

**METHODS AND ORGANISMS FOR CONVERTING SYNTHESIS GAS OR OTHER
GASEOUS CARBON SOURCES AND METHANOL TO 1,3-BUTANEDIOL**

This application claims the benefit of priority of U.S. provisional application serial No.

5 61/285,312, filed December 10, 2009, the entire contents of which are incorporated herein by reference.

BACKGROUND OF THE INVENTION

The present invention relates generally to biosynthetic processes and more specifically to organisms capable of using carbohydrates, methanol, synthesis gas and other gaseous carbon
10 sources in the production of commodity chemicals.

1,3-butanediol (1,3-BDO) is a four carbon diol traditionally produced from acetylene via its hydration. The resulting acetaldehyde is then converted to 3-hydroxybutyraldehyde which is subsequently reduced to form 1,3-BDO. In more recent years, acetylene has been replaced by the less expensive ethylene as a source of acetaldehyde. 1,3-BDO is commonly used as an
15 organic solvent for food flavoring agents. It is also used as a co-monomer for polyurethane and polyester resins and is widely employed as a hypoglycaemic agent. Optically active 1,3-BDO is a useful starting material for the synthesis of biologically active compounds and liquid crystals. Another use of 1,3-butanediol is that its dehydration affords 1,3-butadiene (Ichikawa et al., *J. Molecular Catalysis A-Chemical*, 231:181-189 (2005); Ichikawa et al., *J. Molecular Catalysis
20 A-Chemical*, 256:106-112 (2006)), a chemical used to manufacture synthetic rubbers (e.g. tires), latex, and resins.

Synthesis gas (syngas) is a mixture of primarily H₂ and CO that can be obtained via gasification of any organic feedstock, such as coal, coal oil, natural gas, biomass, or waste organic matter. Numerous gasification processes have been developed, and most designs are based on partial
25 oxidation, where limiting oxygen avoids full combustion, of organic materials at high temperatures (500-1500°C) to provide syngas as a 0.5:1-3:1 H₂/CO mixture. Steam is sometimes added to increase the hydrogen content, typically with increased CO₂ production through the water gas shift reaction. Methanol is most commonly produced industrially from the syngas components, CO, and H₂, via catalysis.

30 Today, coal is the main substrate used for industrial production of syngas, which is traditionally used for heating and power and as a feedstock for Fischer-Tropsch synthesis of methanol and

liquid hydrocarbons. Many large chemical and energy companies employ coal gasification processes on large scale and there is experience in the industry using this technology.

Moreover, technology now exists for cost-effective production of syngas from a plethora of other materials such as biomass, wastes, polymers, and the like, at virtually any location in the world. The benefits of using syngas include flexibility, since syngas can be produced from most organic substances, including biomass. Another benefit is that syngas is inexpensive. In addition, there are known pathways, as in organisms such as *Clostridium* spp., that utilize syngas effectively.

Despite the availability of organisms that utilize syngas, in general the known organisms are poorly characterized and are not well suited for commercial development. For example, *Clostridium* and related bacteria are strict anaerobes that are intolerant to high concentrations of certain products such as butanol, thus limiting titers and commercialization potential. The *Clostridia* also produce multiple products, which presents separations issues in obtaining a desired product. Finally development of facile genetic tools to manipulate *Clostridial* genes is in its infancy; therefore, they are not readily amenable to genetic engineering to improve yield or production characteristics of a desired product.

Increasing the flexibility of inexpensive and readily available feedstocks while minimizing the environmental impact of chemical production are two goals of a sustainable chemical industry. Feedstock flexibility relies on the introduction of methods that enable access and use of a wide range of materials as primary feedstocks for chemical manufacturing. The reliance on petroleum based feedstocks for either acetylene or ethylene warrants the development of a renewable feedstock based route to 1,3-butanediol and to butadiene.

Thus, there exists a need to develop microorganisms and methods of their use to utilize carbohydrates, methanol, syngas or other gaseous carbon sources for the production of 1,3-butanediol. The present invention satisfies this need and provides related advantages as well.

SUMMARY OF THE INVENTION

In some aspects, embodiments disclosed herein relate to a non-naturally occurring microbial organism having a 1,3-butanediol (1,3-BDO) pathway that includes at least one exogenous nucleic acid encoding a 1,3-BDO pathway enzyme or protein expressed in a sufficient amount to produce 1,3-BDO. The 1,3-BDO pathway includes Methanol methyltransferase (MtaB), Corrinoid protein (MtaC), Methyltetrahydrofolate:corrinoid protein methyltransferase (MtaA),

Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), Corrinoid iron-sulfur protein (AcsD), Nickel-protein assembly protein (AcsF & CooC), Ferredoxin (Orf7), Acetyl-CoA synthase (AcsB & AcsC), Carbon monoxide dehydrogenase (AcsA), Hydrogenase (Hyd), Acetoacetyl-CoA thiolase (AtoB), Acetoacetyl-CoA reductase (CoA-dependent, aldehyde forming), 3-oxobutyraldehyde reductase (ketone reducing), 3-hydroxybutyraldehyde reductase, Acetoacetyl-CoA reductase (CoA-dependent, alcohol forming), 3-oxobutyraldehyde reductase (aldehyde reducing), 4-hydroxy,2-butanone reductase, Acetoacetyl-CoA reductase (ketone reducing), 3-hydroxybutyryl-CoA reductase (aldehyde forming), 3-hydroxybutyryl-CoA reductase (alcohol forming), 3-hydroxybutyryl-CoA transferase, 3-hydroxybutyryl-CoA hydrolase, 3-hydroxybutyryl-CoA synthetase, 3-hydroxybutyrate dehydrogenase, 3-hydroxybutyrate reductase, Acetoacetyl-CoA transferase, Acetoacetyl-CoA hydrolase, Acetoacetyl-CoA synthetase, or Acetoacetate reductase.

In some aspects, embodiments disclosed herein relate to a non-naturally occurring microbial organism having a 1,3-butanediol (1,3-BDO) pathway that includes at least one exogenous nucleic acid encoding a 1,3-BDO pathway enzyme or protein expressed in a sufficient amount to produce 1,3-BDO. The 1,3-BDO pathway includes Formate dehydrogenase, Formyltetrahydrofolate synthetase, Methenyltetrahydrofolate cyclohydrolase, Methylenetetrahydrofolate dehydrogenase, Methylenetetrahydrofolate reductase, Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), Corrinoid iron-sulfur protein (AcsD), Nickel-protein assembly protein (AcsF & CooC), Ferredoxin (Orf7), Acetyl-CoA synthase (AcsB & AcsC), Carbon monoxide dehydrogenase (AcsA), Hydrogenase (Hyd), Acetoacetyl-CoA thiolase (AtoB), Acetoacetyl-CoA reductase (CoA-dependent, aldehyde forming), 3-oxobutyraldehyde reductase (ketone reducing), 3-hydroxybutyraldehyde reductase, Acetoacetyl-CoA reductase (CoA-dependent, alcohol forming), 3-oxobutyraldehyde reductase (aldehyde reducing), 4-hydroxy,2-butanone reductase, Acetoacetyl-CoA reductase (ketone reducing), 3-hydroxybutyryl-CoA reductase (aldehyde forming), 3-hydroxybutyryl-CoA reductase (alcohol forming), 3-hydroxybutyryl-CoA transferase, 3-hydroxybutyryl-CoA hydrolase, 3-hydroxybutyryl-CoA synthetase, 3-hydroxybutyrate dehydrogenase, 3-hydroxybutyrate reductase, Acetoacetyl-CoA transferase, Acetoacetyl-CoA hydrolase, Acetoacetyl-CoA synthetase, or Acetoacetate reductase.

In some aspects, embodiments disclosed herein relate to a method for producing 1,3-BDO that includes culturing the aforementioned non-naturally occurring microbial organisms under conditions and for a sufficient period of time to produce 1,3-BDO.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows a diagram depicting the Wood-Ljungdahl pathway and formation routes for acetate and ethanol. The transformations that are characteristic of organisms capable of growth on synthesis gas are 1) CO dehydrogenase, 2) hydrogenase, 3) energy-conserving hydrogenase (ECH), and 4) bi-functional CO dehydrogenase/acetyl-CoA synthase.

Figure 2 shows the complete Wood-Ljungdahl pathway that enables the conversion of gases including CO, CO₂, and/or H₂ to acetyl-CoA which is subsequently converted to cell mass and products such as ethanol or acetate. Abbreviations: 10FTHF: 10-formyltetrahydrofolate, 5MTHF: 5-methyltetrahydrofolate, ACTP: acetyl phosphate, CFeSp: corrinoid iron sulfur protein, FOR: formate, MeOH: methanol, METHF: methyltetrahydrofolate, MLTHF: methenyltetrahydrofolate, THF: tetrahydrofolate.

Figure 3 depicts pathways from acetoacetyl-CoA to 1,3-butanediol. The enzymatic steps are: A) acetoacetyl-CoA reductase (CoA-dependent, aldehyde forming), B) 3-oxobutyraldehyde reductase (ketone reducing), C) 3-hydroxybutyraldehyde reductase, D) acetoacetyl-CoA reductase (CoA-dependent, alcohol forming), E) 3-oxobutyraldehyde reductase (aldehyde reducing), F) 4-hydroxy-2-butanone reductase, G) acetoacetyl-CoA reductase (ketone reducing), H) 3-hydroxybutyryl-CoA reductase (aldehyde forming), I) 3-hydroxybutyryl-CoA reductase (alcohol forming), J) 3-hydroxybutyryl-CoA transferase, hydrolase, or synthetase, K) 3-hydroxybutyrate dehydrogenase, L) 3-hydroxybutyrate reductase, M) acetoacetyl-CoA transferase, hydrolase, or synthetase N) acetoacetate reductase.

Figure 4 shows a biosynthetic metabolic pathway for the conversion of carbohydrates, such as glucose, gases including CO, CO₂, and/or H₂, and methanol to acetyl-CoA and further to 1,3-butanediol. The specific enzymatic transformations that are engineered into a production host are numbered and shown in the figure. Abbreviations: 10FTHF: 10-formyltetrahydrofolate, 5MTHF: 5-methyltetrahydrofolate, ACTP: acetyl phosphate, CFeSp: corrinoid iron sulfur protein, FOR: formate, MeOH: methanol, METHF: methyltetrahydrofolate, MLTHF: methenyltetrahydrofolate, THF: tetrahydrofolate.

Figure 5 shows a biosynthetic metabolic pathway for the conversion of carbohydrates, such as glucose, gases including CO, CO₂, and/or H₂ and methanol to acetyl-CoA, and further to 1,3-butanediol. The specific enzymatic transformations that are engineered into a production host are numbered and shown in the figure. Abbreviations: 10FTHF: 10-formyltetrahydrofolate, 5MTHF: 5-methyltetrahydrofolate, ACTP: acetyl phosphate, CFeSp: corrinoid iron sulfur

protein, FOR: formate, MeOH: methanol, METHF: methyltetrahydrofolate, MLTHF: metheneyltetrahydrofolate, THF: tetrahydrofolate.

Figure 6 shows Western blots of 10 micrograms ACS90 (lane 1), ACS91 (lane 2), Mta98/99 (lanes 3 and 4) cell extracts with size standards (lane 5) and controls of *M. thermoacetica* CODH (lanes 6 and 7) or Mtr (lanes 8 and 9) proteins (50, 150, 250, 350, 450, 500, 750, 900, and 1000 ng).

Figure 7 shows cuvettes used in a methyl viologen assay. A blank is on the right and a cuvette with reduced methyl viologen is on the left. Note, the stoppers and vacuum grease on top of each are used to keep the reactions anaerobic.

Figure 8 shows a spectrogram of ACS90 cell extracts assayed for transfer of CH₃ from added CH₃-THF to purified *M. thermoacetica* corrinoid protein.

Figure 9 shows from left to right, anaerobic growth of recombinant *E. coli* MG1655 in N₂ and CO for 36 hr at 37 °C, empty vector, ACS90, and ACS91.

Figure 10 shows an exemplary flux distribution that can produce a 1.09 mol/mol yield of 1,3-butanediol on glucose by employing the Wood-Ljungdahl pathway enzymes in combination with a 1,3-butanediol biosynthetic pathway.

Figure 11 shows an exemplary flux distribution that can produce a 1.2 mol/mol yield of 1,3-butanediol on glucose by employing enzymes that allow the co-utilization of methanol as a substrate.

DETAILED DESCRIPTION OF THE INVENTION

The present invention relates to developing and using microorganisms capable of utilizing carbohydrates, methanol, syngas and/or other gaseous carbon sources to produce 1,3-butanediol. The invention further relates to expanding the product range of syngas-utilizing microorganisms and generating recombinant organisms capable of utilizing syngas to produce 1,3-butanediol and optimizing their yields, titers, and productivities. Development of a recombinant organism, for example, *Escherichia coli* or other organisms suitable for commercial scale up, that can efficiently utilize syngas as a substrate for growth and for chemical production provides cost-advantaged processes for renewable chemical and fuel manufacturing. The organisms can be optimized and tested rapidly and at reasonable costs.

The potential of syngas as a feedstock resides in its ability to be efficiently and cost-effectively converted into chemicals and fuels of interest. Two main technologies for syngas conversion are Fischer-Tropsch processes and fermentative processes. The Fischer-Tropsch (F-T) technology has been developed since World War II and involves inorganic and metal-based catalysts that allow efficient production of methanol or mixed hydrocarbons as fuels. The drawbacks of F-T processes are: 1) a lack of product selectivity, which results in difficulties separating desired products; 2) catalyst sensitivity to poisoning; 3) high energy costs due to high temperatures and pressures required; and 4) the limited range of products available at commercially competitive costs.

For fermentative processes, syngas has been shown to serve as a carbon and energy source for many anaerobic microorganisms that can convert this material into products such as ethanol, acetate and hydrogen. The main benefits of fermentative conversion of syngas are the selectivity of organisms for production of single products, greater tolerance to syngas impurities, lower operating temperatures and pressures, and potential for a large portfolio of products from syngas. The main drawbacks of fermentative processes are that organisms known to convert syngas tend to generate only a limited range of chemicals, such as ethanol and acetate, and are not efficient producers of other chemicals, the organisms lack established tools for genetic manipulation, and the organisms are sensitive to end products at high concentrations.

The present invention relates to the generation of microorganisms that are effective at producing 1,3-butanediol from syngas or other gaseous carbon sources. The organisms and methods of the present invention allow production of 1,3-butanediol at costs that are significantly advantaged over both traditional petroleum-based products and products derived directly from glucose, sucrose or lignocellulosic sugars. In one embodiment, the invention provides a non-naturally occurring microorganism capable of utilizing syngas or other gaseous carbon sources to produce 1,3-butanediol in which the parent microorganism lacks the natural ability to utilize syngas, as shown in Figure 4 and 5. In such microorganisms, one or more proteins or enzymes are expressed in the microorganism, thereby conferring a pathway to utilize syngas or other gaseous carbon source to produce a desired product. In other embodiments, the invention provides a non-naturally occurring microorganism that has been genetically modified, for example, by expressing one or more exogenous proteins or enzymes that confer an increased efficiency of production of 1,3-butanediol, where the parent microorganism has the ability to utilize syngas or other gaseous carbon source to produce a desired product. Thus, the invention relates to generating a microorganism with a new metabolic pathway capable of utilizing syngas as well as

generating a microorganism with improved efficiency of utilizing syngas or other gaseous carbon source to produce 1,3-butanediol.

The present invention additionally provides a non-naturally occurring microorganism expressing genes encoding enzymes that catalyze reactions associated with the carbonyl-branch of the Wood-Ljungdahl pathway in conjunction with a MtaABC-type methyltransferase system. Such an organism is capable of converting methanol, a relatively inexpensive organic feedstock that can be derived from synthesis gas, and gases comprising CO, CO₂, and/or H₂ into acetyl-CoA, cell mass, and products. The present invention further provides pathways that can achieve an increased yield of 1,3-butanediol on carbohydrate feedstocks over what would be naturally expected, that is about 1 mol 1,3-butanediol/mol of glucose, by providing an efficient mechanism for fixing the carbon present in methanol or carbon dioxide, fed exogenously or produced endogenously, into acetyl-CoA.

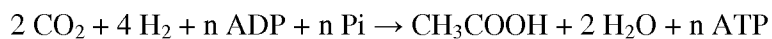
Escherichia coli is an industrial workhorse organism with an unrivaled suite of genetic tools. Engineering the capability to convert CO₂, CO, and/or H₂ into acetyl-CoA, the central metabolite from which all cell mass components and many valuable products can be derived, into a foreign host such as *E. coli* can be accomplished following the expression of exogenous genes that encode various proteins of the Wood-Ljungdahl pathway. This pathway is highly active in acetogenic organisms such as *Moorella thermoacetica* (formerly, *Clostridium thermoaceticum*), which has been the model organism for elucidating the Wood-Ljungdahl pathway since its isolation in 1942 (Fontaine et al., *J Bacteriol.* 43:701-715 (1942)). The Wood-Ljungdahl pathway comprises two branches: the Eastern, or methyl, branch that allows the conversion of CO₂ to methyltetrahydrofolate (Me-THF) and the Western, or carbonyl, branch that allows the conversion of methyl-THF, CO, and Coenzyme-A into acetyl-CoA (see Figures 1 and 2). As disclosed herein, the invention provides a non-naturally occurring microorganism expressing genes that catalyze both branches of the Wood-Ljungdahl pathway, in addition to genes for the production of 1,3-butanediol. Such an organism is capable of converting gases comprising CO, CO₂, and/or H₂ into acetyl-CoA, 1,3-butanediol, cell mass, and products. Such an organism is also capable of producing 1,3-butanediol from carbohydrates at the stoichiometric optimum yield. For example, in combination with any of the acetyl-CoA to 1,3-butanediol pathways, the Wood-Ljungdahl enzymes provide the means to produce 12 moles of 1,3-butanediol for every 11 moles of glucose as opposed to 1 mol 1,3-butanediol/1 mol glucose which would be attainable in the absence of the Wood-Ljungdahl pathway enzymes.

The invention additionally provides a non-naturally occurring microorganism expressing genes encoding enzymes that catalyze the carbonyl-branch of the Wood-Ljungdahl pathway in conjunction with a MtaABC-type methyltransferase system, in addition to genes for the production of 1,3-butanediol. Such an organism is capable of converting methanol, a relatively inexpensive organic feedstock that can be derived from synthesis gas, and gases comprising CO, CO₂, and/or H₂ into acetyl-CoA, 1,3-butanediol, cell mass, and products. In addition to gaseous substrates, the organism can utilize methanol exclusively or in combination with carbohydrate feedstocks such as glucose to produce 1,3-butanediol at high yield.

Synthesis gas, also known as syngas or producer gas, is the major product of gasification of coal and of carbonaceous materials such as biomass materials, including agricultural crops and residues. Syngas is a mixture primarily of H₂ and CO and can be obtained from the gasification of any organic feedstock, including but not limited to coal, coal oil, natural gas, biomass, and waste organic matter. Gasification is generally carried out under a high fuel to oxygen ratio. Although largely H₂ and CO, syngas can also include CO₂ and other gases in smaller quantities. Thus, synthesis gas provides a cost effective source of gaseous carbon such as CO and, additionally, CO₂.

As disclosed herein, gaseous carbon sources such as syngas comprising CO and/or CO₂ can be utilized by non-naturally occurring microorganisms of the invention to produce 1,3-butanediol. Although generally exemplified herein as syngas, it is understood that any source of gaseous carbon comprising CO and/or CO₂ can be utilized by the non-naturally occurring microorganisms of the invention. Thus, the invention relates to non-naturally occurring microorganisms that are capable of utilizing CO and/or CO₂ as a carbon source.

The Wood-Ljungdahl pathway catalyzes the conversion of CO and H₂ to acetyl-CoA and other products such as acetate. Organisms capable of utilizing CO and syngas also generally have the capability of utilizing CO₂ and CO₂/H₂ mixtures through the same basic set of enzymes and transformations encompassed by the Wood-Ljungdahl pathway. H₂-dependent conversion of CO₂ to acetate by microorganisms was recognized long before it was revealed that CO also could be used by the same organisms and that the same pathways were involved. Many acetogens have been shown to grow in the presence of CO₂ and produce compounds such as acetate as long as hydrogen is present to supply the necessary reducing equivalents (see for example, Drake, *Acetogenesis*, pp. 3-60 Chapman and Hall, New York, (1994)). This can be summarized by the following equation:



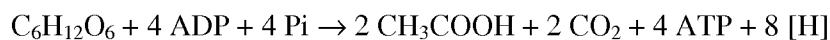
Hence, non-naturally occurring microorganisms possessing the Wood-Ljungdahl pathway can utilize CO₂ and H₂ mixtures as well for the production of acetyl-CoA and other desired products, such as 1,3-butanediol.

5 The Wood-Ljungdahl pathway is well known in the art and consists of 12 reactions which can be separated into two branches: (1) methyl branch and (2) carbonyl branch. The methyl branch converts syngas to methyl-tetrahydrofolate (methyl-THF) whereas the carbonyl branch converts methyl-THF to acetyl-CoA. The reactions in the methyl branch are catalyzed in order by the following enzymes or proteins: ferredoxin oxidoreductase, formate dehydrogenase,
 10 formyltetrahydrofolate synthetase, methenyltetrahydrofolate cyclodehydratase, methylenetetrahydrofolate dehydrogenase and methylenetetrahydrofolate reductase. The reactions in the carbonyl branch are catalyzed in order by the following enzymes or proteins: methyltetrahydrofolate:corrinoid protein methyltransferase (for example, AcsE), corrinoid iron-sulfur protein, nickel-protein assembly protein (for example, AcsF), ferredoxin, acetyl-CoA
 15 synthase, carbon monoxide dehydrogenase and nickel-protein assembly protein (for example, CooC). Following the teachings and guidance provided herein for introducing a sufficient number of encoding nucleic acids to generate an acetyl-CoA pathway, those skilled in the art will understand that the same engineering design also can be performed with respect to introducing at least the nucleic acids encoding the Wood-Ljungdahl enzymes or proteins absent
 20 in the host organism. Therefore, introduction of one or more encoding nucleic acids into the microbial organisms of the invention such that the modified organism contains one branch or the complete Wood-Ljungdahl pathway will confer syngas utilization ability.

Thus, the non-naturally occurring microorganisms of the invention can use syngas or other gaseous carbon sources providing CO and/or CO₂ to produce 1,3-butanediol. In the case of
 25 CO₂, additional sources include, but are not limited to, production of CO₂ as a byproduct in ammonia and hydrogen plants, where methane is converted to CO₂; combustion of wood and fossil fuels; production of CO₂ as a byproduct of fermentation of sugar in the brewing of beer, whisky and other alcoholic beverages, or other fermentative processes; thermal decomposition of limestone, CaCO₃, in the manufacture of lime, CaO; production of CO₂ as byproduct of
 30 sodium phosphate manufacture; and directly from natural carbon dioxide springs, where it is produced by the action of acidified water on limestone or dolomite.

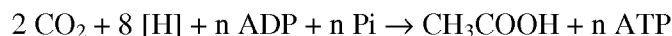
Acetogens, such as *Moorella thermoacetica*, *C. ljungdahlii* and *C. carboxidivorans*, can grow on a number of carbon sources ranging from hexose sugars to carbon monoxide. Hexoses, such as glucose, are metabolized first via Embden-Meyerhof-Parnas (EMP) glycolysis to pyruvate, which is then converted to acetyl-CoA via pyruvate:ferredoxin oxidoreductase (PFOR). Acetyl-

5 CoA can be used to build biomass precursors or can be converted to acetate which produces energy via acetate kinase and phosphotransacetylase. The overall conversion of glucose to acetate, energy, and reducing equivalents is



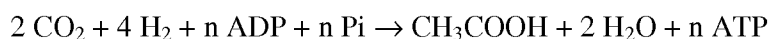
Acetogens extract even more energy out of the glucose to acetate conversion while also

10 maintaining redox balance by further converting the released CO_2 to acetate via the Wood-Ljungdahl pathway:



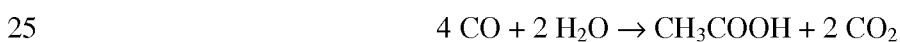
The coefficient n in the above equation signifies that this conversion is an energy generating endeavor, as many acetogens can grow in the presence of CO_2 via the Wood-Ljungdahl pathway

15 even in the absence of glucose as long as hydrogen is present to supply the necessary reducing equivalents.

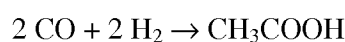


The Wood-Ljungdahl pathway, illustrated in Figure 1, is coupled to the creation of Na^+ or H^+ ion gradients that can generate ATP via an Na^+ - or H^+ - dependant ATP synthase, respectively

20 Muller V. Energy conservation in acetogenic bacteria. *Appl Environ Microbiol* 69:6345-6353 (2003). Based on these known transformations, acetogens also have the capacity to utilize CO as the sole carbon and energy source. Specifically, CO can be oxidized to produce reducing equivalents and CO_2 , or directly assimilated into acetyl-CoA which is subsequently converted to either biomass or acetate.



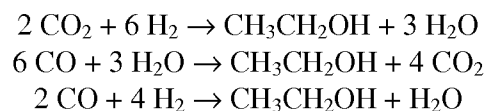
Even higher acetate yields, however, can be attained when enough hydrogen is present to satisfy the requirement for reducing equivalents.



Following from Figure 1, the production of acetate via acetyl-CoA generates one ATP molecule,

30 whereas the production of ethanol from acetyl-CoA does not and requires two reducing

equivalents. Thus, it can be concluded that ethanol production from syngas will not generate sufficient energy for cell growth in the absence of acetate production. However, under certain conditions, *Clostridium ljungdahlii* produces mostly ethanol from synthesis gas (Klasson et al., *Fuel* 72.12:1673-1678 (1993)) indicating that some combination of the pathways does indeed
5 generate enough energy to support cell growth.



Hydrogenic bacteria such as *R. rubrum* can also generate energy from the conversion of CO and
10 water to hydrogen (see Figure 1) (Simpma et al., *Critical Reviews in Biotechnology*, 26.1:41-65 (2006)). A central mechanism is the coordinated action of an energy converting hydrogenase (ECH) and CO dehydrogenase. The CO dehydrogenase supplies electrons from CO which are then used to reduce protons to H₂ by ECH, whose activity is coupled to energy-generating proton translocation. The net result is the generation of energy via the water-gas shift reaction.

15 The processes disclosed herein for the biosynthetic production of 1,3-BDO via syngas involve sustainable manufacturing practices that utilize renewable feedstocks, reduce energy intensity and lower greenhouse gas emissions. Moreover, the dehydration of biobased-1,3-BDO represents a renewable route to butadiene in small end-use facilities where no transport of this flammable and reactive chemical is required.

20 As used herein, the term “non-naturally occurring” when used in reference to a microbial organism or microorganism of the invention is intended to mean that the microbial organism has at least one genetic alteration not normally found in a naturally occurring strain of the referenced species, including wild-type strains of the referenced species. Genetic alterations include, for example, modifications introducing expressible nucleic acids encoding metabolic polypeptides,
25 other nucleic acid additions, nucleic acid deletions and/or other functional disruption of the microbial organism’s genetic material. Such modifications include, for example, coding regions and functional fragments thereof, for heterologous, homologous or both heterologous and homologous polypeptides for the referenced species. Additional modifications include, for example, non-coding regulatory regions in which the modifications alter expression of a gene or
30 operon. Exemplary metabolic polypeptides include enzymes or proteins within a 1,3-butanediol biosynthetic pathway.

A metabolic modification refers to a biochemical reaction that is altered from its naturally occurring state. Therefore, non-naturally occurring microorganisms can have genetic modifications to nucleic acids encoding metabolic polypeptides or, functional fragments thereof. Exemplary metabolic modifications are disclosed herein.

- 5 As used herein, the term “isolated” when used in reference to a microbial organism is intended to mean an organism that is substantially free of at least one component as the referenced microbial organism is found in nature. The term includes a microbial organism that is removed from some or all components as it is found in its natural environment. The term also includes a microbial organism that is removed from some or all components as the microbial organism is
- 10 found in non-naturally occurring environments. Therefore, an isolated microbial organism is partly or completely separated from other substances as it is found in nature or as it is grown, stored or subsisted in non-naturally occurring environments. Specific examples of isolated microbial organisms include partially pure microbes, substantially pure microbes and microbes cultured in a medium that is non-naturally occurring.
- 15 As used herein, the terms “microbial,” “microbial organism” or “microorganism” is intended to mean any organism that exists as a microscopic cell that is included within the domains of archaea, bacteria or eukarya. Therefore, the term is intended to encompass prokaryotic or eukaryotic cells or organisms having a microscopic size and includes bacteria, archaea and eubacteria of all species as well as eukaryotic microorganisms such as yeast and fungi. The term
- 20 also includes cell cultures of any species that can be cultured for the production of a biochemical.

As used herein, the term “CoA” or “coenzyme A” is intended to mean an organic cofactor or prosthetic group (nonprotein portion of an enzyme) whose presence is required for the activity of many enzymes (the apoenzyme) to form an active enzyme system. Coenzyme A functions in

25 certain condensing enzymes, acts in acetyl or other acyl group transfer and in fatty acid synthesis and oxidation, pyruvate oxidation and in other acetylation.

As used herein, the term “substantially anaerobic” when used in reference to a culture or growth condition is intended to mean that the amount of oxygen is less than about 10% of saturation for dissolved oxygen in liquid media. The term also is intended to include sealed chambers of

30 liquid or solid medium maintained with an atmosphere of less than about 1% oxygen.

“Exogenous” as it is used herein is intended to mean that the referenced molecule or the referenced activity is introduced into the host microbial organism. The molecule can be

introduced, for example, by introduction of an encoding nucleic acid into the host genetic material such as by integration into a host chromosome or as non-chromosomal genetic material such as a plasmid. Therefore, the term as it is used in reference to expression of an encoding nucleic acid refers to introduction of the encoding nucleic acid in an expressible form into the microbial organism. When used in reference to a biosynthetic activity, the term refers to an activity that is introduced into the host reference organism. The source can be, for example, a homologous or heterologous encoding nucleic acid that expresses the referenced activity following introduction into the host microbial organism. Therefore, the term “endogenous” refers to a referenced molecule or activity that is present in the host. Similarly, the term when used in reference to expression of an encoding nucleic acid refers to expression of an encoding nucleic acid contained within the microbial organism. The term “heterologous” refers to a molecule or activity derived from a source other than the referenced species whereas “homologous” refers to a molecule or activity derived from the host microbial organism. Accordingly, exogenous expression of an encoding nucleic acid of the invention can utilize either or both a heterologous or homologous encoding nucleic acid.

It is understood that when more than one exogenous nucleic acid is included in a microbial organism that the more than one exogenous nucleic acid refers to the referenced encoding nucleic acid or biosynthetic activity, as discussed above. It is further understood, as disclosed herein, that such more than one exogenous nucleic acids can be introduced into the host microbial organism on separate nucleic acid molecules, on polycistronic nucleic acid molecules, or a combination thereof, and still be considered as more than one exogenous nucleic acid. For example, as disclosed herein a microbial organism can be engineered to express two or more exogenous nucleic acids encoding a desired pathway enzyme or protein. In the case where two exogenous nucleic acids encoding a desired activity are introduced into a host microbial organism, it is understood that the two exogenous nucleic acids can be introduced as a single nucleic acid, for example, on a single plasmid, on separate plasmids, can be integrated into the host chromosome at a single site or multiple sites, and still be considered as two exogenous nucleic acids. Similarly, it is understood that more than two exogenous nucleic acids can be introduced into a host organism in any desired combination, for example, on a single plasmid, on separate plasmids, can be integrated into the host chromosome at a single site or multiple sites, and still be considered as two or more exogenous nucleic acids, for example three exogenous nucleic acids. Thus, the number of referenced exogenous nucleic acids or biosynthetic activities refers to the number of encoding nucleic acids or the number of

biosynthetic activities, not the number of separate nucleic acids introduced into the host organism.

The non-naturally occurring microbial organisms of the invention can contain stable genetic alterations, which refers to microorganisms that can be cultured for greater than five generations without loss of the alteration. Generally, stable genetic alterations include modifications that persist greater than 10 generations, particularly stable modifications will persist more than about 25 generations, and more particularly, stable genetic modifications can be greater than 50 generations, including indefinitely.

Those skilled in the art will understand that the genetic alterations, including metabolic modifications exemplified herein, are described with reference to a suitable host organism such as *E. coli* and their corresponding metabolic reactions or a suitable source organism for desired genetic material such as genes for a desired metabolic pathway. However, given the complete genome sequencing of a wide variety of organisms and the high level of skill in the area of genomics, those skilled in the art will readily be able to apply the teachings and guidance provided herein to essentially all other organisms. For example, the *E. coli* metabolic alterations exemplified herein can readily be applied to other species by incorporating the same or analogous encoding nucleic acid from species other than the referenced species. Such genetic alterations include, for example, genetic alterations of species homologs, in general, and in particular, orthologs, paralogs or nonorthologous gene displacements.

An ortholog is a gene or genes that are related by vertical descent and are responsible for substantially the same or identical functions in different organisms. For example, mouse epoxide hydrolase and human epoxide hydrolase can be considered orthologs for the biological function of hydrolysis of epoxides. Genes are related by vertical descent when, for example, they share sequence similarity of sufficient amount to indicate they are homologous, or related by evolution from a common ancestor. Genes can also be considered orthologs if they share three-dimensional structure but not necessarily sequence similarity, of a sufficient amount to indicate that they have evolved from a common ancestor to the extent that the primary sequence similarity is not identifiable. Genes that are orthologous can encode proteins with sequence similarity of about 25% to 100% amino acid sequence identity. Genes encoding proteins sharing an amino acid similarity less than 25% can also be considered to have arisen by vertical descent if their three-dimensional structure also shows similarities. Members of the serine protease family of enzymes, including tissue plasminogen activator and elastase, are considered to have arisen by vertical descent from a common ancestor.

Orthologs include genes or their encoded gene products that through, for example, evolution, have diverged in structure or overall activity. For example, where one species encodes a gene product exhibiting two functions and where such functions have been separated into distinct genes in a second species, the three genes and their corresponding products are considered to be orthologs. For the production of a biochemical product, those skilled in the art will understand that the orthologous gene harboring the metabolic activity to be introduced or disrupted is to be chosen for construction of the non-naturally occurring microorganism. An example of orthologs exhibiting separable activities is where distinct activities have been separated into distinct gene products between two or more species or within a single species. A specific example is the separation of elastase proteolysis and plasminogen proteolysis, two types of serine protease activity, into distinct molecules as plasminogen activator and elastase. A second example is the separation of mycoplasma 5'-3' exonuclease and *Drosophila* DNA polymerase III activity. The DNA polymerase from the first species can be considered an ortholog to either or both of the exonuclease or the polymerase from the second species and vice versa.

In contrast, paralogs are homologs related by, for example, duplication followed by evolutionary divergence and have similar or common, but not identical functions. Paralogs can originate or derive from, for example, the same species or from a different species. For example, microsomal epoxide hydrolase (epoxide hydrolase I) and soluble epoxide hydrolase (epoxide hydrolase II) can be considered paralogs because they represent two distinct enzymes, co-evolved from a common ancestor, that catalyze distinct reactions and have distinct functions in the same species. Paralogs are proteins from the same species with significant sequence similarity to each other suggesting that they are homologous, or related through co-evolution from a common ancestor. Groups of paralogous protein families include HipA homologs, luciferase genes, peptidases, and others.

A nonorthologous gene displacement is a nonorthologous gene from one species that can substitute for a referenced gene function in a different species. Substitution includes, for example, being able to perform substantially the same or a similar function in the species of origin compared to the referenced function in the different species. Although generally, a nonorthologous gene displacement can be identifiable as structurally related to a known gene encoding the referenced function, less structurally related but functionally similar genes and their corresponding gene products nevertheless will still fall within the meaning of the term as it is used herein. Functional similarity requires, for example, at least some structural similarity in the active site or binding region of a nonorthologous gene product compared to a gene encoding

the function sought to be substituted. Therefore, a nonorthologous gene includes, for example, a paralog or an unrelated gene.

Therefore, in identifying and constructing the non-naturally occurring microbial organisms of the invention having 1,3-butanediol biosynthetic capability, those skilled in the art will understand with applying the teaching and guidance provided herein to a particular species that the identification of metabolic modifications can include identification and inclusion or inactivation of orthologs. To the extent that paralogs and/or nonorthologous gene displacements are present in the referenced microorganism that encode an enzyme catalyzing a similar or substantially similar metabolic reaction, those skilled in the art also can utilize these evolutionally related genes.

Orthologs, paralogs and nonorthologous gene displacements can be determined by methods well known to those skilled in the art. For example, inspection of nucleic acid or amino acid sequences for two polypeptides will reveal sequence identity and similarities between the compared sequences. Based on such similarities, one skilled in the art can determine if the similarity is sufficiently high to indicate the proteins are related through evolution from a common ancestor. Algorithms well known to those skilled in the art, such as Align, BLAST, Clustal W and others compare and determine a raw sequence similarity or identity, and also determine the presence or significance of gaps in the sequence which can be assigned a weight or score. Such algorithms also are known in the art and are similarly applicable for determining nucleotide sequence similarity or identity. Parameters for sufficient similarity to determine relatedness are computed based on well known methods for calculating statistical similarity, or the chance of finding a similar match in a random polypeptide, and the significance of the match determined. A computer comparison of two or more sequences can, if desired, also be optimized visually by those skilled in the art. Related gene products or proteins can be expected to have a high similarity, for example, 25% to 100% sequence identity. Proteins that are unrelated can have an identity which is essentially the same as would be expected to occur by chance, if a database of sufficient size is scanned (about 5%). Sequences between 5% and 24% may or may not represent sufficient homology to conclude that the compared sequences are related. Additional statistical analysis to determine the significance of such matches given the size of the data set can be carried out to determine the relevance of these sequences.

Exemplary parameters for determining relatedness of two or more sequences using the BLAST algorithm, for example, can be as set forth below. Briefly, amino acid sequence alignments can be performed using BLASTP version 2.0.8 (Jan-05-1999) and the following parameters:

Matrix: 0 BLOSUM62; gap open: 11; gap extension: 1; x_dropoff: 50; expect: 10.0; wordsize: 3; filter: on. Nucleic acid sequence alignments can be performed using BLASTN version 2.0.6 (Sept-16-1998) and the following parameters: Match: 1; mismatch: -2; gap open: 5; gap extension: 2; x_dropoff: 50; expect: 10.0; wordsize: 11; filter: off. Those skilled in the art will know what modifications can be made to the above parameters to either increase or decrease the stringency of the comparison, for example, and determine the relatedness of two or more sequences.

In some embodiments, the present invention provides a non-naturally occurring microbial organism having a 1,3-butanediol (1,3-BDO) pathway that includes at least one exogenous nucleic acid encoding a 1,3-BDO pathway enzyme or protein expressed in a sufficient amount to produce 1,3-BDO. The 1,3-BDO pathway includes Methanol methyltransferase (MtaB), Corrinoid protein (MtaC), Methyltetrahydrofolate:corrinoid protein methyltransferase (MtaA), Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), Corrinoid iron-sulfur protein (AcsD), Nickel-protein assembly protein (AcsF & CooC), Ferredoxin (Orf7), Acetyl-CoA synthase (AcsB & AcsC), Carbon monoxide dehydrogenase (AcsA), Hydrogenase (Hyd), Acetoacetyl-CoA thiolase (AtoB), Acetoacetyl-CoA reductase (CoA-dependent, aldehyde forming), 3-oxobutylaldehyde reductase (ketone reducing), 3-hydroxybutylaldehyde reductase, Acetoacetyl-CoA reductase (CoA-dependent, alcohol forming), 3-oxobutylaldehyde reductase (aldehyde reducing), 4-hydroxy,2-butanone reductase, Acetoacetyl-CoA reductase (ketone reducing), 3-hydroxybutyl-CoA reductase (aldehyde forming), 3-hydroxybutyl-CoA reductase (alcohol forming), 3-hydroxybutyl-CoA transferase, 3-hydroxybutyl-CoA hydrolase, 3-hydroxybutyl-CoA synthetase, 3-hydroxybutyrate dehydrogenase, 3-hydroxybutyrate reductase, acetoacetyl-CoA transferase, acetoacetyl-CoA hydrolase, acetoacetyl-CoA synthetase, or acetoacetate reductase.

In some embodiments, the 1,3-BDO pathway enzymes are a set of enzymes selected from: A: 1) Methanol methyltransferase (MtaB), 2) Corrinoid protein (MtaC), 3) Methyltetrahydrofolate:corrinoid protein methyltransferase (MtaA), 4) Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), 5) Corrinoid iron-sulfur protein (AcsD), 6) Nickel-protein assembly protein (AcsF & CooC), 7) Ferredoxin (Orf7), 8) Acetyl-CoA synthase (AcsB & AcsC), 9) Carbon monoxide dehydrogenase (AcsA), 10) Hydrogenase, 11) Acetoacetyl-CoA thiolase (AtoB), 12) Acetoacetyl-CoA reductase (CoA-dependent, aldehyde forming), 13) 3-oxobutylaldehyde reductase (aldehyde reducing), and 14) 4-hydroxy,2-butanone reductase; B: 1) Methanol methyltransferase (MtaB), 2) Corrinoid protein

- (MtaC), 3) Methyltetrahydrofolate:corrinoid protein methyltransferase (MtaA), 4) Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), 5) Corrinoid iron-sulfur protein (AcsD), 6) Nickel-protein assembly protein (AcsF & CooC), 7) Ferredoxin (Orf7), 8) Acetyl-CoA synthase (AcsB & AcsC), 9) Carbon monoxide dehydrogenase (AcsA), 10) Hydrogenase, 11) Acetoacetyl-CoA thiolase (AtoB), 12) Acetoacetyl-CoA reductase (CoA-dependent, aldehyde forming), 13) 3-oxobutyraldehyde reductase (ketone reducing), and 14) 3-hydroxybutyraldehyde reductase; C: 1) Methanol methyltransferase (MtaB), 2) Corrinoid protein (MtaC), 3) Methyltetrahydrofolate:corrinoid protein methyltransferase (MtaA), 4) Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), 5) Corrinoid iron-sulfur protein (AcsD), 6) Nickel-protein assembly protein (AcsF & CooC), 7) Ferredoxin (Orf7), 8) Acetyl-CoA synthase (AcsB & AcsC), 9) Carbon monoxide dehydrogenase (AcsA), 10) Hydrogenase, 11) Acetoacetyl-CoA thiolase (AtoB), 12) Acetoacetyl-CoA reductase (ketone reducing), 13) 3-hydroxybutyryl-CoA reductase (aldehyde forming), and 14) 3-hydroxybutyraldehyde reductase; D: 1) Methanol methyltransferase (MtaB), 2) Corrinoid protein (MtaC), 3) Methyltetrahydrofolate:corrinoid protein methyltransferase (MtaA), 4) Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), 5) Corrinoid iron-sulfur protein (AcsD), 6) Nickel-protein assembly protein (AcsF & CooC), 7) Ferredoxin (Orf7), 8) Acetyl-CoA synthase (AcsB & AcsC), 9) Carbon monoxide dehydrogenase (AcsA), 10) Hydrogenase, 11) Acetoacetyl-CoA thiolase (AtoB), 12) Acetoacetyl-CoA reductase (ketone reducing), and 13) 3-hydroxybutyryl-CoA reductase (alcohol forming); E: 1) Methanol methyltransferase (MtaB), 2) Corrinoid protein (MtaC), 3) Methyltetrahydrofolate:corrinoid protein methyltransferase (MtaA), 4) Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), 5) Corrinoid iron-sulfur protein (AcsD), 6) Nickel-protein assembly protein (AcsF & CooC), 7) Ferredoxin (Orf7), 8) Acetyl-CoA synthase (AcsB & AcsC), 9) Carbon monoxide dehydrogenase (AcsA), 10) Hydrogenase, 11) Acetoacetyl-CoA thiolase (AtoB), 12) Acetoacetyl-CoA reductase (CoA-dependent, alcohol forming), and 14) 4-hydroxy,2-butanone reductase; F: 1) Methanol methyltransferase (MtaB), 2) Corrinoid protein (MtaC), 3) Methyltetrahydrofolate:corrinoid protein methyltransferase (MtaA), 4) Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), 5) Corrinoid iron-sulfur protein (AcsD), 6) Nickel-protein assembly protein (AcsF & CooC), 7) Ferredoxin (Orf7), 8) Acetyl-CoA synthase (AcsB & AcsC), 9) Carbon monoxide dehydrogenase (AcsA), 10) Hydrogenase, 11) Acetoacetyl-CoA thiolase (AtoB), 12) Acetoacetyl-CoA transferase, hydrolase, or synthetase, 13) Acetoacetate reductase, 14) 3-oxobutyraldehyde reductase (ketone reducing), and 15) 3-hydroxybutyraldehyde reductase; G: 1) Methanol methyltransferase

- (MtaB), 2) Corrinoid protein (MtaC), 3) Methyltetrahydrofolate:corrinoid protein methyltransferase (MtaA), 4) Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), 5) Corrinoid iron-sulfur protein (AcsD), 6) Nickel-protein assembly protein (AcsF & CooC), 7) Ferredoxin (Orf7), 8) Acetyl-CoA synthase (AcsB & AcsC), 9) Carbon monoxide dehydrogenase (AcsA), 10) Hydrogenase, 11) Acetoacetyl-CoA thiolase (AtoB), 12) Acetoacetyl-CoA transferase, hydrolase, or synthetase, 13) Acetoacetate reductase, 14) 3-oxobutyraldehyde reductase (aldehyde reducing), and 15) 4-hydroxy,2-butanone reductase; H: 1) Methanol methyltransferase (MtaB), 2) Corrinoid protein (MtaC), 3) Methyltetrahydrofolate:corrinoid protein methyltransferase (MtaA), 4) Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), 5) Corrinoid iron-sulfur protein (AcsD), 6) Nickel-protein assembly protein (AcsF & CooC), 7) Ferredoxin (Orf7), 8) Acetyl-CoA synthase (AcsB & AcsC), 9) Carbon monoxide dehydrogenase (AcsA), 10) Hydrogenase, 11) Acetoacetyl-CoA thiolase (AtoB), 12) Acetoacetyl-CoA reductase (ketone reducing), 13) 3-hydroxybutyryl-CoA transferase, hydrolase, or synthetase, 14) 3-hydroxybutyrate reductase, and 15) 3-hydroxybutyraldehyde reductase; I: 1) Methanol methyltransferase (MtaB), 2) Corrinoid protein (MtaC), 3) Methyltetrahydrofolate:corrinoid protein methyltransferase (MtaA), 4) Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), 5) Corrinoid iron-sulfur protein (AcsD), 6) Nickel-protein assembly protein (AcsF & CooC), 7) Ferredoxin (Orf7), 8) Acetyl-CoA synthase (AcsB & AcsC), 9) Carbon monoxide dehydrogenase (AcsA), 10) Hydrogenase, 11) Acetoacetyl-CoA thiolase (AtoB), 12) Acetoacetyl-CoA transferase, hydrolase, or synthetase, 13) 3-hydroxybutyrate dehydrogenase, 14) 3-hydroxybutyrate reductase, and 15) 3-hydroxybutyraldehyde reductase.

- The non-naturally occurring microbial organism can include two exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme, in some embodiments, three exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme, in other embodiments, four exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme, in other embodiments, five exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme, in other embodiments, six exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme, in other embodiments, seven exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme, in other embodiments, eight exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme, in other embodiments, nine exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme, in other embodiments, 10 exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme, in other embodiments, 11 exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme, in other embodiments, 12 exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme, in other embodiments, 13 exogenous

nucleic acids each encoding a 1,3-BDO pathway enzyme, in other embodiments, 14 exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme, in other embodiments, 15 exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme, in still other embodiments. Any of the at least one exogenous nucleic acid can be a heterologous nucleic acid and the non-naturally occurring microbial organism can be constructed for culturing in a substantially anaerobic culture medium. Such microbial organisms can use a carbon feedstock selected from 1) methanol and CO, 2) methanol, CO₂, and H₂, 3) methanol, CO, CO₂, and H₂, 4) methanol and synthesis gas comprising CO and H₂, 5) methanol and synthesis gas comprising CO, CO₂, and H₂, 6) one or more carbohydrates, 7) methanol and one or more carbohydrates, and 8) methanol. Exemplary carbohydrates include, but are not limited to, glucose, sucrose, xylose, arabinose, and glycerol.

In some embodiments, the present invention provides a non-naturally occurring microbial organism, comprising a microbial organism having a 1,3-butanediol (1,3-BDO) pathway comprising at least one exogenous nucleic acid encoding a 1,3-BDO pathway enzyme or protein expressed in a sufficient amount to produce 1,3-BDO said 1,3-BDO pathway comprising Formate dehydrogenase, Formyltetrahydrofolate synthetase, Methenyltetrahydrofolate cyclohydrolase, Methylenetetrahydrofolate dehydrogenase, Methylenetetrahydrofolate reductase, Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), Corrinoid iron-sulfur protein (AcsD), Nickel-protein assembly protein (AcsF & CooC), Ferredoxin (Orf7), Acetyl-CoA synthase (AcsB & AcsC), Carbon monoxide dehydrogenase (AcsA), Hydrogenase (Hyd), Acetoacetyl-CoA thiolase (AtoB), Acetoacetyl-CoA reductase (CoA-dependent, aldehyde forming), 3-oxobutyraldehyde reductase (ketone reducing), 3-hydroxybutyraldehyde reductase, Acetoacetyl-CoA reductase (CoA-dependent, alcohol forming), 3-oxobutyraldehyde reductase (aldehyde reducing), 4-hydroxy,2-butanone reductase, Acetoacetyl-CoA reductase (ketone reducing), 3-hydroxybutyryl-CoA reductase (aldehyde forming), 3-hydroxybutyryl-CoA reductase (alcohol forming), 3-hydroxybutyryl-CoA transferase, 3-hydroxybutyryl-CoA hydrolase, 3-hydroxybutyryl-CoA synthetase, 3-hydroxybutyrate dehydrogenase, 3-hydroxybutyrate reductase, acetoacetyl-CoA transferase, acetoacetyl-CoA hydrolase, acetoacetyl-CoA synthetase, or acetoacetate reductase.

In some embodiments, the 1,3-BDO pathway enzymes are a set of enzymes selected from: A: 1) Formate dehydrogenase, 2) Formyltetrahydrofolate synthetase, 3) Methenyltetrahydrofolate cyclohydrolase, 4) Methylenetetrahydrofolate dehydrogenase, 5) Methylenetetrahydrofolate reductase, 6) Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), 7) Corrinoid

iron-sulfur protein (AcsD), 8) Nickel-protein assembly protein (AcsF & CooC), 9) Ferredoxin (Orf7), 10) Acetyl-CoA synthase (AcsB & AcsC), 11) Carbon monoxide dehydrogenase (AcsA), 12) Hydrogenase (Hyd), 13) Acetoacetyl-CoA thiolase (AtoB), 14) Acetoacetyl-CoA reductase (CoA-dependent, aldehyde forming), 15) 3-oxobutyraldehyde reductase (aldehyde reducing), and 16) 4-hydroxy,2-butanone reductase; B: 1) Formate dehydrogenase, 2) Formyltetrahydrofolate synthetase, 3) Methenyltetrahydrofolate cyclohydrolase, 4) Methylenetetrahydrofolate dehydrogenase, 5) Methylenetetrahydrofolate reductase, 6) Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), 7) Corrinoid iron-sulfur protein (AcsD), 8) Nickel-protein assembly protein (AcsF & CooC), 9) Ferredoxin (Orf7), 10) Acetyl-CoA synthase (AcsB & AcsC), 11) Carbon monoxide dehydrogenase (AcsA), 12) Hydrogenase (Hyd), 13) Acetoacetyl-CoA thiolase (AtoB), 14) Acetoacetyl-CoA reductase (CoA-dependent, aldehyde forming), 15) 3-oxobutyraldehyde reductase (ketone reducing), and 16) 3-hydroxybutyraldehyde reductase; C: 1) Formate dehydrogenase, 2) Formyltetrahydrofolate synthetase, 3) Methenyltetrahydrofolate cyclohydrolase, 4) Methylenetetrahydrofolate dehydrogenase, 5) Methylenetetrahydrofolate reductase, 6) Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), 7) Corrinoid iron-sulfur protein (AcsD), 8) Nickel-protein assembly protein (AcsF & CooC), 9) Ferredoxin (Orf7), 10) Acetyl-CoA synthase (AcsB & AcsC), 11) Carbon monoxide dehydrogenase (AcsA), 12) Hydrogenase (Hyd), 13) Acetoacetyl-CoA thiolase (AtoB), 14) Acetoacetyl-CoA reductase (ketone reducing), 15) 3-hydroxybutyryl-CoA reductase (aldehyde forming), and 16) 3-hydroxybutyraldehyde reductase; D: 1) Formate dehydrogenase, 2) Formyltetrahydrofolate synthetase, 3) Methenyltetrahydrofolate cyclohydrolase, 4) Methylenetetrahydrofolate dehydrogenase, 5) Methylenetetrahydrofolate reductase, 6) Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), 7) Corrinoid iron-sulfur protein (AcsD), 8) Nickel-protein assembly protein (AcsF & CooC), 9) Ferredoxin (Orf7), 10) Acetyl-CoA synthase (AcsB & AcsC), 11) Carbon monoxide dehydrogenase (AcsA), 12) Hydrogenase (Hyd), 13) Acetoacetyl-CoA thiolase (AtoB), 14) Acetoacetyl-CoA reductase (CoA-dependent, alcohol forming), and 15) 4-hydroxy,2-butanone reductase; E: 1) Formate dehydrogenase, 2) Formyltetrahydrofolate synthetase, 3) Methenyltetrahydrofolate cyclohydrolase, 4) Methylenetetrahydrofolate dehydrogenase, 5) Methylenetetrahydrofolate reductase, 6) Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), 7) Corrinoid iron-sulfur protein (AcsD), 8) Nickel-protein assembly protein (AcsF & CooC), 9) Ferredoxin (Orf7), 10) Acetyl-CoA synthase (AcsB & AcsC), 11) Carbon monoxide dehydrogenase (AcsA), 12) Hydrogenase (Hyd), 13) Acetoacetyl-CoA thiolase (AtoB), 14) Acetoacetyl-CoA reductase (ketone reducing), and 15) 3-hydroxybutyryl-CoA reductase (alcohol forming); F: 1)

Formate dehydrogenase, 2) Formyltetrahydrofolate synthetase, 3) Methenyltetrahydrofolate cyclohydrolase, 4) Methylenetetrahydrofolate dehydrogenase, 5) Methylenetetrahydrofolate reductase, 6) Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), 7) Corrinoid iron-sulfur protein (AcsD), 8) Nickel-protein assembly protein (AcsF & CooC), 9) Ferredoxin (Orf7), 10) Acetyl-CoA synthase (AcsB & AcsC), 11) Carbon monoxide dehydrogenase (AcsA), 12) Hydrogenase (Hyd), 13) Acetoacetyl-CoA thiolase (AtoB), 14) Acetoacetyl-CoA transferase, hydrolase, or synthetase, 15) Acetoacetate reductase, 16) 3-oxobutyraldehyde reductase (ketone reducing), and 17) 3-hydroxybutyraldehyde reductase; G: 1) Formate dehydrogenase, 2) Formyltetrahydrofolate synthetase, 3) Methenyltetrahydrofolate cyclohydrolase, 4) Methylenetetrahydrofolate dehydrogenase, 5) Methylenetetrahydrofolate reductase, 6) Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), 7) Corrinoid iron-sulfur protein (AcsD), 8) Nickel-protein assembly protein (AcsF & CooC), 9) Ferredoxin (Orf7), 10) Acetyl-CoA synthase (AcsB & AcsC), 11) Carbon monoxide dehydrogenase (AcsA), 12) Hydrogenase (Hyd), 13) Acetoacetyl-CoA thiolase (AtoB), 14) Acetoacetyl-CoA transferase, hydrolase, or synthetase, 15) Acetoacetate reductase, 16) 3-oxobutyraldehyde reductase (aldehyde reducing), and 17) 4-hydroxy,2-butanone reductase; H: 1) Formate dehydrogenase, 2) Formyltetrahydrofolate synthetase, 3) Methenyltetrahydrofolate cyclohydrolase, 4) Methylenetetrahydrofolate dehydrogenase, 5) Methylenetetrahydrofolate reductase, 6) Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), 7) Corrinoid iron-sulfur protein (AcsD), 8) Nickel-protein assembly protein (AcsF & CooC), 9) Ferredoxin (Orf7), 10) Acetyl-CoA synthase (AcsB & AcsC), 11) Carbon monoxide dehydrogenase (AcsA), 12) Hydrogenase (Hyd), 13) Acetoacetyl-CoA thiolase (AtoB), 14) Acetoacetyl-CoA reductase (ketone reducing), 15) 3-hydroxybutyryl-CoA transferase, hydrolase, or synthetase, 16) 3-hydroxybutyrate reductase, and 17) 3-hydroxybutyraldehyde reductase; I: 1) Formate dehydrogenase, 2) Formyltetrahydrofolate synthetase, 3) Methenyltetrahydrofolate cyclohydrolase, 4) Methylenetetrahydrofolate dehydrogenase, 5) Methylenetetrahydrofolate reductase, 6) Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), 7) Corrinoid iron-sulfur protein (AcsD), 8) Nickel-protein assembly protein (AcsF & CooC), 9) Ferredoxin (Orf7), 10) Acetyl-CoA synthase (AcsB & AcsC), 11) Carbon monoxide dehydrogenase (AcsA), 12) Hydrogenase (Hyd), 13) Acetoacetyl-CoA thiolase (AtoB), 14) Acetoacetyl-CoA transferase, hydrolase, or synthetase, 15) 3-hydroxybutyrate dehydrogenase, 16) 3-hydroxybutyrate reductase, and 17) 3-hydroxybutyraldehyde reductase.

The non-naturally occurring microbial organism can include two exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme, in some embodiments, three exogenous nucleic acids

each encoding a 1,3-BDO pathway enzyme, in other embodiments, four exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme, in other embodiments, five exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme, in other embodiments, six exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme, in other embodiments, seven exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme, in other embodiments, eight exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme, in other embodiments, nine exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme, in other embodiments, 10 exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme, in other embodiments, 11 exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme, in other embodiments, 12 exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme, in other embodiments, 13 exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme, in other embodiments, 14 exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme, in other embodiments, 15 exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme, 16 exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme, in other embodiments, and 17 exogenous nucleic acids, in still other embodiments. Any of the at least one exogenous nucleic acid can be a heterologous nucleic acid and the non-naturally occurring microbial organism can be constructed for culturing in a substantially anaerobic culture medium. Such microbial organisms can use a carbon feedstock selected from 1) CO, 2) CO₂ and H₂, 3) CO, CO₂, and H₂, 4) synthesis gas comprising CO and H₂, 5) synthesis gas comprising CO, CO₂, and H₂, and 6) one or more carbohydrates. Exemplary carbohydrates include, but are not limited to, glucose, sucrose, xylose, arabinose, and glycerol.

In an additional embodiment, the invention provides a non-naturally occurring microbial organisms having a 1,3-butanediol pathways, wherein the non-naturally occurring microbial organism comprises at least one exogenous nucleic acid encoding an enzyme or protein that converts a substrate to a product selected from the group consisting of carbon dioxide to FOR, FOR to 10FTHF, 10FTHF to METHF, METHF to MLTHF, MLTHF to 5MTHF, methanol to CH₃-MtaC, CH₃-MtaC to 5MTHF, 5MTHF to CFeSp-CH₃, CFeSp-CH₃ to acetyl-CoA, acetyl-CoA to acetoacetyl-CoA, acetoacetyl-CoA to 3-hydroxybutyryl-CoA, 3-hydroxybutyryl-CoA to 3-hydroxybutyraldehyde, 3-hydroxybutyraldehyde to 1,3-BDO, 3-hydroxybutyryl-CoA to 1,3-BDO, acetoacetyl-CoA to 4-hydroxy-2 butanone, acetoacetyl-CoA to 3-oxobutyraldehyde, 3-oxobutyraldehyde to 4-hydroxy-2-butanone, 4-hydroxy-2-butanone to 1,3-BDO, 3-oxobutyraldehyde to 3-hydroxybutyraldehyde, acetoacetyl-CoA to acetoacetate, acetoacetate to 3-oxobutyraldehyde, acetoacetate to 3-hydroxybutyrate, 3-hydroxybutyryl-CoA to 3-hydroxybutyrate, and 3-hydroxybutyrate to 3-hydroxybutyraldehyde. One skilled in the art will

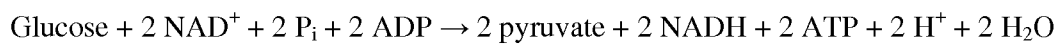
understand that these are merely exemplary and that any of the substrate-product pairs disclosed herein suitable to produce a desired product and for which an appropriate activity is available for the conversion of the substrate to the product can be readily determined by one skilled in the art based on the teachings herein. Thus, the invention provides a non-naturally occurring microbial
5 organism containing at least one exogenous nucleic acid encoding an enzyme or protein, where the enzyme or protein converts the substrates and products of a 1,3-butanediol pathway, such as those shown in Figures 4 and 5.

While generally described herein as a microbial organism that contains a 1,3-butanediol pathway, it is understood that the invention additionally provides a non-naturally occurring
10 microbial organism comprising at least one exogenous nucleic acid encoding a 1,3-butanediol pathway enzyme or protein expressed in a sufficient amount to produce an intermediate of a 1,3-butanediol pathway. For example, as disclosed herein, 1,3-butanediol pathway are exemplified in Figures 4 and 5. Therefore, in addition to a microbial organism containing a 1,3-butanediol pathway that produces 1,3-butanediol, the invention additionally provides a non-naturally
15 occurring microbial organism comprising at least one exogenous nucleic acid encoding a 1,3-butanediol pathway enzyme, where the microbial organism produces a 1,3-butanediol pathway intermediate, for example, acetyl-CoA, acetoacetyl-CoA, acetoacetate, 3-hydroxybutyryl-CoA, 3-hydroxybutyrate, 3-oxobutyraldehyde, 4-hydroxy-2-butanone, and 3-hydroxybutyraldehyde.

It is understood that any of the pathways disclosed herein, as described in the Examples and
20 exemplified in the Figures, including the pathways of Figures 4 and 5, can be utilized to generate a non-naturally occurring microbial organism that produces any pathway intermediate or product, as desired. As disclosed herein, such a microbial organism that produces an intermediate can be used in combination with another microbial organism expressing downstream pathway enzymes to produce a desired product. However, it is understood that a
25 non-naturally occurring microbial organism that produces a 1,3-butanediol pathway intermediate can be utilized to produce the intermediate as a desired product. Thus, the non-naturally occurring organisms of the invention can be used to produce, for example, acetoacetate, 3-hydroxybutyrate, 3-oxobutyraldehyde, 3-hydroxybutyraldehyde, or 4-hydroxy-2-butanone. Thus, any of these intermediate products can be produced in a separate organism utilizing the
30 pathways shown in Figures 4 and 5.

In some embodiments, non-naturally occurring organisms described herein have three capabilities which are depicted in Figure 4: 1) a functional methyltransferase system allowing the production of 5-methyl-tetrahydrofolate (Me-THF) from methanol and THF, 2) the ability to

combine CO, Coenzyme A, and the methyl group of Me-THF to form acetyl-CoA, and 3) the ability to synthesize 1,3-butanediol from acetyl-CoA. The latter can be facilitated by the generation of pyruvate from a carbohydrate via glycolysis. Glycolysis is an anaerobic metabolic pathway that is found in the cytoplasm of cells in all living organisms. The process converts one
 5 molecule of glucose into two molecules of pyruvate, while providing two net molecules of ATP:



Such non-naturally occurring organisms can ‘fix’ carbon from exogenous CO and/or CO₂ and methanol to synthesize acetyl-CoA, cell mass, and products. Note that implementing the pathway to form acetyl-CoA from methanol and syngas is energetically favorable compared to
 10 utilizing the full Wood-Ljungdahl pathway. For example, the direct conversion of synthesis gas to acetate is an energetically neutral process (see Figures 1 and 2). Specifically, one ATP molecule is consumed during the formation of formyl-THF by formyl-THF synthase and one ATP molecule is produced during the production of acetate via acetate kinase. The pathways disclosed herein circumvent the ATP consumption by ensuring that the methyl group on the
 15 methyl branch product, methyl-THF, is obtained from methanol rather than CO₂. This ensures that acetate formation has a positive ATP yield that can help support cell growth and maintenance. A host organism engineered with these capabilities that also possesses the capability for anapleurosis (e.g., *E. coli*) can grow on the methanol and syngas-generated acetyl-CoA in the presence of a suitable external electron acceptor such as nitrate. This electron
 20 acceptor is used to accept electrons from the reduced quinone formed via succinate dehydrogenase. A further benefit of adding an external electron acceptor is that additional energy for cell growth, maintenance, and product formation can be generated from respiration of acetyl-CoA. In some embodiments, another non-naturally occurring microbial organism has a pyruvate ferredoxin oxidoreductase (PFOR) enzyme or other enzymes that facilitate the
 25 conversion of pyruvate into acetyl-CoA or vice versa into the strain to facilitate the synthesis of biomass precursors in the absence of an external electron acceptor. A further characteristic of such non-naturally occurring organisms is the capability for extracting reducing equivalents from molecular hydrogen. This enables a high yield of reduced products such as ethanol, butanol, isobutanol, isopropanol, 1,4-butanediol, 1,3-butanediol, succinic acid, fumaric acid,
 30 malic acid, 4-hydroxybutyric acid, 3-hydroxypropionic acid, lactic acid, adipic acid, methacrylic acid, and acrylic acid.

The organisms described in this invention can produce acetyl-CoA, cell mass, and targeted chemicals, more specifically 1,3-BDO, from: 1) methanol and CO, 2) methanol, CO₂, and H₂, 3)

methanol, CO, CO₂, and H₂, 4) methanol and synthesis gas comprising CO and H₂, 5) methanol and synthesis gas comprising CO, CO₂, and H₂, 6) one or more carbohydrates, 7) methanol and one or more carbohydrates, and 8) methanol. Exemplary carbohydrates include, but are not limited to, glucose, sucrose, xylose, arabinose, and glycerol.

- 5 Successfully engineering these pathways into an organism involves identifying an appropriate set of enzymes, cloning their corresponding genes into a production host, optimizing the stability and expression of these genes, optimizing fermentation conditions, and assaying for product formation following fermentation. A number of enzymes that catalyze each step of the pathways used the conversion of synthesis gas and methanol to acetyl-CoA, and further to 1,3-
- 10 butanediol are described below. To engineer a production host for the utilization of syngas and methanol, one or more exogenous DNA sequence(s) encoding the enzymes of these pathways can be expressed in the microorganism.

Expression of the modified Wood-Ljungdahl pathway in a foreign host, as shown in Figure 4, involves introducing a set of methyltransferases to utilize the carbon and hydrogen provided by

15 methanol and the carbon provided by CO and/or CO₂. A complex of 3 methyltransferase proteins, denoted MtaA, MtaB, and MtaC, perform the desired methanol methyltransferase activity (Sauer et al., *Eur. J. Biochem.* 243.3:670-677 (1997); Naidu and Ragsdale, *J. Bacteriol.* 183.11:3276-3281 (2001); Tallant and Krzycki, *J. Biol. Chem.* 276.6:4485-4493 (2001); Tallant and Krzycki, *J. Bacteriol.* 179.22:6902-6911 (1997); Tallant and Krzycki, *J. Bacteriol.*

20 178.5:1295-1301 (1996); Ragsdale, S.W., *Crit. Rev. Biochem. Mol. Biol.* 39.3:165-195 (2004).

MtaB is a zinc protein that catalyzes the transfer of a methyl group from methanol to MtaC, a corrinoid protein. Exemplary genes encoding MtaB and MtaC can be found in methanogenic archaea such as *Methanosarcina barkeri* (Maeder et al., *J. Bacteriol.* 188.22:7922-7931 (2006)) and *Methanosarcina acetivorans* (Galagan et al., *Genome Res.* 12/4:532-542 (2002)), as well as

25 the acetogen, *Moorella thermoacetica* (Das et al., *Proteins* 67.1:167-176 (2007)). In general, the MtaB and MtaC genes are adjacent to one another on the chromosome as their activities are tightly interdependent. The protein sequences of various MtaB and MtaC encoding genes in *M. barkeri*, *M. acetivorans*, and *M. thermoacetica* can be identified by their following GenBank accession numbers:

<u>Protein</u>	<u>GenBank ID</u>	<u>GI number</u>	<u>Organism</u>
MtaB1	YP_304299	73668284	<i>Methanosarcina barkeri</i>
MtaC1	YP_304298	73668283	<i>Methanosarcina barkeri</i>
MtaB2	YP_307082	73671067	<i>Methanosarcina barkeri</i>
MtaC2	YP_307081	73671066	<i>Methanosarcina barkeri</i>
MtaB3	YP_304612	73668597	<i>Methanosarcina barkeri</i>
MtaC3	YP_304611	73668596	<i>Methanosarcina barkeri</i>
MtaB1	NP_615421	20089346	<i>Methanosarcina acetivorans</i>
MtaB1	NP_615422	20089347	<i>Methanosarcina acetivorans</i>
MtaB2	NP_619254	20093179	<i>Methanosarcina acetivorans</i>
MtaC2	NP_619253	20093178	<i>Methanosarcina acetivorans</i>
MtaB3	NP_616549	20090474	<i>Methanosarcina acetivorans</i>
MtaC3	NP_616550	20090475	<i>Methanosarcina acetivorans</i>
MtaB	YP_430066	83590057	<i>Moorella thermoacetica</i>
MtaC	YP_430065	83590056	<i>Moorella thermoacetica</i>

The MtaB1 and MtaC1 genes, YP_304299 and YP_304298, from *M. barkeri* were cloned into *E. coli* and sequenced (Sauer et al., *Eur. J. Biochem.* 243.3:670-677 (1997)). The crystal
5 structure of this methanol-cobalamin methyltransferase complex is also available Hagemeyer et al., *Proc. Natl. Acad. Sci. U.S.A.* 103.50:18917-18922 (2006)). The MtaB genes, YP_307082 and YP_304612, in *M. barkeri* were identified by sequence homology to YP_304299. In general, homology searches are an effective means of identifying methanol methyltransferases because MtaB encoding genes show little or no similarity to methyltransferases that act on
10 alternative substrates such as trimethylamine, dimethylamine, monomethylamine, or dimethylsulfide. The MtaC genes, YP_307081 and YP_304611, were identified based on their proximity to the MtaB genes and also their homology to YP_304298. The three sets of MtaB and MtaC genes from *M. acetivorans* have been genetically, physiologically, and biochemically characterized Pritchett and Metcalf, *Mol. Microbiol.* 56.5:1183-1194 (2005)). Mutant strains
15 lacking two of the sets were able to grow on methanol, whereas a strain lacking all three sets of MtaB and MtaC genes sets could not grow on methanol. This indicates that each set of genes plays a role in methanol utilization. The *M. thermoacetica* MtaB gene was identified based on homology to the methanogenic MtaB genes and also by its adjacent chromosomal proximity to

the methanol-induced corrinoid protein, MtaC, which has been crystallized (Zhou et al., *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* 61 Pt. 5:537-540 (2005)) and further characterized by Northern hybridization and Western Blotting (Das et al. *Proteins* 67.1:167-76 (2007)).

- 5 MtaA is zinc protein that catalyzes the transfer of the methyl group from MtaC to either Coenzyme M in methanogens or methyltetrahydrofolate in acetogens. MtaA can also utilize methylcobalamin as the methyl donor. Exemplary genes encoding MtaA can be found in methanogenic archaea such as *Methanosarcina barkeri* (Maeder et al., *J. Bacteriol.* 188.22 7922-7931 (2006)) and *Methanosarcina acetivorans* (Galagan et al., *Genome Res.* 12.4:532-542 10 (2002)), as well as the acetogen, *Moorella thermoacetica* (Das et al., *Proteins* 67.1:167-176 (2007)). In general, MtaA proteins that catalyze the transfer of the methyl group from CH₃-MtaC are difficult to identify bioinformatically as they share similarity to other corrinoid protein methyltransferases and are not oriented adjacent to the MtaB and MtaC genes on the chromosomes. Nevertheless, a number of MtaA encoding genes have been characterized. The 15 protein sequences of these genes in *M. barkeri* and *M. acetivorans* can be identified by the following GenBank accession numbers:

<u>Protein</u>	<u>GenBank ID</u>	<u>GI number</u>	<u>Organism</u>
MtaA	YP_304602	73668587	<i>Methanosarcina barkeri</i>
MtaA1	NP_619241	20093166	<i>Methanosarcina acetivorans</i>
MtaA2	NP_616548	20090473	<i>Methanosarcina acetivorans</i>

- The MtaA gene, YP_304602, from *M. barkeri* was cloned, sequenced, and functionally overexpressed in *E. coli* (Harms and Thauer, *Eur. J. Biochem.* 135.3:653-659 (1996). In *M.* 20 *acetivorans*, MtaA1 is used for growth on methanol, whereas MtaA2 is dispensable even though methane production from methanol is reduced in MtaA2 mutants (Bose et al. *J. Bacteriol.* 190.11:4017-4026 (2008). Moreover, there are multiple additional MtaA homologs in *M. barkeri* and *M. acetivorans* that are as yet uncharacterized, but can also catalyze corrinoid protein methyltransferase activity.

- 25 Putative MtaA encoding genes in *M. thermoacetica* were identified by their sequence similarity to the characterized methanogenic MtaA genes. Specifically, three *M. thermoacetica* genes show high homology (>30% sequence identity) to YP_304602 from *M. barkeri*. Unlike methanogenic MtaA proteins that naturally catalyze the transfer of the methyl group from CH₃-MtaC to Coenzyme M, an *M. thermoacetica* MtaA can transfer the methyl group to 30 methyltetrahydrofolate given the similar roles of methyltetrahydrofolate and Coenzyme M in

methanogens and acetogens, respectively. The protein sequences of putative MtaA encoding genes from *M. thermoacetica* can be identified by the following GenBank accession numbers:

<u>Protein</u>	<u>GenBank ID</u>	<u>GI number</u>	<u>Organism</u>
MtaA	YP_430937	83590928	<i>Moorella thermoacetica</i>
MtaA	YP_431175	83591166	<i>Moorella thermoacetica</i>
MtaA	YP_430935	83590926	<i>Moorella thermoacetica</i>

ACS/CODH is the central enzyme of the carbonyl branch of the Wood-Ljungdahl pathway. It catalyzes the reversible reduction of carbon dioxide to carbon monoxide and also the synthesis of acetyl-CoA from carbon monoxide, Coenzyme A, and the methyl group from a methylated corrinoid-iron-sulfur protein. The corrinoid-iron-sulfur-protein is methylated by methyltetrahydrofolate via a methyltransferase. Expression of ACS/CODH in a foreign host involves introducing one or more of the following proteins: Methyltetrahydrofolate:corrinoid, protein methyltransferase (*AcsE*), Corrinoid iron-sulfur protein (*AcsD*), Nickel-protein assembly protein (*AcsF*), Ferredoxin (*Orf7*), Acetyl-CoA synthase (*AcsB* and *AcsC*), Carbon monoxide dehydrogenase (*AcsA*), Nickel-protein assembly protein (*CooC*).

The genes used for carbon-monoxide dehydrogenase/acetyl-CoA synthase activity typically reside in a limited region of the native genome that may be an extended operon (Ragsdale, S.W., *Crit. Rev. Biochem. Mol. Biol.* 39.3:165-195 (2004); Morton et al., *J. Biol. Chem.* 266.35:23824-23828 (1991); Roberts et al., *Proc. Natl. Acad. Sci. U.S.A.* 86.1:32-36 (1989)). Each of the genes in this operon from the acetogen, *M. thermoacetica*, has already been cloned and expressed actively in *E. coli* (Morton et al., *supra* (1991); Roberts et al., *supra* (1989); Lu et al., *J. Biol. Chem.* 268.8:5605-5614 (1993)). The protein sequences of these genes can be identified by the following GenBank accession numbers:

<u>Protein</u>	<u>GenBank ID</u>	<u>GI number</u>	<u>Organism</u>
AcsE	YP_430054	83590045	<i>Moorella thermoacetica</i>
AcsD	YP_430055	83590046	<i>Moorella thermoacetica</i>
AcsF	YP_430056	83590047	<i>Moorella thermoacetica</i>
Orf7	YP_430057	83590048	<i>Moorella thermoacetica</i>
AcsC	YP_430058	83590049	<i>Moorella thermoacetica</i>
AcsB	YP_430059	83590050	<i>Moorella thermoacetica</i>
AcsA	YP_430060	83590051	<i>Moorella thermoacetica</i>
CooC	YP_430061	83590052	<i>Moorella thermoacetica</i>

The hydrogenogenic bacterium, *Carboxydotherrnus hydrogenoformans*, can utilize carbon monoxide as a growth substrate by means of acetyl-CoA synthase (Wu et al. *PLos Genet.*

1.5:e65 (2005)). In strain Z-2901, the acetyl-CoA synthase enzyme complex lacks carbon monoxide dehydrogenase due to a frameshift mutation (Wu et al. *PLoS Genet.* 1.5:e65 (2005)), whereas in strain DSM 6008, a functional unframeshifted full-length version of this protein has been purified (Svetlitchnyi et al., *Proc. Natl. Acad. Sci. U.S.A.* 101.2:446-451 (2004)). The protein sequences of the *C. hydrogenoformans* genes from strain Z-2901 can be identified by the following GenBank accession numbers:

Protein	GenBank ID	GI number	Organism
AcsE	YP_360065	78044202	<i>Carboxydotherrmus hydrogenoformans</i>
AcsD	YP_360064	78042962	<i>Carboxydotherrmus hydrogenoformans</i>
AcsF	YP_360063	78044060	<i>Carboxydotherrmus hydrogenoformans</i>
Orf7	YP_360062	78044449	<i>Carboxydotherrmus hydrogenoformans</i>
AcsC	YP_360061	78043584	<i>Carboxydotherrmus hydrogenoformans</i>
AcsB	YP_360060	78042742	<i>Carboxydotherrmus hydrogenoformans</i>
CooC	YP_360059	78044249	<i>Carboxydotherrmus hydrogenoformans</i>

The methanogenic archaeon, *Methanosarcina acetivorans*, can also grow on carbon monoxide, exhibits acetyl-CoA synthase/carbon monoxide dehydrogenase activity, and produces both acetate and formate (Lessner et al., *Proc. Natl. Acad. Sci. U.S.A.* 103.47:17921-17926 (2006)). This organism contains two sets of genes that encode ACS/CODH activity (Rother and Metcalf, *Proc. Natl. Acad. Sci. U.S.A.* 101.48:16929-16934 (2004)). The protein sequences of both sets of *M. acetivorans* genes can be identified by the following GenBank accession numbers:

Protein	GenBank ID	GI number	Organism
AcsC	NP_618736	20092661	<i>Methanosarcina acetivorans</i>
AcsD	NP_618735	20092660	<i>Methanosarcina acetivorans</i>
AcsF, CooC	NP_618734	20092659	<i>Methanosarcina acetivorans</i>
AcsB	NP_618733	20092658	<i>Methanosarcina acetivorans</i>
AcsEps	NP_618732	20092657	<i>Methanosarcina acetivorans</i>
AcsA	NP_618731	20092656	<i>Methanosarcina acetivorans</i>
AcsC	NP_615961	20089886	<i>Methanosarcina acetivorans</i>
AcsD	NP_615962	20089887	<i>Methanosarcina acetivorans</i>
AcsF, CooC	NP_615963	20089888	<i>Methanosarcina acetivorans</i>
AcsB	NP_615964	20089889	<i>Methanosarcina acetivorans</i>
AcsEps	NP_615965	20089890	<i>Methanosarcina acetivorans</i>
AcsA	NP_615966	20089891	<i>Methanosarcina acetivorans</i>

The AcsC, AcsD, AcsB, AcsE, and AcsA proteins are commonly referred to as the gamma, delta, beta, epsilon, and alpha subunits of the methanogenic CODH/ACS. Homologs to the

epsilon encoding genes are not present in acetogens such as *M. thermoacetica* or hydrogenogenic bacteria such as *C. hydrogenoformans*. Hypotheses for the existence of two active CODH/ACS operons in *M. acetivorans* include catalytic properties (i.e., K_m , V_{max} , k_{cat}) that favor carboxidotrophic or acetoclastic growth or differential gene regulation enabling various stimuli to induce CODH/ACS expression (Rother et al., *Arch. Microbiol.* 188.5:463-472 (2007)).

In both *M. thermoacetica* and *C. hydrogenoformans*, additional CODH encoding genes are located outside of the ACS/CODH operons. These enzymes provide a means for extracting electrons (or reducing equivalents) from the conversion of carbon monoxide to carbon dioxide.

The reducing equivalents are then passed to acceptors such as oxidized ferredoxin, NADP⁺, water, or hydrogen peroxide to form reduced ferredoxin, NADPH, H₂, or water, respectively. In some cases, hydrogenase encoding genes are located adjacent to a CODH. In *Rhodospirillum rubrum*, the encoded CODH/hydrogenase proteins form a membrane-bound enzyme complex that has been indicated to be a site where energy, in the form of a proton gradient, is generated from the conversion of CO to CO₂ and H₂ (Fox et al., *J. Bacteriol.* 178.21:6200-6208 (1996)). The CODH-I of *C. hydrogenoformans* and its adjacent genes have been proposed to catalyze a similar functional role based on their similarity to the *R. rubrum* CODH/hydrogenase gene cluster (Wu et al., *PLoS Genet.* 1.5:e65 (2005)). The *C. hydrogenoformans* CODH-I was also shown to exhibit intense CO oxidation and CO₂ reduction activities when linked to an electrode (Parkin et al., *J. Am. Chem. Soc.*, 129.34:10328-10329 (2007)). The genes encoding the *C. hydrogenoformans* CODH-II and CooF, a neighboring protein, were cloned and sequenced (Gonzalez and Robb, *FEMS Microb. Lett.* 191.2:243-247 (2000)). The resulting complex was membrane-bound, although cytoplasmic fractions of CODH-II were shown to catalyze the formation of NADPH suggesting an anabolic role (Svetlitchnyi et al., *J. Bacteriol.* 183.17:5134-5144 (2001)). The crystal structure of the CODH-II is also available (Dobbek et al., *Science* 293.5533:1281-1285 (2001)). The protein sequences of exemplary CODH and hydrogenase genes can be identified by the following GenBank accession numbers:

Protein	GenBank ID	GI number	Organism
CODH (putative)	YP_430813	83590804	<i>Moorella thermoacetica</i>
CODH-I (CooS-I)	YP_360644	78043418	<i>Carboxydotherrmus hydrogenoformans</i>
CooF	YP_360645	78044791	<i>Carboxydotherrmus hydrogenoformans</i>
HypA	YP_360646	78044340	<i>Carboxydotherrmus hydrogenoformans</i>
CooH	YP_360647	78043871	<i>Carboxydotherrmus hydrogenoformans</i>

CooU	YP_360648	78044023	<i>Carboxydotherrnus hydrogenoformans</i>
CooX	YP_360649	78043124	<i>Carboxydotherrnus hydrogenoformans</i>
CooL	YP_360650	78043938	<i>Carboxydotherrnus hydrogenoformans</i>
CooK	YP_360651	78044700	<i>Carboxydotherrnus hydrogenoformans</i>
CooM	YP_360652	78043942	<i>Carboxydotherrnus hydrogenoformans</i>
CooM	AAC45116	1515466	<i>Rhodospirillum rubrum</i>
CooK	AAC45117	1515467	<i>Rhodospirillum rubrum</i>
CooL	AAC45118	1515468	<i>Rhodospirillum rubrum</i>
CooX	AAC45119	1515469	<i>Rhodospirillum rubrum</i>
CooU	AAC45120	1515470	<i>Rhodospirillum rubrum</i>
CooH	AAC45121	1498746	<i>Rhodospirillum rubrum</i>
CooF	AAC45122	1498747	<i>Rhodospirillum rubrum</i>
CODH (CooS)	AAC45123	1498748	<i>Rhodospirillum rubrum</i>
CooC	AAC45124	1498749	<i>Rhodospirillum rubrum</i>
CooT	AAC45125	1498750	<i>Rhodospirillum rubrum</i>
CooJ	AAC45126	1498751	<i>Rhodospirillum rubrum</i>
CODH-II (CooS-II)	YP_358957	78044574	<i>Carboxydotherrnus hydrogenoformans</i>
CooF	YP_358958	78045112	<i>Carboxydotherrnus hydrogenoformans</i>

Anaerobic growth on synthesis gas and methanol in the absence of an external electron acceptor is conferred upon the host organism with MTR and ACS/CODH activity by allowing pyruvate synthesis via pyruvate ferredoxin oxidoreductase (PFOR). This enzyme allows reversible

5 conversion of pyruvate into acetyl-CoA. The PFOR from *Desulfovibrio africanus* has been cloned and expressed in *E. coli* resulting in an active recombinant enzyme that was stable for several days in the presence of oxygen (Pieulle et al., , *J. Bacteriol.* 179.18:5684-5692 (1997)). Oxygen stability is relatively uncommon in PFORs but can be conferred by a 60 residue extension in the polypeptide chain of the *D. africanus* enzyme. The *M. thermoacetica* PFOR is

10 also well characterized (Menon and Ragsdale *Biochemistry* 36.28:8484-8494 (1997)) and was shown to have high activity in the direction of pyruvate synthesis during autotrophic growth (Furdui and Ragsdale, *J. Biol. Chem.* 275.37:28494-28499 (2000)). Further, *E. coli* possesses an uncharacterized open reading frame, *ydbK*, that encodes a protein that is 51% identical to the *M. thermoacetica* PFOR. Evidence for pyruvate oxidoreductase activity in *E. coli* has been

15 described (Blaschkowski et al., *Eur. J. Biochem.* 123.3:563-569 (1982)). The protein sequences of these exemplary PFOR enzymes can be identified by the following GenBank accession numbers. Several additional PFOR enzymes are described in the following review (Ragsdale, S.W., *Chem. Rev.* 103.6:2333-2346 (2003)):

<u>Protein</u>	<u>GenBank ID</u>	<u>GI number</u>	<u>Organism</u>
Por	CAA70873.1	1770208	<i>Desulfovibrio africanus</i>
Por	YP_428946.1	83588937	<i>Moorella thermoacetica</i>
YdbK	NP_415896.1	16129339	<i>Escherichia coli</i>

The conversion of pyruvate into acetyl-CoA can be catalyzed by several other enzymes or their combinations thereof. For example, pyruvate dehydrogenase can transform pyruvate into acetyl-CoA with the concomitant reduction of a molecule of NAD into NADH. It is a multi-enzyme complex that catalyzes a series of partial reactions which results in acylating oxidative decarboxylation of pyruvate. The enzyme comprises of three subunits: the pyruvate decarboxylase (E1), dihydrolipoamide acyltransferase (E2) and dihydrolipoamide dehydrogenase (E3). This enzyme is naturally present in several organisms, including *E. coli* and *S. cerevisiae*. In the *E. coli* enzyme, specific residues in the E1 component are responsible for substrate specificity (Bisswanger, H., *J. Biol. Chem.* 256:815-82 (1981); Bremer, J., *Eur. J. Biochem.* 8:535-540 (1969); Gong et al., *J. Biol. Chem.* 275:13645-13653 (2000)). Enzyme engineering efforts have improved the *E. coli* PDH enzyme activity under anaerobic conditions (Kim et al., *J. Bacteriol.* 190:3851-3858 (2008); Kim et al., *Appl. Environ. Microbiol.* 73:1766-1771 (2007); Zhou et al., *Biotechnol. Lett.* 30:335-342 (2008)). In contrast to the *E. coli* PDH, the *B. subtilis* complex is active and required for growth under anaerobic conditions (Nakano et al., *J. Bacteriol.* 179:6749-6755 (1997)). The *Klebsiella pneumoniae* PDH, characterized during growth on glycerol, is also active under anaerobic conditions (Menzel et al., *J. Biotechnol.* 56:135-142 (1997)). Crystal structures of the enzyme complex from bovine kidney (Zhou et al., *Proc. Natl. Acad. Sci. U.S.A.* 98:14802-14807 (2001)) and the E2 catalytic domain from *Azotobacter vinelandii* are available (Mattevi et al., *Science* 255:1544-1550 (1992)). Yet another enzyme that can catalyze this conversion is pyruvate formate lyase. This enzyme catalyzes the conversion of pyruvate and CoA into acetyl-CoA and formate. Pyruvate formate lyase is a common enzyme in prokaryotic organisms that is used to help modulate anaerobic redox balance. Exemplary enzymes can be found in *Escherichia coli* encoded by *pflB* (Knappe and Sawers, *FEMS.Microbiol Rev.* 6:383-398 (1990)), *Lactococcus lactis* (Melchiorson et al., *Appl Microbiol Biotechnol* 58:338-344 (2002)), and *Streptococcus mutans* (Takahashi-Abbe et al., *Oral.Microbiol Immunol.* 18:293-297 (2003)). *E. coli* possesses an additional pyruvate formate lyase, encoded by *tdcE*, that catalyzes the conversion of pyruvate or 2-oxobutanoate to acetyl-CoA or propionyl-CoA, respectively (Hesslinger et al., *Mol. Microbiol* 27:477-492 (1998)). Both *pflB* and *tdcE* from *E. coli* require the presence of pyruvate formate lyase activating enzyme, encoded by *pflA*. Further, a short protein encoded by *yfiD* in *E. coli* can

associate with and restore activity to oxygen-cleaved pyruvate formate lyase (Vey et al., *Proc.Natl. Acad. Sci. U.S.A.* 105:16137-16141 (2008). Note that *pflA* and *pflB* from *E. coli* were expressed in *S. cerevisiae* as a means to increase cytosolic acetyl-CoA for butanol production as described in WO/2008/080124]. Additional pyruvate formate lyase and activating enzyme
5 candidates, encoded by *pfl* and *act*, respectively, are found in *Clostridium pasteurianum* (Weidner et al., *J Bacteriol.* 178:2440-2444 (1996)).

Further, different enzymes can be used in combination to convert pyruvate into acetyl-CoA. For example, in *S. cerevisiae*, acetyl-CoA is obtained in the cytosol by first decarboxylating pyruvate to form acetaldehyde; the latter is oxidized to acetate by acetaldehyde dehydrogenase
10 and subsequently activated to form acetyl-CoA by acetyl-CoA synthetase. Acetyl-CoA synthetase is a native enzyme in several other organisms including *E. coli* (Kumari et al., *J. Bacteriol.* 177:2878-2886 (1995)), *Salmonella enterica* (Starai et al., *Microbiology* 151:3793-3801 (2005); Starai et al., *J. Biol. Chem.* 280:26200-26205 (2005)), and *Moorella thermoacetica* (described already). Alternatively, acetate can be activated to form acetyl-CoA by acetate
15 kinase and phosphotransacetylase. Acetate kinase first converts acetate into acetyl-phosphate with the accompanying use of an ATP molecule. Acetyl-phosphate and CoA are next converted into acetyl-CoA with the release of one phosphate by phosphotransacetylase. Both acetate kinase and phosphotransacetylase are well-studied enzymes in several *Clostridia* and *Methanosarcina thermophila*

20 Yet another way of converting pyruvate to acetyl-CoA is via pyruvate oxidase. Pyruvate oxidase converts pyruvate into acetate, using ubiquinone as the electron acceptor. In *E. coli*, this activity is encoded by *poxB*. PoxB has similarity to pyruvate decarboxylase of *S. cerevisiae* and *Zymomonas mobilis*. The enzyme has a thiamin pyrophosphate cofactor (Koland and Gennis, *Biochemistry* 21:4438-4442 (1982)); O'Brien et al., *Biochemistry* 16:3105-3109 (1977); O'Brien
25 and Gennis, *J. Biol. Chem.* 255:3302-3307 (1980)) and a flavin adenine dinucleotide (FAD) cofactor. Acetate can then be converted into acetyl-CoA by either acetyl-CoA synthetase or by acetate kinase and phosphotransacetylase, as described earlier. Some of these enzymes can also catalyze the reverse reaction from acetyl-CoA to pyruvate.

Unlike the redox neutral conversion of CO and MeOH to acetyl-CoA or acetate, the production
30 of more highly reduced products such as ethanol, butanol, isobutanol, isopropanol, 1,4-butanediol, 1,3-butanediol, succinic acid, fumaric acid, malic acid, 4-hydroxybutyric acid, 3-hydroxypropionic acid, lactic acid, adipic acid, methacrylic acid, and acrylic acid at the highest possible yield from gaseous substrates relies on the extraction of additional reducing equivalents

from both CO and H₂ (for example, see ethanol formation in Figure 4). Specifically, reducing equivalents are obtained by the conversion of CO and water to CO₂ via carbon monoxide dehydrogenase or directly from the activity of a hydrogen-utilizing hydrogenase which transfers electrons from H₂ to an acceptor such as ferredoxin, flavodoxin, FAD⁺, NAD⁺, or NADP⁺.

- 5 Native to *E. coli* and other enteric bacteria are multiple genes encoding up to four hydrogenases (Sawers, G., *Antonie van Leeuwenhoek* 66.1:57-88 1994); Sawers et al., *J. Bacteriol.* 168.1:398-404 (1986); Sawers and Boxer, *Eur. J. Biochem.* 156.2:265-275 (1986); Sawers et al., *J. Bacteriol.* 164.3:1324-1331 (1985)). Given the multiplicity of enzyme activities *E. coli* or another host organism can provide sufficient hydrogenase activity to split incoming molecular
- 10 hydrogen and reduce the corresponding acceptor. Among the endogenous hydrogen-lyase enzymes of *E. coli* are hydrogenase 3, a membrane-bound enzyme complex using ferredoxin as an acceptor, and hydrogenase 4 that also uses a ferredoxin acceptor. Hydrogenase 3 and 4 are encoded by the *hyc* and *hyf* gene clusters, respectively. Hydrogenase activity in *E. coli* is also dependent upon the expression of the *hyp* genes whose corresponding proteins are involved in
- 15 the assembly of the hydrogenase complexes (Rangarajan et al. *J. Bacteriol.* 190.4:1447-1458 (2008); Jacobi et al., *Arch. Microbiol.* 158.6:444-451 (1992)). The *M. thermoacetica* hydrogenases are suitable candidates should the production host lack sufficient endogenous hydrogenase activity. *M. thermoacetica* can grow with CO₂ as the exclusive carbon source indicating that reducing equivalents are extracted from H₂ to enable acetyl-CoA synthesis via the
- 20 Wood-Ljungdahl pathway (Drake, H. L., *J. Bacteriol.* 150.2:702-709 (1982); Kellum and Drake, *J. Bacteriol.* 160.1:466-469 (1984); Drake and Daniel *Res. Microbial.* 155.10:869-883 (2004)) (see Figure 2). *M. thermoacetica* has homologs to several *hyp*, *hyc*, and *hyf* genes from *E. coli*. These protein sequences encoded for by these genes can be identified by the following GenBank accession numbers. In addition, several gene clusters encoding hydrogenase functionality are
- 25 present in *M. thermoacetica* and their corresponding protein sequences are also provided below:

<u>Protein</u>	<u>GenBank ID</u>	<u>GI number</u>	<u>Organism</u>
HypA	NP_417206	16130633	<i>Escherichia coli</i>
HypB	NP_417207	16130634	<i>Escherichia coli</i>
HypC	NP_417208	16130635	<i>Escherichia coli</i>
HypD	NP_417209	16130636	<i>Escherichia coli</i>
HypE	NP_417210	226524740	<i>Escherichia coli</i>
HypF	NP_417192	16130619	<i>Escherichia coli</i>

Proteins in *M. thermoacetica* whose genes are homologous to the *E. coli* hyp genes are shown below:

<u>Protein</u>	<u>GenBank ID</u>	<u>GI number</u>	<u>Organism</u>
Moth_2175	YP_431007	83590998	<i>Moorella thermoacetica</i>
Moth_2176	YP_431008	83590999	<i>Moorella thermoacetica</i>
Moth_2177	YP_431009	83591000	<i>Moorella thermoacetica</i>
Moth_2178	YP_431010	83591001	<i>Moorella thermoacetica</i>
Moth_2179	YP_431011	83591002	<i>Moorella thermoacetica</i>
Moth_2180	YP_431012	83591003	<i>Moorella thermoacetica</i>
Moth_2181	YP_431013	83591004	<i>Moorella thermoacetica</i>

Hydrogenase 3 proteins are shown below:

<u>Protein</u>	<u>GenBank ID</u>	<u>GI number</u>	<u>Organism</u>
HycA	NP_417205	16130632	<i>Escherichia coli</i>
HycB	NP_417204	16130631	<i>Escherichia coli</i>
HycC	NP_417203	16130630	<i>Escherichia coli</i>
HycD	NP_417202	16130629	<i>Escherichia coli</i>
HycE	NP_417201	16130628	<i>Escherichia coli</i>
HycF	NP_417200	16130627	<i>Escherichia coli</i>
HycG	NP_417199	16130626	<i>Escherichia coli</i>
HycH	NP_417198	16130625	<i>Escherichia coli</i>
HycI	NP_417197	16130624	<i>Escherichia coli</i>

5 Hydrogenase 4 proteins are shown below:

<u>Protein</u>	<u>GenBank ID</u>	<u>GI number</u>	<u>Organism</u>
HyfA	NP_416976	90111444	<i>Escherichia coli</i>
HyfB	NP_416977	16130407	<i>Escherichia coli</i>
HyfC	NP_416978	90111445	<i>Escherichia coli</i>
HyfD	NP_416979	16130409	<i>Escherichia coli</i>
HyfE	NP_416980	16130410	<i>Escherichia coli</i>
HyfF	NP_416981	16130411	<i>Escherichia coli</i>
HyfG	NP_416982	16130412	<i>Escherichia coli</i>
HyfH	NP_416983	16130413	<i>Escherichia coli</i>
HyfI	NP_416984	16130414	<i>Escherichia coli</i>
HyfJ	NP_416985	90111446	<i>Escherichia coli</i>
HyfR	NP_416986	90111447	<i>Escherichia coli</i>

Proteins in *M. thermoacetica* whose genes are homologous to the *E. coli* *hyc* and/or *hyf* genes are shown below:

<u>Protein</u>	<u>GenBank ID</u>	<u>GI number</u>	<u>Organism</u>
Moth_2182	YP_431014	83591005	<i>Moorella thermoacetica</i>
Moth_2183	YP_431015	83591006	<i>Moorella thermoacetica</i>
Moth_2184	YP_431016	83591007	<i>Moorella thermoacetica</i>
Moth_2185	YP_431017	83591008	<i>Moorella thermoacetica</i>
Moth_2186	YP_431018	83591009	<i>Moorella thermoacetica</i>
Moth_2187	YP_431019	83591010	<i>Moorella thermoacetica</i>
Moth_2188	YP_431020	83591011	<i>Moorella thermoacetica</i>
Moth_2189	YP_431021	83591012	<i>Moorella thermoacetica</i>
Moth_2190	YP_431022	83591013	<i>Moorella thermoacetica</i>
Moth_2191	YP_431023	83591014	<i>Moorella thermoacetica</i>
Moth_2192	YP_431024	83591015	<i>Moorella thermoacetica</i>

Additional hydrogenase-encoding gene clusters in *M. thermoacetica* are shown below:

<u>Protein</u>	<u>GenBank ID</u>	<u>GI number</u>	<u>Organism</u>
Moth_0439	YP_429313	83589304	<i>Moorella thermoacetica</i>
Moth_0440	YP_429314	83589305	<i>Moorella thermoacetica</i>
Moth_0441	YP_429315	83589306	<i>Moorella thermoacetica</i>
Moth_0442	YP_429316	83589307	<i>Moorella thermoacetica</i>
Moth_0809	YP_429670	83589661	<i>Moorella thermoacetica</i>
Moth_0810	YP_429671	83589662	<i>Moorella thermoacetica</i>
Moth_0811	YP_429672	148283119	<i>Moorella thermoacetica</i>
Moth_0814	YP_429674	83589665	<i>Moorella thermoacetica</i>
Moth_0815	YP_429675	83589666	<i>Moorella thermoacetica</i>
Moth_0816	YP_429676	83589667	<i>Moorella thermoacetica</i>
Moth_1193	YP_430050	83590041	<i>Moorella thermoacetica</i>
Moth_1194	YP_430051	83590042	<i>Moorella thermoacetica</i>
Moth_1195	YP_430052	83590043	<i>Moorella thermoacetica</i>
Moth_1196	YP_430053	83590044	<i>Moorella thermoacetica</i>
Moth_1717	YP_430562	83590553	<i>Moorella thermoacetica</i>
Moth_1718	YP_430563	83590554	<i>Moorella thermoacetica</i>
Moth_1719	YP_430564	83590555	<i>Moorella thermoacetica</i>

Moth_1883	YP_430726	83590717	<i>Moorella thermoacetica</i>
Moth_1884	YP_430727	83590718	<i>Moorella thermoacetica</i>
Moth_1885	YP_430728	83590719	<i>Moorella thermoacetica</i>
Moth_1886	YP_430729	83590720	<i>Moorella thermoacetica</i>
Moth_1887	YP_430730	83590721	<i>Moorella thermoacetica</i>
Moth_1888	YP_430731	83590722	<i>Moorella thermoacetica</i>
Moth_1452	YP_430305	83590296	<i>Moorella thermoacetica</i>
Moth_1453	YP_430306	83590297	<i>Moorella thermoacetica</i>
Moth_1454	YP_430307	83590298	<i>Moorella thermoacetica</i>

1,3-butanediol production can be achieved in recombinant *E. coli* by various alternate pathways described in Figure 4 and Figure 5. All pathways first convert two molecules of acetyl-CoA into one molecule of acetoacetyl-CoA employing a thiolase.

- 5 Acetoacetyl-CoA thiolase converts two molecules of acetyl-CoA into one molecule each of acetoacetyl-CoA and CoA. Exemplary acetoacetyl-CoA thiolase enzymes include the gene products of *atoB* from *E. coli* (Martin et al., *Nat. Biotechnol.* 21.7:796-802 (2003)), *thlA* and *thlB* from *C. acetobutylicum* (Hanai et al., *Appl. Environ. Microbiol.* 73.24:7814-7818 (2007); Winzer et al. *J. Mol. Microbiol. Biotechnol.* 2.4:531-541 (2000)), and *ERG10* from *S.*
- 10 *cerevisiae* (Hiser et al., *J. Biol. Chem.* 269.50:31383-31389 (1994)). Information related to these proteins and gene can be found using the information below:

Protein	GenBank ID	GI number	Organism
<i>AtoB</i>	NP_416728	16130161	<i>Escherichia coli</i>
<i>ThlA</i>	NP_349476.1	15896127	<i>Clostridium acetobutylicum</i>
<i>ThlB</i>	NP_149242.1	15004782	<i>Clostridium acetobutylicum</i>
<i>ERG10</i>	NP_015297	6325229	<i>Saccharomyces cerevisiae</i>

- One pathway from acetoacetyl-CoA entails its reduction to 3-hydroxybutyryl-CoA by acetoacetyl-CoA reductase (ketone reducing). This can subsequently be converted to 3-
- 15 hydroxybutyraldehyde via a CoA-dependent aldehyde reductase called 3-hydroxybutyryl-CoA reductase. 3-hydroxybutyraldehyde can eventually be reduced to the product 1,3-BDO by 3-hydroxybutyraldehyde reductase. Alternatively, 3-hydroxybutyryl-CoA can be reduced directly to 1,3-BDO by an alcohol-forming CoA-dependent 3-hydroxybutyryl-CoA reductase. The genes for each of the steps in the pathway are described below.

Acetoacetyl-CoA reductase (ketone reducing) catalyzing the reduction of acetoacetyl-CoA to 3-hydroxybutyryl-CoA participates in the acetyl-CoA fermentation pathway to butyrate in several species of *Clostridia* and has been studied in detail (Jones and Woods, *Microbiol. Rev.* 50.4:484-524 (1986)). The enzyme from *Clostridium acetobutylicum*, encoded by *hbd*, has been cloned and functionally expressed in *E. coli* (Youngleson et al., *J. Bacteriol.* 171.12:6800-6807 (1989)). Additionally, subunits of two fatty acid oxidation complexes in *E. coli*, encoded by *fadB* and *fadJ*, function as 3-hydroxyacyl-CoA dehydrogenases (Binstock and Schulz, *Methods Enzymol.* 71 Pt. C: 403-411 (1981)). Other genes demonstrated to reduce acetoacetyl-CoA to 3-hydroxybutyryl-CoA are *phbB* from *Zoogloea ramigera* (Ploux et al., *Eur. J. Biochem.* 171.1:177-182 (1988)) and *phaB* from *Rhodobacter sphaeroides* (Alber et al., *Mol. Microbiol.* 61.2:297-309 (2006)). The former gene is NADPH-dependent, its nucleotide sequence has been determined (Peoples and Sinskey, *Mol. Microbiol.* 3.3:349-357 (1989)) and the gene has been expressed in *E. coli*. Substrate specificity studies on the gene led to the conclusion that it could accept 3-oxopropionyl-CoA as a substrate besides acetoacetyl-CoA (Ploux et al., *Eur. J. Biochem.* 171.1:177-182 (1988)). Additional genes include *Hbd1* (C-terminal domain) and *Hbd2* (N-terminal domain) in *Clostridium kluyveri* (Hillmer and Gottschalk, *Biochim. Biophys. Acta* 333:12-23 (1974)) and *HSD17B10* in *Bos taurus* (Wakil et al., *J. Biol. Chem.* 207.2:631-638 (1954)). A summary of the gene and protein information are shown below:

<u>Protein</u>	<u>GENBANK ID</u>	<u>GI NUMBER</u>	<u>ORGANISM</u>
<i>fadB</i>	P21177.2	119811	<i>Escherichia coli</i>
<i>fadJ</i>	P77399.1	3334437	<i>Escherichia coli</i>
<i>Hbd2</i>	EDK34807.1	146348271	<i>Clostridium kluyveri</i>
<i>Hbd1</i>	EDK32512.1	146345976	<i>Clostridium kluyveri</i>
<i>hbd</i>	P52041.2		<i>Clostridium acetobutylicum</i>
<i>HSD17B10</i>	O02691.3	3183024	<i>Bos Taurus</i>
<i>phbB</i>	P23238.1	130017	<i>Zoogloea ramigera</i>
<i>phaB</i>	YP_353825.1	77464321	<i>Rhodobacter sphaeroides</i>

A number of similar enzymes have been found in other species of *Clostridia* and in *Metallosphaera sedula* (Berg et al., *Science* 318.5857:1782-1786 (2007)) as shown below:

<u>Protein</u>	<u>GenBank ID</u>	<u>GI number</u>	<u>Organism</u>
<i>Hbd</i>	NP_349314.1	NP_349314.1	<i>Clostridium acetobutylicum</i>
<i>Hbd</i>	AAM14586.1	AAM14586.1	<i>Clostridium beijerinckii</i>
<i>Msed_1423</i>	YP_001191505	YP_001191505	<i>Metallosphaera sedula</i>
<i>Msed_0399</i>	YP_001190500	YP_001190500	<i>Metallosphaera sedula</i>

<i>Msed_0389</i>	YP_001190490	YP_001190490	<i>Metallosphaera sedula</i>
<i>Msed_1993</i>	YP_001192057	YP_001192057	<i>Metallosphaera sedula</i>

Several acyl-CoA dehydrogenases are capable of reducing an acyl-CoA to its corresponding aldehyde and can be used for catalyzing the 3-hydroxybutyryl-CoA reductase (aldehyde forming) activity. Exemplary genes that encode such enzymes include the *Acinetobacter calcoaceticus* *acrI* encoding a fatty acyl-CoA reductase (Reiser and Somerville, *J. Bacteriol.* 179.9:2969-2975 (1997)), the *Acinetobacter* sp. *M-1* fatty acyl-CoA reductase Ishige et al. *Appl. Environ. Microbiol.* 68.3:192-195 (2002)), and a CoA- and NADP- dependent succinate semialdehyde dehydrogenase encoded by the *sucD* gene in *Clostridium kluyveri* (Sohling and Gottschalk, *J. Bacteriol.* 178.3:871-880 (1996); Sohling and Gottschalk *J. Bacteriol.* 178.3:871-880 (1996)). *SucD* of *P. gingivalis* is another succinate semialdehyde dehydrogenase (Takahashi et al., *J. Bacteriol.* 182.17:4704-4710 (2000)). The enzyme acylating acetaldehyde dehydrogenase in *Pseudomonas* sp, encoded by *bphG*, is yet another enzyme demonstrated to oxidize and acylate acetaldehyde, propionaldehyde, butyraldehyde, isobutyraldehyde and formaldehyde (Powlowski et al., *J. Bacteriol.* 175.2:377-385 (1993)). In addition to reducing acetyl-CoA to ethanol, the enzyme encoded by *adhE* in *Leuconostoc mesenteroides* has been shown to oxidize the branched chain compound isobutyraldehyde to isobutyryl-CoA (Kazahaya et al., *J. Gen. Appl. Microbiol.* 18:43-55 (1972); Koo et al., *Biotechnol. Lett.* 27.7:505-510 (2005)). Butyraldehyde dehydrogenase catalyzes a similar reaction, conversion of butyryl-CoA to butyraldehyde, in solventogenic organisms such as *Clostridium saccharoperbutylacetonicum* (Kosaka et al., *Biosci. Biotechnol. Biochem.* 71.1:58-68 (2007)). Information related to these genes and proteins are show below:

<u>Protein</u>	<u>GENBANK ID</u>	<u>GI NUMBER</u>	<u>ORGANISM</u>
<i>acrI</i>	YP_047869.1	50086359	<i>Acinetobacter calcoaceticus</i>
<i>acrI</i>	AAC45217	1684886	<i>Acinetobacter baylyi</i>
<i>acrI</i>	BAB85476.1	18857901	<i>Acinetobacter</i> sp. Strain M-1
<i>sucD</i>	P38947.1	172046062	<i>Clostridium kluyveri</i>
<i>sucD</i>	NP_904963.1	34540484	<i>Porphyromonas gingivalis</i>
<i>bphG</i>	BAA03892.1	425213	<i>Pseudomonas</i> sp
<i>adhE</i>	AAV66076.1	55818563	<i>Leuconostoc mesenteroides</i>
<i>bld</i>	AAP42563.1	31075383	<i>Clostridium saccharoperbutylacetonicum</i>

An additional enzyme type that converts an acyl-CoA to its corresponding aldehyde is malonyl-CoA reductase which transforms malonyl-CoA to malonic semialdehyde. Malonyl-CoA

reductase is a key enzyme in autotrophic carbon fixation via the 3-hydroxypropionate cycle in thermoacidophilic archaeal bacteria (Berg et al., *Science* 318.5857:1782-1786 (2007); Thauer, R.K., *Science* 318.5857:1732-1733). The enzyme utilizes NADPH as a cofactor and has been characterized in *Metallosphaera* and *Sulfolobus* spp (Alber et al., *J. Bacteriol.* 188.24: 8551-8559 (2006); Hugler et al. 2404-10). The enzyme is encoded by *Msed_0709* in *Metallosphaera sedula* ((Alber et al., *J. Bacteriol.* 188.24: 8551-8559 (2006); (Berg et al., *Science* 318.5857:1782-1786 (2007))). A gene encoding a malonyl-CoA reductase from *Sulfolobus tokodaii* was cloned and heterologously expressed in *E. coli* (Alber et al., *J. Bacteriol.* 188.24: 8551-8559 (2006); Alber et al., *Mol. Microbiol.* 61.2:297-309 (2006). This enzyme has also been shown to catalyze the conversion of methylmalonyl-CoA to its corresponding aldehyde. Although the aldehyde dehydrogenase functionality of these enzymes is similar to the bifunctional dehydrogenase from *Chloroflexus aurantiacus*, there is little sequence similarity. Both malonyl-CoA reductase enzyme candidates have high sequence similarity to aspartate-semialdehyde dehydrogenase, an enzyme catalyzing the reduction and concurrent dephosphorylation of aspartyl-4-phosphate to aspartate semialdehyde. Additional genes can be found by sequence homology to proteins in other organisms including *Sulfolobus solfataricus* and *Sulfolobus acidocaldarius* and have been listed below. Yet another gene for CoA-acylating aldehyde dehydrogenase is the *ald* gene from *Clostridium beijerinckii* (Toth et al., *Appl. Environ. Microbiol.* 65.11:4973-4980 (1999)). This enzyme has been reported to reduce acetyl-CoA and butyryl-CoA to their corresponding aldehydes. This gene is very similar to *eutE* that encodes acetaldehyde dehydrogenase of *Salmonella typhimurium* and *E. coli* (Toth et al., *Appl. Environ. Microbiol.* 65.11:4973-4980 (1999)). A summary of relevant gene and protein information is shown below:

PROTEIN	GENBANK ID	GI NUMBER	ORGANISM
<i>MSED_0709</i>	YP_001190808.1	146303492	<i>Metallosphaera sedula</i>
<i>mcr</i>	NP_378167.1	15922498	<i>Sulfolobus tokodaii</i>
<i>asd-2</i>	NP_343563.1	15898958	<i>Sulfolobus solfataricus</i>
<i>Saci_2370</i>	YP_256941.1	70608071	<i>Sulfolobus acidocaldarius</i>
<i>Ald</i>	AAT66436	9473535	<i>Clostridium beijerinckii</i>
<i>eutE</i>	AAA80209	687645	<i>Salmonella typhimurium</i>
<i>eutE</i>	P77445	2498347	<i>Escherichia coli</i>

Enzymes exhibiting 3-hydroxybutyraldehyde reductase activity (EC 1.1.1.61) have been characterized in *Ralstonia eutropha* (Bravo et al., *J. Forensic Sci.* 49.2:379-387 (2004)), *Clostridium kluyveri* (Wolff and Kenealy, *Protein Expr. Purif.* 6.2:206-212 (1995)) and

Arabidopsis thaliana (Breitkreuz et al., *J. Biol. Chem.* 278.42: 41552-41556 (2003)). Yet another gene is the alcohol dehydrogenase *adhI* from *Geobacillus thermoglucosidasius* (Jeon et al., *J. Biotechnol.* 135.2:127-133 (2008)). A summary of gene and protein information is shown below:

<u>PROTEIN</u>	<u>GENBANK ID</u>	<u>GI NUMBER</u>	<u>ORGANISM</u>
<i>4hbd</i>	YP_726053.1	113867564	<i>Ralstonia eutropha</i> H16
<i>4hbd</i>	L21902.1	146348486	<i>Clostridium kluyveri</i> DSM 555
<i>4hbd</i>	Q94B07	75249805	<i>Arabidopsis thaliana</i>
<i>adhI</i>	AAR91477.1	40795502	<i>Geobacillus thermoglucosidasius</i> M10EXG

5

Another exemplary enzyme is 3-hydroxyisobutyrate dehydrogenase which catalyzes the reversible oxidation of 3-hydroxyisobutyrate to methylmalonate semialdehyde. This enzyme participates in valine, leucine and isoleucine degradation and has been identified in bacteria, eukaryotes, and mammals. The enzyme encoded by *P84067* from *Thermus thermophilus* HB8 has been structurally characterized (Lokanath et al., *J. Mol. Biol.* 352.4:905-917 (2005)). The reversibility of the human 3-hydroxyisobutyrate dehydrogenase was demonstrated using isotopically-labeled substrate (Manning and Pollitt, *Biochem. J.* 231.2:481-484 (1985)).

10

Additional genes encoding this enzyme include *3hidh* in *Homo sapiens* (Hawes et al. *Methods Enzymol.* 324:218-228 (2000)) and *Oryctolagus cuniculus* (Hawes et al. *Methods Enzymol.*

15

324:218-228 (2000); Chowdhury et al., *Biosci. Biotechnol. Biochem.* 60.12:2043-2047 (1996)), *mmsb* in *Pseudomonas aeruginosa*, and *dhat* in *Pseudomonas putida* (Aberhart and Hsu, *J. Chem. Soc. (Perkin 1)* 6:1404-1406; Chowdhury et al., *Biosci. Biotechnol. Biochem.* 60.12:2043-2047 (1996)); Chowdhury et al., *Biosci. Biotechnol. Biochem.* 67.2:438-441 (2003)). Information related to these gene and proteins is shown below:

<u>PROTEIN</u>	<u>GENBANK ID</u>	<u>GI NUMBER</u>	<u>ORGANISM</u>
<i>P84067</i>	P84067	75345323	<i>Thermus thermophilus</i>
<i>mmsb</i>	P28811.1	127211	<i>Pseudomonas aeruginosa</i>
<i>dhat</i>	Q59477.1	2842618	<i>Pseudomonas putida</i>
<i>3hidh</i>	P31937.2	12643395	<i>Homo sapiens</i>
<i>3hidh</i>	P32185.1	416872	<i>Oryctolagus cuniculus</i>

20

Other exemplary genes encoding enzymes that catalyze the conversion of an aldehyde to alcohol (i.e., alcohol dehydrogenase or equivalently aldehyde reductase) and can serve as candidates for 3-hydroxybutyraldehyde reductase include *alrA* encoding a medium-chain alcohol dehydrogenase for C2-C14 (Tani et al., *Appl. Environ. Microbiol.* 66.12:5231-5335 (2000)),

ADH2 from *Saccharomyces cerevisiae* (Atsumi et al., *Nature* 451.7174:86-89 (2008)), *yqhD* from *E. coli* which has preference for molecules longer than C3 (Sulzenbacher et al., *J. Mol. Biol.* 342.2:489-502 (2004)), and *bdh* I and *bdh* II from *C. acetobutylicum* which converts butyraldehyde into butanol (Walter et al., *J. Bacteriol.* 174.22:7149;7158 (1992)). The gene product of *yqhD* catalyzes the reduction of acetaldehyde, malondialdehyde, propionaldehyde, butyraldehyde, and acrolein using NADPH as the cofactor (Perez et al., *J. Biol. Chem.* 283.12:7346-7353 (2008)). ADH1 from *Zymomonas mobilis* has been demonstrated to have activity on a number of aldehydes including formaldehyde, acetaldehyde, propionaldehyde, butyraldehyde, and acrolein (Kinoshita et al., *Appl. Microbiol. Biotechnol.* 22:249-254 (1985)).

- 10 The protein sequences for each of these exemplary gene products, if available, can be found using the following GenBank accession numbers:

Protein	GENBANK ID	GI NUMBER	ORGANISM
<i>alrA</i>	BAB12273.1	9967138	<i>Acinetobacter</i> sp. strain M-1
<i>ADH2</i>	NP_014032.1	6323961	<i>Saccharomyces cerevisiae</i>
<i>yqhD</i>	NP_417484.1	16130909	<i>Escherichia coli</i>
<i>bdh</i> I	NP_349892.1	15896543	<i>Clostridium acetobutylicum</i>
<i>bdh</i> II	NP_349891.1	15896542	<i>Clostridium acetobutylicum</i>
<i>adhA</i>	YP_162971.1	56552132	<i>Zymomonas mobilis</i>

The alcohol-forming 3-hydroxybutyryl-CoA reductase can be catalyzed by exemplary two-step oxidoreductases that convert an acyl-CoA to alcohol. These include enzymes that transform substrates such as acetyl-CoA to ethanol (e.g., *adhE* from *E. coli* (Kessler et al., *FEBS Lett.* 281.1-2:59-63 (1991)) and butyryl-CoA to butanol (e.g. *adhE2* from *C. acetobutylicum* (Fontaine et al., *J. Bacteriol.* 184.3:821-830 (2002)). In addition to reducing acetyl-CoA to ethanol, the enzyme encoded by *adhE* in *Leuconostoc mesenteroides* has been shown to oxidize the branched chain compound isobutyraldehyde to isobutyryl-CoA (Kazahaya et al., *J. Gen. App. Microbiol.* 18:43-55 (1972); Koo et al., *Biotechnol. Lett.* 27.7:505-510 (2005)). Relevant information to these genes and proteins are shown below:

Protein	GENBANK ID	GI NUMBER	ORGANISM
<i>adhE</i>	NP_415757.1	16129202	<i>Escherichia coli</i>
<i>adhE2</i>	AAK09379.1	12958626	<i>Clostridium acetobutylicum</i>
<i>adhE</i>	AAV66076.1	55818563	<i>Leuconostoc mesenteroides</i>

Another exemplary enzyme can convert malonyl-CoA to 3-HP. An NADPH-dependent enzyme with this activity has characterized in *Chloroflexus aurantiacus* where it participates in the 3-hydroxypropionate cycle (Hugler et al., *J. Bacteriol.* 184.9:2404-2410 (2002); Strauss and Fuchs, *Eur. J. Biochem.* 215.3:633-643 (1993)). This enzyme, with a mass of 300 kDa, is highly substrate-specific and shows little sequence similarity to other known oxidoreductases (Hugler et al., *J. Bacteriol.* 184.9:2404-2410 (2002)). No enzymes in other organisms have been shown to catalyze this specific reaction, however there is bioinformatic evidence that other organisms can have similar pathways (Klatt et al., *Environ. Microbiol.* 9.8:2067-2078 (2007)). Enzyme candidates in other organisms including *Roseiflexus castenholzii*, *Erythrobacter sp. NAP1* and marine gamma proteobacterium HTCC2080 can be inferred by sequence similarity. Information related to these genes and proteins is shown below:

Protein	GENBANK ID	GI NUMBER	ORGANISM
<i>mcr</i>	AAS20429.1	42561982	<i>Chloroflexus aurantiacus</i>
<i>Rcas_2929</i>	YP_001433009.1	156742880	<i>Roseiflexus castenholzii</i>
<i>NAP1_02720</i>	ZP_01039179.1	85708113	<i>Erythrobacter sp. NAP1</i>
MGP2080_00535	ZP_01626393.1	119504313	marine gamma proteobacterium HTCC2080

Longer chain acyl-CoA molecules can be reduced by enzymes such as the jojoba (*Simmondsia chinensis*) *FAR* which encodes an alcohol-forming fatty acyl-CoA reductase. Its overexpression in *E. coli* resulted in *FAR* activity and the accumulation of fatty alcohol (Metz et al., *Plant Physiol.* 122.3: 635-644 (2000)). Information related to *FAR* is shown below:

Protein	GENBANK ID	GI NUMBER	ORGANISM
<i>FAR</i>	AAD38039.1	5020215	<i>Simmondsia chinensis</i>

A second alternate pathway from acetoacetyl-CoA to 1,3-butanediol proceeds via the reduction of acetoacetyl-CoA into 3-oxobutyraldehyde via the CoA-dependent aldehyde forming acetoacetyl-CoA reductase. 3-oxobutyraldehyde is next reduced to 3-hydroxybutyraldehyde by 3-oxobutyraldehyde reductase (ketone reducing), and eventually, this intermediate is reduced to 1,3-butanediol by a 3-hydroxybutyraldehyde reductase. The enzymes and genes encoding these enzymes for each of these steps are listed below.

Exemplary candidates for acetoacetyl-CoA reductase (CoA-dependent, aldehyde forming) that catalyzes the transformation of acetoacetyl-CoA into 3-oxobutyraldehyde are the same as those described for 3-hydroxybutyryl-CoA reductase (aldehyde forming) described herein above.

There exist several exemplary alcohol dehydrogenases that convert a ketone to a hydroxyl functional group and can be used for catalyzing the 3-oxobutyraldehyde reductase_(ketone-reducing) activity. Two such enzymes from *E. coli* are encoded by malate dehydrogenase (*mdh*) and lactate dehydrogenase (*ldhA*). In addition, lactate dehydrogenase from *Ralstonia eutropha* has been shown to demonstrate high activities on substrates of various chain lengths such as lactate, 2-oxobutyrate, 2-oxopentanoate and 2-oxoglutarate (Steinbuchel and Schlegel, *Eur. J. Biochem.* 130.2:329-334 (1983)). Conversion of the oxo functionality to the hydroxyl group can also be catalyzed by 2-keto1,3-butanediol reductase, an enzyme reported to be found in rat and in human placenta (Suda et al., *Arch. Biochem. Biophys.* 176.2:610-620 (1976); Suda et al., *Biochem. Biophys. Res. Commun.* 342.2:586-591 (1977)). All of these enzymes can be use as a 3-oxobutyraldehyde reductase. An additional enzyme for this step is the mitochondrial 3-hydroxybutyrate dehydrogenase (*bdh*) from the human heart which has been cloned and characterized (Marks et al., *J. Biol. Chem.* 267.22:15459-15463 (1992)). This enzyme is a dehydrogenase that operates on a 3-hydroxyacid. Another exemplary alcohol dehydrogenase that converts acetone to isopropanol as was shown in *C. beijerinckii* (Ismaiel et al., *J. Bacteriol.* 175.16:5097-5105 (1993)) and *T. Brockii* (Lamed and Zeikus, *Biochem. J.* 195.1:183-190 (1981); Peretz and Burstein, *Biochemistry* 28.16:6549-6555 (1989)). Methyl ethyl ketone (MEK) reductase, or alternatively, 2-butanol dehydrogenase, catalyzes the reduction of MEK to form 2-butanol. Exemplary enzymes can be found in *Rhodococcus ruber* (Kosjek et al., *Biotechnol. Bioeng.* 86.1:55-62 (2004)) and *Pyrococcus furiosus* (van der Oost et al., *Eur. J. Biochem.* 268.10:3062-3068 (2001)). Information related to these proteins and genes is shown below:

<u>Protein</u>	<u>GENBANK ID</u>	<u>GINUMBER</u>	<u>ORGANISM</u>
<i>mdh</i>	AAC76268.1	1789632	<i>Escherichia coli</i>
<i>ldhA</i>	NP_415898.1	16129341	<i>Escherichia coli</i>
<i>ldh</i>	YP_725182.1	113866693	<i>Ralstonia eutropha</i>
<i>bdh</i>	AAA58352.1	177198	<i>Homo sapiens</i>
<i>adh</i>	AAA23199.2	60592974	<i>Clostridium beijerinckii</i>
<i>adh</i>	P14941.1	113443	<i>Thermoanaerobacter Brockii</i>
<i>sadh</i>	CAD36475	21615553	<i>Rhodococcus ruber</i>
<i>adhA</i>	3288810	AAC25556	<i>Pyrococcus furiosus</i>

Another pathway from acetoacetyl-CoA proceeds via its reduction to 4-hydroxy,2-butanone by the CoA-dependent, alcohol forming acetoacetyl-CoA reductase. This intermediate is then reduced to 1,3-butanediol by 4-hydroxybutanone reductase. 4-hydroxybutanone can also be

formed from 3-oxobutyraldehyde by an aldehyde reducing 3-oxobutyraldehyde reductase. Acetoacetyl-CoA reductase (CoA-dependent, alcohol forming) can utilize the same enzymes as those for the alcohol-forming 3-hydroxybutyryl-CoA reductase.

4-hydroxybutanone reductase activity can be obtained from the same genes as those described for 3-oxobutyraldehyde reductase. Additionally, a number of organisms can catalyze the reduction of 4-hydroxy,2-butanone to 1,3-butanediol, including those belonging to the genus *Bacillus*, *Brevibacterium*, *Candida*, and *Klebsiella* among others, as described by Matsuyama et al., US Patent No. 5,413,922.

Exemplary genes encoding enzymes that catalyze the conversion of an aldehyde to alcohol (i.e., alcohol dehydrogenase or equivalently aldehyde reductase) include *alrA* encoding a medium-chain alcohol dehydrogenase for C2-C14 (Tani et al., *App. Environ. Microbiol.* 66.12:5231-5235 (2000)), *ADH2* from *Saccharomyces cerevisiae* (Atsumi et al., *Nature* 451.7174:86-89 (2008)), *yqhD* from *E. coli* which has preference for molecules longer than C3 (Sulzenbacher et al., *J. Mol. Biol.* 342.2:489-502 (2004)), and *bdh* I and *bdh* II from *C. acetobutylicum* which converts butyraldehyde into butanol (Walter et al., *J. Bacteriol.* 174.22:7149-7158 (1992)). The gene product of *yqhD* catalyzes the reduction of acetaldehyde, malondialdehyde, propionaldehyde, butyraldehyde, and acrolein using NADPH as the cofactor (Perez et al., *J. Biol. Chem.* 283.12:7346-7353 (2008)). All of these genes can provide the 3-oxobutyraldehyde reductase (aldehyde reducing) activity for converting 3-oxobutyraldehyde into 4-hydroxybutanone. *ADH1* from *Zymomonas mobilis* has been demonstrated to have activity on a number of aldehydes including formaldehyde, acetaldehyde, propionaldehyde, butyraldehyde, and acrolein (Kinoshita et al., *App. Microbiol. Biotechnol.* 22:249-254 (1985)).

Where available, the protein sequences for each of these exemplary gene products, can be found using the following GenBank accession numbers:

Protein	GENBANK ID	GI NUMBER	ORGANISM
<i>alrA</i>	BAB12273.1	9967138	<i>Acinetobacter</i> sp. strain M-1
<i>ADH2</i>	NP_014032.1	6323961	<i>Saccharomyces cerevisiae</i>
<i>yqhD</i>	NP_417484.1	16130909	<i>Escherichia coli</i>
<i>bdh</i> I	NP_349892.1	15896543	<i>Clostridium acetobutylicum</i>
<i>bdh</i> II	NP_349891.1	15896542	<i>Clostridium acetobutylicum</i>
<i>adhA</i>	YP_162971.1	56552132	<i>Zymomonas mobilis</i>

The conversion of acetoacetyl-CoA to acetoacetate can be carried out by a acetoacetyl-CoA transferase which conserves the energy stored in the CoA-ester bond. Several exemplary transferase enzymes capable of catalyzing this transformation are provided below. These enzymes either naturally exhibit the desired acetoacetyl-CoA transferase activity or they can be engineered via directed evolution to accept acetetoacetyl-CoA as a substrate with increased efficiency. Such enzymes, either naturally or following directed evolution, are also suitable for catalyzing the conversion of 3-hydroxybutyryl-CoA to 3-hydroxybutyrate via a transferase mechanism.

Acetoacetyl-CoA:acetyl-CoA transferase naturally converts acetoacetyl-CoA and acetate to acetoacetate and acetyl-CoA. This enzyme can also accept 3-hydroxybutyryl-CoA as a substrate or could be engineered to do so. Exemplary enzymes include the gene products of *atoAD* from *E. coli* (Hanai et al., *Appl Environ Microbiol* 73:7814-7818 (2007)), *ctfAB* from *C. acetobutylicum* (Jojima et al., *Appl Microbiol Biotechnol* 77:1219-1224 (2008)), and *ctfAB* from *Clostridium saccharoperbutylacetonicum* (Kosaka et al., *Biosci. Biotechnol Biochem.* 71:58-68 (2007)). Information related to these proteins and genes is shown below:

<u>Protein</u>	<u>GENBANK ID</u>	<u>GI NUMBER</u>	<u>ORGANISM</u>
<i>AtoA</i>	P76459.1	2492994	<i>Escherichia coli</i>
<i>AtoD</i>	P76458.1	2492990	<i>Escherichia coli</i>
<i>CtfA</i>	NP_149326.1	15004866	<i>Clostridium acetobutylicum</i>
<i>CtfB</i>	NP_149327.1	15004867	<i>Clostridium acetobutylicum</i>
<i>CtfA</i>	AAP42564.1	31075384	<i>Clostridium saccharoperbutylacetonicum</i>
<i>CtfB</i>	AAP42565.1	31075385	<i>Clostridium saccharoperbutylacetonicum</i>

Succinyl-CoA:3-ketoacid-CoA transferase naturally converts succinate to succinyl-CoA while converting a 3-ketoacyl-CoA to a 3-ketoacid. Exemplary succinyl-CoA:3-ketoacid-CoA transferases are present in *Helicobacter pylori* (Corthesy-Theulaz et al., *J. Biol. Chem.* 272:25659-25667 (1997)), *Bacillus subtilis* (Stols et al., *Protein. Expr. Purif.* 53:396-403 (2007)), and *Homo sapiens* (Fukao et al., *Genomics* 68:144-151 (2000); Tanaka et al., *Mol. Hum. Reprod.* 8:16-23 (2002)). Information related to these proteins and genes is shown below:

<u>Protein</u>	<u>GENBANK ID</u>	<u>GI NUMBER</u>	<u>ORGANISM</u>
HPAG1_0676	YP_627417	108563101	<i>Helicobacter pylori</i>
HPAG1_0677	YP_627418	108563102	<i>Helicobacter pylori</i>
ScoA	NP_391778	16080950	<i>Bacillus subtilis</i>
ScoB	NP_391777	16080949	<i>Bacillus subtilis</i>
OXCT1	NP_000427	4557817	<i>Homo sapiens</i>
OXCT2	NP_071403	11545841	<i>Homo sapiens</i>

Additional suitable acetoacetyl-CoA and 3-hydroxybutyryl-CoA transferases are encoded by the gene products of *cat1*, *cat2*, and *cat3* of *Clostridium kluyveri*. These enzymes have been shown to exhibit succinyl-CoA, 4-hydroxybutyryl-CoA, and butyryl-CoA acetyltransferase activity, respectively (Seedorf et al., *Proc. Natl. Acad. Sci. USA* 105:2128-2133 (2008); Sohling and Gottschalk, *J. Bacteriol.* 178:871-880 (1996)). Similar CoA transferase activities are also present in *Trichomonas vaginalis* (van Grinsven et al., *J. Biol. Chem.* 283:1411-1418 (2008)) and *Trypanosoma brucei* (Riviere et al., *J. Biol. Chem.* 279:45337-45346 (2004)). Yet another transferase capable of the desired conversions is butyryl-CoA:acetoacetate CoA-transferase. Exemplary enzymes can be found in *Fusobacterium nucleatum* (Barker et al., *J. Bacteriol.* 152(1):201-7 (1982)), *Clostridium* SB4 (Barker et al., *J. Biol. Chem.* 253(4):1219-25 (1978)), and *Clostridium acetobutylicum* (Wiesenborn et al., *Appl. Environ. Microbiol.* 55(2):323-9 (1989)). Although specific gene sequences have not been indicated for butyryl-CoA:acetoacetate CoA-transferase, the genes FN0272 and FN0273 have been annotated as a butyrate-acetoacetate CoA-transferase (Kapatral et al., *J. Bact.* 184(7) 2005-2018 (2002)). Homologs in *Fusobacterium nucleatum* such as FN1857 and FN1856 can have the desired acetoacetyl-CoA transferase activity. FN1857 and FN1856 are located adjacent to many other genes involved in lysine fermentation and are thus very can encode an acetoacetate:butyrate CoA transferase (Kreimeyer, et al., *J. Biol. Chem.* 282 (10) 7191-7197 (2007)). Additional genes/gene products from *Porphyromonas gingivalis* and *Thermoanaerobacter tengcongensis* can be identified in a similar fashion (Kreimeyer, et al., *J. Biol. Chem.* 282 (10) 7191-7197 (2007)). Information related to these proteins and genes is shown below:

<u>Protein</u>	<u>GENBANK ID</u>	<u>GI NUMBER</u>	<u>ORGANISM</u>
<i>Cat1</i>	P38946.1	729048	<i>Clostridium kluyveri</i>
<i>Cat2</i>	P38942.2	1705614	<i>Clostridium kluyveri</i>
<i>Cat3</i>	EDK35586.1	146349050	<i>Clostridium kluyveri</i>
TVAG_395550	XP_001330176	123975034	<i>Trichomonas vaginalis</i> G3

<i>Tb11.02.0290</i>	XP_828352	71754875	<i>Trypanosoma brucei</i>
<i>FN0272</i>	NP_603179.1	19703617	<i>Fusobacterium nucleatum</i>
<i>FN0273</i>	NP_603180.1	19703618	<i>Fusobacterium nucleatum</i>
<i>FN1857</i>	NP_602657.1	19705162	<i>Fusobacterium nucleatum</i>
<i>FN1856</i>	NP_602656.1	19705161	<i>Fusobacterium nucleatum</i>
<i>PG1066</i>	NP_905281.1	34540802	<i>Porphyromonas gingivalis</i> W83
<i>PG1075</i>	NP_905290.1	34540811	<i>Porphyromonas gingivalis</i> W83
<i>TTE0720</i>	NP_622378.1	20807207	<i>Thermoanaerobacter tengcongensis</i> MB4
<i>TTE0721</i>	NP_622379.1	20807208	<i>Thermoanaerobacter tengcongensis</i> MB4

Acetoacetyl-CoA can be hydrolyzed to acetoacetate by acetoacetyl-CoA hydrolase. Similarly, 3-hydroxybutyryl-CoA can be hydrolyzed to 3-hydroxybutyrate by 3-hydroxybutyryl-CoA hydrolase. Many CoA hydrolases (EC 3.1.2.1) have broad substrate specificity and are suitable enzymes for these transformations either naturally or following enzyme engineering. Though the sequences were not reported, several acetoacetyl-CoA hydrolases were identified in the cytosol and mitochondrion of the rat liver (Aragon and Lowenstein, *J. Biol. Chem.* 258(8):4725-4733 (1983)). Additionally, an enzyme from *Rattus norvegicus* brain (Robinson et al., *Biochem. Biophys. Res. Commun.* 71:959-965 (1976)) can react with butyryl-CoA, hexanoyl-CoA and malonyl-CoA. The *acot12* enzyme from the rat liver was shown to hydrolyze C2 to C6 acyl-CoA molecules (Suematsu et al., *Eur. J. Biochem.* 268:2700-2709 (2001)). Though its sequence has not been reported, the enzyme from the mitochondrion of the pea leaf showed activity on acetyl-CoA, propionyl-CoA, butyryl-CoA, palmitoyl-CoA, oleoyl-CoA, succinyl-CoA, and crotonyl-CoA (Zeihner and Randall, *Plant. Physiol.* 94:20-27 (1990)). Additionally, a glutamate CoA-transferase from *Acidaminococcus fermentans* was transformed by site-directed mutagenesis into an acyl-CoA hydrolase with activity on glutaryl-CoA, acetyl-CoA and 3-butenoyl-CoA (Mack and Buckel, *FEBS Lett.* 405:209-212 (1997)). This indicates that the enzymes encoding succinyl-CoA:3-ketoacid-CoA transferases and acetoacetyl-CoA:acetyl-CoA transferases can also be used as hydrolases with certain mutations to change their function. The acetyl-CoA hydrolase, *ACH1*, from *S. cerevisiae* represents another candidate hydrolase (Buu et al., *J. Biol. Chem.* 278:17203-17209 (2003)). Information related to these proteins and genes is shown below:

<u>Protein</u>	<u>GENBANK ID</u>	<u>GI NUMBER</u>	<u>ORGANISM</u>
<i>Acot12</i>	NP_570103.1	18543355	<i>Rattus norvegicus</i>
<i>GctA</i>	CAA57199	559392	<i>Acidaminococcus fermentans</i>
<i>GctB</i>	CAA57200	559393	<i>Acidaminococcus fermentans</i>
<i>ACH1</i>	NP_009538	6319456	<i>Saccharomyces cerevisiae</i>

Another hydrolase enzyme is the human dicarboxylic acid thioesterase, *acot8*, which exhibits activity on glutaryl-CoA, adipyl-CoA, suberyl-CoA, sebacyl-CoA, and dodecanedioyl-CoA (Westin et al., *J. Biol. Chem.* 280:38125-38132 (2005)) and the closest *E. coli* homolog, *tesB*, which can also hydrolyze a broad range of CoA thioesters (Naggert et al., *J. Biol. Chem.* 266:11044-11050 (1991)) including 3-hydroxybutyryl-CoA (Tseng et al., *Appl. Environ. Microbiol.* 75(10):3137-3145 (2009)). A similar enzyme has also been characterized in the rat liver (Deana, *Biochem. Int.* 26:767-773 (1992)). Other *E. coli* thioester hydrolases include the gene products of *tesA* (Bonner and Bloch, *J. Biol. Chem.* 247:3123-3133 (1972)), *ybgC* (Kuznetsova et al., *FEMS Microbiol. Rev.* 29:263-279 (2005); Zhuang et al., *FEBS Lett.* 516:161-163 (2002)), *paal* (Song et al., *J. Biol. Chem.* 281:11028-11038 (2006)), and *ybdB* (Leduc et al., *J. Bacteriol.* 189:7112-7126 (2007)). Information related to these proteins and genes is shown below:

<u>Protein</u>	<u>GENBANK ID</u>	<u>GI NUMBER</u>	<u>ORGANISM</u>
<i>Acot8</i>	CAA15502	3191970	<i>Homo sapiens</i>
<i>TesB</i>	NP_414986	16128437	<i>Escherichia coli</i>
<i>Acot8</i>	NP_570112	51036669	<i>Rattus norvegicus</i>
<i>TesA</i>	NP_415027	16128478	<i>Escherichia coli</i>
<i>YbgC</i>	NP_415264	16128711	<i>Escherichia coli</i>
<i>Paal</i>	NP_415914	16129357	<i>Escherichia coli</i>
<i>YbdB</i>	NP_415129	16128580	<i>Escherichia coli</i>

15

Additional hydrolase enzymes include 3-hydroxyisobutyryl-CoA hydrolase which has been described to efficiently catalyze the conversion of 3-hydroxyisobutyryl-CoA to 3-hydroxyisobutyrate during valine degradation (Shimomura et al., *J. Biol. Chem.* 269:14248-14253 (1994)). Genes encoding this enzyme include *hibch* of *Rattus norvegicus* (Shimomura et al., *supra* (1994); Shimomura et al., *Methods Enzymol.* 324:229-240 (2000)) and *Homo sapiens* (Shimomura et al., *supra* (1994)). Candidate genes by sequence homology include *hibch* of *Saccharomyces cerevisiae* and *BC_2292* of *Bacillus cereus*. *BC_2292* was shown to demonstrate 3-hydroxybutyryl-CoA hydrolase activity and function as part of a pathway for 3-

20

hydroxybutyrate synthesis when engineered into *Escherichia coli* (Lee et al., *Appl. Microbiol. Biotechnol.* 79:633-641 (2008)). Information related to these proteins and genes is shown below:

<u>Protein</u>	<u>GENBANK ID</u>	<u>GI NUMBER</u>	<u>ORGANISM</u>
<i>Hibch</i>	Q5XIE6.2	146324906	<i>Rattus norvegicus</i>
<i>Hibch</i>	Q6NVY1.2	146324905	<i>Homo sapiens</i>
<i>Hibch</i>	P28817.2	2506374	<i>Saccharomyces cerevisiae</i>
<i>BC_2292</i>	AP09256	29895975	<i>Bacillus cereus</i> ATCC 14579

- 5 An alternative method for removing the CoA moiety from acetoacetyl-CoA or 3-hydroxybutyryl-CoA is to apply a pair of enzymes such as a phosphate-transferring acyltransferase and a kinase to impart acetoacetyl-CoA or 3-hydroxybutyryl-CoA synthetase activity. As used herein, the combination of a phosphotransacylase and a kinase enzyme is referred to as a “synthetase.” This activity allows the net hydrolysis of the CoA-ester of either
- 10 molecule with the simultaneous generation of ATP. For example, the butyrate kinase (*buk*)/phosphotransbutyrylase (*ptb*) system from *Clostridium acetobutylicum* has been successfully applied to remove the CoA group from 3-hydroxybutyryl-CoA when functioning as part of a pathway for 3-hydroxybutyrate synthesis (Tseng et al., *Appl. Environ. Microbiol.* 75(10):3137-3145 (2009)). Specifically, the *ptb* gene from *C. acetobutylicum* encodes an
- 15 enzyme that can convert an acyl-CoA into an acyl-phosphate (Walter et al. *Gene* 134(1): p. 107-11 (1993)); Huang et al. *J Mol Microbiol Biotechnol* 2(1): p. 33-38 (2000). Additional *ptb* genes can be found in butyrate-producing bacterium L2-50 (Louis et al. *J.Bacteriol.* 186:2099-2106 (2004)) and *Bacillus megaterium* (Vazquez et al. *Curr.Microbiol* 42:345-349 (2001)).
- Additional exemplary phosphate-transferring acyltransferases include phosphotransacetylase, encoded by *pta*. The *pta* gene from *E. coli* encodes an enzyme that can convert acetyl-CoA into acetyl-phosphate, and vice versa (Suzuki, T. *Biochim.Biophys.Acta* 191:559-569 (1969)). This enzyme can also utilize propionyl-CoA instead of acetyl-CoA forming propionate in the process (Hesslinger et al. *Mol.Microbiol* 27:477-492 (1998)). Information related to these proteins and genes is shown below:
- 20

<u>Protein</u>	<u>GENBANK ID</u>	<u>GI NUMBER</u>	<u>ORGANISM</u>
<i>Pta</i>	NP_416800.1	16130232	<i>Escherichia coli</i>
<i>Ptb</i>	NP_349676	15896327	<i>Clostridium acetobutylicum</i>
<i>Ptb</i>	AAR19757.1	38425288	butyrate-producing bacterium L2-50
<i>Ptb</i>	CAC07932.1	10046659	<i>Bacillus megaterium</i>

Exemplary kinases include the *E. coli* acetate kinase, encoded by *ackA* (Skarstedt and Silverstein *J.Biol.Chem.* 251:6775-6783 (1976)), the *C. acetobutylicum* butyrate kinases, encoded by *buk1* and *buk2* ((Walter et al. *Gene* 134(1):107-111 (1993); Huang et al. *J Mol Microbiol Biotechnol* 2(1):33-38 (2000)), and the *E. coli* gamma-glutamyl kinase, encoded by *proB* (Smith et al. *J.Bacteriol.* 157:545-551 (1984)). These enzymes phosphorylate acetate, butyrate, and glutamate, respectively. The *ackA* gene product from *E. coli* also phosphorylates propionate (Hesslinger et al. *Mol.Microbiol* 27:477-492 (1998)). Information related to these proteins and genes is shown below:

<u>Protein</u>	<u>GENBANK ID</u>	<u>GI NUMBER</u>	<u>ORGANISM</u>
<i>AckA</i>	NP_416799.1	16130231	<i>Escherichia coli</i>
<i>Buk1</i>	NP_349675	15896326	<i>Clostridium acetobutylicum</i>
<i>Buk2</i>	Q97III	20137415	<i>Clostridium acetobutylicum</i>
<i>ProB</i>	NP_414777.1	16128228	<i>Escherichia coli</i>

The hydrolysis of acetoacetyl-CoA or 3-hydroxybutyryl-CoA can alternatively be carried out by a single enzyme or enzyme complex that exhibits acetoacetyl-CoA or 3-hydroxybutyryl-CoA synthetase activity. This activity enables the net hydrolysis of the CoA-ester of either molecule, and in some cases, results in the simultaneous generation of ATP. For example, the product of the *LSC1* and *LSC2* genes of *S. cerevisiae* and the *sucC* and *sucD* genes of *E. coli* naturally form a succinyl-CoA synthetase complex that catalyzes the formation of succinyl-CoA from succinate with the concomitant consumption of one ATP, a reaction which is reversible in vivo (Gruys et al., US Patent No. 5,958,745, filed September 28, 1999). Information related to these proteins and genes is shown below:

<u>Protein</u>	<u>GENBANK ID</u>	<u>GI NUMBER</u>	<u>ORGANISM</u>
<i>SucC</i>	NP_415256.1	16128703	<i>Escherichia coli</i>
<i>SucD</i>	AAC73823.1	1786949	<i>Escherichia coli</i>
<i>LSC1</i>	NP_014785	6324716	<i>Saccharomyces cerevisiae</i>
<i>LSC2</i>	NP_011760	6321683	<i>Saccharomyces cerevisiae</i>

Additional exemplary CoA-ligases include the rat dicarboxylate-CoA ligase for which the sequence is yet uncharacterized (Vamecq et al., *Biochemical J.* 230:683-693 (1985)), either of the two characterized phenylacetate-CoA ligases from *P. chrysogenum* (Lamas-Maceiras et al., *Biochem. J.* 395:147-155 (2005); Wang et al., *Biochem Biophys Res Commun* 360(2):453-458 (2007)), the phenylacetate-CoA ligase from *Pseudomonas putida* (Martinez-Blanco et al., *J. Biol. Chem.* 265:7084-7090 (1990)), and the 6-carboxyhexanoate-CoA ligase from *Bacillus*

subtilis (Bower et al., *J. Bacteriol.* 178(14):4122-4130 (1996)). Additional candidate enzymes are acetoacetyl-CoA synthetases from *Mus musculus* (Hasegawa et al., *Biochim. Biophys. Acta* 1779:414-419 (2008)) and *Homo sapiens* (Ohgami et al., *Biochem. Pharmacol.* 65:989-994 (2003)), which naturally catalyze the ATP-dependant conversion of acetoacetate into acetoacetyl-CoA. 4-Hydroxybutyryl-CoA synthetase activity has been demonstrated in *Metallosphaera sedula* (Berg et al., *Science* 318:1782-1786 (2007)). This function has been assigned to the Msed_1422 gene. Information related to these proteins and genes is shown below:

<u>Protein</u>	<u>GENBANK ID</u>	<u>GI NUMBER</u>	<u>ORGANISM</u>
<i>Phl</i>	CAJ15517.1	77019264	<i>Penicillium chrysogenum</i>
<i>PhlB</i>	ABS19624.1	152002983	<i>Penicillium chrysogenum</i>
<i>PaaF</i>	AAC24333.2	22711873	<i>Pseudomonas putida</i>
<i>BioW</i>	NP_390902.2	50812281	<i>Bacillus subtilis</i>
<i>AACS</i>	NP_084486.1	21313520	<i>Mus musculus</i>
<i>AACS</i>	NP_076417.2	31982927	<i>Homo sapiens</i>
<i>Msed_1422</i>	YP_001191504	146304188	<i>Metallosphaera sedula</i>

ADP-forming acetyl-CoA synthetase (ACD, EC 6.2.1.13) is another enzyme that couples the conversion of acyl-CoA esters to their corresponding acids with the concurrent synthesis of ATP. Several enzymes with broad substrate specificities have been described in the literature. ACD I from *Archaeoglobus fulgidus*, encoded by AF1211, was shown to operate on a variety of linear and branched-chain substrates including acetyl-CoA, propionyl-CoA, butyryl-CoA, acetate, propionate, butyrate, isobutyryate, isovalerate, succinate, fumarate, phenylacetate, indoleacetate (Musfeldt et al., *J. Bacteriol.* 184:636-644 (2002)). The enzyme from *Haloarcula marismortui* (annotated as a succinyl-CoA synthetase) accepts propionate, butyrate, and branched-chain acids (isovalerate and isobutyrate) as substrates, and was shown to operate in the forward and reverse directions (Brasen et al., *Arch. Microbiol.* 182:277-287 (2004)). The ACD encoded by PAE3250 from hyperthermophilic crenarchaeon *Pyrobaculum aerophilum* showed the broadest substrate range of all characterized ACDs, reacting with acetyl-CoA, isobutyryl-CoA (preferred substrate) and phenylacetyl-CoA (Brasen et al., *supra* (2004)). The enzymes from *A. fulgidus*, *H. marismortui* and *P. aerophilum* have all been cloned, functionally expressed, and characterized in *E. coli* (Musfeldt et al., *supra*; Brasen et al., *supra* (2004)). Information related to these proteins and genes is shown below:

<u>Protein</u>	<u>GENBANK ID</u>	<u>GI NUMBER</u>	<u>ORGANISM</u>
AF1211	NP_070039.1	11498810	<i>Archaeoglobus fulgidus</i> DSM 4304
scs	YP_135572.1	55377722	<i>Haloarcula marismortui</i> ATCC 43049
PAE3250	NP_560604.1	18313937	<i>Pyrobaculum aerophilum</i> str. IM2

The conversion of 3-hydroxybutyrate to 3-hydroxybutyraldehyde can be carried out by a 3-hydroxybutyrate reductase. Similarly, the conversion of acetoacetate to acetoacetaldehyde can be carried out by an acetoacetate reductase. A suitable enzyme for these transformations is the aryl-aldehyde dehydrogenase, or equivalently a carboxylic acid reductase, from *Nocardia iowensis*. Carboxylic acid reductase catalyzes the magnesium, ATP and NADPH-dependent reduction of carboxylic acids to their corresponding aldehydes (Venkitasubramanian et al., *J. Biol. Chem.* 282:478-485 (2007)). This enzyme, encoded by *car*, was cloned and functionally expressed in *E. coli* (Venkitasubramanian et al., *J. Biol. Chem.* 282:478-485 (2007)).

Expression of the *npt* gene product improved activity of the enzyme via post-transcriptional modification. The *npt* gene encodes a specific phosphopantetheine transferase (PPTase) that converts the inactive apo-enzyme to the active holo-enzyme. The natural substrate of this enzyme is vanillic acid, and the enzyme exhibits broad acceptance of aromatic and aliphatic substrates (Venkitasubramanian et al., in *Biocatalysis in the Pharmaceutical and Biotechnology Industries*, ed. R.N. Patel, Chapter 15, pp. 425-440, CRC Press LLC, Boca Raton, FL. (2006)).

Information related to these proteins and genes is shown below:

<u>Protein</u>	<u>GI NUMBER</u>	<u>GENBANK ID</u>	<u>ORGANISM</u>
<i>Car</i>	40796035	AAR91681.1	<i>Nocardia iowensis</i> (sp. NRRL 5646)
<i>Npt</i>	114848891	ABI83656.1	<i>Nocardia iowensis</i> (sp. NRRL 5646)

Additional *car* and *npt* genes can be identified based on sequence homology.

<u>Protein</u>	<u>GENBANK ID</u>	<u>GI NUMBER</u>	<u>ORGANISM</u>
<i>fadD9</i>	YP_978699.1	121638475	<i>Mycobacterium bovis</i> BCG
<i>BCG_2812c</i>	YP_978898.1	121638674	<i>Mycobacterium bovis</i> BCG
<i>nfa20150</i>	YP_118225.1	54023983	<i>Nocardia farcinica</i> IFM 10152
<i>nfa40540</i>	YP_120266.1	54026024	<i>Nocardia farcinica</i> IFM 10152
<i>SGR_6790</i>	YP_001828302.1	182440583	<i>Streptomyces griseus</i> subsp. <i>griseus</i> NBRC 13350
<i>SGR_665</i>	YP_001822177.1	182434458	<i>Streptomyces griseus</i> subsp. <i>griseus</i> NBRC 13350

<i>MSMEG_2956</i>	YP_887275.1	118473501	<i>Mycobacterium smegmatis</i> MC2 155
<i>MSMEG_5739</i>	YP_889972.1	118469671	<i>Mycobacterium smegmatis</i> MC2 155
<i>MSMEG_2648</i>	YP_886985.1	118471293	<i>Mycobacterium smegmatis</i> MC2 155
<i>MAP1040c</i>	NP_959974.1	41407138	<i>Mycobacterium avium</i> subsp. <i>paratuberculosis</i> K-10
<i>MAP2899c</i>	NP_961833.1	41408997	<i>Mycobacterium avium</i> subsp. <i>paratuberculosis</i> K-10
<i>MMAR_2117</i>	YP_001850422.1	183982131	<i>Mycobacterium marinum</i> M
<i>MMAR_2936</i>	YP_001851230.1	183982939	<i>Mycobacterium marinum</i> M
<i>MMAR_1916</i>	YP_001850220.1	183981929	<i>Mycobacterium marinum</i> M
<i>TpauDRAFT_33060</i>	ZP_04027864.1	227980601	<i>Tsukamurella paurometabola</i> DSM 20162
<i>TpauDRAFT_20920</i>	ZP_04026660.1	227979396	<i>Tsukamurella paurometabola</i> DSM 20162
<i>CPCC7001_1320</i>	ZP_05045132.1	254431429	<i>Cyanobium</i> PCC7001
<i>DDBDRAFT_0187729</i>	XP_636931.1	66806417	<i>Dictyostelium discoideum</i> AX4

An additional enzyme candidate found in *Streptomyces griseus* is encoded by the *griC* and *griD* genes. This enzyme is believed to convert 3-amino-4-hydroxybenzoic acid to 3-amino-4-hydroxybenzaldehyde as deletion of either *griC* or *griD* led to accumulation of extracellular 3-acetyl-5-amino-4-hydroxybenzoic acid, a shunt product of 3-amino-4-hydroxybenzoic acid metabolism (Suzuki, et al., *J. Antibiot.* 60(6):380-387 (2007)). Co-expression of *griC* and *griD* with SGR_665, an enzyme similar in sequence to the *Nocardia iowensis* *npt*, can be beneficial. Information related to these proteins and genes is shown below:

<u>Protein</u>	<u>GI NUMBER</u>	<u>GENBANK ID</u>	<u>ORGANISM</u>
<i>griC</i>	182438036	YP_001825755.1	<i>Streptomyces griseus</i> subsp. <i>griseus</i> NBRC 13350
<i>griD</i>	182438037	YP_001825756.1	<i>Streptomyces griseus</i> subsp. <i>griseus</i> NBRC 13350

- 10 An enzyme with similar characteristics, alpha-aminoadipate reductase (AAR, EC 1.2.1.31), participates in lysine biosynthesis pathways in some fungal species. This enzyme naturally reduces alpha-aminoadipate to alpha-aminoadipate semialdehyde. The carboxyl group is first activated through the ATP-dependent formation of an adenylate that is then reduced by NAD(P)H to yield the aldehyde and AMP. Like CAR, this enzyme utilizes magnesium and
- 15 requires activation by a PPTase. Enzyme candidates for AAR and its corresponding PPTase are found in *Saccharomyces cerevisiae* (Morris et al., *Gene* 98:141-145 (1991)), *Candida albicans* (Guo et al., *Mol. Genet. Genomics* 269:271-279 (2003)), and *Schizosaccharomyces pombe* (Ford et al., *Curr. Genet.* 28:131-137 (1995)). The AAR from *S. pombe* exhibited significant activity

when expressed in *E. coli* (Guo et al., *Yeast* 21:1279-1288 (2004)). The AAR from *Penicillium chrysogenum* accepts S-carboxymethyl-L-cysteine as an alternate substrate, but did not react with adipate, L-glutamate or diaminopimelate (Hijarrubia et al., *J. Biol. Chem.* 278:8250-8256 (2003)). The gene encoding the *P. chrysogenum* PPTase has not been identified to date.

5 Information related to these proteins and genes is shown below:

<u>Protein</u>	<u>GI NUMBER</u>	<u>GENBANK ID</u>	<u>ORGANISM</u>
<i>LYS2</i>	171867	AAA34747.1	<i>Saccharomyces cerevisiae</i>
<i>LYS5</i>	1708896	P50113.1	<i>Saccharomyces cerevisiae</i>
<i>LYS2</i>	2853226	AAC02241.1	<i>Candida albicans</i>
<i>LYS5</i>	28136195	AAO26020.1	<i>Candida albicans</i>
<i>Lys1p</i>	13124791	P40976.3	<i>Schizosaccharomyces pombe</i>
<i>Lys7p</i>	1723561	Q10474.1	<i>Schizosaccharomyces pombe</i>
<i>Lys2</i>	3282044	CAA74300.1	<i>Penicillium chrysogenum</i>

Essentially any of these CAR or CAR-like enzymes can exhibit 3-hydroxybutyrate or acetoacetate reductase activity or can be engineered to do so.

The requisite 3-hydroxybutyrate dehydrogenase catalyzes the reduction of acetoacetate to form 3-hydroxybutyrate. Exemplary enzymes can be found in *Rhodococcus ruber* (Kosjek et al., *Biotechnol Bioeng.* 86:55-62 (2004)) and *Pyrococcus furiosus* (van der et al., *Eur.J.Biochem.* 268:3062-3068 (2001)). Additional secondary alcohol dehydrogenase enzymes capable of this transformation include *adh* from *C. beijerinckii* (Hanai et al., *Appl Environ Microbiol* 73:7814-7818 (2007); Jojima et al., *Appl Microbiol Biotechnol* 77:1219-1224 (2008)) and *adh* from *Thermoanaerobacter brockii* (Hanai et al., *Appl Environ Microbiol* 73:7814-7818 (2007); Peretz et al., *Anaerobe* 3:259-270 (1997)). The cloning of the *bdhA* gene from *Rhizobium* (*Sinorhizobium*) *Meliloti* into *E. coli* conferred the ability to utilize 3-hydroxybutyrate as a carbon source (Aneja and Charles, *J. Bacteriol.* 181(3):849-857 (1999)). Additional 3-hydroxybutyrate dehydrogenase can be found in *Pseudomonas fragi* (Ito et al., *J. Mol. Biol.* 355(4) 722-733 (2006)) and *Ralstonia pickettii* (Takanashi et al., *Antonie van Leeuwenhoek*, 95(3):249-262 (2009)). Information related to these proteins and genes is shown below:

<u>Protein</u>	<u>GENBANK ID</u>	<u>GI NUMBER</u>	<u>ORGANISM</u>
<i>Sadh</i>	CAD36475	21615553	<i>Rhodococcus rubber</i>
<i>AdhA</i>	AAC25556	3288810	<i>Pyrococcus furiosus</i>
<i>Adh</i>	P14941.1	113443	<i>Thermoanaerobacter brockii</i>
<i>Adh</i>	AAA23199.2	60592974	<i>Clostridium beijerinckii</i>

<i>BdhA</i>	NP_437676.1	16264884	<i>Rhizobium (Sinorhizobium) Meliloti</i>
<i>PRK13394</i>	BAD86668.1	57506672	<i>Pseudomonas fragi</i>
<i>Bdh1</i>	BAE72684.1	84570594	<i>Ralstonia pickettii</i>
<i>Bdh2</i>	BAE72685.1	84570596	<i>Ralstonia pickettii</i>
<i>Bdh3</i>	BAF91602.1	158937170	<i>Ralstonia pickettii</i>

Engineering the capability to convert CO₂, CO, and/or H₂ into acetyl-CoA, the central metabolite from which all cell mass components and many valuable products can be derived, into a foreign host such as *E. coli* can be accomplished following the expression of exogenous genes that encode various proteins of the Wood-Ljungdahl pathway. This pathway is highly active in acetogenic organisms such as *Moorella thermoacetica* (formerly, *Clostridium thermoaceticum*), which has been the model organism for elucidating the Wood-Ljungdahl pathway since its isolation in 1942 (Fontaine et al., *J. Bacteriol.* 43.6:701-715 (1942)). The Wood-Ljungdahl pathway comprises of two branches: the Eastern (or methyl) branch that enables the conversion of CO₂ to methyltetrahydrofolate (Me-THF) and the Western (or carbonyl) branch that enables the conversion of methyl-THF, CO, and Coenzyme-A into acetyl-CoA (Figure 5). Herein we describe a non-naturally occurring microorganism expressing genes encoding enzymes that catalyze the methyl and carbonyl branches of the Wood-Ljungdahl pathway. Such an organism is capable of converting CO, CO₂, and/or H₂ into acetyl-CoA, cell mass, and products.

In some embodiments, a non-naturally occurring organism utilizing pathways shown in Figure 5 exhibit three capabilities: 1) a functional methyl branch of the Wood-Ljungdahl pathway which enables the conversion of THF and CO₂ to 5-methyl-tetrahydrofolate, 2) the ability to combine CO, Coenzyme A, and the methyl group of Me-THF to form acetyl-CoA, and 3) the ability to synthesize 1,3-butanediol from acetyl-CoA.

Such non-naturally occurring organisms are able to 'fix' carbon from exogenous CO and/or exogenous or endogenously produced CO₂ to synthesize acetyl-CoA, cell mass, and products. A host organism engineered with these capabilities that also naturally possesses the capability for anaplerosis (e.g., *E. coli*) can grow on the syngas-generated acetyl-CoA in the presence of a suitable external electron acceptor such as nitrate. This electron acceptor is required to accept electrons from the reduced quinone formed via succinate dehydrogenase. A further advantage of adding an external electron acceptor is that additional energy for cell growth, maintenance, and

product formation can be generated from respiration of acetyl-CoA. An alternative strategy involves engineering a pyruvate ferredoxin oxidoreductase (PFOR) enzyme or other enzymes that convert pyruvate into acetyl-CoA into the strain to enable synthesis of biomass precursors in the absence of an external electron acceptor. A further characteristic of the engineered organism

5 is the capability for extracting reducing equivalents from molecular hydrogen. This enables a high yield of reduced products such as ethanol, butanol, isobutanol, isopropanol, 1,4-butanediol, 1,3-butanediol, succinic acid, fumaric acid, malic acid, 4-hydroxybutyric acid, 3-hydroxypropionic acid, lactic acid, adipic acid, methacrylic acid, and acrylic acid.

A non-naturally occurring organism can produce acetyl-CoA, cell mass, and targeted chemicals, more specifically 1,3-butanediol, from: 1) CO, 2) CO₂ and H₂, 3) CO, CO₂, and H₂, 4) synthesis gas comprising CO and H₂, 5) synthesis gas comprising CO, CO₂, and H₂, and 6) one or more carbohydrates. Exemplary carbohydrates include, but are not limited to, glucose, sucrose, xylose, arabinose, and glycerol.

The enzymes used for the conversion of CO, CO₂, and/or H₂ to acetyl-CoA are shown in Figure

15 5. To engineer a production host for the utilization of CO, CO₂, and/or H₂, one or more exogenous DNA sequence(s) encoding these enzymes can be expressed in the microorganism. Referring now to Figure 5, described below are enzymes that can be incorporated to utilize CO, CO₂, and/or H₂.

Formate dehydrogenase is a two subunit selenocysteine-containing protein that catalyzes the incorporation of CO₂ into formate in *Moorella thermoacetica* (Andreesen and Ljungdahl, *J. Bacteriol.* 116.2:867-873 (1973); Li et al., *J. Bacteriol.* 92.2:405-412 (1966); Yamamoto et al., *J. Biol. Chem.* 258.3:1826-1832 (1983)). The loci, Moth_2312 and Moth_2313 are actually one gene that is responsible for encoding the alpha subunit of formate dehydrogenase while the beta subunit is encoded by Moth_2314 (Pierce et al. *Environ. Microbiol.* 10:2550-2573 (2008)).

25 Another set of genes encoding formate dehydrogenase activity with a propensity for CO₂ reduction is encoded by Sfum_2703 through Sfum_2706 in *Syntrophobacter fumaroxidans* (Reda et al., *Proc. Natl. Acad. Sci. U.S.A.* 105.31:10654-10658 (2008); de Bok et al., *Eur. J. Biochem.* 270. 11: 2476-2485 (2003)). Similar to their *M. thermoacetica* counterparts, Sfum_2705 and Sfum_2706 are actually one gene. A similar set of genes presumed to carry out

30 the same function are encoded by CHY_0731, CHY_0732, and CHY_0733 in *C. hydrogenoformans* (Wu et al., *PLoS Genet.* 1.5:e65 (2005)). Relevant gene and protein information can be obtained from the information shown below:

<u>Protein</u>	<u>GenBank ID</u>	<u>GI number</u>	<u>Organism</u>
Moth_2312	YP_431142	148283121	<i>Moorella thermoacetica</i>
Moth_2313	YP_431143		<i>Moorella thermoacetica</i>
Moth_2314	YP_431144	83591135	<i>Moorella thermoacetica</i>
Sfum_2703	YP_846816.1	116750129	<i>Syntrophobacter fumaroxidans</i>
Sfum_2704	YP_846817.1	116750130	<i>Syntrophobacter fumaroxidans</i>
Sfum_2705	YP_846818.1	116750131	<i>Syntrophobacter fumaroxidans</i>
Sfum_2706	YP_846819.1	116750132	<i>Syntrophobacter fumaroxidans</i>
CHY_0731	YP_359585.1	78044572	<i>Carboxydotherrmus hydrogenoformans</i>
CHY_0732	YP_359586.1	78044500	<i>Carboxydotherrmus hydrogenoformans</i>
CHY_0733	YP_359587.1	78044647	<i>Carboxydotherrmus hydrogenoformans</i>

Formyltetrahydrofolate synthetase ligates formate to tetrahydrofolate at the expense of one ATP. This reaction is catalyzed by the gene product of Moth_0109 in *M. thermoacetica* (O'brien et al., *Experientia Suppl.*26:249-262 (1976); Lovell et al., *Arch. Microbiol.* 149.4:280-285 (1988);

5 Lovell et al., *Biochemistry* 29.24:5687-5694 (1990)), *FHS* in *Clostridium acidurici* (Whitehead and Rabinowitz, *J. Bacteriol.* 167.1:205-209 (1986); Whitehead and Rabinowitz, *J. Bacteriol.* 170.7:3255-3261 (1988)), and CHY_2385 in *C. hydrogenoformans* (Wu et al., *PLoS Genet.*1.5:e65 (2005)). Relevant gene and protein information can be obtained from the information shown below:

<u>Protein</u>	<u>GenBank ID</u>	<u>GI number</u>	<u>Organism</u>
Moth_0109	YP_428991.1	83588982	<i>Moorella thermoacetica</i>
CHY_2385	YP_361182.1	78045024	<i>Carboxydotherrmus hydrogenoformans</i>
FHS	P13419.1	120562	<i>Clostridium acidurici</i>

10

In *M. thermoacetica*, *E. coli*, and *C. hydrogenoformans*, methenyltetrahydrofolate cyclohydrolase and methylenetetrahydrofolate dehydrogenase are carried out by the bi-functional gene products of Moth_1516, *folD*, and CHY_1878, respectively (Pierce et al. *Environ. Microbiol.* 10:2550-2573 (2008); (Wu et al., *PLoS Genet.*1.5:e65 (2005); D'Ari and

15 Rabinowitz, *J. Biol. Chem.* 266.35:23953-23958 (1991)). Relevant gene and protein information can be obtained from the information shown below:

<u>Protein</u>	<u>GenBank ID</u>	<u>GI number</u>	<u>Organism</u>
Moth_1516	YP_430368.1	83590359	<i>Moorella thermoacetica</i>
<i>folD</i>	NP_415062.1	16128513	<i>Escherichia coli</i>
CHY_1878	YP_360698.1	78044829	<i>Carboxydotherrmus hydrogenoformans</i>

In *M. thermoacetica*, *E. coli*, and *C. hydrogenoformans*, methenyltetrahydrofolate cyclohydrolase and methylenetetrahydrofolate dehydrogenase are carried out by the bi-functional gene products of Moth_1516, *folD*, and CHY_1878, respectively (Pierce et al. *Environ. Microbiol.* 10:2550-2573 (2008); (Wu et al., *PLoS Genet.* 1.5:e65 (2005); D'Ari and Rabinowitz, *J. Biol. Chem.* 266.35:23953-23958 (1991)). Relevant gene and protein information can be obtained from the information shown below:

<u>Protein</u>	<u>GenBank ID</u>	<u>GI number</u>	<u>Organism</u>
Moth_1516	YP_430368.1	83590359	<i>Moorella thermoacetica</i>
<i>folD</i>	NP_415062.1	16128513	<i>Escherichia coli</i>
CHY_1878	YP_360698.1	78044829	<i>Carboxydotherrmus hydrogenoformans</i>

The final step of the methyl branch of the Wood-Ljungdahl pathway is catalyzed by methylenetetrahydrofolate reductase. In *M. thermoacetica*, this enzyme is oxygen-sensitive and contains an iron-sulfur cluster (Clark and Ljungdahl, *J. Biol. Chem.* 259.17:10845-10849 (1984)). This enzyme is encoded by *metF* in *E. coli* (Sheppard et al., *J. Bacteriol.* 181.3:718-725 (1999)) and CHY_1233 in *C. hydrogenoformans* (Wu et al., *PLoS Genet.* 1.5:e65 (2005)). The *M. thermoacetica* genes, and its *C. hydrogenoformans* counterpart, are located near the CODH/ACS gene cluster, separated by putative hydrogenase and heterodisulfide reductase genes.

<u>Protein</u>	<u>GenBank ID</u>	<u>GI number</u>	<u>Organism</u>
Moth_1191	YP_430048.1	83590039	<i>Moorella thermoacetica</i>
<i>metF</i>	NP_418376.1	16131779	<i>Escherichia coli</i>
CHY_1233	YP_360071.1	78044792	<i>Carboxydotherrmus hydrogenoformans</i>

While *E. coli* naturally possesses the capability for some of the required transformations (*i.e.*, methenyltetrahydrofolate cyclohydrolase, methylenetetrahydrofolate dehydrogenase, methylenetetrahydrofolate reductase), the methyl branch enzymes from acetogens can have significantly higher (50 – 100X) specific activities than those from non-acetogens (Morton et al., *Genetics and molecular biology of anaerobic bacteria*, Ed. M. Sebald, New York: Springer Verlag (1992) pages 389-406). Formate dehydrogenase may be specialized for anaerobic conditions (Ljungdahl and Andreessen, *FEBS Lett.* 54.2:279-282 (1975)) (1975). Therefore, various non-native versions of each of these can be expressed in the strain of *E. coli* capable of methanol and CO₂, CO, and/or H₂ utilization. Specifically, these genes can be cloned and combined into an expression vector designed to express them as a set. Initially, a high or

medium copy number vector can be chosen (using ColeE1 or P15A replicons). An exemplary promoter is a strongly constitutive promoter such as lambda pL or an IPTG-inducible version of this, pL-lacO (Lutz and Bujard, *Nucleic Acids Res.* 25.6:1203-1210 (1997)). To make an artificial operon, one 5' terminal promoter is placed upstream of the set of genes and each gene receives a consensus rbs element. The order of genes is based on the natural order whenever possible. Ultimately, the genes are integrated into the *E. coli* chromosome. Enzyme assays are performed as described in (Ljungdahl and Andreessen, *Methods Enzymol.* 53:360-372 (1978); Yamamoto et al., *J. Biol. Chem.* 258.3:1826-1832 (1983); Lovell et al., *Arch. Microbiol.* 149.4:280-285 (1988); de Mata and Rabinowitz, *J. Biol. Chem.* 255.6:2569-2577 (1980); D'Ari and Rabinowitz, *J. Biol. Chem.* 266.35:23953-23958 (1991); Clark and Ljungdahl, 259.17:10845-10849 (1984); Clark and Ljungdahl, *Methods Enzymol.* 122:392-399 (1986)).

After strains of *E. coli* expressing both the carbonyl and methyl branches of the Wood-Ljungdahl pathway are constructed, they are assayed for the ability to utilize CO, CO₂, and/or H₂, for incorporation into acetyl-CoA, cell mass, 1,3-butanediol. Initial conditions employ strictly anaerobically grown cells provided with exogenous glucose. Metabolizing glucose or other carbohydrates to acetyl-CoA provides one potential source of CO₂ that can be fixed via the Wood-Ljungdahl pathway. Alternatively, or in addition to glucose, nitrate can be added to the fermentation broth to serve as an electron acceptor and initiator of growth. Anaerobic growth of *E. coli* on fatty acids, which are ultimately metabolized to acetyl-CoA, has been demonstrated in the presence of nitrate (Campbell et al., *Mol. Microbiol.* 47.3:793-805 (2003)). Oxygen can also be provided as long as its intracellular levels are maintained below any inhibition threshold of the engineered enzymes. 'Synthetic syngas' of a composition suitable for these experiments can also be employed. ¹³C-labeled CO and/or CO₂ are provided to the cells and analytical mass spectrometry is employed to measure incorporation of the labeled carbon into acetate, 1,3-butanediol, and cell mass (e.g., proteinogenic amino acids).

Process considerations for a syngas fermentation include high biomass concentration and good gas-liquid mass transfer (Bredwell et al., *Biotechnol. Prog.* 15.5:834-844 (1999)). The solubility of CO in water is somewhat less than that of oxygen. Continuously gas-sparged fermentations can be performed in controlled fermenters with constant off-gas analysis by mass spectrometry and periodic liquid sampling and analysis by GC and HPLC. The liquid phase can function in batch mode. Fermentation products such as alcohols, organic acids, and residual glucose along with residual methanol are quantified by HPLC (Shimadzu, Columbia MD), for example, using an Aminex® series of HPLC columns (for example, HPX-87 series) (BioRad, Hercules CA),

using a refractive index detector for glucose and alcohols, and a UV detector for organic acids. The growth rate is determined by measuring optical density using a spectrophotometer (600 nm). All piping in these systems is glass or metal to maintain anaerobic conditions. The gas sparging can be performed with glass frits to decrease bubble size and improve mass transfer.

5 Various sparging rates are tested, ranging from about 0.1 to 1 vvm (vapor volumes per minute). To obtain accurate measurements of gas uptake rates, periodic challenges are performed in which the gas flow is temporarily stopped, and the gas phase composition is monitored as a function of time.

In order to achieve the overall target productivity, methods of cell retention or recycle can be
10 employed. One method to increase the microbial concentration is to recycle cells via a tangential flow membrane from a sidestream. Repeated batch culture can also be used, as previously described for production of acetate by *Moorella* (Sakai et al., *J. Biosci. Bioeng.* 99.3:252-258 (2005)). Various other methods can also be used (Bredwell et al., *Biotechnol. Prog.* 15.5:834-844 (1999); Datar et al., *Biotechnol. Bioeng.* 86.5:587-594 (2004)). Additional optimization can
15 be tested such as overpressure at 1.5 atm to improve mass transfer (Najafpour and Younesi, *Enzyme and Microbial Technology* 38:223-228 (2006)).

Once satisfactory performance is achieved using pure H₂/CO as the feed, synthetic gas mixtures can be generated containing inhibitors likely to be present in commercial syngas. For example, a typical impurity profile is 4.5% CH₄, 0.1% C₂H₂, 0.35% C₂H₆, 1.4% C₂H₄, and 150 ppm nitric
20 oxide (Datar et al., *Biotechnol. Bioeng.* 86.5:587-594 (2004)). Tars, represented by compounds such as benzene, toluene, ethylbenzene, p-xylene, o-xylene, and naphthalene, are added at ppm levels to test for any effect on production. For example, it has been shown that 40 ppm NO is inhibitory to *C. carboxidivorans* (Ahmed and Lewis, *Biotechnol. Bioeng.* 97.5:1080-1086 (2007)). Cultures can be tested in shake-flask cultures before moving to a fermentor. Also,
25 different levels of these potential inhibitory compounds are tested to quantify the effect they have on cell growth. This knowledge is used to develop specifications for syngas purity, which is utilized for scale up studies and production. If any particular component is found to be difficult to decrease or remove from syngas used for scale up, an adaptive evolution procedure is utilized to adapt cells to tolerate one or more impurities.

30 The invention is described herein with general reference to the metabolic reaction, reactant or product thereof, or with specific reference to one or more nucleic acids or genes encoding an enzyme associated with or catalyzing, or a protein associated with, the referenced metabolic reaction, reactant or product. Unless otherwise expressly stated herein, those skilled in the art

will understand that reference to a reaction also constitutes reference to the reactants and products of the reaction. Similarly, unless otherwise expressly stated herein, reference to a reactant or product also references the reaction, and reference to any of these metabolic constituents also references the gene or genes encoding the enzymes that catalyze or proteins involved in the referenced reaction, reactant or product. Likewise, given the well known fields of metabolic biochemistry, enzymology and genomics, reference herein to a gene or encoding nucleic acid also constitutes a reference to the corresponding encoded enzyme and the reaction it catalyzes or a protein associated with the reaction as well as the reactants and products of the reaction.

The non-naturally occurring microbial organisms of the invention can be produced by introducing expressible nucleic acids encoding one or more of the enzymes or proteins participating in one or more 1,3-butanediol biosynthetic pathways. Depending on the host microbial organism chosen for biosynthesis, nucleic acids for some or all of a particular 1,3-butanediol biosynthetic pathway can be expressed. For example, if a chosen host is deficient in one or more enzymes or proteins for a desired biosynthetic pathway, then expressible nucleic acids for the deficient enzyme(s) or protein(s) are introduced into the host for subsequent exogenous expression. Alternatively, if the chosen host exhibits endogenous expression of some pathway genes, but is deficient in others, then an encoding nucleic acid is needed for the deficient enzyme(s) or protein(s) to achieve 1,3-butanediol biosynthesis. Thus, a non-naturally occurring microbial organism of the invention can be produced by introducing exogenous enzyme or protein activities to obtain a desired biosynthetic pathway or a desired biosynthetic pathway can be obtained by introducing one or more exogenous enzyme or protein activities that, together with one or more endogenous enzymes or proteins, produces a desired product such as 1,3-butanediol.

Host microbial organisms can be selected from, and the non-naturally occurring microbial organisms generated in, for example, bacteria, yeast, fungus, algae, cyanobacteria, or any of a variety of other microorganisms applicable to fermentation processes. Exemplary bacteria include species selected from *Escherichia coli*, *Klebsiella oxytoca*, *Anaerobiospirillum succiniciproducens*, *Actinobacillus succinogenes*, *Mannheimia succiniciproducens*, *Rhizobium etli*, *Bacillus subtilis*, *Corynebacterium glutamicum*, *Gluconobacter oxydans*, *Zymomonas mobilis*, *Lactococcus lactis*, *Lactobacillus plantarum*, *Streptomyces coelicolor*, *Clostridium acetobutylicum*, *Pseudomonas fluorescens*, and *Pseudomonas putida*. Exemplary yeasts or fungi include species selected from *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*,

- Kluyveromyces lactis*, *Kluyveromyces marxianus*, *Aspergillus terreus*, *Aspergillus niger*, *Pichia pastoris*, *Rhizopus arrhizus*, *Rhizobus oryzae*, and the like. Exemplary cyanobacteria include *Acaryochloris marina* MBIC11017, *Anabaena* sp. PCC 7120, *Anabaena variabilis* ATCC 29413, *Agmenellum quadruplicatum*, *Chlorobium tepidum* TLS, *Cyanothece* sp. ATCC 51142,
- 5 *Gloeobacter violaceus* PCC 7421, *Microcystis aeruginosa* NIES-843, *Nostoc punctiforme* ATCC 29133, *Prochlorococcus marinus* MED4, *Prochlorococcus marinus* MIT9313, *Prochlorococcus marinus* SS120, *Prochlorococcus marinus* str. AS9601, *Prochlorococcus marinus* str. MIT 9211, *Prochlorococcus marinus* str. MIT 9215, *Prochlorococcus marinus* str. MIT 9301, *Prochlorococcus marinus* str. MIT 9303, *Prochlorococcus marinus* str. MIT 9312,
- 10 *Prochlorococcus marinus* str. MIT 9515, *Prochlorococcus marinus* str. NATL1A, *Prochlorococcus marinus* str. NATL2A, *Rhodopseudomonas palustris* CGA009, *Synechococcus elongatus* PCC 6301, *Synechococcus elongatus* PCC 7942, *Synechococcus* sp. CC9311, *Synechococcus* sp. CC9605, *Synechococcus* sp. CC9902, *Synechococcus* sp. JA-2-3B\A(2-13), *Synechococcus* sp. JA-3-3Ab, *Synechococcus* sp. PCC 7002, *Synechococcus* sp. RCC307,
- 15 *Synechococcus* sp. WH 7803, *Synechococcus* sp. WH8102, *Synechocystis* sp. PCC 6803, *Thermosynechococcus elongatus* BP-1, *Trichodesmium erythraeum* IMS101. Exemplary algae include *Botryococcus braunii*, *Chlamydomonas reinhardtii*, *Chlorella* sp., *Cryptocodinium cohnii*, *Cylindrotheca* sp., *Dunaliella primolecta*, *Isochrysis* sp., *Monallanthus salina*, *Nannochloris* sp., *Nannochloropsis* sp., *Neochloris oleoabundans*, *Nitzschia* sp., *Phaeodactylum*
- 20 *tricornutum*, *Schizochytrium* sp., *Tetraselmis sueica*. *E. coli* is a particularly useful host organisms since it is a well characterized microbial organism suitable for genetic engineering. Other particularly useful host organisms include yeast such as *Saccharomyces cerevisiae*. It is understood that any suitable microbial host organism can be used to introduce metabolic and/or genetic modifications to produce a desired product.
- 25 Depending on the 1,3-butanediol biosynthetic pathway constituents of a selected host microbial organism, the non-naturally occurring microbial organisms of the invention will include at least one exogenously expressed 1,3-BDO pathway-encoding nucleic acid and up to all encoding nucleic acids for one or more 1,3-butanediol biosynthetic pathways. For example, 1,3-butanediol biosynthesis can be established in a host deficient in a pathway enzyme or protein
- 30 through exogenous expression of the corresponding encoding nucleic acid. In a host deficient in all enzymes or proteins of a 1,3-butanediol pathway, exogenous expression of all enzyme or proteins in the pathway can be included, although it is understood that all enzymes or proteins of a pathway can be expressed even if the host contains at least one of the pathway enzymes or proteins. For example, exogenous expression of all enzymes or proteins in a pathway for

- production of 1,3-butanediol can be included, such as 1) Methanol methyltransferase (MtaB), 2) Corrinoid protein (MtaC), 3) Methyltetrahydrofolate:corrinoid protein methyltransferase (MtaA), 4) Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), 5) Corrinoid iron-sulfur protein (AcsD), 6) Nickel-protein assembly protein (AcsF & CooC), 7) Ferredoxin (Orf7), 8) Acetyl-CoA synthase (AcsB & AcsC), 9) Carbon monoxide dehydrogenase (AcsA), 10) Hydrogenase, 11) Acetoacetyl-CoA thiolase (AtoB), 12) Acetoacetyl-CoA reductase (CoA-dependent, alcohol forming), 13) 3-oxobutyraldehyde reductase (aldehyde reducing), 14) 4-hydroxy,2-butanone reductase, 15) Acetoacetyl-CoA reductase (CoA-dependent, aldehyde forming), 16) 3-oxobutyraldehyde reductase (ketone reducing), 17) 3-hydroxybutyraldehyde reductase, 18) Acetoacetyl-CoA reductase (ketone reducing), 19) 3-hydroxybutyryl-CoA reductase (aldehyde forming), 20) 3-hydroxybutyryl-CoA reductase (alcohol forming), 21) acetoacetyl-CoA transferase, hydrolase, or synthetase, 22) acetoacetate reductase, 23) 3-hydroxybutyryl-CoA transferase, hydrolase, or synthetase, 24) 3-hydroxybutyrate reductase, and 25) 3-hydroxybutyrate dehydrogenase, as shown in Figure 4.
- 15 Alternatively, exogenous expression of all enzymes or proteins in a pathway for production of 1,3-butanediol, as shown in Figure 5, can be included, such as 1) Formate dehydrogenase, 2) Formyltetrahydrofolate synthetase, 3) Methenyltetrahydrofolate cyclohydrolase, 4) Methylenetetrahydrofolate dehydrogenase, 5) Methylenetetrahydrofolate reductase, 6) Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), 7) Corrinoid iron-sulfur protein (AcsD), 8) Nickel-protein assembly protein (AcsF & CooC), 9) Ferredoxin (Orf7), 10) Acetyl-CoA synthase (AcsB & AcsC), 11) Carbon monoxide dehydrogenase (AcsA), 12) Hydrogenase (Hsd), 13) Acetoacetyl-CoA thiolase (AtoB), 14) Acetoacetyl-CoA reductase (CoA-dependent, alcohol forming), 15) 3-oxobutyraldehyde reductase (aldehyde reducing), 16) 4-hydroxy,2-butanone reductase, 17) Acetoacetyl-CoA reductase (CoA-dependent, aldehyde forming), 18) 3-oxobutyraldehyde reductase (ketone reducing), 19) 3-hydroxybutyraldehyde reductase, 20) Acetoacetyl-CoA reductase (ketone reducing), 21) 3-hydroxybutyryl-CoA reductase (aldehyde forming), 22) 3-hydroxybutyryl-CoA reductase (alcohol forming), 23) acetoacetyl-CoA transferase, hydrolase, or synthetase, 24) acetoacetate reductase, 25) 3-hydroxybutyryl-CoA transferase, hydrolase, or synthetase, 26) 3-hydroxybutyrate reductase, and 27) 3-hydroxybutyrate dehydrogenase .

Given the teachings and guidance provided herein, those skilled in the art will understand that the number of encoding nucleic acids to introduce in an expressible form will, at least, parallel the 1,3-butanediol pathway deficiencies of the selected host microbial organism. Therefore, a

non-naturally occurring microbial organism of the invention can have one, two, three, four, five, six, seven, eight, nine, ten, eleven, twelve, thirteen, fourteen, fifteen, sixteen, seventeen that is, up to all nucleic acids encoding the enzymes or proteins constituting a 1,3-butanediol

biosynthetic pathway disclosed herein and shown in Figure 4 and 5. In some embodiments, the non-naturally occurring microbial organisms also can include other genetic modifications that facilitate or optimize 1,3-butanediol biosynthesis or that confer other useful functions onto the host microbial organism. One such other functionality can include, for example, augmentation of the synthesis of one or more of the 1,3-butanediol pathway precursors such as acetyl-CoA, acetoacetyl-CoA, acetoacetate, 3-hydroxybutyryl-CoA, 3-hydroxybutyrate, 4-hydroxy-2-butanone, 3-oxobutryaldehyde, or 3-hydroxybutryaldehyde.

Generally, a host microbial organism is selected such that it produces the precursor of a 1,3-butanediol pathway, either as a naturally produced molecule or as an engineered product that either provides *de novo* production of a desired precursor or increased production of a precursor naturally produced by the host microbial organism. For example, acetyl-CoA is produced naturally in a host organism such as *E. coli*. A host organism can be engineered to increase production of a precursor, as disclosed herein. In addition, a microbial organism that has been engineered to produce a desired precursor can be used as a host organism and further engineered to express enzymes or proteins of a 1,3-butanediol pathway.

In some embodiments, a non-naturally occurring microbial organism of the invention is generated from a host that contains the enzymatic capability to synthesize 1,3-butanediol. In this specific embodiment it can be useful to increase the synthesis or accumulation of a 1,3-butanediol pathway product to, for example, drive 1,3-butanediol pathway reactions toward 1,3-butanediol production. Increased synthesis or accumulation can be accomplished by, for example, overexpression of nucleic acids encoding one or more of the above-described 1,3-butanediol pathway enzymes or proteins. Over expression the enzyme or enzymes and/or protein or proteins of the 1,3-butanediol pathway can occur, for example, through exogenous expression of the endogenous gene or genes, or through exogenous expression of the heterologous gene or genes. Therefore, naturally occurring organisms can be readily generated to be non-naturally occurring microbial organisms of the invention, for example, producing 1,3-butanediol, through overexpression of one, two, three, four, five, six, seven, eight, nine, ten, eleven, twelve, thirteen, fourteen, fifteen, sixteen, seventeen that is, up to all nucleic acids encoding the enzymes or proteins constituting a 1,3-butanediol biosynthetic pathway disclosed herein and shown in Figure 4 and 5. In addition, a non-naturally occurring organism can be

generated by mutagenesis of an endogenous gene that results in an increase in activity of an enzyme in the 1,3-butanediol biosynthetic pathway.

In particularly useful embodiments, exogenous expression of the encoding nucleic acids is employed. Exogenous expression confers the ability to custom tailor the expression and/or

5 regulatory elements to the host and application to achieve a desired expression level that is controlled by the user. However, endogenous expression also can be utilized in other embodiments such as by removing a negative regulatory effector or induction of the gene's promoter when linked to an inducible promoter or other regulatory element. Thus, an endogenous gene having a naturally occurring inducible promoter can be up-regulated by
10 providing the appropriate inducing agent, or the regulatory region of an endogenous gene can be engineered to incorporate an inducible regulatory element, thereby allowing the regulation of increased expression of an endogenous gene at a desired time. Similarly, an inducible promoter can be included as a regulatory element for an exogenous gene introduced into a non-naturally occurring microbial organism.

15 It is understood that, in methods of the invention, any of the one or more exogenous nucleic acids can be introduced into a microbial organism to produce a non-naturally occurring microbial organism of the invention. The nucleic acids can be introduced so as to confer, for example, a 1,3-butanediol biosynthetic pathway onto the microbial organism. Alternatively, encoding nucleic acids can be introduced to produce an intermediate microbial organism having
20 the biosynthetic capability to catalyze some of the required reactions to confer 1,3-butanediol biosynthetic capability. For example, a non-naturally occurring microbial organism having a 1,3-butanediol biosynthetic pathway can comprise at least two exogenous nucleic acids encoding desired enzymes or proteins. Thus, it is understood that any combination of two or more enzymes or proteins of a biosynthetic pathway can be included in a non-naturally
25 occurring microbial organism of the invention. Similarly, it is understood that any combination of three or more enzymes or proteins of a biosynthetic pathway can be included in a non-naturally occurring microbial organism of the invention. Similarly, any combination of four, or more enzymes or proteins of a biosynthetic pathway as disclosed herein can be included in a non-naturally occurring microbial organism of the invention, as desired, so long as the
30 combination of enzymes and/or proteins of the desired biosynthetic pathway results in production of the corresponding desired product. Likewise, any combination of five, six, seven, eight, nine, ten, eleven, twelve, thirteen, fourteen, fifteen, sixteen, enzymes and/or proteins of a biosynthetic pathway as disclosed herein can be included in a non-naturally occurring microbial

organism of the invention, as desired, so long as the combination of enzymes and/or proteins of the desired biosynthetic pathway results in production of the corresponding desired product.

Exemplary combinations of 17 exogenous enzymes or proteins of a biosynthetic pathway, as disclosed herein, included in a non-naturally occurring microbial organism of the invention

5 include:

A: 1) Formate dehydrogenase, 2) Formyltetrahydrofolate synthetase, 3) Methenyltetrahydrofolate cyclohydrolase, 4) Methylenetetrahydrofolate dehydrogenase, 5) Methylenetetrahydrofolate reductase, 6) Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), 7) Corrinoid iron-sulfur protein (AcsD), 8) Nickel-protein assembly
10 protein (AcsF & CooC), 9) Ferredoxin (Orf7), 10) Acetyl-CoA synthase (AcsB & AcsC), 11) Carbon monoxide dehydrogenase (AcsA), 12) Hydrogenase (Hyd), 13) Acetoacetyl-CoA thiolase (AtoB), 14) Acetoacetyl-CoA transferase, hydrolase, or synthetase, 15) Acetoacetate reductase, 16) 3-oxobutyraldehyde reductase (ketone reducing), 17) 3-hydroxybutyraldehyde reductase;

15 B: 1) Formate dehydrogenase, 2) Formyltetrahydrofolate synthetase, 3) Methenyltetrahydrofolate cyclohydrolase, 4) Methylenetetrahydrofolate dehydrogenase, 5) Methylenetetrahydrofolate reductase, 6) Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), 7) Corrinoid iron-sulfur protein (AcsD), 8) Nickel-protein assembly protein (AcsF & CooC), 9) Ferredoxin (Orf7), 10) Acetyl-CoA synthase (AcsB & AcsC), 11)
20 Carbon monoxide dehydrogenase (AcsA), 12) Hydrogenase (Hyd), 13) Acetoacetyl-CoA thiolase (AtoB), 14) Acetoacetyl-CoA transferase, hydrolase, or synthetase, 15) Acetoacetate reductase, 16) 3-oxobutyraldehyde reductase (aldehyde reducing), 17) 4-hydroxy,2-butanone reductase;

C: 1) Formate dehydrogenase, 2) Formyltetrahydrofolate synthetase, 3) Methenyltetrahydrofolate cyclohydrolase, 4) Methylenetetrahydrofolate dehydrogenase, 5) Methylenetetrahydrofolate reductase, 6) Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), 7) Corrinoid iron-sulfur protein (AcsD), 8) Nickel-protein assembly
25 protein (AcsF & CooC), 9) Ferredoxin (Orf7), 10) Acetyl-CoA synthase (AcsB & AcsC), 11) Carbon monoxide dehydrogenase (AcsA), 12) Hydrogenase (Hyd), 13) Acetoacetyl-CoA thiolase (AtoB), 14) Acetoacetyl-CoA reductase (ketone reducing), 15) 3-hydroxybutyryl-CoA
30 transferase, hydrolase, or synthetase, 16) 3-hydroxybutyrate reductase, 17) 3-hydroxybutyraldehyde reductase;

D: 1) Formate dehydrogenase, 2) Formyltetrahydrofolate synthetase, 3) Methenyltetrahydrofolate cyclohydrolase, 4) Methylenetetrahydrofolate dehydrogenase, 5) Methylenetetrahydrofolate reductase, 6) Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), 7) Corrinoid iron-sulfur protein (AcsD), 8) Nickel-protein assembly protein (AcsF & CooC), 9) Ferredoxin (Orf7), 10) Acetyl-CoA synthase (AcsB & AcsC), 11) Carbon monoxide dehydrogenase (AcsA), 12) Hydrogenase (Hyd), 13) Acetoacetyl-CoA thiolase (AtoB), 14) Acetoacetyl-CoA transferase, hydrolase, or synthetase, 15) 3-hydroxybutyrate dehydrogenase, 16) 3-hydroxybutyrate reductase, 17) 3-hydroxybutyraldehyde reductase.

- 10 Exemplary combinations of 16 exogenous enzymes or proteins of a biosynthetic pathway, as disclosed herein, included in a non-naturally occurring microbial organism of the invention include any combination of 16 of the 17 enzymes disclosed above in A-D or:

E: 1) Formate dehydrogenase, 2) Formyltetrahydrofolate synthetase, 3) Methenyltetrahydrofolate cyclohydrolase, 4) Methylenetetrahydrofolate dehydrogenase, 5) Methylenetetrahydrofolate reductase, 6) Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), 7) Corrinoid iron-sulfur protein (AcsD), 8) Nickel-protein assembly protein (AcsF & CooC), 9) Ferredoxin (Orf7), 10) Acetyl-CoA synthase (AcsB & AcsC), 11) Carbon monoxide dehydrogenase (AcsA), 12) Hydrogenase (Hyd), 13) Acetoacetyl-CoA thiolase (AtoB), 14) Acetoacetyl-CoA reductase (CoA-dependent, aldehyde forming), 15) 3-oxobutyraldehyde reductase (aldehyde reducing), and 16) 4-hydroxy,2-butanone reductase.

- F: 1) Formate dehydrogenase, 2) Formyltetrahydrofolate synthetase, 3) Methenyltetrahydrofolate cyclohydrolase, 4) Methylenetetrahydrofolate dehydrogenase, 5) Methylenetetrahydrofolate reductase, 6) Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), 7) Corrinoid iron-sulfur protein (AcsD), 8) Nickel-protein assembly protein (AcsF & CooC), 9) Ferredoxin (Orf7), 10) Acetyl-CoA synthase (AcsB & AcsC), 11) Carbon monoxide dehydrogenase (AcsA), 12) Hydrogenase (Hyd), 13) Acetoacetyl-CoA thiolase (AtoB), 14) Acetoacetyl-CoA reductase (CoA-dependent, aldehyde forming), 15) 3-oxobutyraldehyde reductase (ketone reducing), and 16) 3-hydroxybutyraldehyde reductase.

- G: 1) Formate dehydrogenase, 2) Formyltetrahydrofolate synthetase, 3) Methenyltetrahydrofolate cyclohydrolase, 4) Methylenetetrahydrofolate dehydrogenase, 5) Methylenetetrahydrofolate reductase, 6) Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), 7) Corrinoid iron-sulfur protein (AcsD), 8) Nickel-protein assembly

protein (AcsF & CooC), 9) Ferredoxin (Orf7), 10) Acetyl-CoA synthase (AcsB & AcsC), 11) Carbon monoxide dehydrogenase (AcsA), 12) Hydrogenase (Hyd), 13) Acetoacetyl-CoA thiolase (AtoB), 14) Acetoacetyl-CoA reductase (ketone reducing), 15) 3-hydroxybutyryl-CoA reductase (aldehyde forming), and 16) 3-hydroxybutyraldehyde reductase.

- 5 Exemplary combinations of 15 exogenous enzymes or proteins of a biosynthetic pathway, as disclosed herein, included in a non-naturally occurring microbial organism of the invention any combination of 15 of the enzymes disclosed above in A-G or:

H: 1) Formate dehydrogenase, 2) Formyltetrahydrofolate synthetase, 3)

Methenyltetrahydrofolate cyclohydrolase, 4) Methylenetetrahydrofolate dehydrogenase, 5)

- 10 Methylenetetrahydrofolate reductase, 6) Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), 7) Corrinoid iron-sulfur protein (AcsD), 8) Nickel-protein assembly protein (AcsF & CooC), 9) Ferredoxin (Orf7), 10) Acetyl-CoA synthase (AcsB & AcsC), 11) Carbon monoxide dehydrogenase (AcsA), 12) Hydrogenase (Hyd), 13) Acetoacetyl-CoA thiolase (AtoB), 14) Acetoacetyl-CoA reductase (CoA-dependent, alcohol forming), and 15) 4-hydroxy,2-butanone reductase.

I: 1) Formate dehydrogenase, 2) Formyltetrahydrofolate synthetase, 3) Methenyltetrahydrofolate cyclohydrolase, 4) Methylenetetrahydrofolate dehydrogenase, 5) Methylenetetrahydrofolate reductase, 6) Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), 7) Corrinoid iron-sulfur protein (AcsD), 8) Nickel-protein assembly protein (AcsF & CooC), 9) Ferredoxin
20 (Orf7), 10) Acetyl-CoA synthase (AcsB & AcsC), 11) Carbon monoxide dehydrogenase (AcsA), 12) Hydrogenase (Hyd), 13) Acetoacetyl-CoA thiolase (AtoB), 14) Acetoacetyl-CoA reductase (ketone reducing), and 15) 3-hydroxybutyryl-CoA reductase (alcohol forming).

J: 1) Methanol methyltransferase (MtaB), 2) Corrinoid protein (MtaC), 3)

Methyltetrahydrofolate:corrinoid protein methyltransferase (MtaA), 4)

- 25 Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), 5) Corrinoid iron-sulfur protein (AcsD), 6) Nickel-protein assembly protein (AcsF & CooC), 7) Ferredoxin (Orf7), 8) Acetyl-CoA synthase (AcsB & AcsC), 9) Carbon monoxide dehydrogenase (AcsA), 10) Hydrogenase, 11) Acetoacetyl-CoA thiolase (AtoB), 12) Acetoacetyl-CoA transferase, hydrolase, or synthetase, 13) Acetoacetate reductase, 14) 3-oxobutyraldehyde reductase (ketone
30 reducing), 15) 3-hydroxybutyraldehyde reductase;

K: 1) Methanol methyltransferase (MtaB), 2) Corrinoid protein (MtaC), 3)

Methyltetrahydrofolate:corrinoid protein methyltransferase (MtaA), 4)

Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), 5) Corrinoid iron-sulfur protein (AcsD), 6) Nickel-protein assembly protein (AcsF & CooC), 7) Ferredoxin (Orf7), 8) Acetyl-CoA synthase (AcsB & AcsC), 9) Carbon monoxide dehydrogenase (AcsA), 10) Hydrogenase, 11) Acetoacetyl-CoA thiolase (AtoB), 12) Acetoacetyl-CoA transferase, 5 hydrolase, or synthetase, 13) Acetoacetate reductase, 14) 3-oxobutylaldehyde reductase (aldehyde reducing), 15) 4-hydroxy,2-butanone reductase;

L: 1) Methanol methyltransferase (MtaB), 2) Corrinoid protein (MtaC), 3) Methyltetrahydrofolate:corrinoid protein methyltransferase (MtaA), 4) Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), 5) Corrinoid iron-sulfur protein (AcsD), 6) Nickel-protein assembly protein (AcsF & CooC), 7) Ferredoxin (Orf7), 8) 10 Acetyl-CoA synthase (AcsB & AcsC), 9) Carbon monoxide dehydrogenase (AcsA), 10) Hydrogenase, 11) Acetoacetyl-CoA thiolase (AtoB), 12) Acetoacetyl-CoA reductase (ketone reducing), 13) 3-hydroxybutyryl-CoA transferase, hydrolase, or synthetase, 14) 3-hydroxybutyrate reductase, 15) 3-hydroxybutylaldehyde reductase;

15 M: 1) Methanol methyltransferase (MtaB), 2) Corrinoid protein (MtaC), 3) Methyltetrahydrofolate:corrinoid protein methyltransferase (MtaA), 4) Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), 5) Corrinoid iron-sulfur protein (AcsD), 6) Nickel-protein assembly protein (AcsF & CooC), 7) Ferredoxin (Orf7), 8) Acetyl-CoA synthase (AcsB & AcsC), 9) Carbon monoxide dehydrogenase (AcsA), 10) 20 Hydrogenase, 11) Acetoacetyl-CoA thiolase (AtoB), 12) Acetoacetyl-CoA transferase, hydrolase, or synthetase, 13) 3-hydroxybutyrate dehydrogenase, 14) 3-hydroxybutyrate reductase, 15) 3-hydroxybutylaldehyde reductase.

Exemplary combinations of 14 exogenous enzymes or proteins of a biosynthetic pathway, as disclosed herein, included in a non-naturally occurring microbial organism of the invention 25 include any combination of 14 of the enzymes disclosed above in A-M or:

N: 1) Methanol methyltransferase (MtaB), 2) Corrinoid protein (MtaC), 3) Methyltetrahydrofolate:corrinoid protein methyltransferase (MtaA), 4) Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), 5) Corrinoid iron-sulfur protein (AcsD), 6) Nickel-protein assembly protein (AcsF & CooC), 7) Ferredoxin (Orf7), 8) 30 Acetyl-CoA synthase (AcsB & AcsC), 9) Carbon monoxide dehydrogenase (AcsA), 10) Hydrogenase, 11) Acetoacetyl-CoA thiolase (AtoB), 12) Acetoacetyl-CoA reductase (CoA-

dependent, aldehyde forming) 13) 3-oxobutyraldehyde reductase (aldehyde reducing), and 14) 4-hydroxy,2-butanone reductase.

O: 1) Methanol methyltransferase (MtaB), 2) Corrinoid protein (MtaC), 3)

Methyltetrahydrofolate:corrinoid protein methyltransferase (MtaA), 4)

- 5 Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), 5) Corrinoid iron-sulfur protein (AcsD), 6) Nickel-protein assembly protein (AcsF & CooC), 7) Ferredoxin (Orf7), 8) Acetyl-CoA synthase (AcsB & AcsC), 9) Carbon monoxide dehydrogenase (AcsA), 10) Hydrogenase, 11) Acetoacetyl-CoA thiolase (AtoB), 12) Acetoacetyl-CoA reductase (CoA-dependent, aldehyde forming), 13) 3-oxobutyraldehyde reductase (ketone reducing), and 14) 3-hydroxybutyraldehyde reductase.

P: 1) Methanol methyltransferase (MtaB), 2) Corrinoid protein (MtaC), 3)

Methyltetrahydrofolate:corrinoid protein methyltransferase (MtaA), 4)

Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), 5) Corrinoid iron-sulfur protein (AcsD), 6) Nickel-protein assembly protein (AcsF & CooC), 7) Ferredoxin (Orf7), 8)

- 15 Acetyl-CoA synthase (AcsB & AcsC), 9) Carbon monoxide dehydrogenase (AcsA), 10) Hydrogenase, 11) Acetoacetyl-CoA thiolase (AtoB), 12) Acetoacetyl-CoA reductase (ketone reducing), 13) 3-hydroxybutyryl-CoA reductase (aldehyde forming), and 14) 3-hydroxybutyraldehyde reductase

Exemplary combinations of 13 exogenous enzymes or proteins of a biosynthetic pathway, as disclosed herein, included in a non-naturally occurring microbial organism of the invention include any combination of 13 of the enzymes disclosed above in A-P or:

Q: 1) Methanol methyltransferase (MtaB), 2) Corrinoid protein (MtaC), 3)

Methyltetrahydrofolate:corrinoid protein methyltransferase (MtaA), 4)

Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), 5) Corrinoid iron-sulfur

- 25 protein (AcsD), 6) Nickel-protein assembly protein (AcsF & CooC), 7) Ferredoxin (Orf7), 8) Acetyl-CoA synthase (AcsB & AcsC), 9) Carbon monoxide dehydrogenase (AcsA), 10) Hydrogenase, 11) Acetoacetyl-CoA thiolase (AtoB), 12) Acetoacetyl-CoA reductase (ketone reducing), and 13) 3-hydroxybutyryl-CoA reductase (alcohol forming),

R: 1) Methanol methyltransferase (MtaB), 2) Corrinoid protein (MtaC), 3)

- 30 Methyltetrahydrofolate:corrinoid protein methyltransferase (MtaA), 4)

Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), 5) Corrinoid iron-sulfur protein (AcsD), 6) Nickel-protein assembly protein (AcsF & CooC), 7) Ferredoxin (Orf7), 8)

Acetyl-CoA synthase (AcsB & AcsC), 9) Carbon monoxide dehydrogenase (AcsA), 10) Hydrogenase, 11) Acetoacetyl-CoA thiolase (AtoB), 12) Acetoacetyl-CoA reductase (CoA-dependent, alcohol forming), and 13) 4-hydroxy,2-butanone reductase.

Exemplary combinations of 12 exogenous enzymes or proteins of a biosynthetic pathway, as disclosed herein, included in a non-naturally occurring microbial organism of the invention include any combination of 12 of the enzymes disclosed above in A-R. Exemplary combinations of 11 exogenous enzymes or proteins of a biosynthetic pathway, as disclosed herein, included in a non-naturally occurring microbial organism of the invention include any combination of 10 of the enzymes disclosed above in A-R. Exemplary combinations of 9 exogenous enzymes or proteins of a biosynthetic pathway, as disclosed herein, included in a non-naturally occurring microbial organism of the invention include any combination of 8 of the enzymes disclosed above in A-R. Exemplary combinations of 7 exogenous enzymes or proteins of a biosynthetic pathway, as disclosed herein, included in a non-naturally occurring microbial organism of the invention include any combination of 6 of the enzymes disclosed above in A-R, and so on down to any combination of 2 of the enzymes disclosed above in A-R.

In addition to the biosynthesis of 1,3-butanediol as described herein, the non-naturally occurring microbial organisms and methods of the invention also can be utilized in various combinations with each other and with other microbial organisms and methods well known in the art to achieve product biosynthesis by other routes. For example, one alternative to produce 1,3-butanediol other than use of the 1,3-butanediol producers is through addition of another microbial organism capable of converting a 1,3-butanediol pathway intermediate to 1,3-butanediol. One such procedure includes, for example, the fermentation of a microbial organism that produces a 1,3-butanediol pathway intermediate. The 1,3-butanediol pathway intermediate can then be used as a substrate for a second microbial organism that converts the 1,3-butanediol pathway intermediate to 1,3-butanediol. The 1,3-butanediol pathway intermediate can be added directly to another culture of the second organism or the original culture of the 1,3-butanediol pathway intermediate producers can be depleted of these microbial organisms by, for example, cell separation, and then subsequent addition of the second organism to the fermentation broth can be utilized to produce the final product without intermediate purification steps.

In other embodiments, the non-naturally occurring microbial organisms and methods of the invention can be assembled in a wide variety of subpathways to achieve biosynthesis of, for example, 1,3-butanediol. In these embodiments, biosynthetic pathways for a desired product of the invention can be segregated into different microbial organisms, and the different microbial

organisms can be co-cultured to produce the final product. In such a biosynthetic scheme, the product of one microbial organism is the substrate for a second microbial organism until the final product is synthesized. For example, the biosynthesis of 1,3-butanediol can be accomplished by constructing a microbial organism that contains biosynthetic pathways for conversion of one pathway intermediate to another pathway intermediate or the product. Alternatively, 1,3-butanediol also can be biosynthetically produced from microbial organisms through co-culture or co-fermentation using two organisms in the same vessel, where the first microbial organism produces a 1,3-butanediol intermediate and the second microbial organism converts the intermediate to 1,3-butanediol.

Given the teachings and guidance provided herein, those skilled in the art will understand that a wide variety of combinations and permutations exist for the non-naturally occurring microbial organisms and methods of the invention together with other microbial organisms, with the co-culture of other non-naturally occurring microbial organisms having subpathways and with combinations of other chemical and/or biochemical procedures well known in the art to produce 1,3-butanediol.

Sources of encoding nucleic acids for a 1,3-butanediol pathway enzyme or protein can include, for example, any species where the encoded gene product is capable of catalyzing the referenced reaction. Such species include both prokaryotic and eukaryotic organisms including, but not limited to, bacteria, including archaea and eubacteria, and eukaryotes, including yeast, plant, algae, insect, animal, and mammal, including human. Exemplary species for such sources include, for example, *Escherichia coli*, as well as other exemplary species disclosed herein or available as source organisms for corresponding genes. However, with the complete genome sequence available for now more than 550 species (with more than half of these available on public databases such as the NCBI), including 395 microorganism genomes and a variety of yeast, fungi, plant, and mammalian genomes, the identification of genes encoding the requisite 1,3-butanediol biosynthetic activity for one or more genes in related or distant species, including for example, homologues, orthologs, paralogs and nonorthologous gene displacements of known genes, and the interchange of genetic alterations between organisms is routine and well known in the art. Accordingly, the metabolic alterations allowing biosynthesis of 1,3-butanediol described herein with reference to a particular organism such as *E. coli* can be readily applied to other microorganisms, including prokaryotic and eukaryotic organisms alike. Given the teachings and guidance provided herein, those skilled in the art will know that a metabolic alteration exemplified in one organism can be applied equally to other organisms.

In some instances, such as when an alternative 1,3-butanediol biosynthetic pathway exists in an unrelated species, 1,3-butanediol biosynthesis can be conferred onto the host species by, for example, exogenous expression of a paralog or paralogs from the unrelated species that catalyzes a similar, yet non-identical metabolic reaction to replace the referenced reaction.

5 Because certain differences among metabolic networks exist between different organisms, those skilled in the art will understand that the actual gene usage between different organisms may differ. However, given the teachings and guidance provided herein, those skilled in the art also will understand that the teachings and methods of the invention can be applied to all microbial organisms using the cognate metabolic alterations to those exemplified herein to construct a
10 microbial organism in a species of interest that will synthesize 1,3-butanediol.

Methods for constructing and testing the expression levels of a non-naturally occurring 1,3-butanediol-producing host can be performed, for example, by recombinant and detection methods well known in the art. Such methods can be found described in, for example, Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Third Ed., Cold Spring Harbor
15 Laboratory, New York (2001); and Ausubel et al., *Current Protocols in Molecular Biology*, John Wiley and Sons, Baltimore, MD (1999).

Exogenous nucleic acid sequences involved in a pathway for production of 1,3-butanediol can be introduced stably or transiently into a host cell using techniques well known in the art including, but not limited to, conjugation, electroporation, chemical transformation,
20 transduction, transfection, and ultrasound transformation. For exogenous expression in *E. coli* or other prokaryotic cells, some nucleic acid sequences in the genes or cDNAs of eukaryotic nucleic acids can encode targeting signals such as an N-terminal mitochondrial or other targeting signal, which can be removed before transformation into prokaryotic host cells, if desired. For example, removal of a mitochondrial leader sequence led to increased expression in *E. coli*
25 (Hoffmeister et al., *J. Biol. Chem.* 280:4329-4338 (2005)). For exogenous expression in yeast or other eukaryotic cells, genes can be expressed in the cytosol without the addition of leader sequence, or can be targeted to mitochondrion or other organelles, or targeted for secretion, by the addition of a suitable targeting sequence such as a mitochondrial targeting or secretion signal suitable for the host cells. Thus, it is understood that appropriate modifications to a nucleic acid
30 sequence to remove or include a targeting sequence can be incorporated into an exogenous nucleic acid sequence to impart desirable properties. Furthermore, genes can be subjected to codon optimization with techniques well known in the art to achieve optimized expression of the proteins.

An expression vector or vectors can be constructed to include one or more 1,3-butanediol biosynthetic pathway encoding nucleic acids as exemplified herein operably linked to expression control sequences functional in the host organism. Expression vectors applicable for use in the microbial host organisms of the invention include, for example, plasmids, phage vectors, viral
5 vectors, episomes and artificial chromosomes, including vectors and selection sequences or markers operable for stable integration into a host chromosome. Additionally, the expression vectors can include one or more selectable marker genes and appropriate expression control sequences. Selectable marker genes also can be included that, for example, provide resistance to antibiotics or toxins, complement auxotrophic deficiencies, or supply critical nutrients not in the
10 culture media. Expression control sequences can include constitutive and inducible promoters, transcription enhancers, transcription terminators, and the like which are well known in the art. When two or more exogenous encoding nucleic acids are to be co-expressed, both nucleic acids can be inserted, for example, into a single expression vector or in separate expression vectors. For single vector expression, the encoding nucleic acids can be operationally linked to one
15 common expression control sequence or linked to different expression control sequences, such as one inducible promoter and one constitutive promoter. The transformation of exogenous nucleic acid sequences involved in a metabolic or synthetic pathway can be confirmed using methods well known in the art. Such methods include, for example, nucleic acid analysis such as Northern blots or polymerase chain reaction (PCR) amplification of mRNA, or
20 immunoblotting for expression of gene products, or other suitable analytical methods to test the expression of an introduced nucleic acid sequence or its corresponding gene product. It is understood by those skilled in the art that the exogenous nucleic acid is expressed in a sufficient amount to produce the desired product, and it is further understood that expression levels can be optimized to obtain sufficient expression using methods well known in the art and as disclosed
25 herein.

In some embodiments, the present invention provides a method for producing 1,3-BDO that includes culturing a non-naturally occurring microbial organism having a 1,3-BDO pathway having at least one exogenous nucleic acid encoding a 1,3-BDO pathway enzyme or protein expressed in a sufficient amount to produce 1,3-BDO, under conditions and for a sufficient
30 period of time to produce 1,3-BDO. In some embodiments, the 1,3-BDO pathway includes Methanol methyltransferase (MtaB), Corrinoid protein (MtaC), Methyltetrahydrofolate:corrinoid protein methyltransferase (MtaA), Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), Corrinoid iron-sulfur protein (AcsD), Nickel-protein assembly protein (AcsF & CooC), Ferredoxin (Orf7), Acetyl-CoA synthase (AcsB & AcsC), Carbon monoxide dehydrogenase

- (AcsA), Hydrogenase (Hyd), Acetoacetyl-CoA thiolase (AtoB), Acetoacetyl-CoA reductase (CoA-dependent, aldehyde forming), 3-oxobutyraldehyde reductase (ketone reducing), 3-hydroxybutyraldehyde reductase, Acetoacetyl-CoA reductase (CoA-dependent, alcohol forming), 3-oxobutyraldehyde reductase (aldehyde reducing), 4-hydroxy,2-butanone reductase,
- 5 Acetoacetyl-CoA reductase (ketone reducing), 3-hydroxybutyryl-CoA reductase (aldehyde forming), 3-hydroxybutyryl-CoA reductase (alcohol forming), 3-hydroxybutyryl-CoA transferase, 3-hydroxybutyryl-CoA hydrolase, 3-hydroxybutyryl-CoA synthetase, 3-hydroxybutyrate dehydrogenase, 3-hydroxybutyrate reductase, acetoacetyl-CoA transferase, acetoacetyl-CoA hydrolase, acetoacetyl-CoA synthetase, or acetoacetate reductase.
- 10 In other embodiments, the present invention provides a method for producing 1,3-BDO that includes culturing a non-naturally occurring microbial organism having a 1,3-BDO pathway having at least one exogenous nucleic acid encoding a 1,3-BDO pathway enzyme or protein expressed in a sufficient amount to produce 1,3-BDO, under conditions and for a sufficient period of time to produce 1,3-BDO. The 1,3-BDO pathway includes Formate dehydrogenase,
- 15 Formyltetrahydrofolate synthetase, Methenyltetrahydrofolate cyclohydrolase, Methylenetetrahydrofolate dehydrogenase, Methylenetetrahydrofolate reductase, Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), Corrinoid iron-sulfur protein (AcsD), Nickel-protein assembly protein (AcsF & CooC), Ferredoxin (Orf7), Acetyl-CoA synthase (AcsB & AcsC), Carbon monoxide dehydrogenase (AcsA), Hydrogenase (Hyd),
- 20 Acetoacetyl-CoA thiolase (AtoB), Acetoacetyl-CoA reductase (CoA-dependent, aldehyde forming), 3-oxobutyraldehyde reductase (ketone reducing), 3-hydroxybutyraldehyde reductase, Acetoacetyl-CoA reductase (CoA-dependent, alcohol forming), 3-oxobutyraldehyde reductase (aldehyde reducing), 4-hydroxy,2-butanone reductase, Acetoacetyl-CoA reductase (ketone reducing), 3-hydroxybutyryl-CoA reductase (aldehyde forming), 3-hydroxybutyryl-CoA
- 25 reductase (alcohol forming), 3-hydroxybutyryl-CoA transferase, 3-hydroxybutyryl-CoA hydrolase, 3-hydroxybutyryl-CoA synthetase, 3-hydroxybutyrate dehydrogenase, 3-hydroxybutyrate reductase, acetoacetyl-CoA transferase, acetoacetyl-CoA hydrolase, acetoacetyl-CoA synthetase, or acetoacetate reductase.

- In some embodiments, culturing the non-naturally occurring microbial organism includes
- 30 culturing under conditions and for a sufficient period of time to produce 1,3-BDO. In some embodiments, culturing is performed in a substantially anaerobic culture medium. In some embodiments, at least one exogenous nucleic acid of the microbial organism is a heterologous nucleic acid. As described above, the cultured non-naturally occurring microbial organisms can

have any number of exogenous nucleic acids in a 1,3-BDO pathway including 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, that is up to all the nucleic acids encoding a 1,3-BDO pathway. Non-naturally occurring microbial organisms of the invention can utilize a carbon feedstock selected from 1) methanol and CO, 2) methanol, CO₂, and H₂, 3) methanol, CO, CO₂, and H₂, 4) methanol and synthesis gas comprising CO and H₂, 5) methanol and synthesis gas comprising CO, CO₂, and H₂, 6) one or more carbohydrates, 7) methanol and one or more carbohydrates, and 8) methanol, in some embodiments or a carbon feedstock selected from 1) CO, 2) CO₂ and H₂, 3) CO, CO₂, and H₂, 4) synthesis gas comprising CO and H₂, 5) synthesis gas comprising CO, CO₂, and H₂, and 6) one or more carbohydrates, in other embodiments.

Suitable purification and/or assays to test for the production of 1,3-butanediol can be performed using well known methods. Suitable replicates such as triplicate cultures can be grown for each engineered strain to be tested. For example, product and byproduct formation in the engineered production host can be monitored. The final product and intermediates, and other organic compounds, can be analyzed by methods such as HPLC (High Performance Liquid Chromatography), GC-MS (Gas Chromatography-Mass Spectroscopy) and LC-MS (Liquid Chromatography-Mass Spectroscopy) or other suitable analytical methods using routine procedures well known in the art. The release of product in the fermentation broth can also be tested with the culture supernatant. Byproducts and residual glucose can be quantified by HPLC using, for example, a refractive index detector for glucose and alcohols, and a UV detector for organic acids (Lin et al., *Biotechnol. Bioeng.* 90:775-779 (2005)), or other suitable assay and detection methods well known in the art. The individual enzyme or protein activities from the exogenous DNA sequences can also be assayed using methods well known in the art.

The 1,3-butanediol can be separated from other components in the culture using a variety of methods well known in the art. Such separation methods include, for example, extraction procedures as well as methods that include continuous liquid-liquid extraction, pervaporation, membrane filtration, membrane separation, reverse osmosis, electrodialysis, distillation, crystallization, centrifugation, extractive filtration, ion exchange chromatography, size exclusion chromatography, adsorption chromatography, and ultrafiltration. All of the above methods are well known in the art.

Any of the non-naturally occurring microbial organisms described herein can be cultured to produce and/or secrete the biosynthetic products of the invention. For example, the 1,3-butanediol producers can be cultured for the biosynthetic production of 1,3-butanediol.

For the production of 1,3-butanediol, the recombinant strains are cultured in a medium with carbon source and other essential nutrients. It is highly desirable to maintain anaerobic conditions in the fermenter to reduce the cost of the overall process. Such conditions can be obtained, for example, by first sparging the medium with nitrogen and then sealing the flasks with a septum and crimp-cap. For strains where growth is not observed anaerobically, microaerobic conditions can be applied by perforating the septum with a small hole for limited aeration. Exemplary anaerobic conditions have been described previously and are well-known in the art. Exemplary aerobic and anaerobic conditions are described, for example, in United States Patent application serial No. 11/891,602, filed August 10, 2007. Fermentations can be performed in a batch, fed-batch or continuous manner, as disclosed herein.

If desired, the pH of the medium can be maintained at a desired pH, in particular neutral pH, such as a pH of around 7 by addition of a base, such as NaOH or other bases, or acid, as needed to maintain the culture medium at a desirable pH. The growth rate can be determined by measuring optical density using a spectrophotometer (600 nm), and the glucose uptake rate by monitoring carbon source depletion over time.

Accordingly, given the teachings and guidance provided herein, those skilled in the art will understand that a non-naturally occurring microbial organism can be produced that secretes the biosynthesized compounds of the invention when grown on a carbon source such as syngas, methanol, or combinations of CO, CO₂, hydrogen, and the like. Such compounds include, for example, 1,3-butanediol and any of the intermediate metabolites in the 1,3-butanediol pathway. All that is required is to engineer in one or more of the required enzyme or protein activities to achieve biosynthesis of the desired compound or intermediate including, for example, inclusion of some or all of the 1,3-butanediol biosynthetic pathways.

The non-naturally occurring microbial organisms of the invention are constructed using methods well known in the art as exemplified herein to exogenously express at least one nucleic acid encoding a 1,3-butanediol pathway enzyme or protein in sufficient amounts to produce 1,3-butanediol. It is understood that the microbial organisms of the invention are cultured under conditions sufficient to produce 1,3-butanediol. Following the teachings and guidance provided herein, the non-naturally occurring microbial organisms of the invention can achieve biosynthesis of 1,3-butanediol resulting in intracellular concentrations between about 0.1-200 mM or more. Generally, the intracellular concentration of 1,3-butanediol is between about 3-150 mM, particularly between about 5-125 mM and more particularly between about 8-100 mM, including about 10 mM, 20 mM, 50 mM, 80 mM, or more. Intracellular concentrations between

and above each of these exemplary ranges also can be achieved from the non-naturally occurring microbial organisms of the invention.

In some embodiments, culture conditions include anaerobic or substantially anaerobic growth or maintenance conditions. Exemplary anaerobic conditions have been described previously and
5 are well known in the art. Exemplary anaerobic conditions for fermentation processes are described herein and are described, for example, in U.S. publication 2009/0047719, filed August 10, 2007. Any of these conditions can be employed with the non-naturally occurring microbial organisms as well as other anaerobic conditions well known in the art. Under such anaerobic conditions, the 1,3-butanediol producers can synthesize 1,3-butanediol at intracellular
10 concentrations of 5-10 mM or more as well as all other concentrations exemplified herein. It is understood that, even though the above description refers to intracellular concentrations, 1,3-butanediol producing microbial organisms can produce 1,3-butanediol intracellularly and/or secrete the product into the culture medium.

In addition to the culturing and fermentation conditions disclosed herein, growth condition for
15 achieving biosynthesis of 1,3-butanediol can include the addition of an osmoprotectant to the culturing conditions. In certain embodiments, the non-naturally occurring microbial organisms of the invention can be sustained, cultured or fermented as described herein in the presence of an osmoprotectant. Briefly, an osmoprotectant refers to a compound that acts as an osmolyte and helps a microbial organism as described herein survive osmotic stress. Osmoprotectants
20 include, but are not limited to, betaines, amino acids, and the sugar trehalose. Non-limiting examples of such are glycine betaine, proline betaine, dimethylthetin, dimethylsulfoniopropionate, 3-dimethylsulfonio-2-methylpropionate, pipercolic acid, dimethylsulfonioacetate, choline, L-carnitine and ectoine. In one aspect, the osmoprotectant is glycine betaine. It is understood to one of ordinary skill in the art that the amount and type of
25 osmoprotectant suitable for protecting a microbial organism described herein from osmotic stress will depend on the microbial organism used. The amount of osmoprotectant in the culturing conditions can be, for example, no more than about 0.1 mM, no more than about 0.5 mM, no more than about 1.0 mM, no more than about 1.5 mM, no more than about 2.0 mM, no more than about 2.5 mM, no more than about 3.0 mM, no more than about 5.0 mM, no more
30 than about 7.0 mM, no more than about 10 mM, no more than about 50 mM, no more than about 100 mM or no more than about 500 mM.

The culture conditions can include, for example, liquid culture procedures as well as fermentation and other large scale culture procedures. As described herein, particularly useful

yields of the biosynthetic products of the invention can be obtained under anaerobic or substantially anaerobic culture conditions

As described herein, one exemplary growth condition for achieving biosynthesis of 1,3-butanediol includes anaerobic culture or fermentation conditions. In certain embodiments, the
5 non-naturally occurring microbial organisms of the invention can be sustained, cultured or fermented under anaerobic or substantially anaerobic conditions. Briefly, anaerobic conditions refer to an environment devoid of oxygen. Substantially anaerobic conditions include, for example, a culture, batch fermentation or continuous fermentation such that the dissolved oxygen concentration in the medium remains between 0 and 10% of saturation. Substantially
10 anaerobic conditions also includes growing or resting cells in liquid medium or on solid agar inside a sealed chamber maintained with an atmosphere of less than 1% oxygen. The percent of oxygen can be maintained by, for example, sparging the culture with an N₂/CO₂ mixture or other suitable non-oxygen gas or gases.

The culture conditions described herein can be scaled up and grown continuously for
15 manufacturing of 1,3-butanediol. Exemplary growth procedures include, for example, fed-batch fermentation and batch separation; fed-batch fermentation and continuous separation, or continuous fermentation and continuous separation. All of these processes are well known in the art. Fermentation procedures are particularly useful for the biosynthetic production of commercial quantities of 1,3-butanediol. Generally, and as with non-continuous culture
20 procedures, the continuous and/or near-continuous production of 1,3-butanediol will include culturing a non-naturally occurring 1,3-butanediol producing organism of the invention in sufficient nutrients and medium to sustain and/or nearly sustain growth in an exponential phase. Continuous culture under such conditions can include, for example, growth for 1 day, 2, 3, 4, 5, 6 or 7 days or more. Additionally, continuous culture can include longer time periods of 1
25 week, 2, 3, 4 or 5 or more weeks and up to several months. Alternatively, organisms of the invention can be cultured for hours, if suitable for a particular application. It is to be understood that the continuous and/or near-continuous culture conditions also can include all time intervals in between these exemplary periods. It is further understood that the time of culturing the microbial organism of the invention is for a sufficient period of time to produce a sufficient
30 amount of product for a desired purpose.

Fermentation procedures are well known in the art. Briefly, fermentation for the biosynthetic production of 1,3-butanediol can be utilized in, for example, fed-batch fermentation and batch separation; fed-batch fermentation and continuous separation, or continuous fermentation and

continuous separation. Examples of batch and continuous fermentation procedures are well known in the art.

In addition to the above fermentation procedures using the 1,3-butanediol producers of the invention for continuous production of substantial quantities of 1,3-butanediol, the 1,3-

5 butanediol producers also can be, for example, simultaneously subjected to chemical synthesis procedures to convert the product to other compounds or the product can be separated from the fermentation culture and sequentially subjected to chemical conversion to convert the product to other compounds, if desired.

To generate better producers, metabolic modeling can be utilized to optimize growth conditions.

10 Modeling can also be used to design gene knockouts that additionally optimize utilization of the pathway (see, for example, U.S. patent publications US 2002/0012939, US 2003/0224363, US 2004/0029149, US 2004/0072723, US 2003/0059792, US 2002/0168654 and US 2004/0009466, and U.S. Patent No. 7,127,379). Modeling analysis allows reliable predictions of the effects on cell growth of shifting the metabolism towards more efficient production of 1,3-butanediol.

15 One computational method for identifying and designing metabolic alterations favoring biosynthesis of a desired product is the OptKnock computational framework (Burgard et al., *Biotechnol. Bioeng.* 84:647-657 (2003)). OptKnock is a metabolic modeling and simulation program that suggests gene deletion or disruption strategies that result in genetically stable microorganisms which overproduce the target product. Specifically, the framework examines
20 the complete metabolic and/or biochemical network of a microorganism in order to suggest genetic manipulations that force the desired biochemical to become an obligatory byproduct of cell growth. By coupling biochemical production with cell growth through strategically placed gene deletions or other functional gene disruption, the growth selection pressures imposed on the engineered strains after long periods of time in a bioreactor lead to improvements in
25 performance as a result of the compulsory growth-coupled biochemical production. Lastly, when gene deletions are constructed there is a negligible possibility of the designed strains reverting to their wild-type states because the genes selected by OptKnock are to be completely removed from the genome. Therefore, this computational methodology can be used to either identify alternative pathways that lead to biosynthesis of a desired product or used in connection
30 with the non-naturally occurring microbial organisms for further optimization of biosynthesis of a desired product.

Briefly, OptKnock is a term used herein to refer to a computational method and system for modeling cellular metabolism. The OptKnock program relates to a framework of models and methods that incorporate particular constraints into flux balance analysis (FBA) models. These constraints include, for example, qualitative kinetic information, qualitative regulatory
5 information, and/or DNA microarray experimental data. OptKnock also computes solutions to various metabolic problems by, for example, tightening the flux boundaries derived through flux balance models and subsequently probing the performance limits of metabolic networks in the presence of gene additions or deletions. OptKnock computational framework allows the construction of model formulations that allow an effective query of the performance limits of
10 metabolic networks and provides methods for solving the resulting mixed-integer linear programming problems. The metabolic modeling and simulation methods referred to herein as OptKnock are described in, for example, U.S. publication 2002/0168654, filed January 10, 2002, in International Patent No. PCT/US02/00660, filed January 10, 2002, and U.S. publication 2009/0047719, filed August 10, 2007.

Another computational method for identifying and designing metabolic alterations favoring biosynthetic production of a product is a metabolic modeling and simulation system termed SimPheny®. This computational method and system is described in, for example, U.S.
15 publication 2003/0233218, filed June 14, 2002, and in International Patent Application No. PCT/US03/18838, filed June 13, 2003. SimPheny® is a computational system that can be used to produce a network model *in silico* and to simulate the flux of mass, energy or charge through
20 the chemical reactions of a biological system to define a solution space that contains any and all possible functionalities of the chemical reactions in the system, thereby determining a range of allowed activities for the biological system. This approach is referred to as constraints-based modeling because the solution space is defined by constraints such as the known stoichiometry
25 of the included reactions as well as reaction thermodynamic and capacity constraints associated with maximum fluxes through reactions. The space defined by these constraints can be interrogated to determine the phenotypic capabilities and behavior of the biological system or of its biochemical components.

These computational approaches are consistent with biological realities because biological
30 systems are flexible and can reach the same result in many different ways. Biological systems are designed through evolutionary mechanisms that have been restricted by fundamental constraints that all living systems must face. Therefore, constraints-based modeling strategy embraces these general realities. Further, the ability to continuously impose further restrictions

on a network model via the tightening of constraints results in a reduction in the size of the solution space, thereby enhancing the precision with which physiological performance or phenotype can be predicted.

Given the teachings and guidance provided herein, those skilled in the art can be able to apply various computational frameworks for metabolic modeling and simulation to design and implement biosynthesis of a desired compound in host microbial organisms. Such metabolic modeling and simulation methods include, for example, the computational systems exemplified above as SimPheny® and OptKnock. For illustration of the invention, some methods are described herein with reference to the OptKnock computation framework for modeling and simulation. Those skilled in the art will know how to apply the identification, design and implementation of the metabolic alterations using OptKnock to any of such other metabolic modeling and simulation computational frameworks and methods well known in the art.

The methods described above will provide one set of metabolic reactions to disrupt. Elimination of each reaction within the set or metabolic modification can result in a desired product as an obligatory product during the growth phase of the organism. Because the reactions are known, a solution to the bilevel OptKnock problem also will provide the associated gene or genes encoding one or more enzymes that catalyze each reaction within the set of reactions. Identification of a set of reactions and their corresponding genes encoding the enzymes participating in each reaction is generally an automated process, accomplished through correlation of the reactions with a reaction database having a relationship between enzymes and encoding genes.

Once identified, the set of reactions that are to be disrupted in order to achieve production of a desired product are implemented in the target cell or organism by functional disruption of at least one gene encoding each metabolic reaction within the set. One particularly useful means to achieve functional disruption of the reaction set is by deletion of each encoding gene. However, in some instances, it can be beneficial to disrupt the reaction by other genetic aberrations including, for example, mutation, deletion of regulatory regions such as promoters or cis binding sites for regulatory factors, or by truncation of the coding sequence at any of a number of locations. These latter aberrations, resulting in less than total deletion of the gene set can be useful, for example, when rapid assessments of the coupling of a product are desired or when genetic reversion is less likely to occur.

To identify additional productive solutions to the above described bilevel OptKnock problem which lead to further sets of reactions to disrupt or metabolic modifications that can result in the biosynthesis, including growth-coupled biosynthesis of a desired product, an optimization method, termed integer cuts, can be implemented. This method proceeds by iteratively solving the OptKnock problem exemplified above with the incorporation of an additional constraint referred to as an integer cut at each iteration. Integer cut constraints effectively prevent the solution procedure from choosing the exact same set of reactions identified in any previous iteration that obligatorily couples product biosynthesis to growth. For example, if a previously identified growth-coupled metabolic modification specifies reactions 1, 2, and 3 for disruption, then the following constraint prevents the same reactions from being simultaneously considered in subsequent solutions. The integer cut method is well known in the art and can be found described in, for example, Burgard et al., *Biotechnol. Prog.* 17:791-797 (2001). As with all methods described herein with reference to their use in combination with the OptKnock computational framework for metabolic modeling and simulation, the integer cut method of reducing redundancy in iterative computational analysis also can be applied with other computational frameworks well known in the art including, for example, SimPheny®.

The methods exemplified herein allow the construction of cells and organisms that biosynthetically produce a desired product, including the obligatory coupling of production of a target biochemical product to growth of the cell or organism engineered to harbor the identified genetic alterations. Therefore, the computational methods described herein allow the identification and implementation of metabolic modifications that are identified by an *in silico* method selected from OptKnock or SimPheny®. The set of metabolic modifications can include, for example, addition of one or more biosynthetic pathway enzymes and/or functional disruption of one or more metabolic reactions including, for example, disruption by gene deletion.

As discussed above, the OptKnock methodology was developed on the premise that mutant microbial networks can be evolved towards their computationally predicted maximum-growth phenotypes when subjected to long periods of growth selection. In other words, the approach leverages an organism's ability to self-optimize under selective pressures. The OptKnock framework allows for the exhaustive enumeration of gene deletion combinations that force a coupling between biochemical production and cell growth based on network stoichiometry. The identification of optimal gene/reaction knockouts requires the solution of a bilevel optimization problem that chooses the set of active reactions such that an optimal growth solution for the

resulting network overproduces the biochemical of interest (Burgard et al., *Biotechnol. Bioeng.* 84:647-657 (2003)).

An *in silico* stoichiometric model of *E. coli* metabolism can be employed to identify essential genes for metabolic pathways as exemplified previously and described in, for example, U.S.

5 patent publications US 2002/0012939, US 2003/0224363, US 2004/0029149, US 2004/0072723, US 2003/0059792, US 2002/0168654 and US 2004/0009466, and in U.S. Patent No. 7,127,379.

As disclosed herein, the OptKnock mathematical framework can be applied to pinpoint gene deletions leading to the growth-coupled production of a desired product. Further, the solution of the bilevel OptKnock problem provides only one set of deletions. To enumerate all meaningful
10 solutions, that is, all sets of knockouts leading to growth-coupled production formation, an optimization technique, termed integer cuts, can be implemented. This entails iteratively solving the OptKnock problem with the incorporation of an additional constraint referred to as an integer cut at each iteration, as discussed above.

As disclosed herein, a nucleic acid encoding a desired activity of a 1,3-butanediol pathway can
15 be introduced into a host organism. In some cases, it can be desirable to modify an activity of a 1,3-butanediol pathway enzyme or protein to increase production of 1,3-butanediol. For example, known mutations that increase the activity of a protein or enzyme can be introduced into an encoding nucleic acid molecule. Additionally, optimization methods can be applied to increase the activity of an enzyme or protein and/or decrease an inhibitory activity, for example,
20 decrease the activity of a negative regulator.

One such optimization method is directed evolution. Directed evolution is a powerful approach that involves the introduction of mutations targeted to a specific gene in order to improve and/or alter the properties of an enzyme. Improved and/or altered enzymes can be identified through the development and implementation of sensitive high-throughput screening assays that allow
25 the automated screening of many enzyme variants (for example, $>10^4$). Iterative rounds of mutagenesis and screening typically are performed to afford an enzyme with optimized properties. Computational algorithms that can help to identify areas of the gene for mutagenesis also have been developed and can significantly reduce the number of enzyme variants that need to be generated and screened. Numerous directed evolution technologies have been developed
30 (for reviews, see Hibbert et al., *Biomol. Eng* 22:11-19 (2005); Huisman and Lalonde, *In Biocatalysis in the pharmaceutical and biotechnology industries* pgs. 717-742 (2007), Patel (ed.), CRC Press; Otten and Quax, *Biomol. Eng* 22:1-9 (2005).; and Sen et al., *Appl Biochem. Biotechnol* 143:212-223 (2007)) to be effective at creating diverse variant libraries, and

these methods have been successfully applied to the improvement of a wide range of properties across many enzyme classes. Enzyme characteristics that have been improved and/or altered by directed evolution technologies include, for example: selectivity/specificity, for conversion of non-natural substrates; temperature stability, for robust high temperature processing; pH
5 stability, for bioprocessing under lower or higher pH conditions; substrate or product tolerance, so that high product titers can be achieved; binding (K_m), including broadening substrate binding to include non-natural substrates; inhibition (K_i), to remove inhibition by products, substrates, or key intermediates; activity (k_{cat}), to increase enzymatic reaction rates to achieve desired flux; expression levels, to increase protein yields and overall pathway flux; oxygen stability, for
10 operation of air sensitive enzymes under aerobic conditions; and anaerobic activity, for operation of an aerobic enzyme in the absence of oxygen.

A number of exemplary methods have been developed for the mutagenesis and diversification of genes to target desired properties of specific enzymes. Such methods are well known to those skilled in the art. Any of these can be used to alter and/or optimize the activity of a 1,3-
15 butanediol pathway enzyme or protein. Such methods include, but are not limited to EpPCR, which introduces random point mutations by reducing the fidelity of DNA polymerase in PCR reactions (Pritchard et al., *J Theor. Biol.* 234:497-509 (2005)); Error-prone Rolling Circle Amplification (epRCA), which is similar to epPCR except a whole circular plasmid is used as the template and random 6-mers with exonuclease resistant thiophosphate linkages on the last 2
20 nucleotides are used to amplify the plasmid followed by transformation into cells in which the plasmid is re-circularized at tandem repeats (Fujii et al., *Nucleic Acids Res.* 32:e145 (2004); and Fujii et al., *Nat. Protoc.* 1:2493-2497 (2006)); DNA or Family Shuffling, which typically involves digestion of two or more variant genes with nucleases such as Dnase I or EndoV to generate a pool of random fragments that are reassembled by cycles of annealing and extension
25 in the presence of DNA polymerase to create a library of chimeric genes (Stemmer, *Proc Natl Acad Sci USA* 91:10747-10751 (1994); and Stemmer, *Nature* 370:389-391 (1994)); Staggered Extension (StEP), which entails template priming followed by repeated cycles of 2 step PCR with denaturation and very short duration of annealing/extension (as short as 5 sec) (Zhao et al., *Nat. Biotechnol.* 16:258-261 (1998)); Random Priming Recombination (RPR), in which random
30 sequence primers are used to generate many short DNA fragments complementary to different segments of the template (Shao et al., *Nucleic Acids Res* 26:681-683 (1998)).

Additional methods include Heteroduplex Recombination, in which linearized plasmid DNA is used to form heteroduplexes that are repaired by mismatch repair (Volkov et al, *Nucleic Acids*

Res. 27:e18 (1999); and Volkov et al., *Methods Enzymol.* 328:456-463 (2000)); Random Chimera-
genesis on Transient Templates (RACHITT), which employs Dnase I fragmentation and
size fractionation of single stranded DNA (ssDNA) (Coco et al., *Nat. Biotechnol.* 19:354-359
(2001)); Recombined Extension on Truncated templates (RETT), which entails template
5 switching of unidirectionally growing strands from primers in the presence of unidirectional
ssDNA fragments used as a pool of templates (Lee et al., *J. Molec. Catalysis* 26:119-129
(2003)); Degenerate Oligonucleotide Gene Shuffling (DOGS), in which degenerate primers are
used to control recombination between molecules; (Bergquist and Gibbs, *Methods Mol.Biol*
352:191-204 (2007); Bergquist et al., *Biomol.Eng* 22:63-72 (2005); Gibbs et al., *Gene* 271:13-20
10 (2001)); Incremental Truncation for the Creation of Hybrid Enzymes (ITCHY), which creates a
combinatorial library with 1 base pair deletions of a gene or gene fragment of interest
(Ostermeier et al., *Proc. Natl. Acad. Sci. USA* 96:3562-3567 (1999); and Ostermeier et al., *Nat.*
Biotechnol. 17:1205-1209 (1999)); Thio-Incremental Truncation for the Creation of Hybrid
Enzymes (THIO-ITCHY), which is similar to ITCHY except that phosphothioate dNTPs are
15 used to generate truncations (Lutz et al., *Nucleic Acids Res* 29:E16 (2001)); SCRATCHY, which
combines two methods for recombining genes, ITCHY and DNA shuffling (Lutz et al., *Proc.*
Natl. Acad. Sci. USA 98:11248-11253 (2001)); Random Drift Mutagenesis (RNDM), in which
mutations made via epPCR are followed by screening/selection for those retaining usable
activity (Bergquist et al., *Biomol. Eng.* 22:63-72 (2005)); Sequence Saturation Mutagenesis
20 (SeSaM), a random mutagenesis method that generates a pool of random length fragments using
random incorporation of a phosphothioate nucleotide and cleavage, which is used as a template
to extend in the presence of “universal” bases such as inosine, and replication of an inosine-
containing complement gives random base incorporation and, consequently, mutagenesis (Wong
et al., *Biotechnol. J.* 3:74-82 (2008); Wong et al., *Nucleic Acids Res.* 32:e26 (2004); and Wong
25 et al., *Anal. Biochem.* 341:187-189 (2005)); Synthetic Shuffling, which uses overlapping
oligonucleotides designed to encode “all genetic diversity in targets” and allows a very high
diversity for the shuffled progeny (Ness et al., *Nat. Biotechnol.* 20:1251-1255 (2002));
Nucleotide Exchange and Excision Technology NexT, which exploits a combination of dUTP
incorporation followed by treatment with uracil DNA glycosylase and then piperidine to
30 perform endpoint DNA fragmentation (Muller et al., *Nucleic Acids Res.* 33:e117 (2005)).

Further methods include Sequence Homology-Independent Protein Recombination (SHIPREC),
in which a linker is used to facilitate fusion between two distantly related or unrelated genes, and
a range of chimeras is generated between the two genes, resulting in libraries of single-crossover
hybrids (Sieber et al., *Nat. Biotechnol.* 19:456-460 (2001)); Gene Site Saturation Mutagenesis™

(GSSM™), in which the starting materials include a supercoiled double stranded DNA (dsDNA) plasmid containing an insert and two primers which are degenerate at the desired site of mutations (Kretz et al., *Methods Enzymol.* 388:3-11 (2004)); Combinatorial Cassette Mutagenesis (CCM), which involves the use of short oligonucleotide cassettes to replace limited regions with a large number of possible amino acid sequence alterations (Reidhaar-Olson et al. *Methods Enzymol.* 208:564-586 (1991); and Reidhaar-Olson et al. *Science* 241:53-57 (1988)); Combinatorial Multiple Cassette Mutagenesis (CMCM), which is essentially similar to CCM and uses epPCR at high mutation rate to identify hot spots and hot regions and then extension by CMCM to cover a defined region of protein sequence space (Reetz et al., *Angew. Chem. Int. Ed Engl.* 40:3589-3591 (2001)); the Mutator Strains technique, in which conditional *ts* mutator plasmids, utilizing the *mutD5* gene, which encodes a mutant subunit of DNA polymerase III, to allow increases of 20 to 4000-X in random and natural mutation frequency during selection and block accumulation of deleterious mutations when selection is not required (Selifonova et al., *Appl. Environ. Microbiol.* 67:3645-3649 (2001)); Low et al., *J. Mol. Biol.* 260:359-3680 (1996)).

Additional exemplary methods include Look-Through Mutagenesis (LTM), which is a multidimensional mutagenesis method that assesses and optimizes combinatorial mutations of selected amino acids (Rajpal et al., *Proc. Natl. Acad. Sci. USA* 102:8466-8471 (2005)); Gene Reassembly, which is a DNA shuffling method that can be applied to multiple genes at one time or to create a large library of chimeras (multiple mutations) of a single gene (Tunable GeneReassembly™ (TGR™) Technology supplied by Verenum Corporation), *in Silico* Protein Design Automation (PDA), which is an optimization algorithm that anchors the structurally defined protein backbone possessing a particular fold, and searches sequence space for amino acid substitutions that can stabilize the fold and overall protein energetics, and generally works most effectively on proteins with known three-dimensional structures (Hayes et al., *Proc. Natl. Acad. Sci. USA* 99:15926-15931 (2002)); and Iterative Saturation Mutagenesis (ISM), which involves using knowledge of structure/function to choose a likely site for enzyme improvement, performing saturation mutagenesis at chosen site using a mutagenesis method such as Stratagene QuikChange (Stratagene; San Diego CA), screening/selecting for desired properties, and, using improved clone(s), starting over at another site and continue repeating until a desired activity is achieved (Reetz et al., *Nat. Protoc.* 2:891-903 (2007); and Reetz et al., *Angew. Chem. Int. Ed Engl.* 45:7745-7751 (2006)).

Any of the aforementioned methods for mutagenesis can be used alone or in any combination. Additionally, any one or combination of the directed evolution methods can be used in conjunction with adaptive evolution techniques, as described herein.

It is understood that modifications which do not substantially affect the activity of the various
5 embodiments of this invention are also included within the definition of the invention provided herein. Accordingly, the following examples are intended to illustrate but not limit the present invention.

EXAMPLE I

Production of 1,3-BDO Using Methanol, CO, and/or CO₂ as Carbon Feedstock

10 This example shows how a non-naturally occurring organism can be constructed to produce 1,3-BDO from methanol, CO, and/or CO₂ as the carbon source.

The first step in the cloning and expression process is to express in *E. coli* the minimal set of genes (e.g., *MtaA*, *MtaB*, and *MtaC*) necessary to produce Me-THF from methanol. These methyltransferase activities require Coenzyme B₁₂ (cobalamin) as a cofactor. In *Moorella*
15 *thermoacetica*, a cascade of methyltransferase proteins mediate incorporation of methanol derived methyl groups into the acetyl-CoA synthase pathway. Recent work (Das et al., *Proteins* 67.1:167-176 (2007)) indicates that *MtaABC* are encoded by Moth_1208-09 and Moth_2346. These genes are cloned via proof-reading PCR and linked together for expression in a high-copy number vector such as pZE22-S under control of the repressible PA1-lacO1 promoter (Lutz and
20 Bujard, *Nucleic Acids Res.* 25.6:1203-1210 (1997)). Cloned genes are verified by PCR and or restriction enzyme mapping to demonstrate construction and insertion of the 3-gene set into the expression vector. DNA sequencing of the presumptive clones is carried out to confirm the expected sequences of each gene. Once confirmed, the final construct is expressed in *E. coli* K-12 (MG1655) cells by addition of IPTG inducer between 0.05 and 1 mM final concentration.
25 Expression of the cloned genes is monitored using SDS-PAGE of whole cell extracts. To optimize levels of soluble vs. pellet (potentially inclusion body origin) protein, the affect of titration of the promoter on these levels can be examined. If no acceptable expression is obtained, higher or lower copy number vectors or variants in promoter strength are tested.

To determine if expression of the *MtaABC* proteins from *M. thermoacetica* confers upon *E. coli*
30 the ability to transfer methyl groups from methanol to tetrahydrofolate (THF) the recombinant strain is fed methanol at various concentrations. Activity of the methyltransferase system is assayed anaerobically as described for vanillate as a methyl source in *M. thermoacetica* (Naidu

and Ragsdale, *J. Bacteriol.* 183.11:3276-3281 (2001)) or for *Methanosarcina barkeri* methanol methyltransferase (Sauer et al., *Eur. J. Biochem.* 243.3:670-677 (1997); Tallant et al., *J. Biol. Chem.* 276.6:4485-4493 (2001); Tallant and Krzycki, *J. Bacteriol.* 179.22:6902-6911 (1997); Tallant and Krzycki, *J. Bacteriol.* 178.5:1295-1301 (1996)). For a positive control, *M.*

5 *thermoacetica* cells are cultured in parallel and assayed anaerobically to confirm endogenous methyltransferase activity. Demonstration of dependence on exogenously added coenzyme B₁₂ confirms methanol:corrinoid methyltransferase activity in *E. coli*.

Once methyltransferase expression is achieved, further work is performed towards optimizing the expression. Titrating the promoter in the expression vector enables the testing of a range of
10 expression levels. This is then used as a guide towards the expression needed in single-copy, or allows the determination of whether or not a single-copy of these genes allows sufficient expression. If so, the methyltransferase genes are integrated into the chromosome as a single, synthetic operon. This entails targeted integration using RecET-based 'recombineering' (Angrand et al., *Nucleic Acids Res.* 27.17:e16 (1999); Muyrers et al., *Nucleic Acids Res.*

15 27.6:1555-1557 (1999); Zhang et al., *Nat. Genet.* 20.2:123-128 (1998)). A potential issue with RecET-based integration of a cassette and removal of a FRT or loxP-bounded selectable marker by FLP or Cre is the production of a recombination scar at each integration site. While problems caused by this can be minimized by a number of methods, other means that do not leave genomic scars are available. The standard alternative is to introduce the desired genes
20 using integrative 'suicide' plasmids coupled to counter-selection such as that allowed by the *Bacillus sacB* gene (Link et al., *J. Bacteriol.* 179.20:6228-6237 (1997)); in this way, markerless and scar less insertions at any location in the *E. coli* chromosome can be generated. The final goal is a strain of *E. coli* K-12 expressing methanol:corrinoid methyltransferase activity under an inducible promoter and in single copy (chromosomally integrated).

25 Using standard PCR methods, entire ACS/CODH operons are assembled into low or medium copy number vectors such as pZA33-S (P15A-based) or pZS13-S (pSC101-based). As described for the methyltransferase genes, the structure and sequence of the cloned genes are confirmed. Expression is monitored via protein gel electrophoresis of whole-cell lysates grown under strictly anaerobic conditions with the requisite metals (Ni, Zn, Fe) and coenzyme B₁₂ provided.
30 As necessary, the gene cluster is modified for *E. coli* expression by identification and removal of any apparent terminators and introduction of consensus ribosomal binding sites chosen from sites known to be effective in *E. coli* (Barrick et al., *Nucleic Acids Res.* 22.7:1287-1295 (1994); Ringquist et al. *Mol. Microbiol.* 6.9:1219-1229). However, each gene cluster is cloned and

expressed in a manner parallel to its native structure and expression. This helps ensure the desired stoichiometry between the various gene products—most of which interact with each other. Once satisfactory expression of the CODH/ACS gene cluster under anaerobic conditions is achieved, the ability of cells expressing these genes to fix CO and/or CO₂ into cellular carbon is assayed. Initial conditions employ strictly anaerobically grown cells provided with exogenous glucose as a carbon and energy source via substrate-level phosphorylation or anaerobic respiration with nitrate as an electron acceptor. Additionally, exogenously provided CH₃-THF can be added to the medium.

The ACS/CODH genes are cloned and expressed in cells also expressing the methanol-methyltransferase system. This can be achieved by introduction of compatible plasmids expressing ACS/CODH into MTR-expressing cells. For added long-term stability, the ACS/CODH and MTR genes can also be integrated into the chromosome. After strains of *E. coli* capable of utilizing methanol to produce Me-THF and of expressing active CODH/ACS gene are made, they are assayed for the ability to utilize both methanol and syngas for incorporation into acetyl-CoA, acetate, and cell mass. Initial conditions employ strictly anaerobically grown cells provided with exogenous glucose as a carbon and energy source. Alternatively, or in addition to glucose, nitrate can be added to the fermentation broth to serve as an electron acceptor and initiator of growth. Anaerobic growth of *E. coli* on fatty acids, which are ultimately metabolized to acetyl-CoA, has been demonstrated in the presence of nitrate (Campbell et al., *Mol. Microbiol.* 47.3:793-805 (2003)). Oxygen can also be provided as long as its intracellular levels are maintained below any inhibition threshold of the engineered enzymes. ¹³C-labeled methanol, ¹³C-labeled bicarbonate or ¹³C-labeled CO are provided to the cells and analytical mass spectrometry is employed to measure incorporation of the labeled carbon into acetate and cell mass (e.g., proteinogenic amino acids).

An alternative or supplement to engineering the methanol methyltransferase system involves engineering the methyl branch of the Wood-Ljungdahl pathway to supply ACS/CODH with the methyl group. While *E. coli* possesses genes encoding enzymes capable of carrying out some of the necessary activities (*fdh*, *metF*, *folD*), it has been indicated that the methyl branch enzymes from acetogens may have significantly higher (50 – 100X) specific activities than those from non-acetogens (Morton et al., *Genetics and molecular biology of anaerobic bacteria* Springer Verlag, New York). The *M. thermoacetica* versions include formate dehydrogenase (*fdhA*, Moth_2312 – Moth_2313 alpha, Moth_2314 beta), formyl-tetrahydrofolate (THF) synthetase (Moth_0109), methenyl-THF cyclohydrolase/methylene-THF dehydrogenase (*folD*,

Moth_1516), methylene-THF reductase (*metF*, Moth_1191), and methyltransferase (*acsE*, Moth_1197). With the exception of the methyltransferase (*acsE*, Moth_1197) that will be cloned as part of the CODH/ACS cluster, these genes are cloned and combined into an expression vector designed to express these as a set. Cloning and expression of the methyl branch genes will be undertaken as described above. Initially, a high or medium copy number vector will be chosen (using ColE1 or P15A replicons). These genes can also be integrated into the *E. coli* chromosome.

The pyruvate ferredoxin oxidoreductase genes from *M. thermoacetica*, *D. africanus*, and *E. coli* can be cloned and expressed in strains exhibiting ACS/CODH activities. Conditions, promoters, etc., are described above. Given the large size of the PFOR genes and oxygen sensitivity of the corresponding enzymes, tests can be performed using low or single-copy plasmid vectors or single-copy chromosomal integrations. Activity assays described in ref. (Furdui and Ragsdale, *J. Biol. Chem.* 275.37:28494-28499 (2000)) can be applied to demonstrate activity. In addition, demonstration of growth on the gaseous carbon sources and methanol in the absence of an external electron acceptor will provide further evidence for PFOR activity *in vivo*.

The endogenous hydrogen-utilizing hydrogenase activity of the host organism can be tested by growing the cells as described above in the presence and absence of hydrogen. If a dramatic shift towards the formation of more reduced products during fermentation is observed (e.g., increased ethanol as opposed to acetate), this indicates that endogenous hydrogenase activity is sufficiently active. In this case, no heterologous hydrogenases are cloned and expressed. If the native enzymes do not have sufficient activity or reduce the needed acceptor, the genes encoding an individual hydrogenase complex can be cloned and expressed in strains exhibiting ACS/CODH activities. Conditions, promoters, etc., are described above.

The nonnative genes needed for 1,3-butanediol synthesis are cloned on expression plasmids as described previously. The host strain also expresses methanol methyltransferase activity, CODH/ACS activity, and possibly PFOR and hydrogenase activities. At this point, these (CODH/ACS, etc.) genes can be integrated into the genome and expressed from promoters that can be used constitutively or with inducers (i.e., PA1-lacO1 is inducible in cells containing *lacI* or is otherwise constitutive). Once expression and yields of 1,3-BDO are optimized, the base strain can be further modified by integration of a single copy of these genes at a neutral locus. Given the relatively limited number of enzymes (at minimum, 3, and at most, 4), one can construct an artificial operon encoding the required genes. This operon can be introduced using integrative plasmids and is coupled to counter-selection methods such as that allowed by the

Bacillus sacB gene (Link et al., *J. Bacteriol.* 179.20:6228-6237 (1997)). In this way, markerless and scar less insertions at any location in the *E. coli* chromosome can be generated.

Optimization involves altering gene order as well as ribosomal binding sites and promoters.

To over express any native genes, for example, the native *atoB* (b2224) gene of *E. coli* which
5 can serve as an alternative to the *C. acetobutylicum* acetyl-coenzyme A [CoA] acetyltransferase required for 1,3-butanediol production, RecET-based methods are applied to integrate a stronger upstream promoter. In the case of *atoB*, this gene is the last in an operon and the next gene downstream (*yfaP*) is both non-essential and in the opposite orientation. Therefore, polarity should not be an issue. A cassette containing a selectable marker such as spectinomycin
10 resistance or chloramphenicol resistance flanked by FRT or loxP sites is used to select for introduction of a strong constitutive promoter (e.g., pL). Once the correct clone is obtained and validated, using qRT-PCR, FLP or Cre expression is used to select for removal of the FRT- or loxP-bounded marker.

EXAMPLE II

15 Cloning And Expression Of *Moorella Thermoacetica* ACS/CODH Encoding Genes

This example describes the creation of *E. coli* plasmids that express the *M. thermoacetica* ACS/CODH operon genes including those used for CODH, ACS, methyltransferase, and the corrinoid iron-sulfur protein. This example further describes the expression these in *E. coli* resulting in observable CO oxidation activity, methyltransferase activity, and corrinoid iron-sulfur protein activity. Finally, this example demonstrates that *E. coli* tolerates high CO
20 concentrations, and may even consume CO when the CO-utilizing gene products from *M. thermoacetica* are expressed.

Expression vectors were chosen from the set described by Lutz and Bujard (Lutz and Bujard, *Nucleic Acids Res.* 25.6:1203-1210 (1997)); these come with compatible replicons that cover a
25 range of copy numbers. Additionally, each contains prA1-lacO1; this T7 early gene promoter is inducible by IPTG and can lead to very high levels of transcription in the presence of IPTG and represses in other conditions. The ACS/CODH-encoding operon was cloned from Moth_1204 (*cooC*) to Moth_1197; a second version containing only Moth_1203 to Moth_1197 was also constructed. Both of these fragments (10 – 11 kbp) were confirmed by DNA sequence analysis.
30 These were constructed in both p15A and ColE1-based vectors for medium to high copy numbers.

To estimate the final concentrations recombinant proteins, SDS-PAGE followed by Western blot analyses were performed on the same cell extracts used in the CO oxidation, ACS, methyltransferase, and corrinoid Fe-S assays. The antisera used were polyclonal to purified *M. thermoacetica* ACS/CODH and Mtr proteins and were visualized using an alkaline phosphatase-linked goat-anti-rabbit secondary antibody. The Westerns Blots are shown in Figure 6. Amounts of CODH in ACS90 and ACS91 were estimated at 50 ng by comparison to the control lanes.

A carbon monoxide oxidation assay (Seravalli et al., *Biochemistry* 43.13:3944-3955 (2004)) was used to test whether or not functional expression of the CODH-encoding genes from *M.*

thermoacetica was achieved. Cultures of *E. coli* MG1655 containing either an empty vector, or the vectors expressing "Acs90" or "Acs91" were grown in Terrific Broth under anaerobic conditions (with supplements of cyanocobalamin, ferrous iron, and reducing agents) until reaching medium to high density at which point, IPTG was added to a final concentration of 0.2 mM to induce the promoter. After 3.5 hrs of growth at 37 °C, the cells were harvested and spun down prior to lysis with lysozyme and mild detergents. There is a benchmark figure of *M. thermoacetica* CODH specific activity, 500 U at 55C or ~60U at 25 °C. This assay employed reduction of methyl viologen in the presence of CO. This is measured at 578 nm in stoppered, anaerobic, glass cuvettes. Reactions positive for CO oxidation by CODH turned a deep violet color (see Figure 7). About 0.5% of the cellular protein was CODH as estimated by Western blotting; therefore, the data in Table 1 are approximately 50X less than the 500 U/mg activity of pure *M. thermoacetica* CODH. Nevertheless, this experiment did clearly demonstrate CO oxidation activity in recombinant *E. coli* with a much smaller amount in the negative controls. The small amount of CO oxidation (CH₃ viologen reduction) seen in the negative controls indicates that *E. coli* may have a limited ability to reduce CH₃ viologen.

Table 1. Crude extract CO Oxidation Activities

ACS90	7.7 mg/ml	ACS91	11.8 mg/ml	
Mta98	9.8 mg/ml	Mta99	11.2 mg/ml	
<u>Extract</u>	<u>Vol</u>	<u>OD/</u>	<u>U/ml</u>	<u>U/mg</u>
ACS90	10 microliters	0.073	.0376	0.049
ACS91	10 microliters	0.096	0.494	0.042
Mta99	10 microliters	0.0031	0.016	0.0014
ACS90	10 microliters	0.099	0.051	0.066
Mta99	25 microliters	0.012	0.025	0.0022
ACS91	25 microliters	0.215	0.443	0.037
Mta98	25 microliters	0.019	0.039	0.004
ACS91	10 microliters	0.129	0.66	0.056
Averages				
ACS90	0.057 U/mg			
ACS91	0.045 U/mg			
Mta99	0.0018 U/mg			

This assay is an *in vitro* reaction that synthesizes acetyl-CoA from methyl-tetrahydrofolate, CO, and CoA using ACS/CODH, methyltransferase, and CFeSP (Raybuck et al., *Biochemistry* 27.20:7698-7702 (1988)). By adding or leaving out each of the enzymes involved, this assay can be used for a wide range of experiments, from testing one or more purified enzymes or cell extracts for activity, to determining the kinetics of the reaction under various conditions or with limiting amounts of substrate or enzyme. Samples of the reaction taken at various time points are quenched with 1M HCl, which liberates acetate from the acetyl-CoA end product. After purification with Dowex columns, the acetate can be analyzed by chromatography, mass spectrometry, or by measuring radioactivity. The exact method can be determined by the specific substrates used in the reaction.

This assay was run in order to determine if the ACS/CODH operon expressed in *E. coli* expresses the Fe-S corrinoid protein activity. Therefore, ¹⁴C-labeled methyl-THF was used as a labeled substrate to measure acetate synthesis by radioactivity incorporation into isolated acetate samples. Six different conditions were tested:

1. Purified ACS/CODH, MeTr, and CFeSP as a positive control
2. Purified ACS/CODH with ACS90 cell extract
3. Purified ACS/CODH with ACS91 cell extract
4. Purified ACS/CODH, MeTr with ACS90 cell extract
- 5 5. Purified ACS/CODH, MeTr with ACS91 cell extract
6. Purified ACS/CODH, MeTr with as much ACS91 cell extract as possible (excluding the MES buffer)

The reaction was assembled in the anaerobic chamber in assay vials filled with CO. The total
 10 reaction volume was small compared to the vial volume, reagents were added prior to filling
 with CO, a gas-tight Hamilton syringe was used and the reagents were kept anaerobic. The
 reaction (~60ul total) consisted of the cell extract (except #1), CoA, Ti(III)citrate, MES (except
 #6), purified ACS/CODH, ¹⁴C-methyl-tetrahydrofolate, methyl-viologen, and ferredoxin.
 Additionally, purified MeTr was added to #1, #4-6 and purified CFeSP was added to #1.

15 The reaction was carried out in the anaerobic chamber in a sand bath at 55°. The final reagent
 added was the ¹⁴C-methyl-tetrahydrofolate, which started the reaction (t = 0s). An initial sample
 was taken immediately, followed by samples at 30 minutes, 1 hour, and 2 hours. These time
 points are not exact, as the 6 conditions were run concurrently (since this experiment was
 primarily a qualitative one). The 15 µL samples were added to 15 µL of 1M HCl in scintillation
 20 vials. After counting the reaction mixtures, it was determined that the corrinoid Fe-S protein in
 ACS90 extracts was active with total activity approaching approximately 1/5 of the positive
 control.

Within the ACS/CODH operon is encoded an essential methyltransferase activity that catalyzes
 the transfer of CH₃ from methyl-tetrahydrofolate to the ACS complex as part of the synthesis of
 25 acetyl-CoA (i.e. this is the step that the methyl and carbonyl paths join together). Within the
 operon in *M. thermoacetica*, the Mtr-encoding gene is Moth_1197 and comes after the main
 CODH and ACS subunits. Therefore, Mtr activity would constitute indirect evidence that the
 more proximal genes can be expressed.

Mtr activity was assayed by spectroscopy. Specifically, methylated CFeSP, with Co(III), has a
 30 small absorption peak at ~450nm, while non-methylated CFeSP, with Co(I), has a large peak at
 ~390nm. This spectrum is due to both the cobalt and iron-sulfur cluster chromophores.
 Additionally, it should be noted that the CFeSP can spontaneously oxidize to Co(II), which
 creates a broad absorption peak at ~470nm (Seravalli et al., *Biochemistry* 38.18:5728-5735
 (1999)). See Figure 8 for the results from *E. coli* cells containing ACS90.

To test whether or not *E. coli* can grow anaerobically in the presence of saturating amounts of CO we made up 120 ml serum bottles with 50 ml of Terrific Broth medium (plus NiCl₂, Fe(II)NH₄SO₄, and cyanocobalamin) in anaerobic conditions. One half of these bottles were equilibrated with nitrogen gas for 30 min. and one half was equilibrated with CO gas for 30 min.

- 5 An empty vector (pZA33) was used as a control and that and both ACS90 and ACS91 were tested with both N₂ and CO. All were grown for 36 hrs with shaking (250 rpm) at 37C. At the end of the 36 period, examination of the flasks showed high amounts of growth in all (Figure 9). The bulk of the observed growth occurred overnight with a long lag of some (low but visible) density. Inocula sizes were ~0.5 ml from *E. coli* stocks.
- 10 The results are shown in Table 2. Growth reached similar levels (by visual inspection) whether or not a strain was cultured in the presence of CO or not. Furthermore, the negative control had a final CO concentration of 930 micromolar vs. 688 and 728 micromolar for the ACS/CODH operon expressing strains. Clearly, the error in these measurements is high given the large standard deviations. Nevertheless, this test does allow two tentative conclusions: 1) *E. coli* can
- 15 tolerate exposure to CO under anaerobic conditions, and 2) *E. coli* cells expressing the ACS/CODH operon might be metabolizing some of the CO. The second conclusion is significantly less certain than the first.

Table 2. Carbon Monoxide Concentrations, 36 hours

<u>Strain and Growth Conditions</u>	<u>Final CO concentration (micromolar)</u>
pZA33-CO	930
ACS90-CO	638
	494
	734
	883
ave	687
SD	164
ACS91-CO	728
	812
	760
	611
ave	728
SD	85

EXAMPLE III

Enhancing The Yield of 1,3-Butanediol on Sugars with Wood-Ljungdahl Pathway Enzymes

In this example, we describe a non-naturally occurring microorganism expressing genes
5 encoding enzymes that catalyze the carbonyl-branch of the Wood-Ljungdahl pathway. Wood-Ljungdahl pathway enzymes assimilate carbon in the form of formate, CO and/or CO₂ into acetyl-CoA, which can subsequently be converted to useful chemical products such as 1,3-butanediol. The Wood-Ljungdahl pathway can also serve as a secondary carbon assimilation pathway during growth on other substrates such as glucose. Specifically, the conversion of one
10 mole of glucose to two moles of acetyl-CoA generates both reducing equivalents and CO₂. The WL pathway enzymes can harness the reducing equivalents to convert the CO₂ and/or formate to additional acetyl-CoA that can be further used for 1,3-butanediol formation.

1,3-Butanediol can be synthesized from acetyl-CoA by several alternate routes described previously in this application and shown in Figures 4 and 5. The maximum achievable 1,3-
15 butanediol yield for any of these pathways from glucose, for example, is 1 mol/mol (0.5 g/g) in the absence of the Wood-Ljungdahl pathway enzymes. Additional assimilation of CO₂ via Wood-Ljungdahl pathway further improves the yield to the stoichiometric theoretical maximum of 1.09 mol/mol (0.545 g/g). A predicted flux distribution for achieving the maximum theoretical yield is shown in Figure 10.

20 Additionally, methanol can be co-fed with a carbohydrate such as glucose to increase the yield of 1,3-butanediol. For example, utilizing glucose and methanol in a 1.0:0.4 ratio affords an increase from 1 mol 1,3-butanediol/mol glucose (0.5 g/g) to 1.2 mol 1,3-butanediol/mol glucose. A predicted flux distribution for achieving this increased yield is shown in Figure 11.

EXAMPLE IV

25 Engineering Cobalamin Synthesis into an Organism

The key enzyme of the Wood-Ljungdahl pathway, ACS/CODH, requires cobalamin (vitamin B₁₂) to function. B₁₂ is synthesized *de novo* in some organisms but must be supplied exogenously to others. Still other organisms such as *S. cerevisiae* lack the ability to efficiently uptake B₁₂. This example describes a strategy for engineering *de novo* B₁₂ synthetic capability
30 into an organism.

B₁₂ biosynthetic pathways have been characterized in several organisms including *Salmonella typhimurium* LT2 (Roth et al., *J. Bacteriol.* 175:3303-3316), *Lactobacillus reuteri* CRL1098 (Harms and Thauer, *Eur. J. Biochem.* 235:653-659 (1996)) and *Bacillus megaterium* (Brey et al., *J. Bacteriol.* 167:623-630 (1986)). Bacterial B₁₂ biosynthesis pathways involve 20-30 genes clustered together in one or more operons. Two cobalamin biosynthesis pathways: late-insertion (aerobic only) and early-insertion (anaerobic) have been described (Scott, A.I., *J. Org. Chem.* 68:2529-2539 (2003)). The final products of the biosynthesis of vitamin B₁₂ are 5'-deoxyadenosylcobalamin (coenzyme B₁₂) and methylcobalamin (MeCbl). Vitamin B₁₂ is defined as cyanocobalamin (CNCbl) which is the form commonly prepared in industry. In this example, B₁₂ refers to all three analogous molecules.

The anaerobic cobalamin biosynthesis pathway has been well-characterized in *Salmonella typhimurium* LT2 (Roth et al., *J. Bacteriol.* 175:3303-3316)). Pathway genes are clustered in a large operon termed the *cob* operon. A plasmid containing the following 20 genes from the *cob* operon (pAR8827) was transformed into *E. coli* and conferred the ability to synthesize cobalamin *de novo* (Raux et al., *J. Bacteriol.* 178:753-767 (1996)). To further improve yield of the cobyric acid precursor, the authors removed the known regulatory elements of *cbiA* and altered the RBS. The genes and corresponding GenBank identifiers and gi numbers are listed below.

<i>cysG</i>	NP_462380.1	16766765	<i>Salmonella typhimurium</i> LT2
<i>cbiK</i>	NP_460970.1	16765355	<i>Salmonella typhimurium</i> LT2
<i>cbiL</i>	NP_460969.1	16765354	<i>Salmonella typhimurium</i> LT2
<i>cbiH</i>	NP_460972.1	16765357	<i>Salmonella typhimurium</i> LT2
<i>cbiF</i>	NP_460974.1	16765359	<i>Salmonella typhimurium</i> LT2
<i>cbiG</i>	NP_460973.1	16765358	<i>Salmonella typhimurium</i> LT2
<i>cbiD</i>	NP_460977.1	16765362	<i>Salmonella typhimurium</i> LT2
<i>cbiJ</i>	NP_460971.1	16765356	<i>Salmonella typhimurium</i> LT2
<i>cbiE</i>	NP_460976.1	16765361	<i>Salmonella typhimurium</i> LT2
<i>cbiT</i>	NP_460975.1	16765360	<i>Salmonella typhimurium</i> LT2
<i>cbiC</i>	NP_460978.1	16765363	<i>Salmonella typhimurium</i> LT2
<i>cbiA</i>	NP_460980.1	16765365	<i>Salmonella typhimurium</i> LT2
<i>fldA</i>	NP_459679.1	16764064	<i>Salmonella typhimurium</i> LT2
<i>cobA</i>	P31570.1	399274	<i>Salmonella typhimurium</i> LT2
<i>cbiP</i>	AAA27268.1	154436	<i>Salmonella typhimurium</i> LT2
<i>cbiB</i>	Q05600.1	543942	<i>Salmonella typhimurium</i> LT2
<i>cobU</i>	NP_460963.1	16765348	<i>Salmonella typhimurium</i> LT2
<i>cobT</i>	NP_460961.1	16765346	<i>Salmonella typhimurium</i> LT2
<i>cobS</i>	AAA27270.1	154438	<i>Salmonella typhimurium</i> LT2
<i>cobC</i>	NP_459635.1	16764020	<i>Salmonella typhimurium</i> LT2

Some organisms unable to synthesize B₁₂ *de novo* are able to catalyze some steps of the pathway. *E. coli*, for example, is unable to synthesize the corrin ring structure but encodes proteins that catalyze several reactions in the pathway (Raux et al. *J. Bacteriol.* 178:753-767 (1996)). The *cysG* gene encodes a functional CysG, a multifunctional enzyme that converts uroporphyrinogen III to precorrin-2 (Hugler et al., *J. Bacteriol.* 184:2404-2410 (2002); Ishige et al., *Appl. Environ. Microbiol.* 68:1192-1195 (2002)). The proteins encoded by *cobTSU* transform cobinamide to cobalamin and introduce the 5'-deoxyadenosyl group (Raux et al., *supra* (1996)).

cobT	NP_416495.1	16129932	Escherichia coli K12 sp. MG1655
cobS	NP_416496.1	16129933	Escherichia coli K12 sp. MG1655
cobU	NP_416497.1	16129934	Escherichia coli K12 sp. MG1655
cysG	NP_417827.1	16131246	Escherichia coli K12 sp. MG1655

S. cerevisiae is not able to synthesize B₁₂ *de novo*, nor is it able to uptake the vitamin at detectable levels. However, the *S. cerevisiae* genome encodes two proteins, Met1p and Met8p, that catalyze several B₁₂ pathway reactions. Met1p is analogous to the uroporphyrinogen III transmethylase CysG of *S. typhimurium*, which catalyzes the first step of B₁₂ biosynthesis from uroporphyrinogen III (Raux et al., *Biochem. J.* 338(pt. 3):701-708 (1999)). The Met8p protein is a bifunctional protein with uroporphyrinogen III transmethylase activity and cobaltochelatase activity analogous to the CysG of *B. megaterium* (Raux et al., *supra* (1999)).

<i>Met1p</i>	NP_012995.1	6322922	<i>Saccharomyces cerevisiae</i>
<i>Met8p</i>	NP_009772.1	6319690	<i>Saccharomyces cerevisiae</i>

Any or all of these genes can be introduced into an organism deficient in one or more components of cobalamin synthesis to enable or increase the efficiency of cobalamin synthesis.

EXAMPLE V

Engineering Enhanced Cobalamin Uptake Capability in an Organism

This example describes engineering B₁₂ uptake capability into a host organism. B₁₂ uptake requires a specific transport system (Sennett et al., *Annu. Rev. Biochem.* 50:1053-1086 (1981)). The B₁₂ transport system of *E. coli* has been extensively studied. High-affinity transport across the outer membrane is calcium-dependent and mediated by a 66 kDa outer membrane porin, BtuB (Heller et al., *J. Bacteriol.* 161:896-903 (1985)). BtuB interacts with the TonB energy transducing system (TonB-ExbB-ExbD), facilitating energy-dependent translocation and

binding to periplasmic binding protein BtuF (WO/2007/141208; Atsumi et al., *Nature* 451:86-89 (2008)). Transport across the inner membrane is facilitated by an ABC type uptake system composed of BtuF, BtuD (ATP binding component) and BtuC (permease) (Binstock et al., *Meth. Enzymol.* 71(pt. C):403-411 (1981)). Crystal structures of the BtuCDF complex are available
 5 (Atsumi et al., *supra* (2008); Binstock et al., *supra* (1981)). An additional protein, BtuE, is coexpressed in the *btuCED* operon, but this protein is not required for B₁₂ transport and its function is unknown (Rioux et al., *Mol. Gen. Genet.* 217:301-308 (1989)). The *btuCED* operon is constitutively expressed. The GenBank identifiers and GI numbers of the genes associated with B₁₂ transport are listed below.

<i>btuB</i>	NP_418401.1	16131804	<i>Escherichia coli K12 sp. MG1655</i>
<i>btuC</i>	NP_416226.1	16129667	<i>Escherichia coli K12 sp. MG1655</i>
<i>btuD</i>	NP_416224.1	16129665	<i>Escherichia coli K12 sp. MG1655</i>
<i>btuF</i>	NP_414700.1	16128151	<i>Escherichia coli K12 sp. MG1655</i>
<i>tonB</i>	NP_415768.1	16129213	<i>Escherichia coli K12 sp. MG1655</i>
<i>exbB</i>	NP_417479.1	16130904	<i>Escherichia coli K12 sp. MG1655</i>
<i>exbD</i>	NP_417478.1	16130903	<i>Escherichia coli K12 sp. MG1655</i>

10 The B₁₂ uptake capability of an organism can be further improved by overexpressing genes encoding the requisite transport proteins, and reducing or eliminating negative regulatory control. Overexpressing the *btuBCDF* genes leads to increased binding of B₁₂ to membranes and increased rate of uptake into cells. Another strategy is to remove regulatory control. The *btuB* mRNA translation is directly repressed by B₁₂ at the 5' UTR (Nahvi et al., *Chem. Biol.*
 15 9:1043 (2002)). This interaction may induce mRNA folding to block ribosome access to the translational start. Mutation or elimination of the B₁₂ binding site removes inhibition and improves the efficiency of B₁₂ uptake (US Patent No. 6,432,686, Bulthuis et al.). These strategies were successfully employed to improve B₁₂ uptake capability in 1,3-PDO producing microorganisms (WO/1999/058686) and (US Patent No. 6,432,686, Bulthuis et al.). A recent
 20 patent application describes improving the efficiency of B₁₂ uptake (WO/2008/152016) by deleting negative regulatory proteins such as *C. glutamicum* btuR2.

S. typhimurium possesses both high and low affinity transporters for B₁₂. The high affinity transporter is encoded by *btuB* (Rioux et al., *J. Bacteriol.* 171:2986-2993 (1989)). Like *E. coli* transport across the periplasmic membrane is predicted to occur via an ABC transport system,
 25 although this has not been characterized to date. The B₁₂ binding protein is encoded by *btuD* and *btuE*, and *btuC* is predicted to encode the permease.

btuB	AAA27031.1	153891	Salmonella typhimurium LT2
btuC	NP_460306.1	16764691	Salmonella typhimurium LT2
btuD	NP_460308.1	16764693	Salmonella typhimurium LT2
btuE	AAL20266.1	16419860	Salmonella typhimurium LT2

Any or all of these genes can be introduced into an organism deficient in one or more components of cobalamin uptake to enable or increase the efficiency of cobalamin uptake.

Method for quantifying B₁₂ in the culture medium. To quantify the amount of B₁₂ in the culture medium, cell free samples are run on HPLC. Cobalamin quantification is achieved by comparing peak area ratios at 278 nm and 361 nm with standards, then applying peak areas to standard curves of cobalamin.

Throughout this application various publications have been referenced within parentheses. The disclosures of these publications in their entireties are hereby incorporated by reference in this application in order to more fully describe the state of the art to which this invention pertains.

Although the invention has been described with reference to the disclosed embodiments, those skilled in the art will readily appreciate that the specific examples and studies detailed above are only illustrative of the invention. It should be understood that various modifications can be made without departing from the spirit of the invention. Accordingly, the invention is limited only by the following claims.

What is claimed is:

1. A non-naturally occurring microbial organism, comprising a microbial organism having a 1,3-butanediol (1,3-BDO) pathway comprising at least one exogenous nucleic acid encoding a 1,3-BDO pathway enzyme or protein expressed in a sufficient amount to produce

5 1,3-BDO said 1,3-BDO pathway comprising Formate dehydrogenase, Formyltetrahydrofolate synthetase, Methenyltetrahydrofolate cyclohydrolase, Methylenetetrahydrofolate dehydrogenase, Methylenetetrahydrofolate reductase, Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), Corrinoid iron-sulfur protein (AcsD), Nickel-protein assembly protein (AcsF & CooC), Ferredoxin (Orf7), Acetyl-CoA synthase (AcsB & AcsC), Carbon monoxide

10 dehydrogenase (AcsA), Hydrogenase (Hyd), Acetoacetyl-CoA thiolase (AtoB), Acetoacetyl-CoA reductase (CoA-dependent, aldehyde forming), 3-oxobutyraldehyde reductase (ketone reducing), 3-hydroxybutyraldehyde reductase, Acetoacetyl-CoA reductase (CoA-dependent, alcohol forming), 3-oxobutyraldehyde reductase (aldehyde reducing), 4-hydroxy,2-butanone reductase, Acetoacetyl-CoA reductase (ketone reducing), 3-hydroxybutyryl-CoA reductase

15 (aldehyde forming), 3-hydroxybutyryl-CoA reductase (alcohol forming), 3-hydroxybutyryl-CoA transferase, 3-hydroxybutyryl-CoA hydrolase, 3-hydroxybutyryl-CoA synthetase, 3-hydroxybutyrate dehydrogenase, 3-hydroxybutyrate reductase, Acetoacetyl-CoA transferase, Acetoacetyl-CoA hydrolase, Acetoacetyl-CoA synthetase, or Acetoacetate reductase;

wherein said 1,3-BDO pathway enzymes are a set of enzymes selected from: A: 1) 3-hydroxybutyryl-CoA reductase (aldehyde forming), 2) Acetyl-CoA synthase (AcsB & AcsC), 3) Formate dehydrogenase, 4) Formyltetrahydrofolate synthetase, 5) Methenyltetrahydrofolate cyclohydrolase, 6) Methylenetetrahydrofolate dehydrogenase, 7) Methylenetetrahydrofolate reductase, 8) Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), 9) Corrinoid iron-sulfur protein (AcsD), 10) Nickel-protein assembly protein (AcsF & CooC), 11) Ferredoxin

25 (Orf7), 12) Carbon monoxide dehydrogenase (AcsA), 13) Hydrogenase (Hyd), 14) Acetoacetyl-CoA thiolase (AtoB), 15) Acetoacetyl-CoA reductase (ketone reducing), and 16) 3-hydroxybutyraldehyde reductase; B: 1) 3-hydroxybutyrate reductase, 2) Acetyl-CoA synthase (AcsB & AcsC), 3) Formate dehydrogenase, 4) Formyltetrahydrofolate synthetase, 5) Methenyltetrahydrofolate cyclohydrolase, 6) Methylenetetrahydrofolate dehydrogenase, 7) Methylenetetrahydrofolate reductase, 8) Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), 9) Corrinoid iron-sulfur protein (AcsD), 10) Nickel-protein assembly protein (AcsF & CooC), 11) Ferredoxin (Orf7), 12) Carbon monoxide dehydrogenase (AcsA), 13) Hydrogenase (Hyd), 14) Acetoacetyl-CoA thiolase (AtoB), 15) Acetoacetyl-CoA reductase

30

(ketone reducing), 16) 3-hydroxybutyryl-CoA transferase, hydrolase, or synthetase, and 17) 3-hydroxybutyraldehyde reductase; C: 1) 3-hydroxybutyryl-CoA reductase (alcohol forming); 2) Acetyl-CoA synthase (AcsB & AcsC), 3) Formate dehydrogenase, 4) Formyltetrahydrofolate synthetase, 5) Methenyltetrahydrofolate cyclohydrolase, 6) Methylenetetrahydrofolate dehydrogenase, 7) Methylenetetrahydrofolate reductase, 8) Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), 9) Corrinoid iron-sulfur protein (AcsD), 10) Nickel-protein assembly protein (AcsF & CooC), 11) Ferredoxin (Orf7), 12) Carbon monoxide dehydrogenase (AcsA), 13) Hydrogenase (Hyd), 14) Acetoacetyl-CoA thiolase (AtoB), and 15) Acetoacetyl-CoA reductase (ketone reducing), D: 1) Formate dehydrogenase, 2) Formyltetrahydrofolate synthetase, 3) Methenyltetrahydrofolate cyclohydrolase, 4) Methylenetetrahydrofolate dehydrogenase, 5) Methylenetetrahydrofolate reductase, 6) Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), 7) Corrinoid iron-sulfur protein (AcsD), 8) Nickel-protein assembly protein (AcsF & CooC), 9) Ferredoxin (Orf7), 10) Acetyl-CoA synthase (AcsB & AcsC), 11) Carbon monoxide dehydrogenase (AcsA), 12) Hydrogenase (Hyd), 13) Acetoacetyl-CoA thiolase (AtoB), 14) Acetoacetyl-CoA reductase (CoA-dependent, aldehyde forming), 15) 3-oxobutyraldehyde reductase (aldehyde reducing), and 16) 4-hydroxy,2-butanone reductase; E: 1) Formate dehydrogenase, 2) Formyltetrahydrofolate synthetase, 3) Methenyltetrahydrofolate cyclohydrolase, 4) Methylenetetrahydrofolate dehydrogenase, 5) Methylenetetrahydrofolate reductase, 6) Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), 7) Corrinoid iron-sulfur protein (AcsD), 8) Nickel-protein assembly protein (AcsF & CooC), 9) Ferredoxin (Orf7), 10) Acetyl-CoA synthase (AcsB & AcsC), 11) Carbon monoxide dehydrogenase (AcsA), 12) Hydrogenase (Hyd), 13) Acetoacetyl-CoA thiolase (AtoB), 14) Acetoacetyl-CoA reductase (CoA-dependent, aldehyde forming), 15) 3-oxobutyraldehyde reductase (ketone reducing), and 16) 3-hydroxybutyraldehyde reductase; F: 1) Formate dehydrogenase, 2) Formyltetrahydrofolate synthetase, 3) Methenyltetrahydrofolate cyclohydrolase, 4) Methylenetetrahydrofolate dehydrogenase, 5) Methylenetetrahydrofolate reductase, 6) Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), 7) Corrinoid iron-sulfur protein (AcsD), 8) Nickel-protein assembly protein (AcsF & CooC), 9) Ferredoxin (Orf7), 10) Acetyl-CoA synthase (AcsB & AcsC), 11) Carbon monoxide dehydrogenase (AcsA), 12) Hydrogenase (Hyd), 13) Acetoacetyl-CoA thiolase (AtoB), 14) Acetoacetyl-CoA reductase (CoA-dependent, alcohol forming), and 15) 4-hydroxy,2-butanone reductase; G: 1) Formate dehydrogenase, 2) Formyltetrahydrofolate synthetase, 3) Methenyltetrahydrofolate cyclohydrolase, 4) Methylenetetrahydrofolate dehydrogenase, 5) Methylenetetrahydrofolate reductase, 6) Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), 7) Corrinoid

iron-sulfur protein (AcsD), 8) Nickel-protein assembly protein (AcsF & CooC), 9) Ferredoxin (Orf7), 10) Acetyl-CoA synthase (AcsB & AcsC), 11) Carbon monoxide dehydrogenase (AcsA), 12) Hydrogenase (Hyd), 13) Acetoacetyl-CoA thiolase (AtoB), 14) Acetoacetyl-CoA transferase, hydrolase, or synthetase, 15) Acetoacetate reductase, 16) 3-oxobutyraldehyde reductase (ketone reducing), and 17) 3-hydroxybutyraldehyde reductase; H: 1) Formate dehydrogenase, 2) Formyltetrahydrofolate synthetase, 3) Methenyltetrahydrofolate cyclohydrolase, 4) Methylenetetrahydrofolate dehydrogenase, 5) Methylenetetrahydrofolate reductase, 6) Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), 7) Corrinoid iron-sulfur protein (AcsD), 8) Nickel-protein assembly protein (AcsF & CooC), 9) Ferredoxin (Orf7), 10) Acetyl-CoA synthase (AcsB & AcsC), 11) Carbon monoxide dehydrogenase (AcsA), 12) Hydrogenase (Hyd), 13) Acetoacetyl-CoA thiolase (AtoB), 14) Acetoacetyl-CoA transferase, hydrolase, or synthetase, 15) Acetoacetate reductase, 16) 3-oxobutyraldehyde reductase (aldehyde reducing), and 17) 4-hydroxy,2-butanone reductase; and I: 1) Formate dehydrogenase, 2) Formyltetrahydrofolate synthetase, 3) Methenyltetrahydrofolate cyclohydrolase, 4) Methylenetetrahydrofolate dehydrogenase, 5) Methylenetetrahydrofolate reductase, 6) Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), 7) Corrinoid iron-sulfur protein (AcsD), 8) Nickel-protein assembly protein (AcsF & CooC), 9) Ferredoxin (Orf7), 10) Acetyl-CoA synthase (AcsB & AcsC), 11) Carbon monoxide dehydrogenase (AcsA), 12) Hydrogenase (Hyd), 13) Acetoacetyl-CoA thiolase (AtoB), 14) Acetoacetyl-CoA transferase, hydrolase, or synthetase, 15) 3-hydroxybutyrate dehydrogenase, 16) 3-hydroxybutyrate reductase, and 17) 3-hydroxybutyraldehyde reductase.

2. The non-naturally occurring microbial organism of claim 1, wherein said microbial organism comprises two exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

3. The non-naturally occurring microbial organism of claim 1, wherein said microbial organism comprises three exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

4. The non-naturally occurring microbial organism of claim 1, wherein said microbial organism comprises four exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

5. The non-naturally occurring microbial organism of claim 1, wherein said microbial organism comprises five exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

6. The non-naturally occurring microbial organism of claim 1, wherein said microbial organism comprises six exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

7. The non-naturally occurring microbial organism of claim 1, wherein said microbial organism comprises seven exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

8. The non-naturally occurring microbial organism of claim 1, wherein said microbial organism comprises eight exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

5 9. The non-naturally occurring microbial organism of claim 1, wherein said microbial organism comprises nine exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

10. The non-naturally occurring microbial organism of claim 1, wherein said microbial organism comprises 10 exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

10 11. The non-naturally occurring microbial organism of claim 1, wherein said microbial organism comprises 11 exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

12. The non-naturally occurring microbial organism of claim 1, wherein said microbial organism comprises 12 exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

13. The non-naturally occurring microbial organism of claim 1, wherein said microbial organism comprises 13 exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

15 14. The non-naturally occurring microbial organism of claim 1, wherein said microbial organism comprises 14 exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

15. The non-naturally occurring microbial organism of claim 1, wherein said microbial organism comprises 15 exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

20 16. The non-naturally occurring microbial organism of claim 1, wherein said microbial organism comprises 16 exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

17. The non-naturally occurring microbial organism of claim 1, wherein said microbial organism comprises 17 exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

18. The non-naturally occurring microbial organism of claim 1, wherein said at least one exogenous nucleic acid is a heterologous nucleic acid.

25 19. The non-naturally occurring microbial organism of claim 1, wherein said non-naturally occurring microbial organism is in a substantially anaerobic culture medium.

20. The non-naturally occurring microbialorganism of claim 1, wherein said microbial organism uses a carbon feedstock selected from 1) CO, 2) CO₂ and H₂, 3) CO, CO₂, and H₂, 4) synthesis gas comprising CO and H₂, 5) synthesis gas comprising CO, CO₂, and H₂, and 6) one or more carbohydrates.

5 21. A method for producing 1,3-BDO, comprising culturing the non-naturally occurring microbial organism of claim 1 under conditions and for a sufficient period of time to produce 1,3-BDO.

22. The method of claim 21, wherein said non-naturally occurring microbial organism is in a substantially anaerobic culture medium.

10 23. The method of claim 21, wherein said microbial organism comprises two exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

24. The method of claim 21, wherein said microbial organism comprises three exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

15 25. The method of claim 21, wherein said microbial organism comprises four exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

26. The method of claim 21, wherein said microbial organism comprises five exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

27. The method of claim 21, wherein said microbial organism comprises six exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

20 28. The method of claim 21, wherein said microbial organism comprises seven exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

29. The method of claim 21, wherein said microbial organism comprises eight exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

25 30. The method of claim 21, wherein said microbial organism comprises nine exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

31. The method of claim 21, wherein said microbial organism comprises 10 exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

32. The method of claim 21, wherein said microbial organism comprises 11 exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

33. The method claim 21, wherein said microbial organism comprises 12 exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

5 34. The method claim 21, wherein said microbial organism comprises 13 exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

35. The method of claim 21, wherein said microbial organism comprises 14 exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

10 36. The method of claim 21, wherein said microbial organism comprises 15 exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

37. The method of claim 21, wherein said microbial organism comprises 16 exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

38. The method of claim 21, wherein said microbial organism comprises 17 exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

15 39. The method of claim 21, wherein said at least one exogenous nucleic acid is a heterologous nucleic acid.

40. The method of claim 21, wherein said microbial organism uses a carbon feedstock selected from 1) CO, 2) CO₂ and H₂, 3) CO, CO₂, and H₂, 4) synthesis gas comprising CO and H₂, 5) synthesis gas comprising CO, CO₂, and H₂, and 6) one or more carbohydrates.

20 41. A non-naturally occurring microbial organism, comprising a microbial organism having a 1,3-butanediol (1,3-BDO) pathway comprising at least one exogenous nucleic acid encoding a 1,3-BDO pathway enzyme or protein expressed in a sufficient amount to produce 1,3-BDO said 1,3-BDO pathway comprising Methanol methyltransferase (MtaB), Corrinoid protein (MtaC), Methyltetrahydrofolate:corrinoid protein methyltransferase (MtaA),
25 Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), Corrinoid iron-sulfur protein (AcsD), Nickel-protein assembly protein (AcsF & CooC), Ferredoxin (Orf7), Acetyl-CoA synthase (AcsB & AcsC), Carbon monoxide dehydrogenase (AcsA), Hydrogenase (Hyd), Acetoacetyl-CoA thiolase (AtoB), Acetoacetyl-CoA reductase (CoA-dependent, aldehyde forming), 3-oxobutylaldehyde reductase (ketone reducing), 3-hydroxybutylaldehyde reductase,

Acetoacetyl-CoA reductase (CoA-dependent, alcohol forming) , 3-oxobutyraldehyde reductase (aldehyde reducing), 4-hydroxy,2-butanone reductase, Acetoacetyl-CoA reductase (ketone reducing), 3-hydroxybutyryl-CoA reductase (aldehyde forming), 3-hydroxybutyryl-CoA reductase (alcohol forming), 3-hydroxybutyryl-CoA transferase, 3-hydroxybutyryl-CoA
 5 hydrolase, 3-hydroxybutyryl-CoA synthetase, 3-hydroxybutyrate dehydrogenase, 3-hydroxybutyrate reductase, Acetoacetyl-CoA transferase, Acetoacetyl-CoA hydrolase, Acetoacetyl-CoA synthetase, or Acetoacetate reductase.

42. The non-naturally occurring microbial organism of claim 41, wherein said 1,3-BDO pathway enzymes are a set of enzymes selected from: A: 1) Methanol methyltransferase (MtaB),
 10 2) Corrinoid protein (MtaC), 3) Methyltetrahydrofolate:corrinoid protein methyltransferase (MtaA), 4) Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), 5) Corrinoid iron-sulfur protein (AcsD), 6) Nickel-protein assembly protein (AcsF & CooC), 7) Ferredoxin (Orf7), 8) Acetyl-CoA synthase (AcsB & AcsC), 9) Carbon monoxide dehydrogenase (AcsA), 10) Hydrogenase, 11) Acetoacetyl-CoA thiolase (AtoB), 12) Acetoacetyl-CoA reductase (CoA-
 15 dependent, aldehyde forming), 13) 3-oxobutyraldehyde reductase (aldehyde reducing), and 14) 4-hydroxy,2-butanone reductase; B: 1) Methanol methyltransferase (MtaB), 2) Corrinoid protein (MtaC), 3) Methyltetrahydrofolate:corrinoid protein methyltransferase (MtaA), 4) Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), 5) Corrinoid iron-sulfur protein (AcsD), 6) Nickel-protein assembly protein (AcsF & CooC), 7) Ferredoxin (Orf7), 8)
 20 Acetyl-CoA synthase (AcsB & AcsC), 9) Carbon monoxide dehydrogenase (AcsA), 10) Hydrogenase, 11) Acetoacetyl-CoA thiolase (AtoB), 12) Acetoacetyl-CoA reductase (CoA-dependent, aldehyde forming), 13) 3-oxobutyraldehyde reductase (ketone reducing), and 14) 3-hydroxybutyraldehyde reductase; C: 1) Methanol methyltransferase (MtaB), 2) Corrinoid protein (MtaC), 3) Methyltetrahydrofolate:corrinoid protein methyltransferase (MtaA), 4)
 25 Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), 5) Corrinoid iron-sulfur protein (AcsD), 6) Nickel-protein assembly protein (AcsF & CooC), 7) Ferredoxin (Orf7), 8) Acetyl-CoA synthase (AcsB & AcsC), 9) Carbon monoxide dehydrogenase (AcsA), 10) Hydrogenase, 11) Acetoacetyl-CoA thiolase (AtoB), 12) Acetoacetyl-CoA reductase (ketone reducing), 13) 3-hydroxybutyryl-CoA reductase (aldehyde forming), and 14) 3-
 30 hydroxybutyraldehyde reductase; D: 1) Methanol methyltransferase (MtaB), 2) Corrinoid protein (MtaC), 3) Methyltetrahydrofolate:corrinoid protein methyltransferase (MtaA), 4) Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), 5) Corrinoid iron-sulfur protein (AcsD), 6) Nickel-protein assembly protein (AcsF & CooC), 7) Ferredoxin (Orf7), 8) Acetyl-CoA synthase (AcsB & AcsC), 9) Carbon monoxide dehydrogenase (AcsA), 10)

Hydrogenase, 11) Acetoacetyl-CoA thiolase (AtoB), 12) Acetoacetyl-CoA reductase (ketone reducing), and 13) 3-hydroxybutyryl-CoA reductase (alcohol forming); E: 1) Methanol methyltransferase (MtaB), 2) Corrinoid protein (MtaC), 3) Methyltetrahydrofolate:corrinoid protein methyltransferase (MtaA), 4) Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), 5) Corrinoid iron-sulfur protein (AcsD), 6) Nickel-protein assembly protein (AcsF & CooC), 7) Ferredoxin (Orf7), 8) Acetyl-CoA synthase (AcsB & AcsC), 9) Carbon monoxide dehydrogenase (AcsA), 10) Hydrogenase, 11) Acetoacetyl-CoA thiolase (AtoB), 12) Acetoacetyl-CoA reductase (CoA-dependent, alcohol forming), and 13) 4-hydroxy,2-butanone reductase; F: 1) Methanol methyltransferase (MtaB), 2) Corrinoid protein (MtaC), 3) Methyltetrahydrofolate:corrinoid protein methyltransferase (MtaA), 4) Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), 5) Corrinoid iron-sulfur protein (AcsD), 6) Nickel-protein assembly protein (AcsF & CooC), 7) Ferredoxin (Orf7), 8) Acetyl-CoA synthase (AcsB & AcsC), 9) Carbon monoxide dehydrogenase (AcsA), 10) Hydrogenase, 11) Acetoacetyl-CoA thiolase (AtoB), 12) Acetoacetyl-CoA transferase, hydrolase, or synthetase, 13) Acetoacetate reductase, 14) 3-oxobutyraldehyde reductase (ketone reducing), and 15) 3-hydroxybutyraldehyde reductase; G: 1) Methanol methyltransferase (MtaB), 2) Corrinoid protein (MtaC), 3) Methyltetrahydrofolate:corrinoid protein methyltransferase (MtaA), 4) Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), 5) Corrinoid iron-sulfur protein (AcsD), 6) Nickel-protein assembly protein (AcsF & CooC), 7) Ferredoxin (Orf7), 8) Acetyl-CoA synthase (AcsB & AcsC), 9) Carbon monoxide dehydrogenase (AcsA), 10) Hydrogenase, 11) Acetoacetyl-CoA thiolase (AtoB), 12) Acetoacetyl-CoA transferase, hydrolase, or synthetase, 13) Acetoacetate reductase, 14) 3-oxobutyraldehyde reductase (aldehyde reducing), and 15) 4-hydroxy,2-butanone reductase; H: 1) Methanol methyltransferase (MtaB), 2) Corrinoid protein (MtaC), 3) Methyltetrahydrofolate:corrinoid protein methyltransferase (MtaA), 4) Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), 5) Corrinoid iron-sulfur protein (AcsD), 6) Nickel-protein assembly protein (AcsF & CooC), 7) Ferredoxin (Orf7), 8) Acetyl-CoA synthase (AcsB & AcsC), 9) Carbon monoxide dehydrogenase (AcsA), 10) Hydrogenase, 11) Acetoacetyl-CoA thiolase (AtoB), 12) Acetoacetyl-CoA reductase (ketone reducing), 13) 3-hydroxybutyryl-CoA transferase, hydrolase, or synthetase, 14) 3-hydroxybutyrate reductase, and 15) 3-hydroxybutyraldehyde reductase; I: 1) Methanol methyltransferase (MtaB), 2) Corrinoid protein (MtaC), 3) Methyltetrahydrofolate:corrinoid protein methyltransferase (MtaA), 4) Methyltetrahydrofolate:corrinoid protein methyltransferase (AcsE), 5) Corrinoid iron-sulfur protein (AcsD), 6) Nickel-protein assembly protein (AcsF &

CooC), 7) Ferredoxin (Orf7), 8) Acetyl-CoA synthase (AcsB & AcsC), 9) Carbon monoxide dehydrogenase (AcsA), 10) Hydrogenase, 11) Acetoacetyl-CoA thiolase (AtoB), 12) Acetoacetyl-CoA transferase, hydrolase, or synthetase, 13) 3-hydroxybutyrate dehydrogenase, 14) 3-hydroxybutyrate reductase, and 15) 3-hydroxybutyraldehyde reductase.

5 43. The non-naturally occurring microbial organism of claim 41, wherein said microbial organism comprises two exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

 44. The non-naturally occurring microbial organism of claim 41, wherein said microbial organism comprises three exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

 45. The non-naturally occurring microbial organism of claim 41, wherein said microbial
10 organism comprises four exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

 46. The non-naturally occurring microbial organism of claim 41, wherein said microbial organism comprises five exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

 47. The non-naturally occurring microbial organism of claim 41, wherein said microbial organism comprises six exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

15 48. The non-naturally occurring microbial organism of claim 41, wherein said microbial organism comprises seven exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

 49. The non-naturally occurring microbial organism of claim 41, wherein said microbial organism comprises eight exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

 50. The non-naturally occurring microbial organism of claim 41, wherein said microbial
20 organism comprises nine exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

 51. The non-naturally occurring microbial organism of claim 41, wherein said microbial organism comprises 10 exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

 52. The non-naturally occurring microbial organism of claim 41, wherein said microbial organism comprises 11 exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

25 53. The non-naturally occurring microbial organism of claim 41, wherein said microbial organism comprises 12 exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

54. The non-naturally occurring microbial organism of claim 41, wherein said microbial organism comprises 13 exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

55. The non-naturally occurring microbial organism of claim 41, wherein said microbial organism comprises 14 exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

5 56. The non-naturally occurring microbial organism of claim 41, wherein said microbial organism comprises 15 exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

57. The non-naturally occurring microbial organism of claim 41, wherein said at least one exogenous nucleic acid is a heterologous nucleic acid.

10 58. The non-naturally occurring microbial organism of claim 41, wherein said non-naturally occurring microbial organism is in a substantially anaerobic culture medium.

59. The non-naturally occurring microbial organism of claim 41, wherein said microbial organism uses a carbon feedstock selected from 1) methanol and CO, 2) methanol, CO₂, and H₂, 3) methanol, CO, CO₂, and H₂, 4) methanol and synthesis gas comprising CO and H₂, 5) methanol and synthesis gas comprising CO, CO₂, and H₂, 6) one or more carbohydrates, 7) 15 methanol and one or more carbohydrates, and 8) methanol.

60. A method for producing 1,3-BDO comprising culturing the non-naturally occurring microbial organism of claim 41 or 42 under conditions and for a sufficient period of time to produce 1,3-BDO.

20 61. The method of claim 60, wherein said non-naturally occurring microbial organism is in a substantially anaerobic culture medium.

62. The method of claim 60, wherein said microbial organism comprises two exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

63. The method of claim 60, wherein said microbial organism comprises three exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

25 64. The method of claim 60, wherein said microbial organism comprises four exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

65. The method of claim 60, wherein said microbial organism comprises five exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

66. The method of claim 60, wherein said microbial organism comprises six exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

67. The method of claim 60, wherein said microbial organism comprises seven exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

5 68. The method of claim 60, wherein said microbial organism comprises eight exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

69. The method of claim 60, wherein said microbial organism comprises nine exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

10 70. The method of claim 60, wherein said microbial organism comprises 10 exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

71. The method of claim 60, wherein said microbial organism comprises 11 exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

72. The method claim 60, wherein said microbial organism comprises 12 exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

15 73. The method claim 60, wherein said microbial organism comprises 13 exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

74. The method of claim 60, wherein said microbial organism comprises 14 exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

20 75. The method of claim 60, wherein said microbial organism comprises 15 exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

76. The method of claim 60, wherein said microbial organism comprises 16 exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

77. The method of claim 60, wherein said microbial organism comprises 17 exogenous nucleic acids each encoding a 1,3-BDO pathway enzyme.

25 78. The method of claim 60, wherein said at least one exogenous nucleic acid is a heterologous nucleic acid.

79. The method of claim 60, wherein said microbial organism uses a carbon feedstock selected from 1) methanol and CO, 2) methanol, CO₂, and H₂, 3) methanol, CO, CO₂, and H₂, 4) methanol and synthesis gas comprising CO and H₂, 5) methanol and synthesis gas comprising CO, CO₂, and H₂, 6) one or more carbohydrates, 7) methanol and one or more carbohydrates, and 8) methanol.
- 5