes, to produce a combination of propanediol, isopropanol, glycerol and ethanol. Trace amounts of microbially produced propanediol were first detected in 1954 during cultivation of *Clostridium thermobutyricum*. See Enebo, L. 1954, "Studies in cellulose decomposition by an anaerobic thermophilic bacterium and two associated non-cellulolytic species," p. 94-96. Viktor Pettersons Bokindustrie Aktiebolag, Stockholm. Since then, reports have

indicated native production of propanediol from common sugars during fermentations of *C. sphenoides* and *T. thermosaccharolyticum*. See Tran-Din, K., & Gottschalk, G., 1985, *Arch. Microbiol.* 142, 87-92; Cameron, D.C., & Clooney, C., 1986, *Bio/Technology* 4, 651-654. Recombinant *E. coli* strains have been developed that produce propanediol from dihydroxyacetone phosphate, an intermediate of sugar metabolism, using multiple recombinant genes. See Altaras, N.E., & Cameron, D.C., 1999, *Appl Environ Microbiol.* 65(3), 1180-5; U.S. Patent No. 6,303,352.

[0412] The objective of this example is to provide new pathways for the production of high yields of mixed alcohols in bacteria and yeast. The bacterial CBP platforms comprise microorganisms that are in the same family as *C. sphenoides* and *T. thermosaccharolyticum*, which contain native genes for propanediol production and, unlike the literature, do not rely on expression of recombinant activities to convert dihydroxyacetone phosphate to propanediol. For example, *T. saccharolyticum* is able to ferment L-Rhamnose to equimolar amounts of propanediol and a mixture of ethanol, acetic acid, lactic acid, H₂ and CO₂. See Lee *et al.*, *International Journal of Systematic Bacteriology*, 43(1): 41-51 (1993). However, in the past, the exploitation of thermophilic clostridia for production of propanediol was not feasible due to a lack of genetically tractable systems required for stable genetic engineering. The successful genetic engineering of thermophilic clostridia and thermoanaerobacter and thermoanaerobacterium strains now makes such exploitation for metabolic engineering possible. See U.S. Prov. Appl. No. 61/232,648, filed August 10, 2009. Further, production of propanediol in yeast has been observed by the expression of a single gene, methylglyoxal synthase (*mgs*), indicating that additional activities necessary to convert methylglyoxal to propanediol are endogenous to yeast. See Lee, W., & DaSilva, N. A., 2006, *Metabolic Eng.* 8, 58-65.

[0413] The 1,2-propandiol produced using these platforms can be used as a valuable intermediate or converted to propionate and propanol using microbes such as *Lactobacillus reuteri* strain isolated from sourdough that is known to do this reaction. See Sriramulu, D.D., *et al.*, 2008, *J Bacteriol.* 190(13):4559-67. Chemical routes might also exist for direct conversion of propanediol to propanol or even propylene.

[0414] Isopropanol can be produced by the addition of a pathway to produce acetone and a dehydrogenase capable of utilizing acetone as a substrate. The best known and studied

acetone production route is from the metabolism of *Clostridium acetobutylicum*. All enzymes in this pathway have been sequenced and cloned into other hosts such as *E. coli*. See Bermejo, L.L., *et al.*, 1998, *Appl Environ Microbiol.* 64(3), 1079-85. *C. acetobutylicum* has been used in industrial fermentations beginning in the early 1900's and the acetone produced was used as a major source for gunpowder during the First World War. The fermentation was widely used until the 1960's when the process was no longer able to compete with the emergent petrochemical process due to rising costs of fermentable sugars. The bacterial and yeast CBP platforms makes the production of isopropanol readily tractable.

1.2 Pathway Definition and Stoichiometric Calculations for Production of Mixed Alcohols

[0415] The combined production of propanediol and isopropanol from glucose is outlined in the pathways of Figure 1 and requires the activity of several distinct enzymes (Table 2).

Table 2. List of native and non-native gene candidates pertaining to engineering of mixed alcohols in bacteria and yeast CBP platforms.

Activity	EC	Cthe	Tsacch	Yeast	Non-native-bacteria	Non-native-yeast
methylglyoxal synthase	4.2.3.3	95	or2316			<i>Oryza sativa</i> <i>mgs</i>
aldo-keto reductase (methylglyoxal to acetol)	1.1.1.-	152 236 283	or1401 or1402 or785 or414 or2491			
aldo-keto reductase (acetol to propanediol)	1.1.1.-	101 394 423 2445 2579	or1043 or2289 or411 or2426 or0286			
phosphotransacetylase	2.3.1.8	1029	or1741			<i>Tsacch</i> or1741
acetate kinase	2.7.2.1	1028	or1742			<i>Tsacch</i> or1742
thiolase	2.3.1.9			ERG10	<i>C. acetobutylicum</i>	
coA transferase	2.8.3.8				<i>C. acetobutylicum</i>	<i>C. acetobutylicum</i>

acetoacetate decarboxylase	4.1.1.4				<i>C.</i> <i>acetobutylicum</i>	<i>C.</i> <i>acetobutylicum</i>
isopropanol dehydrogenase	1.1.1.80	101 394 423 2445 2579	or1411 or1043 or2426 or2289 or0286	ADH1 ADH2 ADH3 ADH4 ADH5 ADH6 ADH7 BDH2 SFA1 YPL088 W		
alcohol dehydrogenase	1.1.1.1	423	or411			
PFOR (oxidoreductase)	1.2.7.1	2390- 3	or0047			
fructose 1,6- biphosphate aldolase	4.1.2.13	0349 1019	or0260 or0330	FBA1		
triose-phosphate isomerase	5.3.1.1	0139	or2687	TPI1		
glycerol-3-phosphate dehydrogenase	1.1.1.8			GPD2		
glycerol-3- phosphatase	3.1.3.21			GPP1		
pyruvate formate- lyase	2.3.1.54					<i>E. coli</i> pflA / pflB
formate dehydrogenase	1.2.1.2			FDH1		
aldehyde reductase	1.1.1.21	101 394 423 2445 2579	or1043 or2289 or411 or2426 or0286	GRE3		
glyoxylate reductase	1.1.1.79	152 236 283	or1401 or1402 or785 or414 or2491	GOR1 YPL113 C		
methylglyoxal dehydrogenase	1.2.1.49	152 236 283	or1401 or1402 or785 or414 or2491	GCY1 ALD2 ALD3 ALD4 ALD5 ALD6 HFD1		

Genes to KO						
Activity	EC	Cthe	Tsacch	Yeast		
lactate dehydrogenase	1.1.1.27	1053	or180			
pyruvate decarboxylase	4.1.1.1			PDC1 PDC5 PDC6		

[0416] The branched metabolic pathways can be subdivided into distinct production routes as follows:

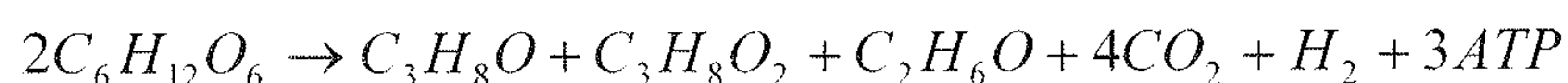
[0417] (i) the conversion of dihydroxyacetone phosphate into propanediol

[0418] (ii) the conversion of pyruvate into isopropanol

[0419] (iii) the conversion of pyruvate into ethanol (bacterial CBP platform only)

[0420] (iv) the conversion of dihydroxyacetone phosphate into glycerol (yeast CBP platform only).

[0421] The combined production of isopropanol, propanediol, and ethanol (routes (i), (ii), and (iii)) from two glucose molecules during bacterial metabolism is governed by the overall stoichiometric equation with a theoretical yield of one propanol, one propanediol, and one ethanol per two glucose, as follows:



[0422] The theoretical yield of propanediol, propanol, and ethanol on hexose and pentose sugar for the above pathway is:

Hexose

0.21 g propanediol/g sugar

0.17 g isopropanol/g sugar

0.13 g ethanol/g sugar

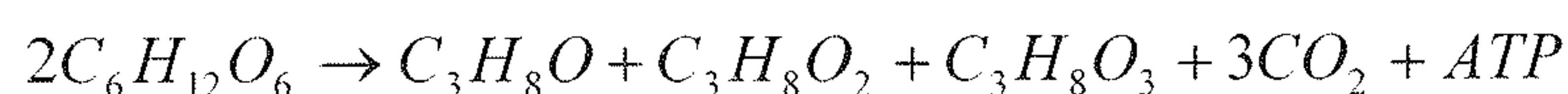
Pentose

0.21 g propanediol/g sugar

0.17 g isopropanol /g sugar

0.13 g ethanol /g sugar

[0423] The combined production of isopropanol, propanediol, and glycerol in yeast, *S. cerevisiae*, (routes (i), (ii), and (iv)) results in the net gain of one ATP, and is governed by the overall stoichiometric equation:



[0424] The co-production of isopropanol and propanediol together with the loss of carbon to glycerol and CO₂ are necessary to maintain the redox balance. The theoretical yield of propanediol, propanol, and glycerol on hexose and pentose sugar for the above pathway is:

Hexose

0.21 g propanediol/g sugar

0.17 g isopropanol/g sugar

0.26 g glycerol/g sugar

Pentose

0.21 g propanediol/g sugar

0.17 g isopropanol /g sugar

0.26 g glycerol /g sugar

[0425] The above stoichiometric equations were calculated using a hexose as a carbohydrate source; however, pentose sugars, including but not limited to xylose, can be readily utilized as well. When a pentose sugar is used as the carbohydrate source, six pentose sugars are required as the equivalent for five hexose sugars.

1.3 Production Routes for Mixed Alcohols and Corresponding Enzymology

[0426] Bacterial CBP Platforms

[0427] The combined production of propanediol, isopropanol, and ethanol from glucose in a bacterial CBP platform can be subdivided into the following distinct production routes: (i) the conversion of dihydroxyacetone phosphate into propanediol; (ii) the conversion of pyruvate into isopropanol; and (iii) the conversion of pyruvate into ethanol (Figure 1). The microbial hosts utilize carbohydrate sources, shown as glucose in Figure 1, to produce the mixed alcohols, but as mentioned above, pentose sugars such as xylose can be readily utilized as well, requiring six pentose sugars as equivalent for five hexose sugars. The first step in the pathway uses the microbial host's cellular metabolism to metabolize the carbohydrate source, employing, *e.g.*, the Embden-Meyerhof-Parnas (EMP) pathway to produce dihydroxyacetone phosphate and glyceraldehyde 3-phosphate (Figure 1). These metabolites can be interchanged using triosephosphate isomerase (E.C. 5.3.1.1).

[0428] During route (i), dihydroxyacetone phosphate is converted to methylglyoxal by methylglyoxal synthase (E.C. 4.2.3.3). Methylglyoxal is subsequently converted to either acetol by an oxidoreductase, which is to be identified from EC 1.1.1.- (*see* Table 2), or lactaldehyde by a keto-reductase (E.C. 1.1.1.79, 1.2.1.49). These intermediates are further reduced to propanediol by, oxidoreductases (E.C. 1.1.1.-) for acetol or (E.C. 1.1.1.2) 1 lactaldehyde.

[0429] For route (ii), glyceraldehyde 3-phosphate is further metabolized to pyruvate through standard glycolysis reactions, producing ATP to power the cellular reactions and the required reducing equivalents needed to reduce the carbon end-products. During bacterial metabolism, pyruvate is metabolized to acetyl-CoA, reduced ferredoxin, and

CO₂ by pyruvate ferredoxin oxidoreductase (E.C. 1.2.7.1) (Figure 1, light gray box). NADH and H₂ are subsequently produced during the oxidation of ferredoxin. Acetyl-CoA is then converted to acetate by phosphate acetyltransferase (E.C. 2.3.1.8) and acetate kinase (E.C. 2.7.2.1) in an ATP generating reaction. Two acetyl-CoA molecules are converted to acetoacetyl-CoA by thiolase (E.C. 2.3.1.9). Acetoacetyl-CoA is then converted to acetoacetate by CoA enzyme transferase (E.C. 2.8.3.8), where the CoA species is transferred from acetoacetyl-CoA to acetate, replenishing the acetyl-CoA consumed during the thiolase reaction. Acetoacetate is then converted to acetone by acetoacetate decarboxylase (E.C. 4.1.1.4). The reduction of acetone to isopropanol can be accomplished by an alcohol dehydrogenase (E.C. 1.1.1.80).

[0430] In route (iii), acetyl-CoA is converted to ethanol by acetaldehyde dehydrogenase (EC 1.2.1.3) and an alcohol dehydrogenase (E.C. 1.1.1.1), or through a bi-functional enzyme catalyzing both steps.

[0431] All the required enzymatic activities have been demonstrated in *C. thermosaccharolyticum* (see Cameron, D.C., & Clooney, C., 1986, *Bio/Technology* 4, 651-654) and relevant endogenous enzymes in the bacteria CBP platform production strains that exhibit high levels of homology to the desired enzymatic domains have been identified (see Table 2). The enzymes catalyzing the production of acetone from acetyl-CoA have been identified in the literature, and activities associated with (E.C. 2.3.1.9), (E.C. 2.8.3.8), and (E.C. 4.1.1.4) can be engineered using genes from *C. acetobutylicum*. See Bermejo, L.L., *et al.*, 1998, *Appl Environ Microbiol.* 64(3), 1079-85.

[0432] The conversion of acetone to isopropanol has been shown by multiple alcohol dehydrogenases and endogenous enzymes from the microbial CBP hosts can be screened for their capability to accept acetone as a substrate. Additional efforts must be made to readily control the flux through the different metabolic branch points through the modulation of enzyme levels and regulation. To this end, the deletion of *ldh* (E.C. 1.1.1.27) will prevent flow of carbon from pyruvate to lactic acid (see Table 2, "Genes to KO").

[0433] Yeast CBP Platforms

[0434] The combined production of propanediol, isopropanol, and glycerol from glucose in a yeast CBP platform can be subdivided into the following distinct production routes: (i) the conversion of dihydroxyacetone phosphate into propanediol; (ii) the conversion of

pyruvate into isopropanol; and (iv) the conversion of dihydroxyacetone phosphate into glycerol (Figure 1). As described above, the microbial hosts utilize carbohydrate sources, such as glucose as shown in Figure 1, or a pentose sugar such as xylose. The first step in the pathway uses the microbial host's cellular metabolism to metabolize the carbohydrate source, employing, *e.g.*, the Embden-Meyerhof-Parnas (EMP) pathway to produce dihydroxyacetone phosphate and glyceraldehyde phosphate (Figure 1). These metabolites can be interchanged using triosephosphate isomerase (E.C. 5.3.1.1).

[0435] Route (i) is proposed in the yeast CBP platform in a similar manner as route (i) in the bacteria CBP platform, converting dihydroxyacetone phosphate to methylglyoxal and using the two alternate pathways presented to generate propanediol from methylglyoxal. *See* Figure 1. However, based on current yeast literature, only a third route might be available, in part, because all result in the same redox change. All three begin with the production of methylglyoxal from dihydroxyacetone phosphate by methylglyoxal synthase, MGS, (E.C. 4.2.3.3) which can be obtained from one of several potential sources. The introduction of the *mgs* gene alone in yeast has been shown to result in the production of propanediol, but at relatively low titers; subsequent introduction of a glycerol dehydrogenase (E.C. 1.1.1.6) doubled the amount of propanediol formed. *See* Hoffman, M.L., 1999, Metabolic engineering of 1,2-propanediol production in *Saccharomyces cerevisiae*. Ph.D. Dissertation, University of Wisconsin-Madison. Alternatively, aldehyde reductase (E.C. 1.1.1.21) may be capable of converting methylglyoxal to lactaldehyde and then subsequently to propanediol. The native yeast aldehyde reductase, GRE3, can be overexpressed to test this possibility. In addition methylglyoxal could potentially be converted to lactaldehyde by glyoxylate reductase (E.C. 1.1.1.79) or by methylglyoxal dehydrogenase (E.C. 1.2.1.49). These enzymatic activities have not been reported in *S. cerevisiae*, but there are a number of endogenous genes which may contain these activities: two potential glyoxylate reductases (GOR1 and YPL113C), a glycerol dehydrogenase (GCY1), six aldehyde dehydrogenases (ALD2-6 and HFD1), and the ten alcohol dehydrogenases mentioned below. *See* Table 2. It might be desirable to engineer in a combination of the two alternate pathways outlined above for producing propanediol from methylglyoxal to reach a desirable titer for propanediol.

[0436] For route (ii), glyceraldehyde 3-phosphate is further metabolized to pyruvate through standard glycolysis reactions, as described above for bacteria CBP platforms. In

yeast metabolism, acetyl-CoA and formate is produced from pyruvate by pyruvate formate lyase (E.C. 2.3.1.8) (Figure 1, dark gray box). Formate is further metabolized to CO₂, NADH, and H₂ by formate dehydrogenase (E.C. 1.2.1.2) (Figure 1, dark gray box). Production of isopropanol from acetyl-CoA is performed as described above for the bacteria CBP platform.

[0437] Five enzymatic activities can be engineered into yeast for route (ii). The pyruvate formate lyase (PFL) (E.C. 2.3.1.8) is required for the formation of acetyl-CoA in the cytosol, because in a majority of yeast species the endogenously produced acetyl-CoA is sequestered in the mitochondria. Enzymatically active PFL has been expressed in yeast for the production of formate. See Waks, Z., & Silver, P. A., 2009, *Appl. Env. Microbiol.* 75, 1867-1875. *S. cerevisiae* has an endogenous formate dehydrogenase (E.C. 1.2.1.2) to convert the formate generated to CO₂ and H₂. The cytosolic acetyl-CoA generated is subsequently converted to acetone by the introduction of the *C. acetobutylicum* pathway, as described above for the bacteria CBP platform, working together with the yeast acetyl-CoA acetyltransferase, ERG10, (E.C. 2.3.1.9). An alcohol dehydrogenase executes the final reaction in this section, acetone to isopropanol. The *S. cerevisiae* genome encodes for ten alcohol dehydrogenases (ADH1-7, BDH2, SFA1, and YPL088W), which can be assayed for the capability of converting acetone to isopropanol. See Table 2. If necessary an exogenous alcohol dehydrogenase can be engineered into *S. cerevisiae*. Three pyruvate decarboxylase genes (E.C. 4.1.1.1) can be deleted: PDC1, PDC5, and PDC6. The presence of these three enzymes would result in the loss of significant pyruvate to acetaldehyde.

[0438] In route (iv), dihydroxyacetone phosphate is converted to glycerol by glycerol-3-phosphate dehydrogenase (E.C. 1.1.1.8) and glycerol-3-phosphatase (E.C. 3.1.3.21) (Figure 1, dark gray boxes). The enzymes required for route (iv) are already present in *S. cerevisiae*.

EXAMPLE 2

2.1 Production of n-Propanol and Isopropanol in Bacterial CBP Platforms

[0439] All current native and recombinant propanol producing metabolic pathways have at most a theoretical yield of 0.33 g propanol/g carbohydrate. Yan Y. & Liao J. 2009, *J*

Indus Microbiol and Biotech 36(4):471-479. This yield, corresponding to one mole isopropanol per mole glucose, incorporates into isopropanol only 75% of the free energy available from glucose during anaerobic fermentation. The additional 25% of the free energy, also referred to as available electrons, must be incorporated into a co-product during anaerobic fermentation, or consumed by oxygen during aerobic fermentation.

[0440] The present example proposes a new pathway for propanol production from lignocellulosic carbohydrates at a yield of 0.44 g/g carbohydrate, and incorporates 100% of the free energy available from carbohydrate conversion. In order to produce propanol at this theoretical maximum yield using biochemical pathways found in nature, production of both n- and iso- forms are required. In the metabolic pathway described here, isopropanol production serves in an ATP generating capacity, while n-propanol production serves as an electron sink to balance the anaerobic fermentation. This pathway allows for a balanced fermentation equation that is thermodynamically feasible.

[0441] Both products can be recovered from the fermentation broth via distillation, reducing downstream processing complexity. Isopropanol is a product natively produced by solventogenic *Clostridia*, and is rapidly produced by *Thermoanaerobacter* species when fed with acetone, indicating the presence of a native alcohol dehydrogenase with high activity for the desired reaction. See Lamed RJ and Zeikus JG. 1981, *The Biochemical J* 195(1):183-190. Acetone production has been extensively studied, and the Clostridial pathway has been heterologously expressed in *E. coli* as described above. See Bermejo, L.L., et al., 1998, *Appl. Environ. Microbiol.* 64(3), 1079-85. n-propanol is a natural product of propanediol degradation, with many microorganisms reported to perform this catalysis under anaerobic conditions. Recently, the genes involved in this conversion have been identified in one species, *Listeria innocua*, which will facilitate the expression of this pathway in the bacterial CBP organisms. See Xue J. et al., 2008, *Applied and Environmental Microbiol.* 74(22):7073-7079. Propanediol, a key intermediate of the n-propanol pathway, is a natural fermentation product of thermophilic bacteria. *T. thermosaccharolyticum* HG-8, the organism reported to produce the highest titer of propanediol, can be engineered for the production of n-propanol.

2.2 Pathway Definition and Stoichiometric Calculations for Production of Propanols

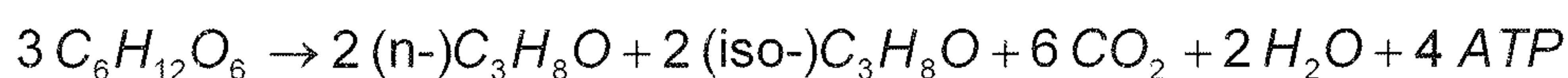
[0442] The combined production of n-propanol and isopropanol from glucose or xylose is outlined in the pathways of Figure 2 and requires the activity of several distinct enzymes (Table 3).

Table 3. List of native and non-native gene candidates pertaining to engineering of n-propanol and isopropanol in the CBP bacterial platform.

Activity	EC	<i>C. thermocellum</i>	<i>T. saccharolyticum</i>	Non-native bacteria
triose phosphate isomerase	5.3.1.1	139	or2687	
methylglyoxal synthase	4.2.3.3	95	or2316	
aldo-keto reductase (methylglyoxal to acetol)	1.1.1.-	152 236 283	or1401 or1402 or785 or414 or2491	
aldo-keto reductase (acetol to propanediol)	1.1.1.-	101 394 423 2445 2579	or1043 or2289 or411 or2426 or0286	
propanediol dehydratase	4.2.1.28		or0222, or0224-or0226	<i>T. sacch</i> genes can be expressed in <i>C. therm</i>
propanaldehyde dehydrogenase	1.1.1.202	101 394 423 2579	0411 1043 2426 2289 0286	
phosphotransacetylase	2.3.1.8	1029	or1741	
acetate kinase	2.7.2.1	1028	or1742	
thiolase	2.3.1.9			<i>C. acetobutylicum</i>
coA transferase	2.8.3.8			<i>C. acetobutylicum</i> CtfAB
acetoacetate decarboxylase	4.1.1.4			<i>C. acetobutylicum</i> Adc, Aad
PFOR (oxidoreductase)	1.2.7.1	2390-93	or0047	

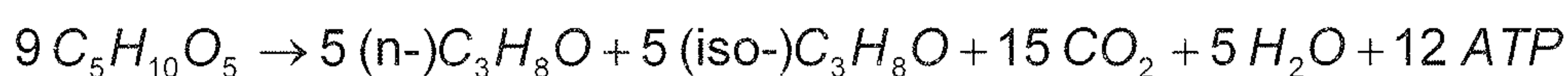
Genes to KO				
Activity	EC	C.the	T.sacch	Non-native-bacteria
alcohol dehydrogenase	1.1.1.1	423	or411	
lactate dehydrogenase	1.1.1.27	1053	or180	
hydrogenase	1.12.7.2	425-31	or1545-48	

[0443] The combined production of n-propanol and isopropanol from 3 glucose molecules during bacterial metabolism is governed by the overall stoichiometric equation:



[0444] The theoretical yield of propanols on a hexose sugar for the above pathway is 0.44 g propanols/g hexose.

[0445] The combined production of n-propanol and isopropanol from 9 xylose molecules during bacterial metabolism is governed by the overall stoichiometric equation:



[0446] The theoretical yield of propanols on a pentose sugar for the above pathway is 0.44 g propanols/g hexose.

[0447] For this metabolic pathway, product yields are identical for hexose, *e.g.*, glucose, and pentose, *e.g.*, xylose, carbohydrates due to the activity of triosephosphate isomerase (*tpi*) (E.C. 5.3.1.1). Pentose fermentation produces more of the isomer glyceraldehyde 3-phosphate (GAP) than dihydroxyacetone phosphate (DHAP) compared to hexose fermentation, which produces equimolar ratios of the two compounds. However, *tpi* allows for the conversion of GAP to DHAP and vice-versa, creating equal product yields for both carbohydrates.

2.3 Production Routes for Propanols and Corresponding Enzymology

[0448] The metabolic pathways for the production of n-propanol and isopropanol can be subdivided into two distinct production routes: (i) the conversion of dihydroxyacetone phosphate into n-propanol; and (ii) the conversion of pyruvate into isopropanol.

[0449] For the n-propanol route, route (i), dihydroxyacetone phosphate is converted to methylglyoxal by methylglyoxal synthase (E.C. 4.2.3.3). Methylglyoxal is subsequently converted to acetol by an oxidoreductase (E.C. 1.1.1.-) or to lactaldehyde by a keto-

reductase (1.1.1.79 or 1.2.1.49). These intermediates are then further reduced to propanediol by enzymes from (E.C. 1.1.1.-). Propanediol is then dehydrated to propanal by a diol-hydrolase (E.C. 4.2.1.28) and reduced to n-propanol by a dehydrogenase (E.C. 1.1.1.202). See Figure 2.

[0450] All the required enzymatic activities for the production of propanediol have been demonstrated in *C. thermosaccharolyticum*, a strain that can be genetically engineered. Cameron, D.C., *et al.*, 1998, *Biotechnol. Prog.* 14, 116–125. Relevant endogenous enzymes in the bacterial CBP platform production strains that exhibit high levels of homology to the desired enzymatic domains have also been identified (Table 3). The enzymes leading to propanediol in the bacterial CBP platform production strains can be characterized for implementation in route (i).

[0451] For the isopropanol route, route (ii), glyceraldehyde 3-phosphate is further metabolized to pyruvate through standard glycolysis reactions, producing ATP to power cellular reactions and reducing equivalents needed to balance n-propanol production during anaerobic fermentation. Pyruvate is then metabolized to acetyl-CoA, reduced ferredoxin, and CO₂ by pyruvate ferredoxin oxidoreductase (E.C. 1.2.7.1). NADH and H₂ are subsequently produced during the oxidation of ferredoxin. See Figure 2.

[0452] Acetyl-CoA is then converted to acetate by phosphate acetyltransferase (EC 2.3.1.8) and acetate kinase (E.C. 2.7.2.1) in an ATP generating reaction. Two acetyl-CoA molecules are converted to acetoacetyl-CoA by thiolase (E.C. 2.3.1.9). Acetoacetyl-CoA is then converted to acetoacetate by CoA enzyme transferase (E.C. 2.8.3.8), where the CoA species is transferred from acetoacetyl-CoA to acetate, replenishing the acetyl-CoA consumed during the thiolase reaction. Acetoacetate is then converted to acetone by acetoacetate decarboxylase (E.C. 4.1.1.4). The reduction of acetone to isopropanol can be accomplished by alcohol dehydrogenases (E.C. 1.1.1.80).

[0453] The enzymes catalyzing the production of acetone from acetyl-CoA have been identified in the literature from *C. acetobutylicum*. See Bermejo, L.L., *et al.*, 1998, *Appl Environ Microbiol.* 64(3), 1079-85. The conversion of acetone to isopropanol has been shown by multiple alcohol dehydrogenases and endogenous bacterial enzymes can be screened for their capability to accept acetone as a substrate.

[0454] Gene deletions will also be required to achieve high yields of propanol production. These include deletion of L-lactate dehydrogenase, *ldh* (E.C. 1.1.1.27); hydrogenase, *hyd* (E.C. 1.12.7.2); and acetaldehyde dehydrogenase, *acdh* (E.C. 1.2.1.10).

Example 3

3.1 Production of Isopropanol and Methanol in Bacterial CBP Platforms

[0455] Co-production of isopropanol and methanol from lignocellulosic carbohydrates allows for a balanced fermentation equation that is thermodynamically feasible. Isopropanol is theoretically produced at 0.33 g/g carbohydrate and incorporates 75% of the electrons available from carbohydrate conversion. Both isopropanol and methanol can be recovered from the fermentation broth via distillation, reducing downstream processing complexity. Further, methanol is a natural product of pectin degradation, and many characterized methylotropic organisms contain genes for methanol metabolism.

3.2 Pathway Definition and Stoichiometric Calculations for Production of Isopropanol and Methanol

[0456] The production of isopropanol and methanol from carbohydrates is outlined in the pathways in Figure 3 and requires the activity of several distinct enzymes (*see* Table 4).

Table 4. List of native and non-native gene candidates pertaining to engineering of isopropanol and methanol in the CBP bacterial platform.

Activity	EC	<i>C.the</i>	<i>T.sacch</i>	Non-native-bacterial
pyruvate formate lyase	2.3.1.54	505	or0628	
phosphotransacetylase	2.3.1.8	1029	or1741	
acetate kinase	2.7.2.1	1028	or1742	
formaldehyde dehydrogenase	1.2.1.46	218	2445, 0388	Pput_0350 <i>P. putida</i>
methanol dehydrogenase	1.1.1.244	101 394 423 2445 2579	or1411 or1043 or2426 or2289 or286	
formate dehydrogenase	1.2.1.43	342 430	or2328	Moth_2312 <i>M. thermoacetica</i>

		3004 3003		
thiolase	2.3.1.9			<i>C. acetobutylicum</i> ThlA
coA transferase	2.8.3.8			<i>C. acetobutylicum</i> CtfAB
acetoacetate decarboxylase	4.1.1.4			<i>C. acetobutylicum</i> Adc, Aad
oxidoreductase	1.1.1.80	101 394 423 2445 2579	or1411 or1043 or2426 or2289 or0286	
Genes to KO				
	EC	C.the	T.sacch	
lactate dehydrogenase	1.1.1.27	1053	or180	
alcohol dehydrogenase	1.1.1.1	423	or411	
methylglyoxal synthase	4.2.2.3	95	or2316	
PFOR (oxidoreductase)	1.2.7.1	2390- 93	or0047	

[0457] The combined production of isopropanol and methanol from one glucose molecule during bacterial metabolism is governed by the overall stoichiometric equation, with a theoretical yield of one propanol and one methanol per glucose, as follows:



[0458] The theoretical yield of isopropanol and methanol on hexose and pentose sugar for the above pathways (see Figure 3) are:

Hexose yield:

0.33 g isopropanol /g hexose

0.18 g methanol/ g hexose

Pentose yield:

0.33 g isopropanol/g pentose

0.18 g methanol/g pentose

[0459] During cellular metabolism, the microbial hosts can utilize hexose or pentose carbohydrate sources, with six pentose sugars equivalent to five hexose sugars, employing, e.g., the Embden-Meyerhof-Parnas (EMP) pathway to produce

dihydroxyacetone phosphate and glyceraldehyde 3-phosphate. These metabolites can be interchanged using the triosephosphate isomerase (E.C. 5.3.1.1).

3.3 Production Routes for Isopropanol and Methanol and Corresponding Enzymology

[0460] The branched metabolic pathways for the combined production of isopropanol and methanol from carbohydrates can be subdivided into the following production routes: (i) the conversion of pyruvate into isopropanol; and (ii) the conversion of formate into CO₂ and methanol.

[0461] As described above, glyceraldehyde 3-phosphate is metabolized to pyruvate through standard glycolysis reactions, producing ATP to power the cellular reactions and the required reducing equivalents needed to reduce the carbon end-products. From pyruvate, acetyl-CoA and formate are produced by pyruvate formate lyase (E.C. 2.3.1.54). For isopropanol production, route (i), acetyl-CoA is converted to acetate by phosphate acetyltransferase (E.C. 2.3.1.8) and acetate kinase (E.C. 2.7.2.1) in an ATP generating reaction. Two acetyl-CoA molecules are converted to acetoacetyl-CoA by thiolase (E.C. 2.3.1.9). Acetoacetyl-CoA is then converted to acetoacetate by CoA enzyme transferase (E.C. 2.8.3.8), where the CoA species is transferred from acetoacetyl-CoA to acetate, replenishing the acetyl-CoA consumed during the thiolase reaction. Acetoacetate is then converted to acetone by acetoacetate decarboxylase (E.C. 4.1.1.4). The reduction of acetone to isopropanol can be accomplished by alcohol dehydrogenases (E.C. 1.1.1.80).

[0462] As described above, the enzymes catalyzing the production of acetone from acetyl-CoA have been identified in the literature from *C. acetobutylicum*. See Bermejo, L.L., *et al.*, 1998, *Appl Environ Microbiol.* 64(3), 1079-85. The conversion of acetone to isopropanol has been shown by multiple alcohol dehydrogenases and endogenous bacterial enzymes can be screened for their capability to accept acetone as a substrate.

[0463] In route (ii), formate is further metabolized via two pathways in an equimolar ratio first leading to CO₂ and NADPH by formate dehydrogenase (E.C. 1.2.1.43), and the second leading to methanol with the incorporation of two NADH and production of water by the combined action of formaldehyde dehydrogenase (E.C. 1.2.1.46) and methanol dehydrogenase (E.C. 1.1.1.244).

[0464] The production of CO₂ and NADPH via formate is a well characterized pathway with a large body of literature. However, the production of methanol via formate is a less well characterized pathway. The majority of characterized organisms that have methanol metabolism pathways consume methanol, rather than produce it. Methanol production from formate is thermodynamically feasible under anaerobic conditions. The most likely route for engineering a high yielding pathway is to introduce enzymes that natively catalyze the net reaction in the reverse direction and then use evolutionary engineering techniques to select for strains with increased flux towards methanol formation. This strategy for pathway flux improvement has been successfully employed both in the engineering of other metabolic pathways and is anticipated to work for this pathway due to the thermodynamic favorability of the net reaction.

EXAMPLE 4

4.1 Anaerobic Production of Propanediol and Acetone in Bacterial and Yeast CBP Platforms

[0465] The native microbial production of propanediol has been well documented in *Clostridium thermosaccharolyticum* by Cameron, D.C., & Clooney, C., 1986 *Bio/Technology* 4, 651-654, although the endogenous enzymes have yet to be identified and cloned. The native enzymes can be identified from the bacterial CBP platform microbes and utilized in the bacterial CBP platform hosts eliminating the need for "recombinant" genes (e.g., *Thermoanaerobacter saccharolyticum* and *Clostridium thermocellum*) and/or readily transferred to the yeast CBP platform hosts.

[0466] The theoretical maximum yield for anaerobic propanediol production that includes ATP generation requires the production of a co-fermentation product such as acetate. See U.S. Patent No. 6,303,352. The pathways presented in this Example achieve the anaerobic maximum theoretical yield and use acetate as an intermediate during the generation of acetone as the co-fermentation product. Acetone was chosen as a co-fermentation product because it is potentially a chemical of value and a less toxic fermentation product to the microorganisms relative to acetate. The simultaneous production of propanediol and acetone represents a novel fermentation process. In addition, relatively little is known about the enzymology converting methylglyoxal to propanediol, but as described above, can now be ascertained.

4.2 Pathway Definition and Stoichiometric Calculations for Production of Propanediol and Acetone

[0467] The anaerobic production of propanediol and acetone from carbohydrates is outlined in the pathways in Figure 4 and requires the activity of several distinct enzymes (see Table 5).

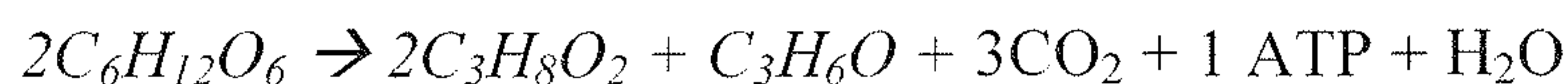
Table 5. List of native and non-native gene candidates pertaining to engineering of propanediol and acetone in the CBP bacterial and CBP yeast platforms.

Activity	EC	<i>C.th</i> <i>e</i>	<i>T.sacc</i> <i>h</i>	Yeast	Non-native- bacteria	non-native- yeast
methylglyoxal synthase	4.2.3.3	95	or2316			<i>Oryza sativa mgs</i>
aldo-keto reductase (methylglyoxal to acetol)	1.1.1.-	152 236 283	or1401 or1402 or785 or414 or2491			<i>P. putida gldA</i>
aldo-keto reductase (acetol to propanediol)	1.1.1.-	101 394 423 2445 2579	or1043 or2289 or411 or2426 or0286			
phosphotransacetylase	2.3.1.8	1029	or1741			<i>Tsacch</i> or1741
acetate kinase	2.7.2.1	1028	or1742			<i>Tsacch</i> or1742
thiolase	2.3.1.9			ERG10	<i>C. acetobutylicum</i> ThlA	
coA transferase	2.8.3.8				<i>C. acetobutylicum</i> CtfAB	<i>C. acetobutylicum</i> CtfAB
acetoacetate decarboxylase	4.1.1.4				<i>C. acetobutylicum</i> Adc, Aad	<i>C. acetobutylicum</i> Adc, Aad
alcohol dehydrogenase	1.1.1.1			ADH1 ADH2 ADH3 ADH4 ADH5 ADH6 ADH7 BDH2		

				SFA1 YPL088 W		
PFOR (oxidoreductase)	1.2.7.1	2390 -3	or0047			
fructose 1,6- biphosphate aldolase	4.1.2.13	0349 1019	or0260 or0330	FBA1		
triose-phosphate isomerase	5.3.1.1	0139	or2687	TPI1		
pyruvate formate- lyase	2.3.1.54					<i>E. coli</i> pflA / pflB
formate dehydrogenase	1.2.1.2			FDH1		
aldehyde reductase	1.1.1.21	101 394 423 2445 2579	or1043 or2289 or411 or2426 or0286	GRE3		
glyoxylate reductase	1.1.1.79	101 394 423 2445 2579	or1043 or2289 or411 or2426 or0286	GOR1 YPL113C		
methylglyoxal dehydrogenase	1.2.1.49	152 236 283	or1401 or1402 or785 or414 or2491	GCY1 ALD2 ALD3 ALD4 ALD5 ALD6 HFD1		
glucokinase	2.7.1.2	0390 2938	or0272	GLK1		
glucose 6 phosphate isomerase	5.3.1.9	0217	or1389	PGI1		
6-phosphofructokinase	2.7.1.11	1261	or2875	PFK1		
				PFK2		
Genes to KO						
Activity	EC	<i>C.the</i>	<i>T.sac</i> <i>ch</i>	Yeast		
lactate dehydrogenase	1.1.1.27	1053	or180			
alcohol dehydrogenase	1.1.1.1	423	or411			
pyruvate decarboxylase	4.1.1.1			PDC1		
				PDC5		
				PDC6		
glycerol-3-phosphate dehydrogenase	1.1.1.8			GPD2		
glycerol-3-	3.1.3.21			GPP1		

phosphatase						
-------------	--	--	--	--	--	--

[0468] The combined production of propanediol and acetone from two glucose molecules during bacterial or yeast anaerobic metabolism is governed by the overall stoichiometric equation, resulting in overall redox balance and the net gain of one ATP, as follows:



[0469] The theoretical yield of propanediol and acetone on hexose and pentose sugar for the above pathway are:

Hexose

0.42 g propanediol/g hexose

0.16 g acetone/g hexose

Pentose

0.42 g propanediol/g pentose

0.16 g acetone/g pentose

[0470] During cellular metabolism, the microbial hosts can utilize hexose or pentose carbohydrate sources, with six pentose sugars equivalent to five hexose sugars, employing the Embden-Meyerhof-Parnas (EMP) pathway to produce dihydroxyacetone phosphate and glyceraldehyde 3-phosphate. These metabolites can be interchanged using the triosephosphate isomerase (EC 5.3.1.1).

4.3 Anaerobic Production Routes for Propanediol and Acetone and Corresponding Enzymology

[0471] The co-production of propanediol and acetone from hexose and pentose sugars in thermophilic clostridia and *S. cerevisiae* can be broken down into three routes: (i) the production of dihydroxyacetone phosphate and glyceraldehyde 3-phosphate from glucose; (ii) the subsequent generation of propanediol from dihydroxyacetone phosphate; and (iii) the generation of acetone from glyceraldehyde 3 phosphate. *See* Figure 4.

[0472] For the bacterial and yeast CBP platforms, the enzyme activities required for route (i), production of dihydroxyacetone phosphate and glyceraldehyde 3-phosphate from glucose, are part of the native glycolytic pathway, *e.g.*, the EMP pathway, as described above. *See* Table 5.

[0473] For route (ii), the subsequent generation of propanediol from dihydroxyacetone phosphate, two alternative routes are presented, in part because both result in the same redox balance and *a priori* the best route is not known. Both begin with the production of methylglyoxal from dihydroxyacetone phosphate by methylglyoxal synthase, *mgs* (E.C.

4.2.3.3). *See* Figure 4. This gene is endogenous to the bacterial CBP platform organisms, however for yeast it will have to be obtained from one of several potential sources.

[0474] For the bacterial CBP platform, which comprises thermophilic bacteria, acetol is the likely intermediate from methylglyoxal to propanediol, as has been shown in *T. thermosaccharolyticum*. *See* Cameron, D.C., & Clooney, C., 1986, *Bio/Technology* 4, 651-654. In *E. coli*, various aldo-keto reductases have been shown to catalyze the conversion of methylglyoxal to acetol (E.C. 1.1.1.-). *See* Ko, J., et al., 2005, *J Bacteriol.* 187(16), 5782-9. The list of endogenous aldo-keto reductases for the bacterial platform organisms are shown in Table 5. These genes can be over-expressed and/or deleted to determine their role in propanediol production. It is also possible that lactaldehyde, produced by a glyoxylate reductase (E.C. 1.1.1.79) and a methylglyoxal dehydrogenase (E.C. 1.2.1.49) is an intermediate. To determine if acetol or lactaldehyde is the primary intermediate during conversion of methylglyoxal to propanediol, analytical chemistry procedures such as HPLC can be used to identify these intermediates in fermentation samples. *See e.g.*, Cameron, D.C., & Clooney, C., 1986, *Bio/Technology* 4, 651-654; Altaras, N.E., & Cameron, D.C., 1999, *Appl Environ Microbiol.* 65(3), 1180-5. Alternatively, cells can be fed acetol or lactaldehyde to determine which intermediate is more effectively converted to propanediol. To determine which genes are responsible for the production of propanediol from acetol or lactaldehyde, the native alcohol dehydrogenases and aldo-keto reductases listed in Table 5 can be deleted and/or over-expressed while propanediol production is monitored.

[0475] For the yeast CBP platform, multiple routes from methylglyoxal to propanediol also exist. *See* Figure 4. One route through lactaldehyde involves introduction of a glycerol dehydrogenase (E.C. 1.1.1.-), which doubled the amount of propanediol formed. *See* Hoffman, M.L., 1999, Metabolic engineering of 1,2-propanediol production in *Saccharomyces cerevisiae*. Ph.D. Dissertation, University of Wisconsin-Madison. Alternatively, aldehyde reductase (E.C. 1.1.1.21) may be capable of converting methylglyoxal to lactaldehyde and then subsequently to propanediol—the native yeast aldehyde reductase, GRE3, can be overexpressed to test this possibility. In addition, methylglyoxal could potentially be converted to lactaldehyde by glyoxylate reductase (E.C. 1.1.1.79) or to lactaldehyde by methylglyoxal dehydrogenase (E.C. 1.2.1.49). The presence of these alcohol dehydrogenase activities can be screened among the ten native

alcohol dehydrogenases. See Table 5. It might be necessary to engineer in a combination of the two pathways outlined above to reach a desirable titer for propanediol.

[0476] The enzymes that convert methylglyoxal to propanediol are oxidoreductases, of which there are examples using either NADH or NADPH as a co-factor. Knowledge of the co-factor is important for producing propanediol in the yeast platform because the compartmentalization of the cell, and the relative difficulty of inter-converting NADH to NADPH, limit the cell's ability to deal with an imbalance in these cofactors. For the anaerobic production of propanediol, an enzyme (or enzymes) that are linked to NADH would be required, since these are the reducing equivalents generated during the production of CO₂ and acetone from glyceraldehyde 3-phosphate. Several of the enzymes identified in bacterial systems have this characteristic.

[0477] For route (iii), the generation of acetone from glyceraldehydes 3-phosphate, the engineering of non-native enzymatic activities into both the bacterial and yeast platforms is required. The bacterial organisms have a native enzyme activity (E.C. 1.2.7.1) that converts pyruvate to acetyl-CoA (Figure 4, light gray box), while the yeast platform requires the expression of a non-native activity (E.C. 2.3.1.54) to convert pyruvate to acetyl-CoA (Figure 9a, dark gray box).

[0478] To convert acetyl-CoA to acetone in the bacterial platform, activities associated with (E.C. 2.3.1.9), (E.C. 2.8.3.8), and (E.C. 4.1.1.4) can be engineered using genes from *C. acetobutylicum*, while activities associated with (E.C. 1.2.7.1), (E.C. 2.3.1.8), and (E.C. 2.7.2.1) are in fact endogenous (Figure 4). See Bermejo, L.L., et al., 1998, *Appl Environ Microbiol.* 64(3), 1079-85. Taken together, these activities will allow the formation of acetone from two molecules of pyruvate. For the yeast platform three enzymatic activities can be engineered into yeast. The pyruvate formate lyase, PFL (E.C. 2.3.1.54), is required for the formation of acetyl-CoA in the cytosol, because the majority of yeast endogenously produced acetyl-CoA is sequestered in the mitochondria. Enzymatically active PFL has been expressed in yeast for the production of formate. Waks, Z., & Silver, P. A., 2009, *Appl. Env. Microbiol.* 75, 1867-1875. *S. cerevisiae* has an endogenous formate dehydrogenase (E.C. 1.2.1.2) to convert the formate generated to CO₂ and H⁺. The cytosolic acetyl-CoA generated can be subsequently converted to acetone by the introduction of the *C. acetobutylicum* pathway (E.C. 2.8.3.8) and (E.C.

4.1.1.4), as described above, working together with the yeast acetyl-CoA acetyltransferase, ERG10 (E.C. 2.3.1.9).

[0479] The description of the above pathways describes native and non-native genes required to direct carbon flow from sugars to propanediol and acetone. In addition, to prevent decreases in product yield, *i.e.*, carbon from flowing away from desired end products, various genes can be deleted from each platform. For the bacterial CBP system, these genes are shown in Table 5. The deletion of *adh* (E.C. 1.1.1.1) will prevent flow from acetyl-CoA to acetaldehyde while the deletion of *ldh* (E.C. 1.1.1.27) will prevent flow of carbon from pyruvate to lactic acid. Deleting the hydrogenase genes (E.C. 1.12.7.2) will ensure that reducing equivalents generated during glycolysis can be used to make reduced end products such as 1,2-propanediol and not the more oxidized couple of H₂ and acetate. For the yeast CBP platform, genes to be deleted are listed in Table 5. Genes encoding activity associated with (E.C. 4.1.1.1) can be deleted to prevent carbon flow from pyruvate to acetaldehyde. In addition, genes associated with (E.C. 1.1.1.8) and (E.C. 3.1.3.21) activity can be deleted to prevent carbon loss from dihydroxyacetone phosphate as glycerol.

Example 5

Aerobic Production of Propanediol in Yeast CBP Platforms

[0480] The purpose of the present Example is to provide a novel pathway for the aerobic production of propanediol in yeast CBP platforms. Aerobic production of propanediol provides some benefits in terms of ATP production. For example, the advantages of aerobic production are discussed in Cameron *et al.*, "Metabolic engineering of propanediol pathways," *Biotechnology Progress*, 14(1): 116-125 (1998), where a yield of 0.61 g propanediol/g can be achieved in a non-compartmentalized organism. Indeed, the commercial production of 1,3-propanediol is done via an aerobic process. Although not as high as 0.61 g propanediol/g in a non-compartmentalized organism, the present pathway provides for a high yield of propanediol in a compartmentalized organism as discussed below.

[0481] The 1,2-propandiol produced using this platform can be used as a valuable intermediate or converted to propionate and propanol using microbes such as

Lactobacillus reuteri strain isolated from sourdough that is known to do this reaction. See Sriramulu, D.D., *et al.*, 2008, *J. Bacteriol.* 190(13):4559-67. Chemical routes might also exist for direct conversion of propanediol to propanol or even propylene.

Pathway Definition and Stoichiometric Calculations for Production of Propanediol

[0482] The aerobic production of propanediol from carbohydrates is outlined in the pathways in Figure 5 and requires the activity of several distinct enzymes (*see* Table 5).

[0483] The production of propanediol, which is the only soluble product of the reaction, from 6 glucose molecules during yeast aerobic metabolism is governed by the overall stoichiometric equation:



[0484] In order to balance the redox in the cytosol, 1 molecule of glucose 6-phosphate must be completely oxidized by the pentose phosphate pathway (PPP) for every molecule of propanediol produced. In addition, a positive ATP balance is generated via oxidation of the glyceraldehyde 3-phosphate in the TCA cycle and the electron transport chain. See Figure 4.

[0485] The theoretical yield of propanediol on hexose sugar for the above pathway is 0.42 g propanediol/g hexose. 100% xylose could not be converted via this pathway, but a glucose/xylose mixture could convert with a yield similar to glucose alone. Although not as high of a yield as for a non-compartmentalized organism, the proposed pathway provides a high yield for propanediol. Further, the possibility of shuttling NADH to the cytosol from the mitochondrial matrix cannot be ruled out since such a shuttle has been demonstrated. See Bakker, B.M., *et al.*, 2000, *Appl. Env. Micro.* 182, 4730-4737. This would potentially allow higher yields in *S. cerevisiae*. In *Kluyveromyces* type yeasts, yields might also be increased due to shuttling of reducing equivalents to the cytoplasm, and the enhanced activity of the pentose phosphate pathway in these organisms.

5.3 Aerobic Production Routes for Propanediol and Corresponding Enzymology

- [0486] For the production of dihydroxyacetone phosphate and glyceraldehyde 3-phosphate from glucose, the enzyme activities are part of the native glycolytic pathway, *e.g.*, the EMP pathway, as described above. *See* Table 5 and Figure 5.
- [0487] For the subsequent generation of propanediol from dihydroxyacetone phosphate, two alternative routes are presented as in Example 4 (*see* Figure 4), in part because both result in the same redox balance and *a priori* the best route is not known. Both begin with the production of methylglyoxal from dihydroxyacetone phosphate by methylglyoxal synthase, *mgs* (E.C. 4.2.3.3). *See* Figure 4. For yeast, this gene will have to be obtained from one of several potential sources.
- [0488] As described above in Example 4, multiple routes from methylglyoxal to propanediol exist in yeast. *See* Figure 4. One route through lactaldehyde involves introduction of a glycerol dehydrogenase (E.C. 1.1.1.-), which doubled the amount of propanediol formed. *See* Hoffman, M.L., 1999, Metabolic engineering of 1,2-propanediol production in *Saccharomyces cerevisiae*. Ph.D. Dissertation, University of Wisconsin-Madison. Alternatively, aldehyde reductase (E.C. 1.1.1.21) may be capable of converting methylglyoxal to lactaldehyde and then subsequently to propanediol—the native yeast aldehyde reductase, GRE3, can be overexpressed to test this possibility. In addition, methylglyoxal could potentially be converted to lactaldehyde by glyoxylate reductase (E.C. 1.1.1.79) or to lactaldehyde by methylglyoxal dehydrogenase (E.C. 1.2.1.49). The presence of these alcohol dehydrogenase activities can be screened among the ten native alcohol dehydrogenases. *See* Table 5. It might be necessary to engineer in a combination of the two pathways outlined above to reach a desirable titer for propanediol.
- [0489] As described above, the enzymes that convert methylglyoxal to propanediol are oxidoreductases, of which there are examples using either NADH or NADPH as a co-factor. Knowledge of the co-factor is important for producing propanediol in the yeast platform because the compartmentalization of the cell, and the relative difficulty of inter-converting NADH to NADPH, limit the cell's ability to deal with an imbalance in these cofactors. In the aerobic production of propanediol, the NADPH linked versions of an enzyme (or enzymes) are required, since the production of reducing equivalents in the

form of NADPH is accomplished in the pentose phosphate pathway. The *S. cerevisiae* *gre3* gene is a good example (and candidate) for use in the aerobic system.

[0490] To convert the carbohydrate source to propanediol in yeast using an aerobic process, control of the flux of carbon down particular pathways will be needed. Redox balance is obtained by controlling flux to the PPP and propanediol, while optimal product yield is obtained when the flux to the TCA cycle and electron transport chain is held to a minimal level. Controlling flux to the PPP involves manipulating the expression level of *zwf1*, which converts glucose 6-phosphate to D-glucono-1,5-lactone 6-phosphate, relative to the activity of *pgi*, which converts glucose 6-phosphate to fructose 6-phosphate. In order to control the amount of flux to the TCA cycle and the electron transport chain, one of two methods could be used. One would be to down-regulate PDH, and thereby reduce the amount of pyruvate being converted to acetyl-CoA in the mitochondria. The other would be to control the oxygen flux in the fermentation vessel to limit the amount of oxygen available for the electron transport chain. The former genetic approach has an advantage in that it alleviates the necessity of careful process control for aeration at large scale.

Example 6

Identification and Characterization of *T. saccharolyticum* pdu Gene Cluster

[0491] Several microorganisms metabolize propanediol to propanol anaerobically. Examples of propanediol utilization can be found among various bacterial species including *Thermoanaerobacteria*, *Salmonella*, *Listeria*, and *Clostridia*. In some microorganisms, e.g., *Listeria* spp. and *Salmonella* spp., the genes required for propanediol utilization (pdu) are clustered on the genome. See generally Scott, K.P., et al., *J. Bacteriol.* 188(12):4340-49 (2006); Bobik, T.A., et al., *J. Bacteriol.* 181(19):5967-75; Xue, J., et al., *Appl. Env. Microbiol.* 74(22):7073-79 (2008).

[0492] Two enzyme activities required for conversion of propanediol to propanol include:

- 1) diol dehydratase (encoded by pduCDE) and
- 2) dehydrogenase (encoded by pduQ).

See Table 3 above. In several microorganisms, such as *Salmonella* spp. or *Listeria* spp., the first enzyme activity often involves catalysis via a heteromeric diol dehydratase

enzyme that is dependent on vitamin B12. The *pdu* gene clusters are often found to include or be associated with the enzymes required for the synthesis of vitamin B12. Some of the *pdu* gene clusters include genes for 1) B12 synthesis, 2) AraC type transcription activator, 3) two-component response regulator, 4) an alcohol and aldehyde dehydrogenase, or 5) rnfC homolog. See Scott, K.P., *et al.*, *J. Bacteriol.* 188(12):4340-49 (2006); Bobik, T.A., *et al.*, *J. Bacteriol.* 181(19):5967-75; Xue, J., *et al.*, *Appl. Env. Microbiol.* 74(22):7073-79 (2008).

[0493] Thus far, no *pdu* gene clusters have been identified in thermophilic anaerobic bacteria. This Example provides the identification and characterization of the *T. saccharolyticum pdu* gene cluster for its use in conversion of propanediol to propanol, following, *e.g.*, the scheme described in Example 2.

[0494] The *pdu* gene organization in *T. saccharolyticum* is shown in Figure 6 and includes several of the genes found in *pdu* gene clusters from other microorganisms. The *T. saccharolyticum pdu* genes include ABC-sugar transporter components (or201, or202, or203), a two-component response regulator (or206, or207), rhamnose isomerase (or209) rhamulokinase (or208), rhamnulose-1-phosphate lactaldehyde lyase (or227), a putative propanediol:NAD⁺ oxidoreductase (or211), micro-compartment proteins (*pduJ*, *pduL*, etc. or212, or214, or215, or216, or217), an aldehyde dehydrogenase (or219), an alcohol dehydrogenase (or218), a phosphotransacetylase (or213), B12 accessory enzymes (or223, or222, or221, or220), a B-12 dependent diol dehydratase *pduCDE* (or226, or225, or224), and a transcriptional regulator (or228). The activities of these genes can be characterized, *e.g.*, through various gene deletion studies, growth on rhamnose, and/or expression into heterologous systems such as *T. thermosaccharolyticum* and *C. thermocellum*.

[0495] The ability of *T. saccharolyticum*, which harbors the above-identified *pdu* gene cluster, to produce detectable levels of n-propanol was determined. The wild-type *T. saccharolyticum* YS485 strain was grown in TSC1 medium (Table 6) with 10 g/L CaCO₃ and a starting pH of 5.8 at 55 °C and 200 rpm under anaerobic conditions. The medium was supplemented with 0.001 g/L vitamin B12.

Table 6. Composition of TSC1 medium.

Components	Concentration (g/L)
------------	------------------------

(NH ₄) ₂ SO ₄	1.85
FeSO ₄ *7H ₂ O	0.05
KH ₂ PO ₄	0.5
MgSO ₄	1
CaCl ₂ *2H ₂ O	0.05
Trisodium citrate * 2 H ₂ O	2
Yeast Extract	8.5
CaCO ₃	10
L-rhamnose	18

[0496] Batch fermentation was done and samples were drawn at various time points shown in Table 7. The samples were analyzed by HPLC to detect remaining L-rhamnose and end products, including lactic acid (LA), acetic acid (AA), ethanol (Etoh), 1,2-propanediol (1,2 PD), and n-propanol. The results are depicted in Table 7.

Table 7. Production of 1,2-Propanediol and n-Propanol in *T. saccharolyticum* Grown on L-rhamnose

Time (hr)	L-rhamnose (g/L)	LA (g/L)	AA (g/L)	Etoh (g/L)	1,2 PD (g/L)	n-propanol (g/L)
0	17.820	0.000	0.118	0.000	0.277	0.000
17.5	11.440	0.242	2.054	0.315	2.248	0.363
24.5	2.522	0.346	4.289	0.437	4.623	1.072
41.25	0.679	0.384	5.024	0.527	5.073	1.525
69.5	0.427	0.407	5.135	0.567	5.134	1.638

[0497] These results demonstrate that *T. saccharolyticum* has the native ability to produce 1,2-propanediol (up to 5.1 g/L) and n-propanol (1.6 g/L) when grown on L-rhamnose. The *pdu* gene cluster includes some rhamnose utilization and sugar uptake genes indicating that those are likely to be involved in this process. This provides the first example of a thermophilic anaerobic bacterium shown to be capable of producing n-propanol.

Example 7

Production of Propanol Via Propanediol Using a B12-independent Diol Dehydratase in Yeast

[0498] As described above, one of the two enzyme activities required for conversion of propanediol to propanol includes a diol dehydratase enzyme, which in several microorganisms is dependent on vitamin B12. Yeast lack the metabolic machinery to synthesize vitamin B12, and thus, it is not possible to engineer a vitamin B12-dependent enzyme in yeast without also providing, *e.g.*, the enzyme activities to synthesize vitamin B12. There have been a few reports of propanediol dehydratase enzymes that do not require vitamin B12. *See* Raynaud, C., *et al.*, *PNAS (USA)* 100(9):5010-15 (2003); Scott, K.P., *et al.*, *J. Bacteriol.* 188(12):4340-49 (2006); Hartmanis, M.G., and Stadtman, T.C., *Arch. Biochem. Biophys.* 245(1):144-52 (1986).

[0499] Because of the requirement for vitamin B12, the anaerobic conversion of propanediol to propanol was thought to be impossible due to the requirement of a vitamin B12-dependent enzyme. Recent reports describing the B12-independent diol dehydratase provide a source and incentive to screen for existing B12-independent diol dehydratases in nature and express them into yeast. *See* Raynaud, C., *et al.*, *PNAS (USA)* 100(9):5010-15 (2003); Scott, K.P., *et al.*, *J. Bacteriol.* 188(12):4340-49 (2006); Hartmanis, M.G., and Stadtman, T.C., *Arch. Biochem. Biophys.* 245(1):144-52 (1986). If successfully done, this would be the first n-propanol producing yeast engineered so far. The purpose of this Example is to identify and engineer a vitamin B12-independent diol dehydratase, as well as other necessary enzymes, in yeast, *e.g.*, *Saccharomyces cerevisiae*, to anaerobically convert propanediol to propanol.

[0500] The metabolic pathway for generating propanol from, *e.g.*, a carbohydrate source, in yeast is similar to the route described above in Example 2 and as shown in Figure 2. In order to successfully achieve this conversion of glucose, several enzyme activities need to be engineered in yeast. Conversion of glucose to pyruvate and dihydroxyacetone-P are achieved via the endogenous enzyme activities in yeast. Those activities which need to be engineered are highlighted in Figure 7 and are as follows:

[0501] 1) The conversion of pyruvate to acetyl-CoA and formate via pyruvate-formate lyase (PFL) (E.C. 2.3.1.8) has been successfully engineered and demonstrated. *See* Waks, Z. and Silver, P.A., *Appl. Env. Microbiol.* 75(7):1867-75 (2009). This is an

important step to generate a pool of acetyl-CoA in the yeast cytosol for its subsequent conversion into isopropanol. Simultaneously, the flux of pyruvate to acetyl-CoA via pyruvate decarboxylase (PDC) needs to be avoided for which the PDC1, PDC5 and PDC6 need to be knocked out. The conversion of formate to carbon dioxide is catalyzed by an endogenous enzyme, formate dehydrogenase (E.C. 1.2.1.2).

[0502] 2) Acetyl-CoA is further converted to acetate by phosphate acetyltransferase (E.C. 2.3.1.8) and acetate kinase (E.C. 2.7.2.1) in an ATP generating reaction. Two acetyl-CoA molecules are converted to acetoacetyl-CoA by thiolase (E.C. 2.3.1.9). Acetoacetyl-CoA is then converted to acetoacetate by CoA enzyme transferase (E.C. 2.8.3.8), where the CoA species is transferred from acetoacetyl-CoA to acetate, replenishing the acetyl-CoA consumed during the thiolase reaction. Acetoacetate is then converted to acetone by acetoacetate decarboxylase (E.C. 4.1.1.4). The reduction of acetone to isopropanol can be accomplished by alcohol dehydrogenases (E.C. 1.1.1.80).

[0503] 3) Synthesis of methylglyoxal from dihydroxyacetone-P can be achieved by expression of heterologous methylglyoxal synthase (mgs) and glycerol dehydrogenase (gldA) as has been previously demonstrated. See Lee, W. and DaSilva, N.A., *Metabolic Eng.* 8(1):58-65 (2006).

[0504] 4) The conversion of propanediol to propanol requires two enzyme activities as described above, involving a diol dehydratase and a dehydrogenase. Although several microorganisms can convert 1,2-propandiol to propanol using a vitamin B12-dependent diol dehydratase, reaction via a vitamin B12-dependent diol dehydratase is not feasible in yeast due to the B12 dependency. The few recently discovered examples of vitamin B12-independent diol dehydratase include those identified from *Clostridium butyricum*, *Roseburia inulinivorans*, *Clostridium glycolicum* and *Klebsiella* spp. The *C. butyricum* enzyme is extensively characterized and shown to be functional independent of B12 and in a heterologous system (*E. coli*). See Tang, X., et al., *Appl. Env. Microbiol.* 75(6):1628-34 (2009). The results obtained with the *C. butyricum* B12-independent diol dehydratase activity suggest that the enzyme can be engineered into a heterologous system such as yeast.

[0505] In addition to the incorporation of these enzymatic activities, the flux of carbon from pyruvate to ethanol must be disrupted in yeast. This can be accomplished via the deletion of *pdcl*, *pdcs*, and *pdcs*. PDC deletion strains are slow growing and require a

small amount of added ethanol or acetate to be viable; however, these issues can be overcome via an evolutionary based approach. *See, e.g., van Maris, A.J.A., et al., Appl. Env. Microbiol. 70(1):159-66 (2004).* The fact that such strains produce pyruvate at high levels indicates that this compound would be available for subsequent conversion to propanol via the proposed pathway above.

[0506] In order to identify additional B12-independent diol dehydratases for engineering in part 4 above, other B12-independent diol dehydratase enzymes existing in nature can be identified. Suitable methods for identifying can include, but are not limited to, alignment searches based on homology to known B12-independent diol dehydratases, an enzymatic activity assay combined with protein purification and protein sequencing, and whole-genome transcriptional analysis of 1,2 propanediol utilizing organisms. *See, e.g., Scott, K. P. et al., J. Bact 188(12):4340-4349 (2006), and Raynaud, C. et al., PNAS 100(9):5010-5015 (2003).*

[0507] Once identified and isolated, the gene responsible for the activity is cloned into yeast along with other enzyme activities as described above. Optimization of expression of the B12-independent diol dehydratase and analytical assays for production of propanol is subsequently followed.

Incorporation by Reference

[0508] All of the U.S. patents and U.S. published patent applications cited herein are hereby incorporated by reference.

Equivalents

[0509] Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following claims.

CLAIMS

1. A recombinant microorganism;

wherein said microorganism expresses one or more native and/or heterologous enzymes;

wherein said one or more enzymes function in one or more engineered metabolic pathways to convert a carbohydrate source to 1,2-propanediol, isopropanol and ethanol or to 1,2-propanediol, isopropanol and glycerol; and

wherein said one or more native and/or heterologous enzymes is activated, upregulated or downregulated.
2. The recombinant microorganism of claim 1, wherein said carbohydrate source is lignocellulose.
3. The recombinant microorganism of claim 1 or 2, wherein said carbohydrate source is converted to dihydroxyacetone phosphate and glyceraldehyde phosphate.
4. The recombinant microorganism of any one of claims 1-3, wherein one of said engineered metabolic pathways comprises the conversion of dihydroxyacetone phosphate to 1,2-propanediol.
5. The recombinant microorganism of any one of claims 1-4, wherein one of said engineered metabolic pathways comprises conversion of pyruvate to isopropanol.
6. The recombinant microorganism of any of claims 1-5, wherein one of said engineered metabolic pathway comprises conversion of pyruvate to ethanol.
7. The recombinant microorganism of any of claims 1-5, wherein one of said engineered metabolic pathway comprises conversion of dihydroxyacetone phosphate to glycerol.
8. The recombinant microorganism of any one of claims 1-7, wherein the engineered metabolic pathway that produces 1, 2 propanediol comprises the following steps: (a) conversion of dihydroxyacetone phosphate to methylglyoxal; (b) conversion of methylglyoxal to acetol or lactaldehyde; and (c) reduction of acetol or lactaldehyde to propanediol.

9. The recombinant microorganism of any one of claims 1-8, wherein the engineered metabolic pathway that produces isopropanol comprises: (a) conversion of acetyl-CoA to acetate; (b) conversion of acetyl-CoA to acetoacetyl-CoA; (c) conversion of acetoacetyl-CoA to acetoacetate; (d) conversion of acetoacetate to acetone; and (e) reduction of acetone to isopropanol.

10. The recombinant microorganism of any one of claims 1-6 or 8-9, wherein the engineered metabolic pathway that produces ethanol comprises: (a) conversion of acetyl-CoA to ethanol.

11. The recombinant microorganism of any one of claims 1-5 or 7-9, wherein the engineered metabolic pathway that produces glycerol comprises: (a) conversion of dihydroxyacetone phosphate to glycerol.

12. The recombinant microorganism of any one of claims 1-6 or 8-10, wherein said microorganism produces 1,2-propanediol, isopropanol and ethanol.

13. The recombinant microorganism of claim 12, wherein the combined production of 1,2-propanediol, isopropanol and ethanol is governed by the stoichiometric equation $2C_6H_{12}O_6 \rightarrow C_3H_8O + C_3H_8O_2 + C_2H_6O + 4CO_2 + H_2 + 3ATP$.

14. The recombinant microorganism of any one of claims 1-6, 8-10 or 12-13, wherein said microorganism is a thermophilic or mesophilic bacterium.

15. The recombinant microorganism of claim 14, wherein said thermophilic or mesophilic bacterium is a species of the genera *Thermoanaerobacterium*, *Thermoanaerobacter*, *Clostridium*, *Geobacillus*, *Saccharococcus*, *Paenibacillus*, *Bacillus*, *Caldicellulosiruptor*, *Anaerocellum*, or *Anoxybacillus*.

16. The recombinant microorganism of claim 15, wherein said microorganism is a bacterium selected from the group consisting of: *Thermoanaerobacterium thermosulfurigenes*, *Thermoanaerobacterium aotearoense*, *Thermoanaerobacterium polysaccharolyticum*, *Thermoanaerobacterium zeae*, *Thermoanaerobacterium xylanolyticum*, *Thermoanaerobacterium saccharolyticum*, *Thermoanaerobium brockii*, *Thermoanaerobacterium thermosaccharolyticum*, *Thermoanaerobacter thermohydrosulfuricus*, *Thermoanaerobacter ethanolicus*,

Thermoanaerobacter brocki, *Clostridium thermocellum*, *Clostridium cellulolyticum*, *Clostridium phytofermentans*, *Clostridium straminosolvens*, *Geobacillus thermoglucosidasius*, *Geobacillus stearothermophilus*, *Saccharococcus caldoxylosilyticus*, *Saccharococcus thermophilus*, *Paenibacillus campinasensis*, *Bacillus flavothermus*, *Anoxybacillus kamchatkensis*, *Anoxybacillus gonensis*, *Caldicellulosiruptor acetigenus*, *Caldicellulosiruptor saccharolyticus*, *Caldicellulosiruptor kristjanssonii*, *Caldicellulosiruptor owensensis*, *Caldicellulosiruptor lactoaceticus*, and *Anaerocellum thermophilum*.

17. The recombinant microorganism of claim 16, wherein said microorganism is selected from the group consisting of *Clostridium thermocellum*, and *Thermoanaerobacterium saccharolyticum*.

18. The recombinant microorganism of any one of claims 1-5, 7-9 or 11, wherein said microorganism produces 1,2-propanediol, isopropanol and glycerol.

19. The recombinant microorganism of claim 18, wherein the combined production of 1,2-propanediol, isopropanol and ethanol is governed by the stoichiometric equation $2C_6H_{12}O_6 \rightarrow C_3H_8O + C_3H_8O_2 + C_3H_8O_3 + 3CO_2 + ATP$.

20. The recombinant microorganism of any one of claims 1-5, 7-9, 11 or 18-19, wherein said microorganism is selected from the group consisting of *Saccharomyces cerevisiae*, *Kluyveromyces lactis*, *Kluyveromyces marxianus*, *Pichia pastoris*, *Yarrowia lipolytica*, *Hansenula polymorpha*, *Phaffia rhodozyma*, *Candida utliis*, *Arxula adeninivorans*, *Pichia stipitis*, *Debaryomyces hansenii*, *Debaryomyces polymorphus*, *Schizosaccharomyces pombe*, *Candida albicans*, and *Schwanniomyces occidentalis*.

21. The recombinant microorganism of claim 20, wherein said microorganism is *Saccharomyces cerevisiae*.

22. The recombinant microorganism of any one of claims 3-21, wherein said dihydroxyacetone phosphate is converted to methylglyoxal by methylglyoxal synthase (EC 4.2.3.3); wherein said methylglyoxal is converted to acetol by an aldo-keto reductase or converted to lactaldehyde by a glyoxylate reductase, a methylglyoxal dehydrogenase or an aldehyde dehydrogenase; and wherein said acetol is converted to propanediol by an aldo-keto reductase or said lactaldehyde is converted to propanediol by an aldehyde reductase.

23. The recombinant microorganism of claim 22, wherein said propanediol is converted to propionate and propanol biologically or chemically.

24. The recombinant microorganism of claim 23, wherein said biological conversion of propanediol is achieved using a microbe.

25. The recombinant microorganism of claim 24, wherein said microbe is *Lactobacillus reuteri*.

26. The recombinant microorganism of any one of claims 3-21, wherein said glyceraldehyde phosphate is converted to pyruvate; wherein said pyruvate is converted to acetyl-CoA by pyruvate formate lyase (E.C. 2.3.1.8); wherein said acetyl-CoA is converted to acetoacetyl-CoA by thiolase ERG10 (E.C. 2.3.1.9); wherein said acetoacetyl-CoA is converted to acetoacetate by a CoA transferase; wherein said acetoacetate is converted to acetone by an acetoacetate decarboxylase; and wherein said acetone is converted to isopropanol by isopropanol or alcohol dehydrogenase.

27. The recombinant microorganism of claim 26, wherein said alcohol dehydrogenase is selected from the group consisting of ADH1-7, BDH2, SFA1 and YPL088W.

28. The recombinant microorganism of claim 26 or 27, wherein said organism is a PDC1, PDC5 and/or PDC6 knockout strain.

29. The recombinant microorganism of claim 22 or 26, wherein said methylglyoxal synthase, aldo-keto reductase, glyoxylate reductase, methylglyoxal dehydrogenase, aldehyde reductase, pyruvate formate lyase, thiolase, CoA transferase, acetoacetate decarboxylase, isopropanol, aldehyde dehydrogenase and/or alcohol dehydrogenase is selected from any one of the corresponding enzymes listed in Table 2.

30. The microorganism of any one of claims 1-6, 8-10 or 12-17, which generates at least 10%, at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, or at least 95% of the theoretical yield of propanediol, propanol and ethanol, wherein the theoretical yield is 0.21 g propanediol/g carbohydrate, 0.17 g propanol/g carbohydrate and 0.13 g ethanol/g carbohydrate.

31. The microorganism of any of claims 1-5, 7-9, 11, or 18-21 which generates at least 10%, at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, or at least 95% of the theoretical yield of propanediol, propanol and glycerol, wherein the theoretical yield is 0.21 g propanediol/g of carbohydrate, 0.17 g propanol/g carbohydrate and 0.26 g glycerol/g carbohydrate.

32. A recombinant microorganism;

wherein said microorganism expresses one or more native and/or heterologous enzymes;

wherein said one or more enzymes function in one or more engineered metabolic pathways to convert a carbohydrate source to n-propanol and isopropanol; and

wherein said one or more native and/or heterologous enzymes is activated, upregulated or downregulated.

33. The recombinant microorganism of claim 33, wherein said carbohydrate source is a lignocellulose.

34. The recombinant microorganism of claim 33, wherein the production of n-propanol and isopropanol is governed by the stoichiometric equation $3C_6H_{12}O_6 \rightarrow 2(n-)C_3H_8O + 2(iso-)C_3H_8O + 6CO_2 + 2H_2O + 4ATP$.

35. The recombinant microorganism of claim 33, wherein the production of n-propanol and isopropanol is governed by the stoichiometric equation $9C_5H_{10}O_5 \rightarrow 5(n-)C_3H_8O + 5(iso-)C_3H_8O + 15CO_2 + 5H_2O + 12ATP$.

36. The recombinant microorganism of any one of claims 32-35, which yields at least 0.35 g propanol/g carbohydrate source.

37. The recombinant microorganism of any one of claims 32-35, which yields at least 0.40 g propanol/g carbohydrate source.

38. The recombinant microorganism of anyone of claims 32-37, which generates at least 10%, at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at

least 80%, at least 90%, or at least 95% of the theoretical yield of propanol, wherein the theoretical yield is 0.44 g propanol/g carbohydrate.

39. The recombinant microorganism of any of claims 32-38, wherein said carbohydrate source is converted to dihydroxyacetone phosphate and glyceraldehyde phosphate.

40. The recombinant microorganism of claim 39, wherein one of said engineered metabolic pathways comprises the conversion of dihydroxyacetone phosphate to n-propanol.

41. The recombinant microorganism of claim 39 or 40, wherein one of said engineered metabolic pathways comprises the conversion of pyruvate to isopropanol.

42. The recombinant microorganism of any one of claims 39-41, wherein said dihydroxyacetone phosphate is converted to methylglyoxal by methylglyoxal synthase (E.C. 4.2.3.3); wherein said methylglyoxal is converted to acetol by an aldo-keto reductase or converted to lactaldehyde by a glyoxylate reductase, a methylglyoxal dehydrogenase or an aldehyde dehydrogenase; wherein said acetol is converted to propanediol by an aldo-keto reductase or said lactaldehyde is converted to propanediol by an aldehyde reductase; wherein said propanediol is dehydrated to propanal by diol-dehydratase (E.C. 4.2.1.28); and wherein said propanal is reduced to n-propanol by a dehydrogenase (E.C. 1.1.1.202).

43. The recombinant microorganism of any one of claims 39-41, wherein said glyceraldehyde phosphate is converted to pyruvate; wherein said pyruvate is converted to acetyl-CoA by pyruvate formate lyase (E.C. 2.3.1.8); wherein said acetyl-CoA is converted to acetoacetyl-CoA by thiolase ERG10 (E.C. 2.3.1.9); wherein said acetoacetyl-CoA is converted to acetoacetate by a CoA transferase; wherein said acetoacetate is converted to acetone by an acetoacetate decarboxylase; and wherein said acetone is converted to isopropanol by isopropanol or alcohol dehydrogenase.

44. The recombinant microorganism of claim 42 or 43, wherein said methylglyoxal synthase, aldo-keto reductase, glyoxylate reductase, methylglyoxal dehydrogenase, aldehyde reductase, pyruvate formate lyase, thiolase, CoA transferase, acetoacetate decarboxylase, isopropanol, aldehyde dehydrogenase, alcohol dehydrogenase, diol-dehydratase and/or dehydrogenase is selected from any one of the corresponding enzymes listed in Table 3.

45. The recombinant microorganism of any of claims 32-44, wherein said microorganism is a lactate dehydrogenase (*ldh*), hydrogenase (*hyd*) and/or acetaldehyde dehydrogenase (*acdH*) knockout strain.

46. The recombinant microorganism of any of claims 32-45, wherein said microorganism is a species of the genera *Thermoanaerobacterium*, *Thermoanaerobacter*, *Clostridium*, *Geobacillus*, *Saccharococcus*, *Paenibacillus*, *Bacillus*, *Caldicellulosiruptor*, *Anaerocellum*, or *Anoxybacillus*.

47. The recombinant microorganism of claim 46, wherein said microorganism is a bacterium selected from the group consisting of: *Thermoanaerobacterium thermosulfurigenes*, *Thermoanaerobacterium aotearoense*, *Thermoanaerobacterium polysaccharolyticum*, *Thermoanaerobacterium zeae*, *Thermoanaerobacterium xylanolyticum*, *Thermoanaerobacterium saccharolyticum*, *Thermoanaerobacterium brockii*, *Thermoanaerobacterium thermosaccharolyticum*, *Thermoanaerobacter thermohydrosulfuricus*, *Thermoanaerobacter ethanolicus*, *Thermoanaerobacter brocki*, *Clostridium thermocellum*, *Clostridium cellulolyticum*, *Clostridium phytofermentans*, *Clostridium straminosolvens*, *Geobacillus thermoglucosidasius*, *Geobacillus stearothermophilus*, *Saccharococcus caldoxylosilyticus*, *Saccharococcus thermophilus*, *Paenibacillus campinasensis*, *Bacillus flavothermus*, *Anoxybacillus kamchatkensis*, *Anoxybacillus gonensis*, *Caldicellulosiruptor acetigenus*, *Caldicellulosiruptor saccharolyticus*, *Caldicellulosiruptor kristjanssonii*, *Caldicellulosiruptor owensensis*, *Caldicellulosiruptor lactoaceticus*, and *Anaerocellum thermophilum*.

48. The recombinant microorganism of claim 47, wherein said microorganism is selected from the group consisting of *Clostridium thermocellum*, and *Thermoanaerobacterium saccharolyticum*.

49. The recombinant microorganism of any one of claims 32-45, wherein said microorganism is selected from the group consisting of *Saccharomyces cerevisiae*, *Kluyveromyces lactis*, *Kluyveromyces marxianus*, *Pichia pastoris*, *Yarrowia lipolytica*, *Hansenula polymorpha*, *Phaffia rhodozyma*, *Candida utliis*, *Arxula adeninivorans*, *Pichia stipitis*, *Debaryomyces hansenii*, *Debaryomyces polymorphus*, *Schizosaccharomyces pombe*, *Candida albicans*, and *Schwanniomyces occidentalis*.

50. A recombinant microorganism;

wherein said microorganism expresses one or more native and/or heterologous enzymes;

wherein said one or more enzymes function in one or more engineered metabolic pathways to convert a carbohydrate source to isopropanol and methanol; and

wherein said one or more native and/or heterologous enzymes is activated, upregulated or downregulated.

51. The recombinant microorganism of claim 50, wherein said carbohydrate source is a lignocellulose.

52. The recombinant microorganism of claim 51, wherein the production of n-propanol and isopropanol is governed by the stoichiometric equation $C_6H_{12}O_6 \rightarrow C_3H_8O + CH_4O + 2CO_2 + 3ATP$.

53. The recombinant microorganism of any one of claims 50-52, which yields at least 0.25 g propanol/g carbohydrate source.

54. The recombinant microorganism of any one of claims 50-53, which yields at least 0.30 g propanol/g carbohydrate source.

55. The recombinant microorganism of any one of claims 50-52, which generates at least 10%, at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, or at least 95% of the theoretical yield of propanol, wherein the theoretical yield is 0.33 g isopropanol/g carbohydrate and 0.18 g methanol/g carbohydrate.

56. The recombinant microorganism of any of claims 50-55, wherein said carbohydrate source is converted to pyruvate; and wherein said pyruvate is converted to acetyl-CoA and formate by pyruvate formate lyase (E.C. 2.3.1.54).

57. The recombinant microorganism of any of claims 50-56, wherein one of said engineered metabolic pathways comprises the conversion of pyruvate to isopropanol.

58. The recombinant microorganism of any of claims 50-57, wherein one of said engineered metabolic pathways comprises the conversion of formate to CO₂ and methanol.

59. The recombinant microorganism of any of claims 56-58, wherein said acetyl-CoA is converted to acetate by phosphotransacetylase (E.C. 2.3.1.8) and to acetoacetyl-CoA by thiolase (E.C. 2.3.1.9); wherein said acetoacetyl-CoA is converted to acetoacetate by a CoA transferase; wherein said acetoacetate is converted to acetone by an acetoacetate decarboxylase (E.C. 4.1.1.4); and wherein said acetone is reduced to isopropanol by an oxidoreductase or an alcohol dehydrogenase.

60. The recombinant microorganism of any one of claims 56-59, wherein said formate is methabolized via two pathways, the first resulting in the production of CO₂ and NADH by formate dehydrogenase and the second resulting in the production of methanol by formaldehyde dehydrogenase and methanol dehydrogenase.

61. The recombinant microorganism of claim 59 or 60, wherein said phosphotransacetylase, thiolase, CoA transferase, acetoacetate decarboxylase, oxidoreductase, alcohol dehydrogenase, formate dehydrogenase, formaldehyde dehydrogenase or methanol dehydrogenase is selected from any one of the corresponding enzymes listed in Table 4.

62. The recombinant microorganism of any of claims 50-61, wherein said microorganism is a species of the genera *Thermoanaerobacterium*, *Thermoanaerobacter*, *Clostridium*, *Geobacillus*, *Saccharococcus*, *Paenibacillus*, *Bacillus*, *Caldicellulosiruptor*, *Anaerocellum*, or *Anoxybacillus*.

63. The recombinant microorganism of claim 62, wherein said microorganism is a bacterium selected from the group consisting of: *Thermoanaerobacterium thermosulfurigenes*, *Thermoanaerobacterium aotearoense*, *Thermoanaerobacterium polysaccharolyticum*, *Thermoanaerobacterium zeae*, *Thermoanaerobacterium xylanolyticum*, *Thermoanaerobacterium saccharolyticum*, *Thermoanaerobium brockii*, *Thermoanaerobacterium thermosaccharolyticum*, *Thermoanaerobacter thermohydrosulfuricus*, *Thermoanaerobacter ethanolicus*, *Thermoanaerobacter brocki*, *Clostridium thermocellum*, *Clostridium cellulolyticum*, *Clostridium phytofermentans*, *Clostridium straminosolvens*, *Geobacillus thermoglucosidasius*, *Geobacillus stearothermophilus*, *Saccharococcus caldocxylosilyticus*, *Saccharococcus thermophilus*, *Paenibacillus campinasensis*, *Bacillus flavothermus*, *Anoxybacillus kamchatkensis*,

Anoxybacillus gonensis, *Caldicellulosiruptor acetigenus*, *Caldicellulosiruptor saccharolyticus*, *Caldicellulosiruptor kristjanssonii*, *Caldicellulosiruptor owensensis*, *Caldicellulosiruptor lactoaceticus*, and *Anaerocellum thermophilum*.

64. The recombinant microorganism of claim 63, wherein said microorganism is selected from the group consisting of *Clostridium thermocellum*, and *Thermoanaerobacterium saccharolyticum*.

65. A recombinant microorganism;

wherein said microorganism expresses one or more native and/or heterologous enzymes;

wherein said one or more enzymes function in one or more engineered metabolic pathways to convert a carbohydrate source to propanediol and acetone; and

wherein said one or more native and/or heterologous enzymes is activated, upregulated or downregulated.

66. The recombinant microorganism of claim 65, wherein said carbohydrate source is lignocellulose.

67. The recombinant microorganism of claim 65 or 66, wherein said carbohydrate source is converted to dihydroxyacetone phosphate and glyceraldehyde phosphate.

68. The recombinant microorganism of anyone of claims 65-67, wherein one of said engineered metabolic pathways comprises the conversion of dihydroxyacetone phosphate to propanediol.

69. The recombinant microorganism of anyone of claims 65-68, wherein one of said engineered metabolic pathways comprises conversion of pyruvate to acetone.

70. The recombinant microorganism of anyone of claims 65-69, wherein the engineered metabolic pathway that produces propanediol comprises the following steps: (a) conversion of

dihydroxyacetone phosphate to methylglyoxal; (b) conversion of methylglyoxal to acetol or lactaldehyde; and (c) reduction of acetol or lactaldehyde to propanediol.

71. The recombinant microorganism of anyone of claims 65-70, wherein the engineered metabolic pathway that produces acetone comprises: (a) conversion of acetyl-CoA to acetate; (b) conversion of acetyl-CoA to acetoacetyl-CoA; (c) conversion of acetoacetyl-CoA to acetoacetate; and (d) conversion of acetoacetate to acetone.

72. The recombinant microorganism of anyone of claims 65-71, wherein the engineered metabolic pathway that produces acetone comprises: (a) conversion of acetyl-CoA to acetone.

73. The recombinant microorganism of anyone of claims 65-72, wherein the combined production of propanediol and acetone is governed by the stoichiometric equation $C_6H_{12}O_6 \rightarrow 2C_3H_8O + C_3H_6O_8 + 3CO_2 + 1ATP$.

74. The recombinant microorganism of anyone of claims 65-73, wherein said microorganism is a thermophilic or mesophilic bacterium.

75. The recombinant microorganism of claim 74, wherein said thermophilic or mesophilic bacterium is a species of the genera *Thermoanaerobacterium*, *Thermoanaerobacter*, *Clostridium*, *Geobacillus*, *Saccharococcus*, *Paenibacillus*, *Bacillus*, *Caldicellulosiruptor*, *Anaerocellum*, or *Anoxybacillus*.

76. The recombinant microorganism of claim 75, wherein said microorganism is a bacterium selected from the group consisting of: *Thermoanaerobacterium thermosulfurigenes*, *Thermoanaerobacterium aotearoense*, *Thermoanaerobacterium polysaccharolyticum*, *Thermoanaerobacterium zae*, *Thermoanaerobacterium xylanolyticum*, *Thermoanaerobacterium saccharolyticum*, *Thermoanaerobium brockii*, *Thermoanaerobacterium thermosaccharolyticum*, *Thermoanaerobacter thermohydrosulfuricus*, *Thermoanaerobacter ethanolicus*, *Thermoanaerobacter brocki*, *Clostridium thermocellum*, *Clostridium cellulolyticum*, *Clostridium phytofermentans*, *Clostridium straminosolvens*, *Geobacillus thermoglucosidasius*, *Geobacillus stearothermophilus*, *Saccharococcus caldxylosilyticus*, *Saccharococcus thermophilus*,

Paenibacillus campinasensis, *Bacillus flavothermus*, *Anoxybacillus kamchatkensis*, *Anoxybacillus gonensis*, *Caldicellulosiruptor acetigenus*, *Caldicellulosiruptor saccharolyticus*, *Caldicellulosiruptor kristjanssonii*, *Caldicellulosiruptor owensensis*, *Caldicellulosiruptor lactoaceticus*, and *Anaerocellum thermophilum*.

77. The recombinant microorganism of claim 76, wherein said microorganism is selected from the group consisting of *Clostridium thermocellum* and *Thermoanaerobacterium saccharolyticum*.

78. The recombinant microorganism of anyone of claims 65-73, wherein said microorganism IS selected from the group consisting of *Saccharomyces cerevisiae*, *Kluyveromyces lactis*, *Kluyveromyces marxianus*, *Pichia pastoris*, *Yarrowia lipolytica*, *Hansenula polymorpha*, *Phaffia rhodozyma*, *Candida utliis*, *Arxula adeninivorans*, *Pichia stipitis*, *Debaryomyces hansenii*, *Debaryomyces polymorphus*, *Schizosaccharomyces pombe*, *Candida albicans*, and *Schwanniomyces occidentalis*.

79. The recombinant microorganism of claim 78, wherein said microorganism IS *Saccharomyces cerevisiae* or *Kluyveromyces lactis*.

80. The recombinant microorganism of anyone of claims 65-79, which generates at least 10%, at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, or at least 95% of the theoretical yield of propanol, wherein the theoretical yield is 0.42 g propanediol/g carbohydrate and 0.16 g acetone/ g carbohydrate.

81. The recombinant microorganism of claim 78 or 79, wherein production of propanediol is aerobic.

82. The recombinant microorganism of claim 81, wherein production of propanediol is governed by the stoichiometric equation $6 \text{ glucose} + 12\text{O}_2 \rightarrow 6 \text{ X propanediol} + 12\text{H}_2\text{O} + 18\text{CO}_2 + 26\text{ATP}$.

83. The recombinant microorganism of anyone of claims 65-82, wherein said dihydroxyacetone phosphate is converted to methyglyoxal by a methyglyoxal synthase (EC 4.2.3.3); wherein said methyglyoxal is converted to acetol by an aldo-keto reductase or converted to lactaldehyde by a glyoxylate reductase, a methyglyoxal dehydrogenase or an aldehyde dehydrogenase; and wherein said acetol is converted to propanediol by an aldo-keto reductase or said lactaldehyde is converted to propanediol by an aldehyde reductase.

84. The recombinant microorganism of anyone of claims 65-82, wherein said glyceraldehyde phosphate is converted to pyruvate; wherein said pyruvate is converted to acetyl-CoA by a pyruvate formate lyase (E.C. 2.3.1.8); wherein said acetyl-CoA is converted to acetoacetyl-CoA by a thiolase (E.C.2.3.1.9); wherein said acetoacetyl-CoA is converted to acetoacetate by a CoA transferase; and wherein said acetoacetate is converted to acetone by an acetoacetate decarboxylase.

85. The recombinant microorganism of any of claims 65-83, wherein said organism is a a lactate dehydrogenase (*ldh*), hydrogenase (*hyd*), alcohol dehydrogenase (*adh*), pyruvate decarboxylase (*pdh*), glycerol-3-phosphate dehydrogenase and/or glycerol-3-phosphate knockout strain.

86. The recombinant microorganism of claim 83 or 84, wherein said methyglyoxal synthase, aldo-keto reductase, glyoxylate reductase, methyglyoxal dehydrogenase, aldehyde reductase, pyruvate formate lyase, thiolase, CoA transferase, acetoacetate decarboxylase, isopropanol, aldehyde dehydrogenase and/or alcohol dehydrogenase is selected from anyone of the corresponding enzymes listed in Table 5.

87. A process for converting lignocellulosic biomass to 1,2-propanediol, isopropanol, methanol, n-propanol, or acetone comprising contacting lignocellulosic biomass with a recombinant microorganism according to anyone of claims 1-86.

88. The process of claim 87, wherein said lignocellulosic biomass is selected from the group consisting of grass, switch grass, cord grass, rye grass, reed canary grass, mixed prairie grass, miscanthus, sugar-processing residues, sugarcane bagasse, sugarcane straw, agricultural wastes,

rice straw, rice hulls, barley straw, corn cobs, cereal straw, wheat straw, canola straw, oat straw, oat hulls, corn fiber, stover, soybean stover, corn stover, forestry wastes, recycled wood pulp fiber, paper sludge, sawdust, hardwood, softwood, and combinations thereof.

89. An engineered metabolic pathway for converting lignocellulosic biomass to 1,2-propanediol, isopropanol, methanol, n-propanol, or acetone according to anyone of claims 1-86.

90. The recombinant microorganism of claim 42, wherein said propanediol is dehydrated to propanal by diol dehydratase from *T. saccharolyticum* (pduCDE).

91. The recombinant microorganism of claim 42, wherein said propanediol is dehydrated to propanal by a vitamin B12-independent diol dehydratase.

92. The recombinant microorganism of claim 91, wherein said recombinant microorganism is a yeast selected from the group consisting of *Saccharomyces cerevisiae*, *Kluyveromyces lactis*, *Kluyveromyces marxianus*, *Pichia pastoris*, *Yarrowia lipolytica*, *Hansenula polymorpha*, *Phaffia rhodozyma*, *Candida utilis*, *Arxula adeninivorans*, *Pichia stipitis*, *Debaryomyces hansenii*, *Debaryomyces polymorphus*, *Schizosaccharomyces pombe*, *Candida albicans*, and *Schwanniomyces occidentalis*.

93. The recombinant microorganism of claim 42, wherein said vitamin B12-independent diol-dehydratase is a vitamin B12-independent diol dehydratase from *Clostridium butyricum*, *Roseburia inulinivorans*, *Clostridium glycolicum*, or *Klebsiella* spp.

94. An isolated propanediol utilization (pdu) gene cluster from *T. saccharolyticum*.

95. The isolated propanediol utilization (pdu) gene cluster of claim 94, wherein said gene cluster comprises polynucleotides that encode one or more of an ABC-sugar transporter, a two-component response regulator, a rhamnose utilization enzyme, a micro-compartment protein pduJ, a micro-compartment protein pduL, an alcohol or an aldehyde dehydrogenase, a B12 enzyme, and a diol dehydratase pduCDE.

96. The isolated propanediol utilization (pdu) gene cluster of claim 94, wherein said gene cluster comprises one or more polynucleotides selected from the group consisting of SEQ ID NO:63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, and 119.

97. The isolated propanediol utilization (pdu) gene cluster of claim 94, wherein said gene cluster comprises one or more polynucleotides that encode polypeptides selected from the group consisting of SEQ ID NO:62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, and 118.

98. A recombinant microorganism comprising the isolated propanediol utilization (pdu) gene cluster of any one of claims 94-97.

99. A method for identifying a vitamin B12-independent diol dehydratase that converts propanediol to propanal comprising:

(a) measuring a property of a polypeptide or a polynucleotide and a reference vitamin B12-independent diol dehydratase polypeptide or polynucleotide and

(b) determining if the property of said polypeptide or said polynucleotide and the property of said reference vitamin B12-independent diol dehydratase polypeptide or polynucleotide are similar,

thereby identifying whether said polypeptide or said polynucleotide is a vitamin B12-independent diol dehydratase.

100. The method of claim 99, wherein said property comprises a physical or chemical property of the polypeptide or the polynucleotide, wherein said physical or chemical property includes genomic, transcriptomic, proteomic, and metabolomic properties.

101. The method of claim 99, wherein said reference vitamin B12-independent diol dehydratase polypeptide is a vitamin B12-independent diol dehydratase derived from the group consisting of *Clostridium butyricum*, *Roseburia inulinivorans*, *Clostridium glycolicum*, and *Klebsiella* spp.

102. The method of claim 99, wherein said measuring step involves obtaining and aligning a sequence of the polypeptide or polynucleotide with a sequence of a reference vitamin B12-independent diol dehydratase polypeptide or polynucleotide.

103. The method of claim 99, wherein said measuring step involves performing an enzymatic assay.

104. The method of claim 99, wherein said measuring step involves performing a whole-genome transcriptional assay.

105. A recombinant microorganism comprising the isolated vitamin B12-independent diol dehydratase polypeptide of any one of claims 99-104.

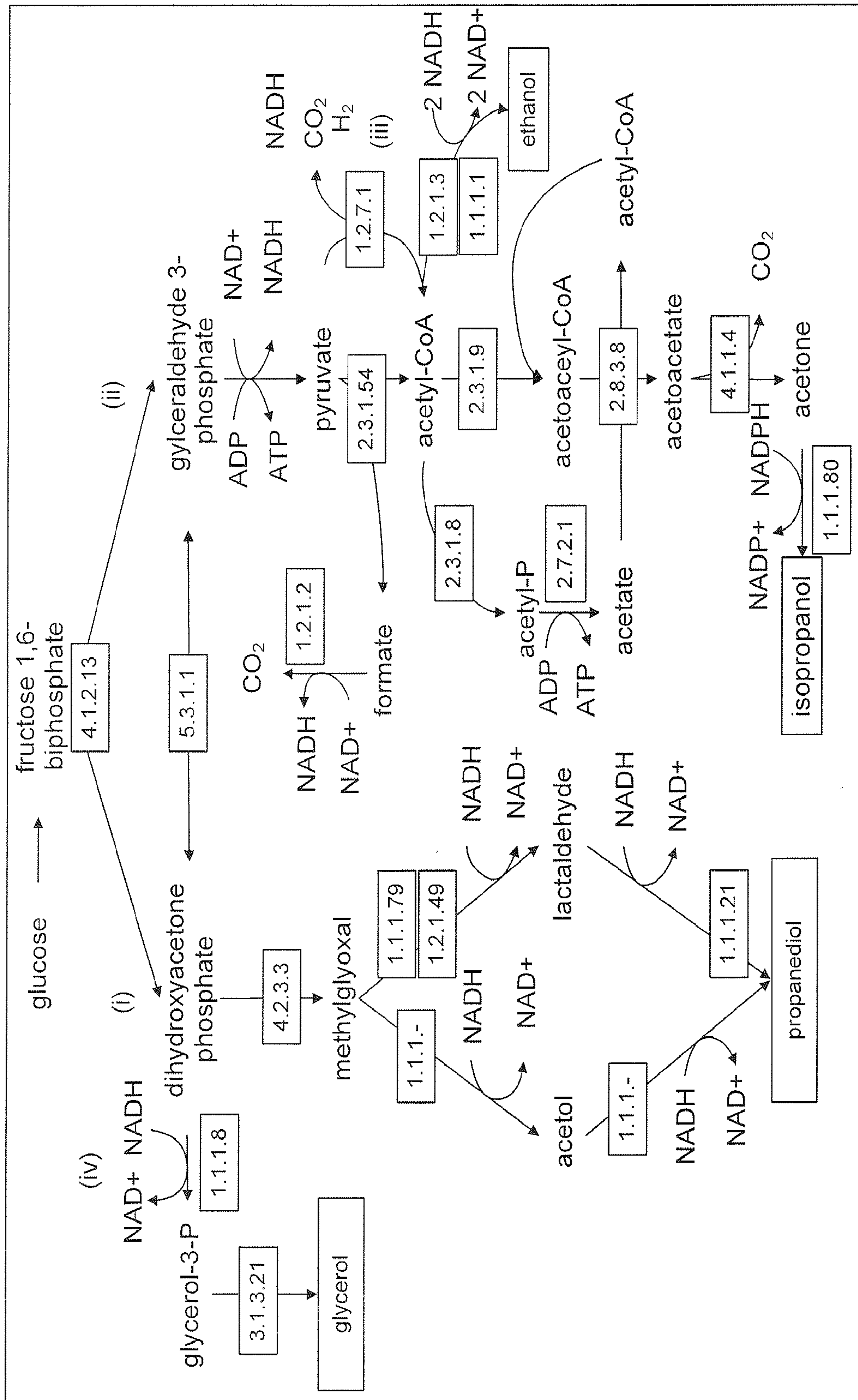


FIG. 1

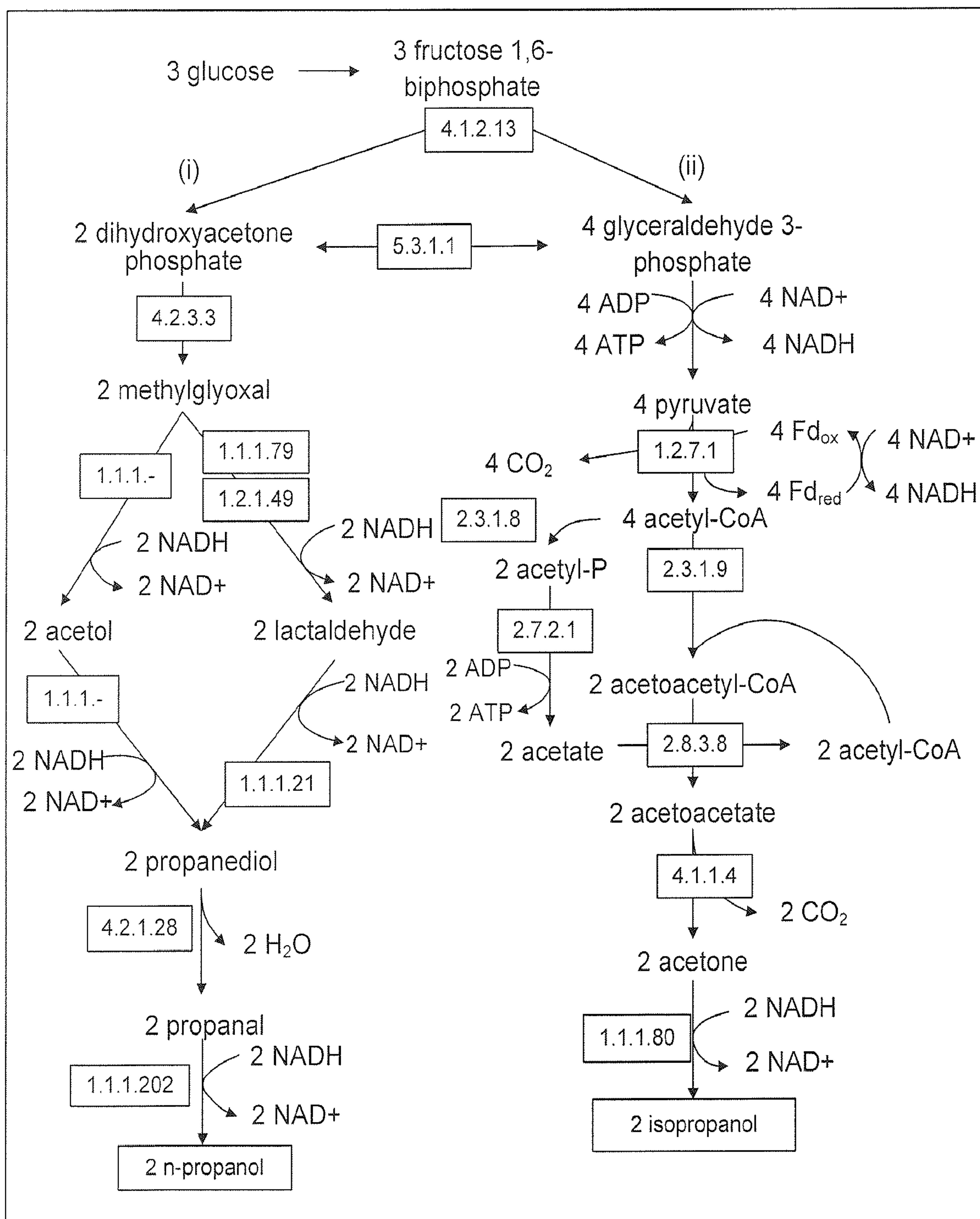


FIG. 2

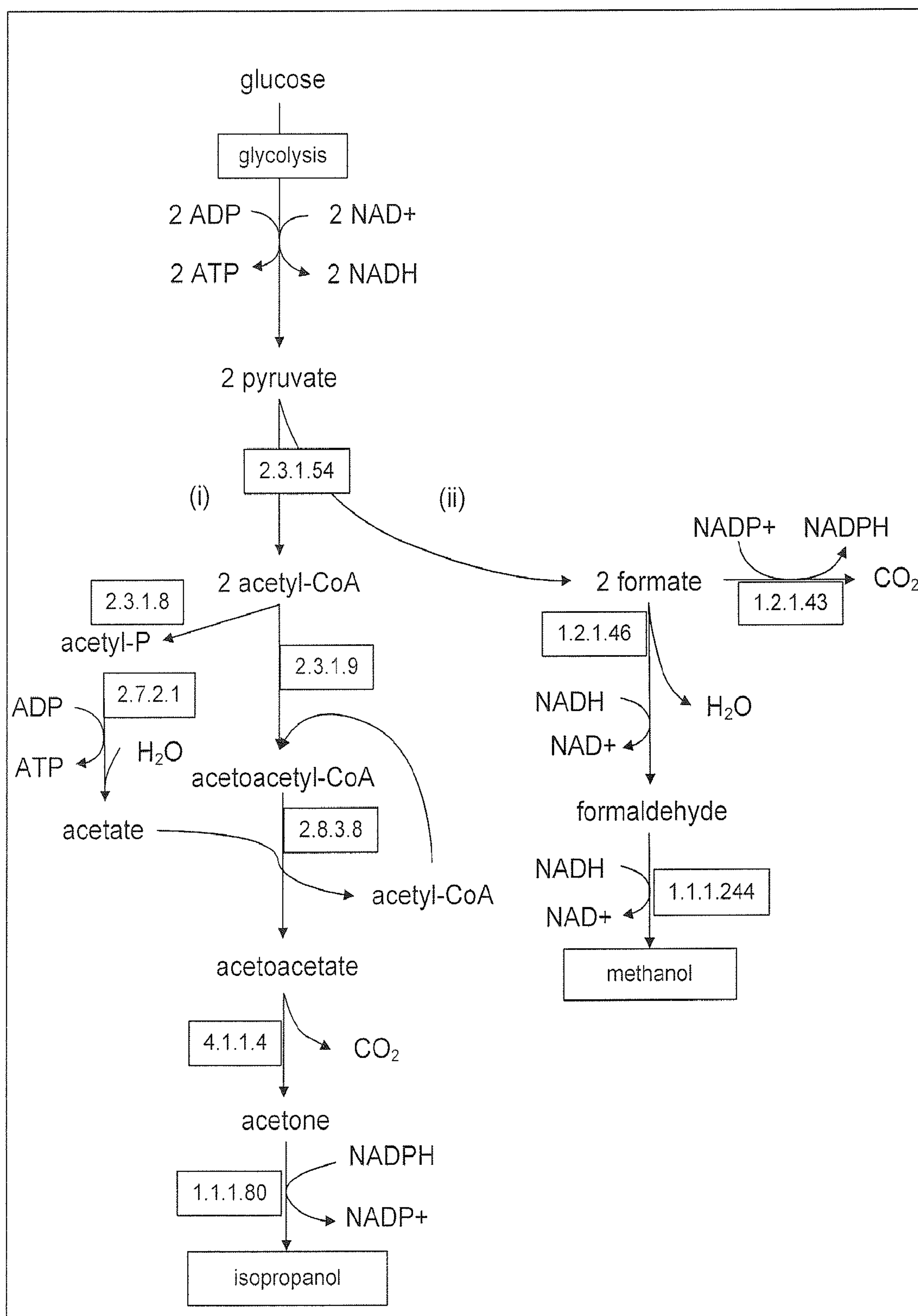


FIG. 3

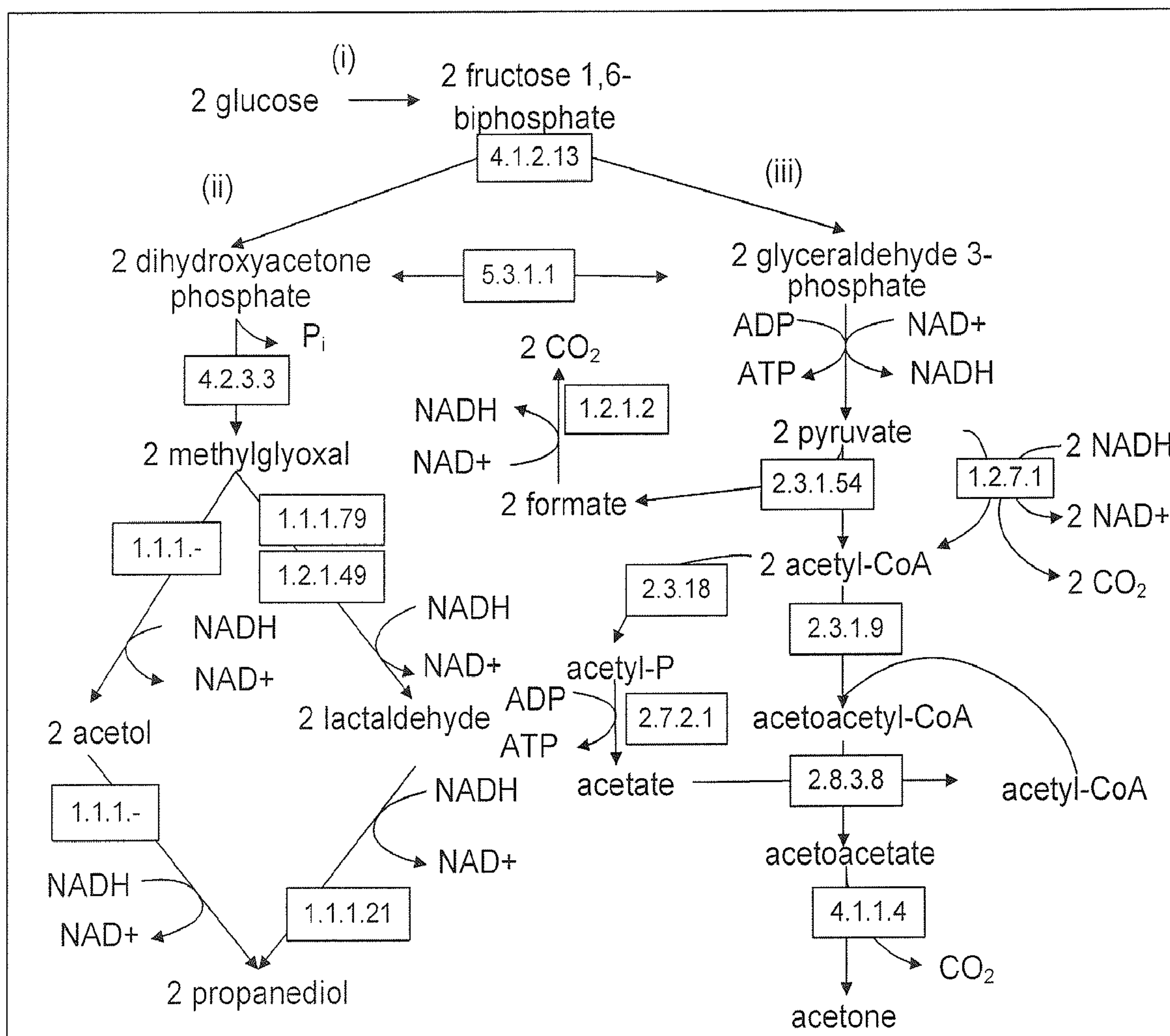


FIG. 4

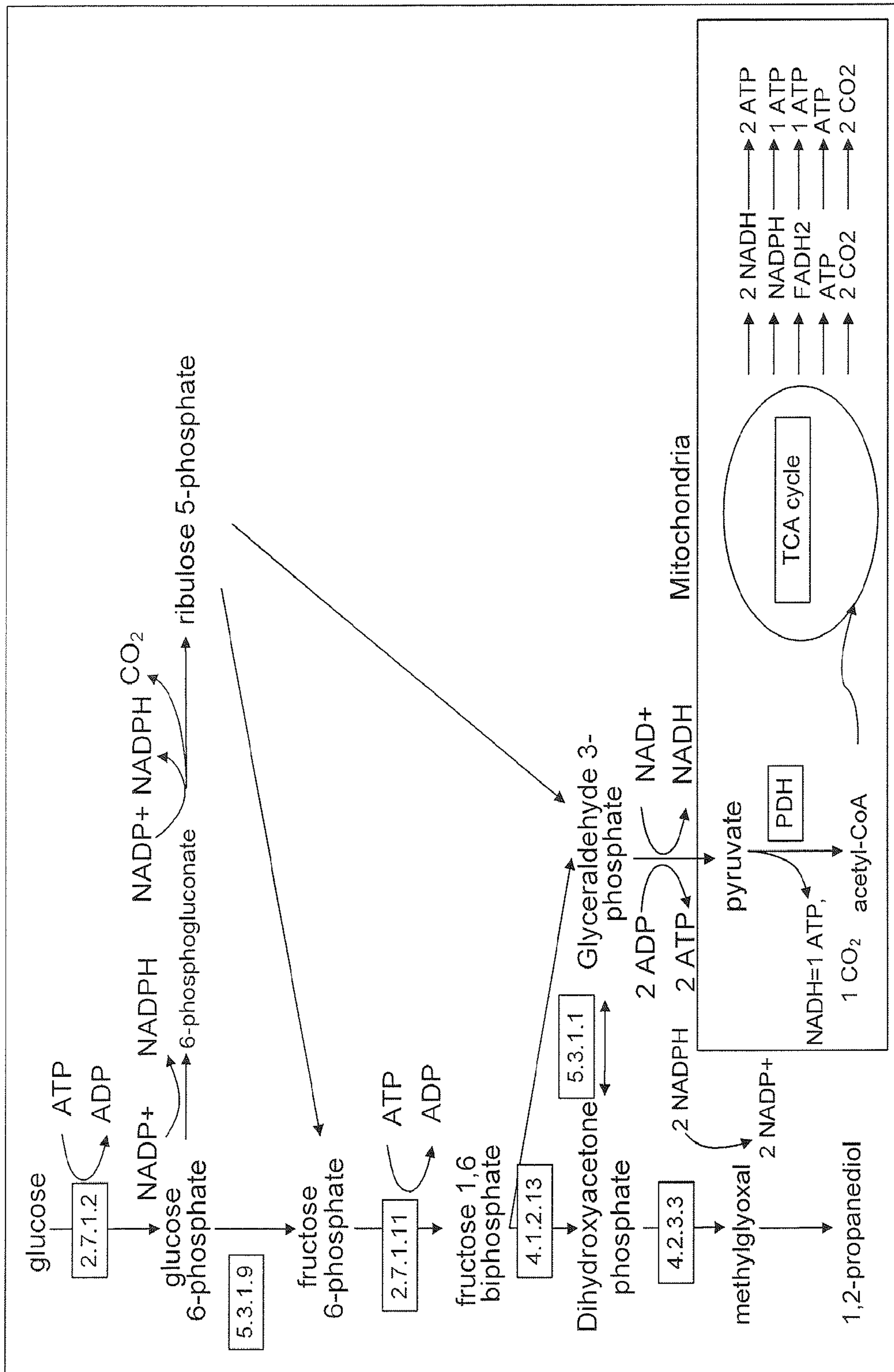
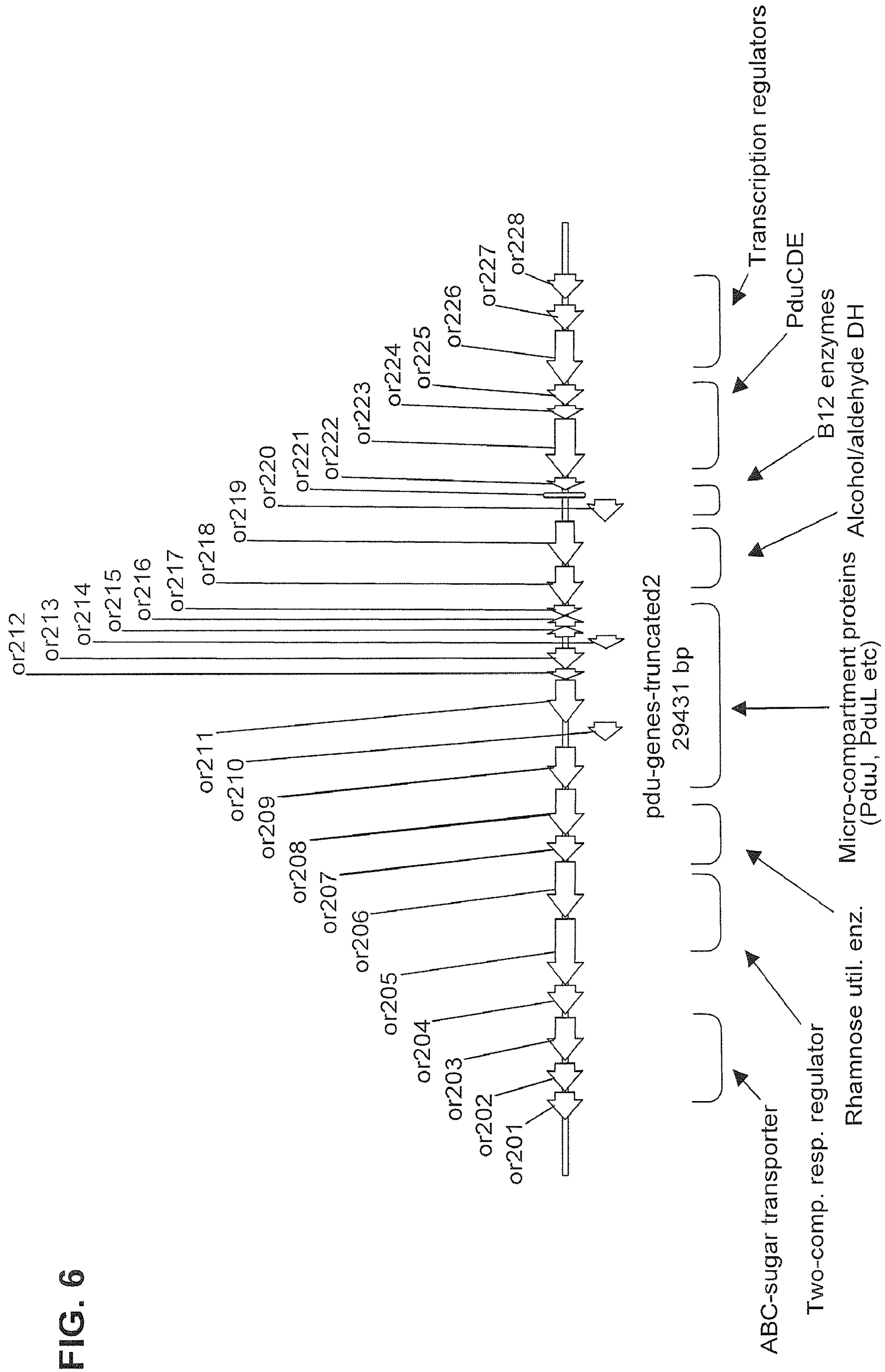


FIG. 5



7/7

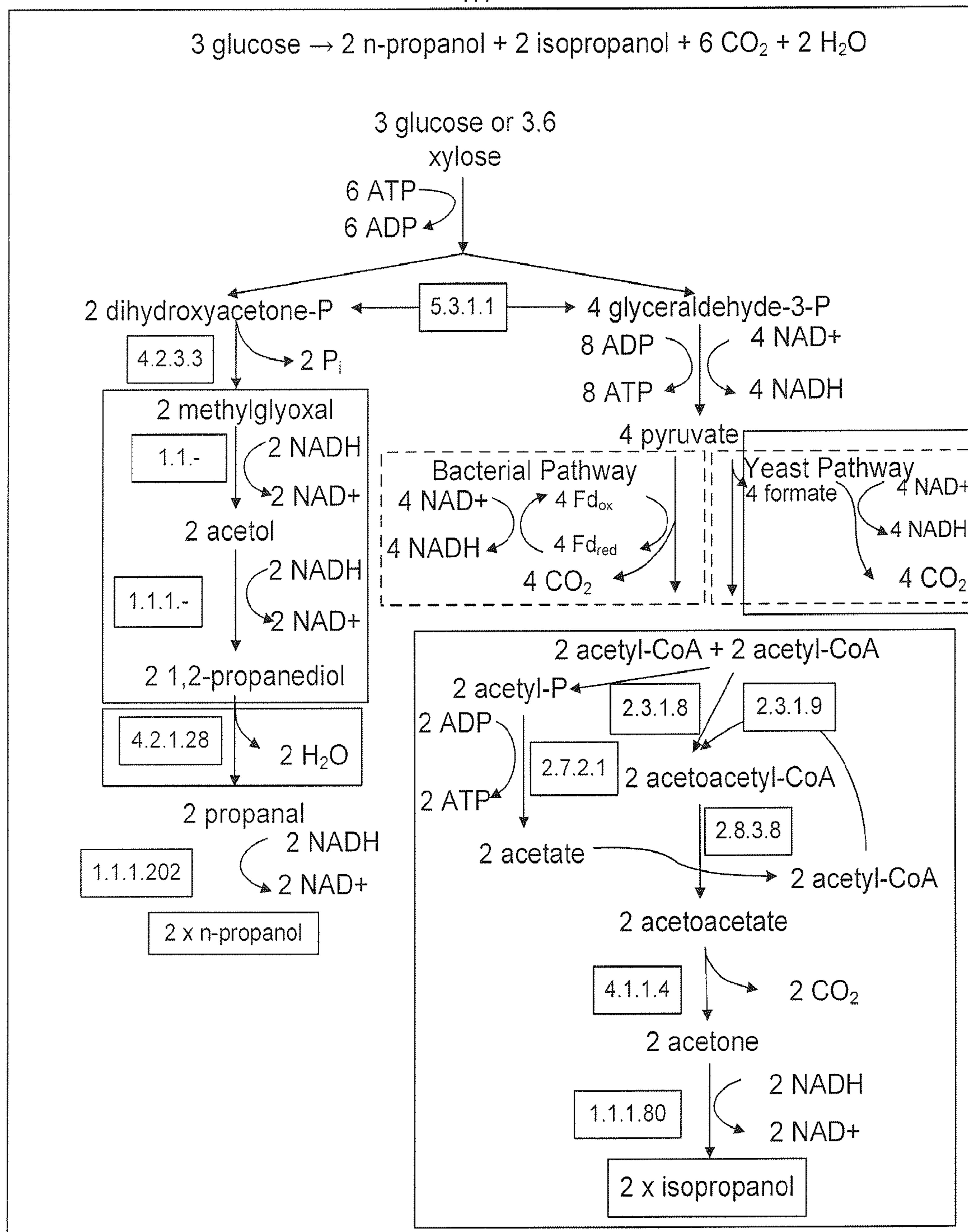


FIG. 7

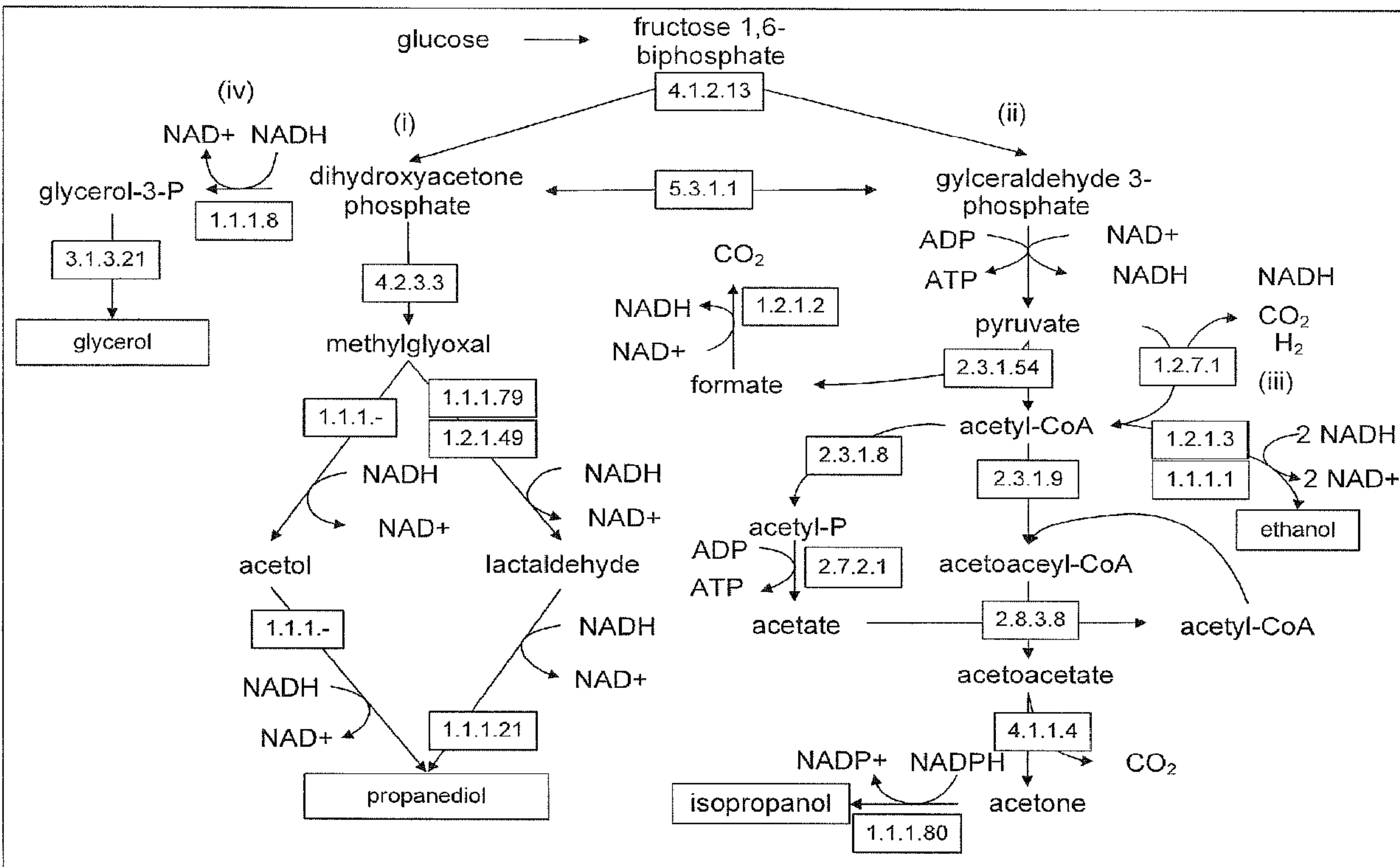


FIG. 1