



(51) International Patent Classification:

C12P 5/02 (2006.01) C12N 1/00 (2006.01)
C12N 15/74 (2006.01)

(21) International Application Number:

PCT/US2012/043293

(22) International Filing Date:

20 June 2012 (20.06.2012)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

61/500,121 22 June 2011 (22.06.2011) US
61/502,817 29 June 2011 (29.06.2011) US

(71) Applicant (for all designated States except US): **GENOMATICA, INC.** [US/US]; 10520 Wateridge Circle, San Diego, CA 92121 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **BURGARD, Anthony, P.** [US/US]; 10520 Wateridge Circle, San Diego, CA 92121 (US). **OSTERHOUT, Robin, E.** [US/US]; 10520 Wateridge Circle, San Diego, CA 92121 (US). **SUN, Jun** [CN/US]; 10520 Wateridge Circle, San Diego, CA 92121 (US). **PHARKYA, Priti** [IN/US]; 10520 Wateridge Circle, San Diego, CA 92121 (US).

(74) Agents: **GAY, David, A.** et al.; Jones Day, 222 East 41st Street, New York, NY 10017-6702 (US).

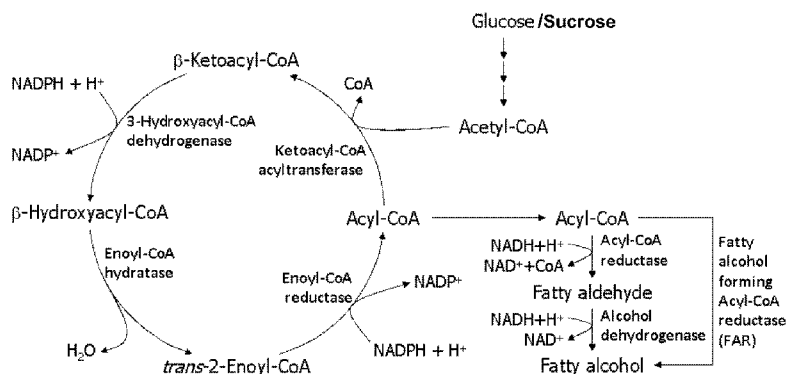
(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG,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

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

Published:

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

(54) Title: MICROORGANISM FOR PRODUCING PRIMARY ALCOHOLS AND RELATED COMPOUNDS AND METHODS RELATED THERETO



(57) Abstract: The invention provides a non-naturally occurring microbial organism having a microbial organism having at least one exogenous gene insertion and/or one or more gene disruptions that confer production of primary alcohols and at least one exogenous nucleic acid that encodes an enzyme that increases the yields of the primary alcohol by (i) enhancing carbon fixation via the reductive TCA cycle, and/or (ii) accessing additional reducing equivalents from gaseous carbon sources and/or syngas components such as CO, CO₂, and/or H₂. A method for producing long chain alcohols includes culturing these non-naturally occurring microbial organisms.



**MICROORGANISM FOR PRODUCING PRIMARY ALCOHOLS AND RELATED
COMPOUNDS AND METHODS RELATED THERETO**

This application claims the benefit of priority of United States Provisional application serial No. 61/500,121, filed June 22, 2011, United States Provisional application serial No.

5 61/502,817, filed June 29, 2011, the entire contents of which are incorporated herein by reference.

The instant application contains a Sequence Listing which has been submitted via EFS-Web and is hereby incorporated by reference in its entirety. Said ASCII copy, created June 18, 2012, is named 871943-228150_PCT_Sequence_Listing.txt and is 77,825 bytes in size

10 **BACKGROUND OF THE INVENTION**

This invention relates generally to biosynthetic processes and, more specifically to organisms having primary alcohol biosynthetic capability.

Primary alcohols are a product class of compounds having a variety of industrial applications which include a variety of biofuels and specialty chemicals. Primary alcohols also can be
15 used to make a large number of additional industrial products including polymers and surfactants. For example, higher primary alcohols (C₄-C₂₀) and their ethoxylates are used as surfactants in many consumer detergents, cleaning products and personal care products worldwide such as laundry powders and liquids, dishwashing liquid and hard surface cleaners. They are also used in the manufacture of a variety of industrial chemicals and in
20 lubricating oil additives. Long-chain primary alcohols, such as octanol and hexanol, have useful organoleptic properties and have long been employed as fragrance and flavor materials. Smaller chain (C₄-C₈) higher primary alcohols (e.g., butanol) are used as chemical intermediates for production of derivatives such as acrylates used in paints, coatings, and adhesives applications.

25 Primary alcohols are currently produced from, for example, hydrogenation of fatty acids, hydroformylation of terminal olefins, partial oxidation of n-paraffins and the Al- catalyzed polymerization of ethylene. Unfortunately, it is not commercially viable to produce primary alcohols directly from the oxidation of petroleum-based linear hydrocarbons (n-paraffins). This impracticality is because the oxidation of n-paraffins produces primarily secondary
30 alcohols, tertiary alcohols or ketones, or a mixture of these compounds, but does not produce

high yields of primary alcohols. Additionally, currently known methods for producing primary alcohols suffer from the disadvantage that they are restricted to feedstock which is relatively expensive, notably ethylene, which is produced via the thermal cracking of petroleum. In addition, current methods require several steps, and several catalyst types.

- 5 LCA production by microorganisms involves fatty acid synthesis followed by acyl-reduction steps. The universal fatty acid biosynthesis pathway found in most cells has been investigated for production of LCAs and other fatty acid derivatives. There is currently a great deal of improvement that can be achieved to provide more efficient biosynthesis pathways for LCA production with significantly higher theoretical product and energy yields.
- 10 Thus, there exists a need for alternative means for effectively producing commercial quantities of primary alcohols. 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
- 15 organism having a microbial organism having a malonyl-CoA-independent fatty acid synthesis (FAS) pathway and an acyl-reduction pathway having at least one exogenous nucleic acid encoding a malonyl-CoA-independent FAS pathway enzyme expressed in sufficient amounts to produce a primary alcohol and at least one exogenous nucleic acid that encodes an enzyme that increases the yields of the primary alcohol by (i) enhancing carbon
- 20 fixation via the reductive TCA cycle, and/or (ii) accessing additional reducing equivalents from gaseous carbon sources and/or syngas components such as CO, CO₂, and/or H₂, the malonyl-CoA-independent FAS pathway having ketoacyl-CoA acyltransferase or ketoacyl-CoA thiolase, 3-hydroxyacyl-CoA dehydrogenase, enoyl-CoA hydratase and enoyl-CoA reductase, the acyl-reduction pathway having an acyl-CoA reductase and an alcohol
- 25 dehydrogenase.

- In other aspects, embodiments disclosed herein relate to a method for producing a primary alcohol. The method includes culturing a non-naturally occurring microbial organism have having a malonyl-CoA-independent fatty acid synthesis (FAS) pathway and an acyl-reduction pathway having at least one exogenous nucleic acid encoding a malonyl-CoA-independent
- 30 FAS pathway enzyme expressed in sufficient amounts to produce a primary alcohol and at least one exogenous nucleic acid that encodes an enzyme that increases the yields of the

primary alcohol by (i) enhancing carbon fixation via the reductive TCA cycle, and/or (ii) accessing additional reducing equivalents from gaseous carbon sources and/or syngas components such as CO, CO₂, and/or H₂, under substantially anaerobic conditions for a sufficient period of time to produce the primary alcohol, the malonyl-CoA-independent FAS pathway having ketoacyl-CoA acyltransferase or ketoacyl-CoA thiolase, 3-hydroxyacyl-CoA dehydrogenase, enoyl-CoA hydratase and enoyl-CoA reductase, the acyl-reduction pathway having an acyl-CoA reductase and an alcohol dehydrogenase.

In some aspects, embodiments disclosed herein relate to a non-naturally occurring microbial organism that includes one or more gene disruptions occurring in genes encoding enzymes that couple long-chain alcohol (LCA) production to growth of the non-naturally occurring microbial organism. In other embodiments, LCA production can be accomplished during non-growth phases using the same disruption strategies. The one or more gene disruptions reduce the activity of the enzyme, whereby the gene disruptions confer production of LCA onto the non-naturally occurring microbial organism.

In other aspects, embodiments disclosed herein relate to a method for producing LCA that includes culturing a non-naturally occurring microbial organism having one or more gene disruptions. The one or more gene disruptions occur in genes encoding an enzyme that confers LCA production in the organism.

In some aspects, embodiments disclosed herein relate to a non-naturally occurring eukaryotic organism, that includes one or more gene disruptions. The one or more gene disruptions occur in genes that encode enzymes such as a cytosolic pyruvate decarboxylase, a mitochondrial pyruvate dehydrogenase, a cytosolic ethanol-specific alcohol dehydrogenase and a mitochondrial ethanol-specific alcohol dehydrogenase. These disruptions confer production of long chain alcohols in the cytosol of the organism.

In some aspects, embodiments disclosed herein relate to a non-naturally occurring eukaryotic organism that includes one or more gene disruptions. The one or more gene disruptions occur in genes encoding enzymes such as a cytosolic pyruvate decarboxylase, a cytosolic ethanol-specific alcohol dehydrogenase, and a mitochondrial ethanol-specific alcohol dehydrogenase. These disruptions confer production of long chain alcohols in the mitochondrion of said organism.

In other aspects, embodiments disclosed herein relate to a method for producing long chain alcohols, including culturing these non-naturally occurring eukaryotic organisms.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows the malonyl-CoA-independent fatty acid synthesis and reduction (MI-LCA) pathway to produce LCAs.

Figure 2 shows the contrasted hypothetical production envelopes of an OptKnock-designed strain against a typical non-growth-coupled production strain. The potential evolutionary trajectories of the OptKnock strain lead to a high producing phenotype.

Figure 3 shows the growth-coupled LCA production characteristics of strain design I (alternating dotted and dashed) compared with those of wild-type *E. coli* (black). A glucose uptake rate of 10 mmol/gDW/hr is assumed.

Figure 4 shows the growth-coupled LCA production characteristics of strain designs II (alternating dotted and dashed), III-V (dashed), and VI-XI (dotted) compared with those of wild-type *E. coli* (black). A glucose uptake rate of 10 mmol/gDW/hr is assumed.

Figure 5 shows the growth-coupled LCA production characteristics of strain designs XII (alternating dotted and dashed) and XIII-XV (dashed) compared with those of wild-type *E. coli* (black). A glucose uptake rate of 10 mmol/gDW/hr is assumed.

Figure 6 shows the growth-coupled LCA production characteristics of strain designs XVI-XVIII (alternating dotted and dashed) and XIX-XXI (dashed) compared with those of wild-type *E. coli* (black). A glucose uptake rate of 10 mmol/gDW/hr is assumed.

Figure 7 shows the growth-coupled LCA production characteristics of Designs I (alternating dotted and dashed), V (dashed), and V_A (dotted) compared with those of wild-type *E. coli* (black). A glucose uptake rate of 10 mmol/gDW/hr is assumed. Point A refers to the dodecanol production rate at maximum growth of a strain engineered according to design V_A and point B refers to the minimal dodecanol production rate required for growth.

Figure 8 shows the growth-coupled LCA production characteristics of Designs I (alternating dotted and dashed), XII (long dashed), XII_A (short dashed), and XII_B (dotted) compared with those of wild-type *E. coli* (black). A glucose uptake rate of 10 mmol/gDW/hr is

assumed. Point A refers to the dodecanol production rate at maximum growth of a strain engineered according to design XII_B and point B refers to the minimal dodecanol production rate required for growth.

Figure 9a shows the formation of dodecanol in the cytosol by relying on the AMP-forming acetyl CoA synthetase for the formation of acetyl CoA for dodecanol production. The dotted arrows depict the flow of the majority of the carbon flux in this production scenario.

Figure 9b shows the growth-coupled production envelopes for the production of dodecanol in *S. cerevisiae* in the scenario where acetyl CoA synthetase is used for acetyl CoA production in the cytosol. The black curve shows the production envelope for the wild-type network under aerobic conditions, and the dark gray curve shows the growth-coupled production characteristics for the mutant network. A glucose uptake rate of 10 mmol/gDCW.hr is assumed.

Figure 10a shows the formation of dodecanol in the cytosol by relying on the ADP-forming acetate CoA ligase for the formation of acetyl CoA for dodecanol production. The gray arrow represents the addition of a heterologous enzyme. The dotted arrows depict the flow of the majority of the carbon flux in this production scenario.

Figure 10b shows the growth-coupled production envelopes for the production of dodecanol in *S. cerevisiae* in the scenario where acetate CoA ligase is employed for acetyl-CoA production in the cytosol. The black curve shows the production envelope for the wild-type network under aerobic conditions. The light gray curve shows the increase in feasible space after acetate CoA ligase is added to the network and the dark gray curve shows the growth-coupled production characteristics for the mutant network in the presence of oxygen. A glucose uptake rate of 10 mmol/gDCW.hr is assumed.

Figure 11a shows the formation of dodecanol in the cytosol by relying on the acylating acetaldehyde dehydrogenase for the formation of acetyl CoA for dodecanol production. The gray arrow shows a heterologous enzyme. The dotted arrows depict the flow of the majority of the carbon flux in this production scenario.

Figure 11b shows the growth-coupled production envelopes for the anaerobic production of dodecanol in *S. cerevisiae*. The black curve shows the production capabilities for the wild-type network, the light gray dotted curve shows the production characteristics when acylating

acetaldehyde dehydrogenase is added to the network and the dark gray curve shows the growth-coupling when alcohol dehydrogenase is deleted from the augmented network. Note the increase in the theoretical maximum when acylating acetaldehyde dehydrogenase is functional. A glucose uptake rate of 10 mmol/gDCW.hr is assumed.

- 5 Figure 12 shows the formation of dodecanol in the cytosol by relying on a cytosolic pyruvate dehydrogenase for acetyl CoA and NADH production. This can be accomplished by introducing a heterologous cytosolic enzyme (shown in gray) or by retargeting the native mitochondrial enzyme to the cytosol. The dotted arrows depict the flow of the majority of the carbon flux in this production scenario.
- 10 Figure 13 shows the formation of dodecanol in the cytosol by relying on a cytosolic pyruvate:NADP oxidoreductase for acetyl CoA and NADH production. This can be accomplished by introducing a heterologous enzyme in the cytosol (shown in gray). The dotted arrows depict the flow of the majority of the carbon flux in this production scenario.
- Figure 14 shows the formation of dodecanol in the cytosol by the introduction of a
- 15 heterologous pyruvate formate lyase (shown in gray) in the cytosol. The dotted arrows depict the flow of the majority of the carbon flux in this production scenario.
- Figure 15a shows the formation of dodecanol in the mitochondrion by using the pyruvate dehydrogenase for the formation of acetyl-CoA. The dotted arrows depict the flow of the majority of the carbon flux in this production scenario.
- 20 Figure 15b shows the growth-coupled production envelopes for the production of dodecanol in *S. cerevisiae* mitochondrion. The black curve shows the production capabilities for the wild-type network under anaerobic conditions and the dark gray curve shows the production characteristics in the absence of oxygen when pyruvate decarboxylase is deleted from the network. A glucose uptake rate of 10 mmol/gDCW.hr is assumed.
- 25 Figure 16 shows the formation of dodecanol in the mitochondrion by using the pyruvate:NADP oxidoreductase for formation of acetyl CoA. The gray arrow shows the heterologous enzyme and the dotted arrows depict the flow of the majority of the carbon flux in this production scenario.

Figure 17 shows the formation of dodecanol in the mitochondrion by using the pyruvate formate lyase for formation of acetyl CoA. The gray arrow shows the heterologous enzyme and the dotted arrows depict the flow of the majority of the carbon flux in this production scenario.

- 5 Figure 18 shows the formation of dodecanol in the mitochondrion by adding the mitochondrial acylating acetaldehyde dehydrogenase. The gray arrow shows the heterologous enzyme(s) and the dotted arrows depict the flow of the majority of the carbon flux in this production scenario.

- 10 Figure 19a shows the formation of dodecanol in the mitochondrion by using the acetyl CoA synthetase for formation of acetyl CoA. The gray arrow shows the heterologous enzyme(s) and the dotted arrows depict the flow of the majority of the carbon flux in this production scenario.

- 15 Figure 19b shows the growth-coupled production envelopes for the production of dodecanol in *S. cerevisiae* mitochondrion when acetyl-CoA is formed through the mitochondrial acetyl-CoA synthetase. The black curve shows the production envelope for the wild-type network under aerobic conditions, the light dark gray curve shows the production characteristics when the deletions have been imposed upon the network. The growth coupling can be improved further (dark gray curve) when flux through the oxidative part of the pentose phosphate pathway is decreased. A glucose uptake rate of 10 mmol/gDCW.hr is assumed.

- 20 Figure 20 shows the formation of dodecanol in the mitochondrion by using the acetate CoA ligase for formation of acetyl CoA. The gray arrows show the heterologous enzyme(s) and the dotted arrows depict the flow of the majority of the carbon flux in this production scenario.

- 25 Figure 21 shows the reverse TCA cycle for fixation of CO₂ on carbohydrates as substrates. The enzymatic transformations are carried out by the enzymes as shown.

Figure 22 shows the pathway for the reverse TCA cycle coupled with carbon monoxide dehydrogenase and hydrogenase for the conversion of syngas to acetyl-CoA.

Figure 23 shows Western blots of 10 micrograms ACS90 (lane 1), ACS91 (lane2), Mta98/99 (lanes 3 and 4) cell extracts with size standards (lane 5) and controls of *M. thermoacetica*

CODH (Moth_1202/1203) or Mtr (Moth_1197) proteins (50, 150, 250, 350, 450, 500, 750, 900, and 1000 ng).

Figure 24 shows CO oxidation assay results. Cells (*M. thermoacetica* or *E. coli* with the CODH/ACS operon; ACS90 or ACS91 or empty vector: pZA33S) were grown and extracts
5 prepared. Assays were performed at 55°C at various times on the day the extracts were prepared. Reduction of methylviologen was followed at 578 nm over a 120 sec time course.

Figure 25, panels A and B, show exemplary pathways for fixation of CO₂ to acetyl-CoA using the reductive TCA cycle and exemplary pathways for the biosynthesis of long chain acyl-CoA's, acids, aldehydes, and alcohols from acetyl-CoA.

10 Figure 26, panels A and B, show exemplary pathways for fixation of CO₂ to acetyl-CoA using the reductive TCA cycle and exemplary pathways for the biosynthesis of long chain acyl-ACP's, acyl-CoA's, acids, aldehydes, and alcohols from acetyl-CoA.

Figure 27A shows the nucleotide sequence (SEQ ID NO: 1) of carboxylic acid reductase from *Nocardia iowensis* (GNM_720), and Figure 27B shows the encoded amino acid
15 sequence (SEQ ID NO: 2).

Figure 28A shows the nucleotide sequence (SEQ ID NO: 3) of phosphopantetheine transferase, which was codon optimized, and Figure 28B shows the encoded amino acid sequence (SEQ ID NO: 4).

Figure 29A shows the nucleotide sequence (SEQ ID NO: 5) of carboxylic acid reductase
20 from *Mycobacterium smegmatis* mc(2)155 (designated 890), and Figure 29B shows the encoded amino acid sequence (SEQ ID NO: 6).

Figure 30A shows the nucleotide sequence (SEQ ID NO: 7) of carboxylic acid reductase from *Mycobacterium avium* subspecies paratuberculosis K-10 (designated 891), and Figure 30B shows the encoded amino acid sequence (SEQ ID NO: 8).

25 Figure 31A shows the nucleotide sequence (SEQ ID NO: 9) of carboxylic acid reductase from *Mycobacterium marinum* M (designated 892), and Figure 31B shows the encoded amino acid sequence (SEQ ID NO: 10).

Figure 32A shows the nucleotide sequence (SEQ ID NO: 11) of carboxylic acid reductase designated 891GA, and Figure 32B shows the encoded amino acid sequence (SEQ ID NO: 12).

DETAILED DESCRIPTION OF THE INVENTION

5 The invention is directed, in part, to recombinant microorganisms capable of synthesizing the primary alcohols using a malonyl-CoA-independent fatty acid synthesis and reduction pathway. The modified microorganisms of the invention also are capable of secreting the resultant primary alcohol into the culture media or fermentation broth for further manipulation or isolation. Recombinant microorganisms of the invention can be engineered
10 to produce commercial quantities of a variety of different primary alcohols having different chain lengths between 4 (C4) and 24 (C24) or more carbon atoms. Production of primary alcohols through the modified pathways of the invention is particularly useful because it results in higher product and ATP yields than through naturally occurring biosynthetic pathways such as the well-documented malonyl-CoA dependent fatty acid synthesis pathway.
15 Using acetyl-CoA as a C2 extension unit instead of malonyl-acyl carrier protein (malonyl-ACP) saves one ATP molecule per unit flux of acetyl-CoA entering the elongation cycle. The elongation cycle results in acyl-CoA instead of acyl-ACP, and precludes the need of the ATP-consuming acyl-CoA synthase reactions for the production of octanol and other primary alcohols. The primary alcohol producing organisms of the invention can additionally allow
20 the use of biosynthetic processes to convert low cost renewable feedstock for the manufacture of chemical products.

In one specific embodiment, the invention utilizes a heterologous malonyl-CoA-independent fatty acid synthesis pathway coupled with an acyl-CoA reduction pathway to form primary alcohol species. The coupling of these two pathways will convert a carbon or energy source
25 into acetyl-CoA, which is used as both primer and extension unit in biosynthetic elongation cycle. The elongation cycle includes ketoacyl-CoA thiolase (or ketoacyl-CoA acyltransferase), 3-hydroxyacyl-CoA dehydrogenase, enoyl-CoA hydratase and enoyl-CoA reductase. Each cycle results in the formation of an acyl-CoA extended by one C2 unit compared to the acyl-CoA substrate entering the elongation cycle. Carbon chain-length of
30 the primary alcohols can be controlled by chain-length specific enoyl-CoA reductase, ketoacyl-CoA thiolase and/or acyl-CoA reductase. Acyl-CoA products with desired chain-lengths are funneled into a reduction pathway and reduced through the combination of acyl-

CoA reductase and alcohol dehydrogenase or the fatty alcohol forming acyl-CoA reductase to form desired primary alcohol. These reduction steps serve as another mechanism for control of chain length, for example, through the use of chain-length specific acyl-CoA reductases.

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, 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 operon. Exemplary metabolic polypeptides include enzymes or proteins within a malonyl-CoA-independent fatty acid biosynthetic pathway and enzymes within an acyl-reduction 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.

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 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.

As used herein, the terms “microbial,” “microbial organism” or “microorganism” are 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 also includes cell cultures of any species that can be cultured for the production of a biochemical.

As used herein, the term “primary alcohol” is intended to mean an alcohol which has the hydroxyl radical connected to a primary carbon. The term includes an alcohol that possesses the group $-CH_2OH$ which can be oxidized so as to form a corresponding aldehyde and acid having the same number of carbon atoms. Alcohols include any of a series of hydroxyl compounds, the simplest of which are derived from saturated hydrocarbons, have the general formula $C_nH_{2n+1}OH$, and include ethanol and methanol. Exemplary primary alcohols include butanol, hexanol, heptanol, octanol, nonanol, decanol, dodecanol, tetradecanol, and hexadecanol.

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, for example, in 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 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 acids 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.

As used herein, the term “growth-coupled” when used in reference to the production of a biochemical is intended to mean that the biosynthesis of the referenced biochemical is a product produced during the growth phase of a microorganism. “Non-growth-coupled” when used in reference to the production of a biochemical is intended to mean that the biosynthesis of the referenced biochemical is a product produced during a non-growth phase of a microorganism. Production of a biochemical product can be optionally obligatory to the growth of the organism.

As used herein, the term “metabolic modification” is intended to refer to a biochemical reaction that is altered from its naturally occurring state. Metabolic modifications can include, for example, elimination of a biochemical reaction activity by functional disruptions of one or more genes encoding an enzyme participating in the reaction. Sets of exemplary metabolic modifications are illustrated in Table 1. Individual reactions specified by such metabolic modifications and their corresponding gene complements are exemplified in Table 2 for *Escherichia coli*. Reactants and products utilized in these reactions are exemplified in Table 3.

As used herein, the term “gene disruption,” or grammatical equivalents thereof, is intended to mean a genetic alteration that renders the encoded gene product inactive. The genetic alteration can be, for example, deletion of the entire gene, deletion of a regulatory sequence required for transcription or translation, deletion of a portion of the gene which results in a truncated gene product or by any of various mutation strategies that inactivate the encoded gene product. One particularly useful method of gene disruption is complete gene deletion because it reduces or eliminates the occurrence of genetic reversions in the non-naturally occurring microorganisms of the invention. The term “gene disruption” is also intended to mean a genetic alteration that lowers the activity of a given gene product relative to its activity in a wild-type organism. This attenuation of activity can be due to, for example, a deletion in a portion of the gene which results in a truncated gene product or any of various mutation strategies that render the encoded gene product less active than its natural form, replacement or mutation of the promoter sequence leading to lower or less efficient expression of the gene, culturing the organism under a condition where the gene is less highly expressed than under normal culture conditions, or introducing antisense RNA molecules that interact with complementary mRNA molecules of the gene and alter its expression.

As used herein, the term “stable” when used in reference to growth-coupled production of a biochemical product is intended to refer to microorganism that can be cultured for greater than five generations without loss of the coupling between growth and biochemical synthesis. Generally, stable growth-coupled biochemical production will be greater than 10 generations, particularly stable growth-coupled biochemical production will be greater than about 25 generations, and more particularly, stable growth-coupled biochemical production will be greater than 50 generations, including indefinitely. Stable growth-coupled production of a biochemical can be achieved, for example, by disruption of a gene encoding an enzyme catalyzing each reaction within a set of metabolic modifications. The stability of growth-coupled production of a biochemical can be enhanced through multiple disruptions, significantly reducing the likelihood of multiple compensatory reversions occurring for each disrupted activity.

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.

As used herein, the term “confers production” refers not only to organisms that lack operational metabolic pathways for the production of LCAs, but also to organisms that may have some level of production of LCAs. Thus, an organism that already generates LCAs can benefit from improved production conferred onto the organism by the disruption of one or more genes.

As used herein, the term “eukaryotic organism” refers to any organism having a cell type having specialized organelles in the cytoplasm and a membrane-bound nucleus enclosing genetic material organized into chromosomes. The term is intended to encompass all

eukaryotic organisms including eukaryotic microbial organisms such as yeast and fungi. The term also includes cell cultures of any eukaryotic species that can be cultured for the production of a biochemical where the eukaryotic species need not be a microbial organism. A “eukaryotic microbial organism,” “microbial organism” or “microorganism” is intended to mean any eukaryotic organism that exists as a microscopic cell that is included within the domain of eukarya.

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 two 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.

- 5 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
- 10 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.
- 15 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 will be identifiable as structurally related to a known gene
- 20 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
- 25 includes, for example, a paralog or an unrelated gene.

Therefore, in identifying and constructing the non-naturally occurring microbial organisms of the invention having primary alcohol, fatty acyl-CoA, fatty ester, wax, acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol biosynthetic capability, those skilled in the art will understand with applying the teaching and guidance provided herein to a particular species

30 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 to ensure that any functional redundancy in enzymatic activities do not short circuit the designed metabolic modifications.

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.

The non-naturally occurring microbial organisms of the invention can contain stable genetic alterations, which refer to microorganisms that can be cultured for greater than five
5 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 will be greater than 50 generations, including indefinitely.

Those skilled in the art will understand that the genetic alterations, including metabolic
10 modifications exemplified herein are described with reference to *Euglena gracilis*, *E. coli* and *S. cerevisiae* genes and their corresponding metabolic reactions. 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*
15 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.

20 In some embodiments, the invention provides a non-naturally occurring microbial organism having a malonyl-CoA-independent fatty acid synthesis (FAS) pathway and an acyl-reduction pathway having at least one exogenous nucleic acid encoding a malonyl-CoA-independent FAS pathway enzyme expressed in sufficient amounts to produce a primary alcohol, said malonyl-CoA-independent FAS pathway comprising ketoacyl-CoA acyltransferase or
25 ketoacyl-CoA thiolase, 3-hydroxyacyl-CoA dehydrogenase, enoyl-CoA hydratase and enoyl-CoA reductase, said acyl-reduction pathway comprising an acyl-CoA reductase and an alcohol dehydrogenase.

Malonyl-CoA-independent fatty acid synthesis is a metabolic process used by photosynthetic flagellate such as *Euglena gracilis* (Inui et al., *Euro. J. Biochem.* 96:931-34 (1984). These
30 single cell organisms exhibit both algae and protozoan characteristics and, depending on conditions, can utilize either light energy (photosynthesis) or chemical energy (eating) for

biochemical processes. Under anaerobic conditions, *E. gracilis* converts paramylon, the reserve beta-1,2-glucan polysaccharide, into wax ester with concomitant generation of ATP, a phenomenon named wax ester fermentation (Inui et al., *supra*, 1982; Inui et al., *Agricultural and Biological Chemistry* 47:2669-2671 (1983)). Fatty acid synthesis through the malonyl-CoA-independent pathway results in a net gain of ATP, whereas other fatty acid synthesis systems can not support the net gain of ATP. ATP also can be produced under aerobic conditions (Inui et al., *Archives Biochemistry and Biophysics* 237:423-29 (1985)).

In the absence of oxygen, acetyl-CoA is generated from pyruvate via an oxygen-sensitive pyruvate:NADP⁺ oxidoreductase (Inui et al., *supra*, 1984; Inui et al., *supra*, 1985; Inui et al., *Archives of Biochemistry and Biophysics* 280:292-98 (1990); Inui et al., *Journal of Biological Chemistry* 262:9130-35 (1987)), and serves as the terminal electron acceptor of glucose oxidation via the malonyl-CoA-independent fatty acid synthesis to form wax ester (Inui et al., *supra*, (1985)). *E. gracilis* contains five different systems of fatty acid synthesis, including four fatty acid synthesis systems located in different compartments, and the mitochondrial malonyl-CoA-independent FAS system involved in anaerobic wax ester fermentation (Hoffmeister et al., *J. of Biological Chemistry* 280:4329-38 (2005)). The malonyl-CoA-independent FAS system has been shown to produce C8-C18 fatty acids. A fatty acid is reduced to alcohol, esterified with another fatty acid, and deposited in the cytosol as waxes (Inui et al., *Febs Letters* 150:89-93 (1982); Inui et al., *European Journal of Biochemistry* 142:121-126 (1984)). The wax can constitute approximately 50% of the total lipid in dark grown cells (Rosenberg, A., *Biochemistry* 2:1148 (1963)). A particularly useful embodiment of the invention harness the malonyl-CoA-independent fatty acid synthesis (FAS) system under anaerobic conditions to produce large quantities of alcohols using the modified biosynthetic pathways described herein.

The malonyl-CoA-independent fatty acid synthesis pathway is similar to the reversal of fatty acid oxidation and is referred as the fatty acid synthesis in mitochondria or acyl-carrier protein (ACP)-independent fatty acid synthesis as it is known in the art. Compared to the malonyl-CoA-dependent fatty acid synthesis (a.k.a. ACP dependent fatty acid synthesis; Smith et al., *Progress in Lipid Research* 42:289-317 (2003); White et al., *Annual Review of Biochemistry* 74:791-831 (2005)), there are several differences. First, acetyl-CoA is used as the extension unit instead of malonyl-ACP. Utilization of acetyl-CoA as elongation substrate in the malonyl-CoA-independent pathway eliminates the need for acetyl-CoA carboxylase

complex (ACC), which converts acetyl-CoA to malonyl-CoA, and thus conserves one ATP molecule per unit flux of acetyl-CoA entering the elongation cycle. Second, all of the intermediates in the elongation cycle are attached to CoA instead of ACP. The elongation cycle can include (i) ketoacyl-CoA acyltransferase (or ketoacyl-CoA thiolase, EC 2.3.1.16),
5 (ii) 3-hydroxyacyl-CoA dehydrogenase (EC 1.1.1.35 and 1.1.1.211), (iii) enoyl-CoA hydratase (EC 4.2.1.17 and 4.2.1.74), and (iv) enoyl-CoA reductase (EC 1.3.1.44 and 1.3.1.38). Third, the product from the elongation cycle is acyl-CoA, which can be utilized directly by acyl-CoA reductase, followed by a dehydrogenase for conversion to alcohol, or by fatty acid forming acyl-CoA reductase (FAR), which converts acyl-CoA directly to alcohol.
10 Therefore, thioesterase and acyl-CoA synthase are not required for the production of primary alcohols, as is the case with the malonyl-CoA-dependent pathways.

For example, the microorganisms of the invention utilize the malonyl-CoA-independent fatty acid synthesis pathway coupled with the reduction of the fatty acid to form primary alcohol as illustrated in Figure 1. The microorganism can additionally be modified to convert, for
15 example, renewable feedstock to acetyl-CoA. In the bioengineered pathways of the invention, acetyl-CoA can be used as both a primer and an extension unit in the elongation cycle described above. At the end of each elongation cycle, an acyl-CoA is formed that is one C2 unit longer than the acyl-CoA entering the elongation cycle. Coupling the above synthesis pathway to a reduction pathway yields the primary alcohol products of the
20 invention. Particularly useful is the coupling of acyl-CoA having a desired chain-length to a reduction pathway that uses the combination of chain-length specific acyl-CoA reductase (EC 1.2.1.50) and alcohol dehydrogenase (1.1.1.1) or the fatty alcohol forming acyl-CoA reductase (FAR, EC 1.1.1) to form desired primary alcohol. Carbon chain-length of the primary alcohols can be controlled by chain-length specific enoyl-CoA reductase, ketoacyl-
25 CoA thiolase and/or acyl-CoA reductase.

The microorganisms of the invention having the coupled biosynthetic pathways described above can produce primary alcohols at very high levels. For example, the maximum theoretical yield for octanol using the malonyl-CoA-independent fatty acid biosynthetic pathway and the associated energetics were calculated by adding the malonyl-CoA-
30 independent fatty acid synthesis, acyl-CoA reductase and alcohol dehydrogenase reactions to a predictive *E. coli* metabolic stoichiometric network using the *in silico* metabolic modeling system known in the art as SimPheny™ (see, for example, U.S. Patent Application Serial No.

10/173,547, filed June 14, 2002, and in International Patent Application No.

PCT/US03/18838, filed June 13, 2003). The model assumes that the secretion of octanol does not require energy. Table 4 shows the maximum theoretical yield for octanol under both aerobic and anaerobic conditions. The malonyl-CoA-independent fatty acid biosynthetic pathway is much more energy-efficient than the malonyl-CoA-dependent fatty acid synthesis pathways, and allows for a maximum theoretical yield of 0.5 mole octanol/mole of glucose and maximum ATP yield of 2.125 mole/mole of glucose under both aerobic and anaerobic conditions.

Table 4: Comparison of the maximum theoretical yield of octanol using (1) the malonyl-CoA-independent fatty acid synthesis and acyl-reduction pathway and (2) the ACP-dependent fatty acid synthesis and pathway.

	Malonyl-CoA-independent fatty acid biosynthetic and reduction pathway		Typical fatty acid biosynthetic and reduction pathway	
	Anaerobic	Aerobic	Anaerobic	Aerobic
Octanol Yield (mole/mole glucose)	0.5	0.5	0.375	0.48
Max ATP Yield @ max octanol yield (mole/mole glucose)	2.125	2.125	0	0

A non-naturally occurring microbial organism of the invention employs combinations of metabolic reactions for biosynthetically producing a target primary alcohol or a target mixture of primary alcohols of the invention. The combination of metabolic reactions can be engineered in a variety of different alternatives to achieve exogenous expression of a malonyl-CoA-independent FAS pathway in sufficient amounts to produce a primary alcohol. The non-naturally occurring microbial organisms will express at least one exogenous nucleic acid encoding a malonyl-CoA-independent FAS pathway enzyme. In certain embodiments, the non-naturally occurring microbial organisms of the invention will be engineered to exogenously express more than one, including all, nucleic acids encoding some or all of the enzymes for the complete pathway of malonyl-CoA independent FAS pathway enzymes. Some or all of the enzymes for acyl-reduction also can be exogenously expressed. Exogenous expression should be at levels sufficient to produce metabolically utilizable gene product and result in the production of a target primary alcohol or set of alcohols.

The biochemical reactions for formation of primary alcohols from a carbon or other energy source through a malonyl-CoA independent FAS pathway is shown in Figure 1. The malonyl-CoA independent FAS pathway produces acyl-CoA. Concomitant utilization of this intermediate product to produce target primary alcohols by an acyl-reduction pathway also is shown in Figure 1.

This invention is also directed, in part to engineered biosynthetic pathways to improve carbon flux through a central metabolism intermediate en route to a primary alcohol, a fatty acyl-CoA, a fatty ester, or a wax. The present invention provides non-naturally occurring microbial organisms having one or more exogenous genes encoding enzymes that can catalyze various enzymatic transformations en route to a primary alcohol, a fatty acyl-CoA, a fatty ester, or a wax. In some embodiments, these enzymatic transformations are part of the reductive tricarboxylic acid (RTCA) cycle and are used to improve product yields, including but not limited to, from carbohydrate-based carbon feedstock.

In numerous engineered pathways, realization of maximum product yields based on carbohydrate feedstock is hampered by insufficient reducing equivalents or by loss of reducing equivalents and/or carbon to byproducts. In accordance with some embodiments, the present invention increases the yields of a primary alcohol, a fatty acyl-CoA, a fatty ester, or a wax by (i) enhancing carbon fixation via the reductive TCA cycle, and/or (ii) accessing additional reducing equivalents from gaseous carbon sources and/or syngas components such as CO, CO₂, and/or H₂. In addition to syngas, other sources of such gases include, but are not limited to, the atmosphere, either as found in nature or generated.

The CO₂-fixing reductive tricarboxylic acid (RTCA) cycle is an endergenic anabolic pathway of CO₂ assimilation which uses reducing equivalents and ATP (Figure 21). One turn of the RTCA cycle assimilates two moles of CO₂ into one mole of acetyl-CoA, or four moles of CO₂ into one mole of oxaloacetate. This additional availability of acetyl-CoA improves the maximum theoretical yield of product molecules derived from carbohydrate-based carbon feedstock. Exemplary carbohydrates include but are not limited to glucose, sucrose, xylose, arabinose and glycerol.

In some embodiments, the reductive TCA cycle, coupled with carbon monoxide dehydrogenase and/or hydrogenase enzymes, can be employed to allow syngas, CO₂, CO, H₂, and/or other gaseous carbon source utilization by microorganisms. Synthesis gas (syngas), in

particular is a mixture of primarily H₂ and CO, sometimes including some amounts of CO₂, 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 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. In addition to coal, biomass of many types has been used for syngas production and represents an inexpensive and flexible feedstock for the biological production of renewable chemicals and fuels. Carbon dioxide can be provided from the atmosphere or in condensed form, for example, from a tank cylinder, or via sublimation of solid CO₂. Similarly, CO and hydrogen gas can be provided in reagent form and/or mixed in any desired ratio. Other gaseous carbon forms can include, for example, methanol or similar volatile organic solvents.

The components of synthesis gas and/or other carbon sources can provide sufficient CO₂, reducing equivalents, and ATP for the reductive TCA cycle to operate. One turn of the RTCA cycle assimilates two moles of CO₂ into one mole of acetyl-CoA and requires 2 ATP and 4 reducing equivalents. CO and/or H₂ can provide reducing equivalents by means of carbon monoxide dehydrogenase and hydrogenase enzymes, respectively. Reducing equivalents can come in the form of NADH, NADPH, FADH, reduced quinones, reduced ferredoxins, thioredoxins, and reduced flavodoxins. The reducing equivalents, particularly NADH, NADPH, and reduced ferredoxin, can serve as cofactors for the RTCA cycle enzymes, for example, malate dehydrogenase, fumarate reductase, alpha-ketoglutarate:ferredoxin oxidoreductase (alternatively known as 2-oxoglutarate:ferredoxin oxidoreductase, alpha-ketoglutarate synthase, or 2-oxoglutarate synthase), pyruvate:ferredoxin oxidoreductase and isocitrate dehydrogenase. The electrons from these reducing equivalents can alternatively pass through an ion-gradient producing electron transport chain where they are passed to an acceptor such as oxygen, nitrate, oxidized metal ions, protons, or an electrode. The ion-gradient can then be used for ATP generation via an ATP synthase or similar enzyme.

The reductive TCA cycle was first reported in the green sulfur photosynthetic bacterium *Chlorobium limicola* (Evans et al., *Proc. Natl. Acad. Sci. U.S.A.* 55:928-934 (1966)). Similar pathways have been characterized in some prokaryotes (proteobacteria, green sulfur bacteria and thermophilic Knallgas bacteria) and sulfur-dependent archaea (Hugler et al., *J.*

Bacteriol. 187:3020-3027 (2005; Hugler et al., *Environ. Microbiol.* 9:81-92 (2007)). In some cases, reductive and oxidative (Krebs) TCA cycles are present in the same organism (Hugler et al., *supra* (2007); Siebers et al., *J. Bacteriol.* 186:2179-2194 (2004)). Some methanogens and obligate anaerobes possess incomplete oxidative or reductive TCA cycles that may
 5 function to synthesize biosynthetic intermediates (Ekiel et al., *J. Bacteriol.* 162:905-908 (1985); Wood et al., *FEMS Microbiol. Rev.* 28:335-352 (2004)).

The key carbon-fixing enzymes of the reductive TCA cycle are alpha-ketoglutarate:ferredoxin oxidoreductase, pyruvate:ferredoxin oxidoreductase and isocitrate dehydrogenase. Additional carbon may be fixed during the conversion of
 10 phosphoenolpyruvate to oxaloacetate by phosphoenolpyruvate carboxylase or phosphoenolpyruvate carboxykinase or by conversion of pyruvate to malate by malic enzyme..

Many of the enzymes in the TCA cycle are reversible and can catalyze reactions in the reductive and oxidative directions. However, some TCA cycle reactions are irreversible *in vivo* and thus different enzymes are used to catalyze these reactions in the directions required for the reverse TCA cycle. These reactions are: (1) conversion of citrate to oxaloacetate and acetyl-CoA, (2) conversion of fumarate to succinate, and (3) conversion of succinyl-CoA to alpha-ketoglutarate. In the TCA cycle, citrate is formed from the condensation of
 15 oxaloacetate and acetyl-CoA. The reverse reaction, cleavage of citrate to oxaloacetate and acetyl-CoA, is ATP-dependent and catalyzed by ATP-citrate lyase, or citryl-CoA synthetase and citryl-CoA lyase. Alternatively, citrate lyase can be coupled to acetyl-CoA synthetase, an acetyl-CoA transferase, or phosphotransacetylase and acetate kinase to form acetyl-CoA and oxaloacetate from citrate. The conversion of succinate to fumarate is catalyzed by
 20 succinate dehydrogenase while the reverse reaction is catalyzed by fumarate reductase. In the TCA cycle succinyl-CoA is formed from the NAD(P)⁺ dependent decarboxylation of alpha-ketoglutarate by the alpha-ketoglutarate dehydrogenase complex. The reverse reaction is
 25 catalyzed by alpha-ketoglutarate:ferredoxin oxidoreductase.

An organism capable of utilizing the reverse tricarboxylic acid cycle to enable production of acetyl-CoA-derived products on 1) CO, 2) CO₂ and H₂, 3) CO and CO₂, 4) synthesis gas
 30 comprising CO and H₂, and 5) synthesis gas or other gaseous carbon sources comprising CO, CO₂, and H₂ can include any of the following enzyme activities: ATP-citrate lyase, a citrate lyase, a citryl-CoA synthetase, a citryl-CoA lyase, aconitase, isocitrate dehydrogenase, alpha-

ketoglutarate:ferredoxin oxidoreductase, succinyl-CoA synthetase, succinyl-CoA transferase, fumarate reductase, fumarase, malate dehydrogenase, acetate kinase, phosphotransacetylase, acetyl-CoA synthetase, acetyl-CoA transferase, pyruvate:ferredoxin oxidoreductase, NAD(P)H:ferredoxin oxidoreductase, carbon monoxide dehydrogenase, hydrogenase, and
5 ferredoxin (see Figure 23). Enzymes and the corresponding genes required for these activities are described herein above.

Carbon from syngas or other gaseous carbon sources can be fixed via the reverse TCA cycle and components thereof. Specifically, the combination of certain carbon gas-utilization pathway components with the pathways for formation of a primary alcohol, a fatty acyl-CoA,
10 a fatty ester, or a wax from acetyl-CoA results in high yields of these products by providing an efficient mechanism for fixing the carbon present in carbon dioxide, fed exogenously or produced endogenously from CO, into acetyl-CoA.

In some embodiments, a primary alcohol, a fatty acyl-CoA, a fatty ester, or a wax pathway in a non-naturally occurring microbial organism of the invention can utilize any combination of
15 (1) CO, (2) CO₂, (3) H₂, or mixtures thereof to enhance the yields of biosynthetic steps involving reduction, including addition to driving the reductive TCA cycle.

In some embodiments a non-naturally occurring microbial organism having a primary alcohol, a fatty acyl-CoA, a fatty ester, or a wax pathway includes at least one exogenous nucleic acid encoding a reductive TCA pathway enzyme. The at least one exogenous nucleic acid is selected from an ATP-citrate lyase, a citrate lyase, a citryl-CoA synthetase, a citryl-CoA lyase, a fumarate reductase, isocitrate dehydrogenase, aconitase, and an alpha-ketoglutarate:ferredoxin oxidoreductase; and at least one exogenous enzyme selected from a carbon monoxide dehydrogenase, a hydrogenase, a NAD(P)H:ferredoxin oxidoreductase, and a ferredoxin, expressed in a sufficient amount to allow the utilization of (1) CO, (2) CO₂, (3)
20 H₂, (4) CO₂ and H₂, (5) CO and CO₂, (6) CO and H₂, or (7) CO, CO₂, and H₂.

In some embodiments a method includes culturing a non-naturally occurring microbial organism having a primary alcohol, a fatty acyl-CoA, a fatty ester, or a wax pathway also comprising at least one exogenous nucleic acid encoding a reductive TCA pathway enzyme. The at least one exogenous nucleic acid is selected from an ATP-citrate lyase, a citrate lyase, a citryl-CoA synthetase, a citryl-CoA lyase, a fumarate reductase, isocitrate dehydrogenase, aconitase, and an alpha-ketoglutarate:ferredoxin oxidoreductase. Additionally, such an
30

organism can also include at least one exogenous enzyme selected from a carbon monoxide dehydrogenase, a hydrogenase, a NAD(P)H:ferredoxin oxidoreductase, and a ferredoxin, expressed in a sufficient amount to allow the utilization of (1) CO, (2) CO₂, (3) H₂, (4) CO₂ and H₂, (5) CO and CO₂, (6) CO and H₂, or (7) CO, CO₂, and H₂ to produce a product.

- 5 In some embodiments a non-naturally occurring microbial organism having a primary alcohol, a fatty acyl-CoA, a fatty ester, or a wax pathway further includes at least one exogenous nucleic acid encoding a reductive TCA pathway enzyme expressed in a sufficient amount to enhance carbon flux through acetyl-CoA. The at least one exogenous nucleic acid is selected from an ATP-citrate lyase, a citrate lyase, a citryl-CoA synthetase, a citryl-CoA
10 lyase, a fumarate reductase, a pyruvate:ferredoxin oxidoreductase, isocitrate dehydrogenase, aconitase, and an alpha-ketoglutarate:ferredoxin oxidoreductase.

- In some embodiments a non-naturally occurring microbial organism having a primary alcohol, a fatty acyl-CoA, a fatty ester, or a wax pathway includes at least one exogenous nucleic acid encoding an enzyme expressed in a sufficient amount to enhance the availability
15 of reducing equivalents in the presence of carbon monoxide and/or hydrogen, thereby increasing the yield of redox-limited products via carbohydrate-based carbon feedstock. The at least one exogenous nucleic acid is selected from a carbon monoxide dehydrogenase, a hydrogenase, an NAD(P)H:ferredoxin oxidoreductase, and a ferredoxin. In some embodiments, the present invention provides a method for enhancing the availability of
20 reducing equivalents in the presence of carbon monoxide or hydrogen thereby increasing the yield of redox-limited products via carbohydrate-based carbon feedstock, such as sugars or gaseous carbon sources, the method includes culturing this non-naturally occurring microbial organism under conditions and for a sufficient period of time to produce a primary alcohol, a fatty acyl-CoA, a fatty ester, or a wax.

- 25 In some embodiments, the non-naturally occurring microbial organism having a primary alcohol, a fatty acyl-CoA, a fatty ester, or a wax pathway includes two exogenous nucleic acids, each encoding a reductive TCA pathway enzyme. In some embodiments, the non-naturally occurring microbial organism having a primary alcohol, a fatty acyl-CoA, a fatty ester, or a wax pathway includes three exogenous nucleic acids each encoding a reductive
30 TCA pathway enzyme. In some embodiments, the non-naturally occurring microbial organism includes three exogenous nucleic acids encoding an ATP-citrate lyase, a fumarate reductase, and an alpha-ketoglutarate:ferredoxin oxidoreductase. In some embodiments, the

non-naturally occurring microbial organism includes three exogenous nucleic acids encoding a citrate lyase, a fumarate reductase, and an alpha-ketoglutarate:ferredoxin oxidoreductase.

In some embodiments, the non-naturally occurring microbial organism includes four exogenous nucleic acids encoding a pyruvate:ferredoxin oxidoreductase; a

- 5 phosphoenolpyruvate carboxylase or a phosphoenolpyruvate carboxykinase, a CO dehydrogenase; and an H₂ hydrogenase. In some embodiments, the non-naturally occurring microbial organism includes two exogenous nucleic acids encoding a CO dehydrogenase and an H₂ hydrogenase.

- 10 In some embodiments, the non-naturally occurring microbial organisms having a primary alcohol, a fatty acyl-CoA, a fatty ester, or a wax pathway further include an exogenous nucleic acid encoding an enzyme selected from a pyruvate:ferredoxin oxidoreductase, an aconitase, an isocitrate dehydrogenase, a succinyl-CoA synthetase, a succinyl-CoA transferase, a fumarase, a malate dehydrogenase, an acetate kinase, a phosphotransacetylase, an acetyl-CoA synthetase, an NAD(P)H:ferredoxin oxidoreductase, and combinations
15 thereof.

In some embodiments, the non-naturally occurring microbial organism having a primary alcohol, a fatty acyl-CoA, a fatty ester, or a wax pathway further includes an exogenous nucleic acid encoding an enzyme selected from carbon monoxide dehydrogenase, acetyl-CoA synthase, ferredoxin, NAD(P)H:ferredoxin oxidoreductase and combinations thereof.

- 20 In some embodiments, the non-naturally occurring microbial organism having a primary alcohol, a fatty acyl-CoA, a fatty ester, or a wax pathway utilizes a carbon feedstock selected from (1) CO, (2) CO₂, (3) CO₂ and H₂, (4) CO and H₂, or (5) CO, CO₂, and H₂. In some embodiments, the non-naturally occurring microbial organism having a primary alcohol, a fatty acyl-CoA, a fatty ester, or a wax pathway utilizes hydrogen for reducing equivalents. In
25 some embodiments, the non-naturally occurring microbial organism having a primary alcohol, a fatty acyl-CoA, a fatty ester, or a wax pathway utilizes CO for reducing equivalents. In some embodiments, the non-naturally occurring microbial organism having a primary alcohol, a fatty acyl-CoA, a fatty ester, or a wax pathway utilizes combinations of CO and hydrogen for reducing equivalents.

- 30 In some embodiments, the non-naturally occurring microbial organism having a primary alcohol, a fatty acyl-CoA, a fatty ester, or a wax pathway further includes one or more

nucleic acids encoding an enzyme selected from a phosphoenolpyruvate carboxylase, a phosphoenolpyruvate carboxykinase, a pyruvate carboxylase, and a malic enzyme.

In some embodiments, the non-naturally occurring microbial organism having a primary alcohol, a fatty acyl-CoA, a fatty ester, or a wax pathway further includes one or more

5 nucleic acids encoding an enzyme selected from a malate dehydrogenase, a fumarase, a fumarate reductase, a succinyl-CoA synthetase, and a succinyl-CoA transferase.

In some embodiments, the non-naturally occurring microbial organism having a primary alcohol, a fatty acyl-CoA, a fatty ester, or a wax pathway further includes at least one exogenous nucleic acid encoding a citrate lyase, an ATP-citrate lyase, a citryl-CoA

10 synthetase, a citryl-CoA lyase, an aconitase, an isocitrate dehydrogenase, a succinyl-CoA synthetase, a succinyl-CoA transferase, a fumarase, a malate dehydrogenase, an acetate kinase, a phosphotransacetylase, an acetyl-CoA synthetase, and a ferredoxin.

It is understood by those skilled in the art that the above-described pathways for increasing product yield can be combined with any of the pathways disclosed herein, including those pathways depicted in the figures. One skilled in the art will understand that, depending on the pathway to a desired product and the precursors and intermediates of that pathway, a particular pathway for improving product yield, as discussed herein above and in the examples, or combination of such pathways, can be used in combination with a pathway to a desired product to increase the yield of that product or a pathway intermediate.

20 In one embodiment, the invention provides a non-naturally occurring microbial organism, comprising a microbial organism having a primary alcohol pathway comprising at least one exogenous nucleic acid encoding a primary alcohol pathway enzyme expressed in a sufficient amount to produce a primary alcohol; said non-naturally occurring microbial organism further comprising: (i) a reductive TCA pathway comprising at least one exogenous nucleic acid encoding a reductive TCA pathway enzyme, wherein said at least one exogenous nucleic acid is selected from an ATP-citrate lyase, a citrate lyase, a citryl-CoA synthetase, a citryl-CoA lyase, a fumarate reductase, and an alpha-ketoglutarate:ferredoxin oxidoreductase; (ii) a reductive TCA pathway comprising at least one exogenous nucleic acid encoding a reductive TCA pathway enzyme, wherein said at least one exogenous nucleic acid is selected from a pyruvate:ferredoxin oxidoreductase, a phosphoenolpyruvate carboxylase, a phosphoenolpyruvate carboxykinase, a CO dehydrogenase, and an H₂ hydrogenase; or (iii) at

25

30

least one exogenous nucleic acid encodes an enzyme selected from a CO dehydrogenase, an H₂ hydrogenase, and combinations thereof; and wherein said primary alcohol pathway comprises a malonyl-CoA-independent FAS pathway comprising ketoacyl-CoA acyltransferase or ketoacyl-CoA thiolase, 3-hydroxyacyl-CoA dehydrogenase, enoyl-CoA hydratase and enoyl-CoA reductase; and an acyl-reduction pathway comprising an acyl-CoA reductase and an alcohol dehydrogenase. In a specific embodiment, said microbial organism comprising (i) further comprises an exogenous nucleic acid encoding an enzyme selected from a pyruvate:ferredoxin oxidoreductase, an aconitase, an isocitrate dehydrogenase, a succinyl-CoA synthetase, a succinyl-CoA transferase, a fumarase, a malate dehydrogenase, an acetate kinase, a phosphotransacetylase, an acetyl-CoA synthetase, an NAD(P)H:ferredoxin oxidoreductase, ferredoxin, and combinations thereof. In another specific embodiment, said microbial organism comprising (ii) further comprises an exogenous nucleic acid encoding an enzyme selected from an aconitase, an isocitrate dehydrogenase, a succinyl-CoA synthetase, a succinyl-CoA transferase, a fumarase, a malate dehydrogenase, and combinations thereof. In another specific embodiment, said microbial organism comprises two, three, four, five or six exogenous nucleic acids each encoding a primary alcohol, a fatty acyl-CoA, a fatty ester or a wax pathway enzyme. In another specific embodiment, said microbial organism comprises four exogenous nucleic acids encoding malonyl-CoA-independent FAS pathway enzymes comprising ketoacyl-CoA acyltransferase or ketoacyl-CoA thiolase, 3-hydroxyacyl-CoA dehydrogenase, enoyl-CoA hydratase and enoyl-CoA reductase; and two exogenous nucleic acids encoding acyl-reduction pathway enzymes comprising an acyl-CoA reductase and an alcohol dehydrogenase. In another specific embodiment, said microbial organism comprising (i) comprises two, three or four exogenous nucleic acids each encoding a reductive TCA pathway enzyme. In another specific embodiment, said microbial organism comprising (ii) comprises two, three or four exogenous nucleic acids each encoding a reductive TCA pathway enzyme. In another specific embodiment, said at least one exogenous nucleic acid in said non-naturally occurring microbial organism is a heterologous nucleic acid. In another specific embodiment, said non-naturally occurring microbial organism is in a substantially anaerobic culture medium. In another specific embodiment, said exogenous nucleic acid encoding an acyl-reduction pathway enzyme in said non-naturally occurring microbial organism comprises an enzyme having acyl-CoA reductase and alcohol dehydrogenase activity, e.g., fatty alcohol forming acyl-CoA reductase (FAR). In another embodiment, said non-naturally occurring microbial organism further comprises an acyl-reduction pathway comprising an acyl-CoA hydrolase, an

acyl-CoA transferase or an acyl-CoA ligase; a carboxylic acid reductase and an alcohol dehydrogenase. In another specific embodiment, said primary alcohol is produced in amounts at least 10% greater levels compared to a microbial organism lacking said exogenous nucleic acid encoding a malonyl-CoA-independent FAS pathway enzyme. In

5 another specific embodiment, said primary alcohol comprises an alcohol having between 4-24 carbon atoms, e.g., said primary alcohol is selected from butanol, hexanol, heptanol, octanol, nonanol, decanol, dodecanol, tetradecanol or hexadecanol.

In another embodiment, the invention provides a non-naturally occurring microbial organism, comprising a microbial organism having a fatty acyl-CoA pathway comprising at least one
10 exogenous nucleic acid encoding a fatty acyl-CoA pathway enzyme expressed in a sufficient amount to produce a fatty acyl-CoA; said non-naturally occurring microbial organism further comprising: (i) a reductive TCA pathway comprising at least one exogenous nucleic acid encoding a reductive TCA pathway enzyme, wherein said at least one exogenous nucleic acid is selected from an ATP-citrate lyase, a citrate lyase, a citryl-CoA synthetase, a citryl-CoA
15 lyase, a fumarate reductase, and an alpha-ketoglutarate:ferredoxin oxidoreductase; (ii) a reductive TCA pathway comprising at least one exogenous nucleic acid encoding a reductive TCA pathway enzyme, wherein said at least one exogenous nucleic acid is selected from a pyruvate:ferredoxin oxidoreductase, a phosphoenolpyruvate carboxylase, a phosphoenolpyruvate carboxykinase, a CO dehydrogenase, and an H₂ hydrogenase; or (iii) at
20 least one exogenous nucleic acid encodes an enzyme selected from a CO dehydrogenase, an H₂ hydrogenase, and combinations thereof; and wherein said fatty acyl-CoA pathway comprises a malonyl-CoA-independent FAS pathway comprising ketoacyl-CoA acyltransferase or ketoacyl-CoA thiolase, 3-hydroxyacyl-CoA dehydrogenase, enoyl-CoA hydratase and enoyl-CoA reductase. In a specific embodiment, said microbial organism
25 comprising (i) further comprises an exogenous nucleic acid encoding an enzyme selected from a pyruvate:ferredoxin oxidoreductase, an aconitase, an isocitrate dehydrogenase, a succinyl-CoA synthetase, a succinyl-CoA transferase, a fumarase, a malate dehydrogenase, an acetate kinase, a phosphotransacetylase, an acetyl-CoA synthetase, an NAD(P)H:ferredoxin oxidoreductase, ferredoxin, and combinations thereof. In another
30 specific embodiment, said microbial organism comprising (ii) further comprises an exogenous nucleic acid encoding an enzyme selected from an aconitase, an isocitrate dehydrogenase, a succinyl-CoA synthetase, a succinyl-CoA transferase, a fumarase, a malate dehydrogenase, and combinations thereof. In another specific embodiment, said microbial

organism comprises two, three, four, five or six exogenous nucleic acids each encoding a primary alcohol, a fatty acyl-CoA, a fatty ester or a wax pathway enzyme. In another specific embodiment, said microbial organism comprises four exogenous nucleic acids encoding malonyl-CoA-independent FAS pathway enzymes comprising ketoacyl-CoA acyltransferase or ketoacyl-CoA thiolase, 3-hydroxyacyl-CoA dehydrogenase, enoyl-CoA hydratase and enoyl-CoA reductase. In another specific embodiment, said microbial organism comprising (i) comprises two, three or four exogenous nucleic acids each encoding a reductive TCA pathway enzyme. In another specific embodiment, said microbial organism comprising (ii) comprises two, three or four exogenous nucleic acids each encoding a reductive TCA pathway enzyme. In another specific embodiment, said at least one exogenous nucleic acid in said non-naturally occurring microbial organism is a heterologous nucleic acid. In another specific embodiment, said non-naturally occurring microbial organism is in a substantially anaerobic culture medium.

In another embodiment, the invention provides a non-naturally occurring microbial organism, comprising a microbial organism having a fatty ester pathway comprising at least one exogenous nucleic acid encoding a fatty ester pathway enzyme expressed in a sufficient amount to produce a fatty ester; said non-naturally occurring microbial organism further comprising: (i) a reductive TCA pathway comprising at least one exogenous nucleic acid encoding a reductive TCA pathway enzyme, wherein said at least one exogenous nucleic acid is selected from an ATP-citrate lyase, a citrate lyase, a citryl-CoA synthetase, a citryl-CoA lyase, a fumarate reductase, and an alpha-ketoglutarate:ferredoxin oxidoreductase; (ii) a reductive TCA pathway comprising at least one exogenous nucleic acid encoding a reductive TCA pathway enzyme, wherein said at least one exogenous nucleic acid is selected from a pyruvate:ferredoxin oxidoreductase, a phosphoenolpyruvate carboxylase, a phosphoenolpyruvate carboxykinase, a CO dehydrogenase, and an H₂ hydrogenase; or (iii) at least one exogenous nucleic acid encodes an enzyme selected from a CO dehydrogenase, an H₂ hydrogenase, and combinations thereof; and wherein said fatty ester pathway comprises a pathway selected from: (a) a malonyl-CoA-independent fatty acid synthesis (FAS) pathway and a wax ester synthase, wherein said malonyl-CoA-independent FAS pathway comprises ketoacyl-CoA acyltransferase or ketoacyl-CoA thiolase, 3-hydroxyacyl-CoA dehydrogenase, enoyl-CoA hydratase and enoyl-CoA reductase; and (b) a malonyl-CoA-independent fatty acid synthesis (FAS) pathway and an alcohol acetyltransferase, wherein said malonyl-CoA-independent FAS pathway comprises ketoacyl-CoA acyltransferase or ketoacyl-CoA

thiolase, 3-hydroxyacyl-CoA dehydrogenase, enoyl-CoA hydratase and enoyl-CoA reductase.

In a specific embodiment, said microbial organism comprising (i) further comprises an exogenous nucleic acid encoding an enzyme selected from a pyruvate:ferredoxin

oxidoreductase, an aconitase, an isocitrate dehydrogenase, a succinyl-CoA synthetase, a

5 succinyl-CoA transferase, a fumarase, a malate dehydrogenase, an acetate kinase, a phosphotransacetylase, an acetyl-CoA synthetase, an NAD(P)H:ferredoxin oxidoreductase, ferredoxin, and combinations thereof. In another specific embodiment, said microbial

organism comprising (ii) further comprises an exogenous nucleic acid encoding an enzyme selected from an aconitase, an isocitrate dehydrogenase, a succinyl-CoA synthetase, a

10 succinyl-CoA transferase, a fumarase, a malate dehydrogenase, and combinations thereof. In

another specific embodiment, said microbial organism comprises two, three, four, five or six exogenous nucleic acids each encoding a primary alcohol, a fatty acyl-CoA, a fatty ester or a wax pathway enzyme. In another specific embodiment, said microbial organism comprises

(A) four exogenous nucleic acids encoding malonyl-CoA-independent fatty acid synthesis

15 (FAS) pathway enzymes and one exogenous nucleic acid encoding a wax ester synthase,

wherein said malonyl-CoA-independent FAS pathway enzymes comprise a ketoacyl-CoA acyltransferase or a ketoacyl-CoA thiolase, a 3-hydroxyacyl-CoA dehydrogenase, an enoyl-CoA hydratase and an enoyl-CoA reductase; or (B) four exogenous nucleic acids encoding

malonyl-CoA-independent fatty acid synthesis (FAS) pathway enzymes and one exogenous

20 nucleic acid encoding an alcohol acetyltransferase, wherein said malonyl-CoA-independent FAS pathway enzymes comprise a ketoacyl-CoA acyltransferase or a ketoacyl-CoA thiolase, a 3-hydroxyacyl-CoA dehydrogenase, an enoyl-CoA hydratase and an enoyl-CoA reductase.

In another specific embodiment, said microbial organism comprising (i) comprises two, three or four exogenous nucleic acids each encoding a reductive TCA pathway enzyme. In another

25 specific embodiment, said microbial organism comprising (ii) comprises two, three or four exogenous nucleic acids each encoding a reductive TCA pathway enzyme. In another

specific embodiment, said at least one exogenous nucleic acid in said non-naturally occurring microbial organism is a heterologous nucleic acid. In another specific embodiment, said non-naturally occurring microbial organism is in a substantially anaerobic culture medium.

30 In another embodiment, the invention provides a non-naturally occurring microbial organism, comprising a microbial organism having a wax pathway comprising at least one exogenous nucleic acid encoding a wax pathway enzyme expressed in a sufficient amount to produce a wax; said non-naturally occurring microbial organism further comprising: (i) a reductive

TCA pathway comprising at least one exogenous nucleic acid encoding a reductive TCA pathway enzyme, wherein said at least one exogenous nucleic acid is selected from an ATP-citrate lyase, a citrate lyase, a citryl-CoA synthetase, a citryl-CoA lyase, a fumarate reductase, and an alpha-ketoglutarate:ferredoxin oxidoreductase; (ii) a reductive TCA pathway

5 comprising at least one exogenous nucleic acid encoding a reductive TCA pathway enzyme, wherein said at least one exogenous nucleic acid is selected from a pyruvate:ferredoxin oxidoreductase, a phosphoenolpyruvate carboxylase, a phosphoenolpyruvate carboxykinase, a CO dehydrogenase, and an H₂ hydrogenase; or (iii) at least one exogenous nucleic acid encodes an enzyme selected from a CO dehydrogenase, an H₂ hydrogenase, and

10 combinations thereof; and wherein said wax pathway comprises a pathway selected from: (a) a malonyl-CoA-independent fatty acid synthesis (FAS) pathway and a wax ester synthase, wherein said malonyl-CoA-independent FAS pathway comprises ketoacyl-CoA acyltransferase or ketoacyl-CoA thiolase, 3-hydroxyacyl-CoA dehydrogenase, enoyl-CoA hydratase and enoyl-CoA reductase; and (b) a malonyl-CoA-independent fatty acid synthesis

15 (FAS) pathway and an alcohol acetyltransferase, wherein said malonyl-CoA-independent FAS pathway comprises ketoacyl-CoA acyltransferase or ketoacyl-CoA thiolase, 3-hydroxyacyl-CoA dehydrogenase, enoyl-CoA hydratase and enoyl-CoA reductase. In a specific embodiment, said microbial organism comprising (i) further comprises an exogenous nucleic acid encoding an enzyme selected from a pyruvate:ferredoxin oxidoreductase, an

20 aconitase, an isocitrate dehydrogenase, a succinyl-CoA synthetase, a succinyl-CoA transferase, a fumarase, a malate dehydrogenase, an acetate kinase, a phosphotransacetylase, an acetyl-CoA synthetase, an NAD(P)H:ferredoxin oxidoreductase, ferredoxin, and combinations thereof. In another specific embodiment, said microbial organism comprising

(ii) further comprises an exogenous nucleic acid encoding an enzyme selected from an

25 aconitase, an isocitrate dehydrogenase, a succinyl-CoA synthetase, a succinyl-CoA transferase, a fumarase, a malate dehydrogenase, and combinations thereof. In another specific embodiment, said microbial organism comprises two, three, four, five or six exogenous nucleic acids each encoding a primary alcohol, a fatty acyl-CoA, a fatty ester or a wax pathway enzyme. In another specific embodiment, said microbial organism comprises

30 (A) four exogenous nucleic acids encoding malonyl-CoA-independent fatty acid synthesis (FAS) pathway enzymes and one nucleic acid encoding a wax ester synthase, wherein said malonyl-CoA-independent FAS pathway enzymes comprise a ketoacyl-CoA acyltransferase or a ketoacyl-CoA thiolase, a 3-hydroxyacyl-CoA dehydrogenase, an enoyl-CoA hydratase and enoyl-CoA reductase; or (B) four exogenous nucleic acids encoding malonyl-CoA-

independent fatty acid synthesis (FAS) pathway enzymes and one nucleic acid encoding an alcohol acetyltransferase, wherein said malonyl-CoA-independent FAS pathway enzymes comprise a ketoacyl-CoA acyltransferase or a ketoacyl-CoA thiolase, a 3-hydroxyacyl-CoA dehydrogenase, an enoyl-CoA hydratase and an enoyl-CoA reductase. In another specific embodiment, said microbial organism comprising (i) comprises two, three or four exogenous nucleic acids each encoding a reductive TCA pathway enzyme. In another specific embodiment, said microbial organism comprising (ii) comprises two, three or four exogenous nucleic acids each encoding a reductive TCA pathway enzyme. In another specific embodiment, said at least one exogenous nucleic acid in said non-naturally occurring microbial organism is a heterologous nucleic acid. In another specific embodiment, said non-naturally occurring microbial organism is in a substantially anaerobic culture medium.

In another embodiment, the invention provides a non-naturally occurring microbial organism, comprising a microbial organism having an acyl-ACP pathway comprising at least one exogenous nucleic acid encoding an acyl-ACP pathway enzyme expressed in a sufficient amount to produce acyl-ACP; said non-naturally occurring microbial organism further comprising: (i) a reductive TCA pathway comprising at least one exogenous nucleic acid encoding a reductive TCA pathway enzyme, wherein said at least one exogenous nucleic acid is selected from an ATP-citrate lyase, a citrate lyase, a citryl-CoA synthetase, a citryl-CoA lyase, a fumarate reductase, and an alpha-ketoglutarate:ferredoxin oxidoreductase; (ii) a reductive TCA pathway comprising at least one exogenous nucleic acid encoding a reductive TCA pathway enzyme, wherein said at least one exogenous nucleic acid is selected from a pyruvate:ferredoxin oxidoreductase, a phosphoenolpyruvate carboxylase, a phosphoenolpyruvate carboxykinase, a CO dehydrogenase, and an H₂ hydrogenase; or (iii) at least one exogenous nucleic acid encodes an enzyme selected from a CO dehydrogenase, an H₂ hydrogenase, and combinations thereof; and wherein said acyl-ACP pathway comprises a pathway selected from: (A) acetyl-CoA carboxylase and a fatty acid synthase; and (B) a fatty acid synthase. In a specific embodiment, said non-naturally occurring microbial organism further comprises a fatty acid pathway comprising an exogenous nucleic acid encoding a thioesterase expressed in a sufficient amount to produce fatty acid. In another specific embodiment, said non-naturally occurring microbial organism further comprises an acyl-CoA pathway comprising an exogenous nucleic acid encoding an acyl-CoA synthetase or an acyl-CoA ligase expressed in a sufficient amount to produce acyl-CoA. In another specific embodiment, said non-naturally occurring microbial organism further comprises a

fatty aldehyde pathway comprising an exogenous nucleic acid encoding an acyl-CoA reductase (aldehyde forming) or carboxylic acid reductase expressed in a sufficient amount to produce a fatty aldehyde and/or further comprising a fatty alcohol pathway comprising an exogenous nucleic acid encoding an acyl-CoA reductase (alcohol forming) expressed in a sufficient amount to produce a fatty alcohol. In another specific embodiment, said non-naturally occurring microbial organism further comprises a fatty alcohol pathway comprising an exogenous nucleic acid encoding an aldehyde reductase expressed in a sufficient amount to produce a fatty alcohol. In another specific embodiment, said microbial organism comprising (i) further comprises an exogenous nucleic acid encoding an enzyme selected from a pyruvate:ferredoxin oxidoreductase, an aconitase, an isocitrate dehydrogenase, a succinyl-CoA synthetase, a succinyl-CoA transferase, a fumarase, a malate dehydrogenase, an acetate kinase, a phosphotransacetylase, an acetyl-CoA synthetase, an NAD(P)H:ferredoxin oxidoreductase, ferredoxin, and combinations thereof. In another specific embodiment, said microbial organism comprising (ii) further comprises an exogenous nucleic acid encoding an enzyme selected from an aconitase, an isocitrate dehydrogenase, a succinyl-CoA synthetase, a succinyl-CoA transferase, a fumarase, a malate dehydrogenase, and combinations thereof. In another specific embodiment, said microbial organism comprises two, three, four, five, or six exogenous nucleic acids each encoding a acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol pathway enzyme. In another specific embodiment, said microbial organism comprising (i) comprises two, three or four exogenous nucleic acids each encoding a reductive TCA pathway enzyme. In another specific embodiment, said microbial organism comprising (ii) comprises two, three or four exogenous nucleic acids each encoding a reductive TCA pathway enzyme. In another specific embodiment, said said at least one exogenous nucleic acid is a heterologous nucleic acid. In another specific embodiment, said non-naturally occurring microbial organism is in a substantially anaerobic culture medium.

In some embodiments, the carbon feedstock and other cellular uptake sources such as phosphate, ammonia, sulfate, chloride and other halogens can be chosen to alter the isotopic distribution of the atoms present in a primary alcohol, a fatty acyl-CoA, a fatty ester, a wax, an acyl-ACP, a fatty acid, acyl-CoA, a fatty aldehyde or a fatty alcohol or any primary alcohol, fatty acyl-CoA, fatty ester, wax, acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol pathway intermediate. The various carbon feedstock and other uptake sources enumerated above will be referred to herein, collectively, as “uptake sources.” Uptake

sources can provide isotopic enrichment for any atom present in the product primary alcohol, fatty acyl-CoA, fatty ester, wax, acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol or primary alcohol, fatty acyl-CoA, fatty ester, wax, acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol pathway intermediate, or for side products generated in reactions diverging away from a primary alcohol, fatty acyl-CoA, fatty ester, wax, acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol pathway. Isotopic enrichment can be achieved for any target atom including, for example, carbon, hydrogen, oxygen, nitrogen, sulfur, phosphorus, chloride or other halogens.

In some embodiments, the uptake sources can be selected to alter the carbon-12, carbon-13, and carbon-14 ratios. In some embodiments, the uptake sources can be selected to alter the oxygen-16, oxygen-17, and oxygen-18 ratios. In some embodiments, the uptake sources can be selected to alter the hydrogen, deuterium, and tritium ratios. In some embodiments, the uptake sources can be selected to alter the nitrogen-14 and nitrogen-15 ratios. In some embodiments, the uptake sources can be selected to alter the sulfur-32, sulfur-33, sulfur-34, and sulfur-35 ratios. In some embodiments, the uptake sources can be selected to alter the phosphorus-31, phosphorus-32, and phosphorus-33 ratios. In some embodiments, the uptake sources can be selected to alter the chlorine-35, chlorine-36, and chlorine-37 ratios.

In some embodiments, the isotopic ratio of a target atom can be varied to a desired ratio by selecting one or more uptake sources. An uptake source can be derived from a natural source, as found in nature, or from a man-made source, and one skilled in the art can select a natural source, a man-made source, or a combination thereof, to achieve a desired isotopic ratio of a target atom. An example of a man-made uptake source includes, for example, an uptake source that is at least partially derived from a chemical synthetic reaction. Such isotopically enriched uptake sources can be purchased commercially or prepared in the laboratory and/or optionally mixed with a natural source of the uptake source to achieve a desired isotopic ratio. In some embodiments, a target atom isotopic ratio of an uptake source can be achieved by selecting a desired origin of the uptake source as found in nature. For example, as discussed herein, a natural source can be a biobased derived from or synthesized by a biological organism or a source such as petroleum-based products or the atmosphere. In some such embodiments, a source of carbon, for example, can be selected from a fossil fuel-derived carbon source, which can be relatively depleted of carbon-14, or an environmental or

atmospheric carbon source, such as CO₂, which can possess a larger amount of carbon-14 than its petroleum-derived counterpart.

The unstable carbon isotope carbon-14 or radiocarbon makes up for roughly 1 in 10¹² carbon atoms in the earth's atmosphere and has a half-life of about 5700 years. The stock of carbon is replenished in the upper atmosphere by a nuclear reaction involving cosmic rays and ordinary nitrogen (¹⁴N). Fossil fuels contain no carbon-14, as it decayed long ago. Burning of fossil fuels lowers the atmospheric carbon-14 fraction, the so-called "Suess effect".

Methods of determining the isotopic ratios of atoms in a compound are well known to those skilled in the art. Isotopic enrichment is readily assessed by mass spectrometry using techniques known in the art such as accelerated mass spectrometry (AMS), Stable Isotope Ratio Mass Spectrometry (SIRMS) and Site-Specific Natural Isotopic Fractionation by Nuclear Magnetic Resonance (SNIF-NMR). Such mass spectral techniques can be integrated with separation techniques such as liquid chromatography (LC), high performance liquid chromatography (HPLC) and/or gas chromatography, and the like.

In the case of carbon, ASTM D6866 was developed in the United States as a standardized analytical method for determining the biobased content of solid, liquid, and gaseous samples using radiocarbon dating by the American Society for Testing and Materials (ASTM) International. The standard is based on the use of radiocarbon dating for the determination of a product's biobased content. ASTM D6866 was first published in 2004, and the current active version of the standard is ASTM D6866-11 (effective April 1, 2011). Radiocarbon dating techniques are well known to those skilled in the art, including those described herein.

The biobased content of a compound is estimated by the ratio of carbon-14 (¹⁴C) to carbon-12 (¹²C). Specifically, the Fraction Modern (Fm) is computed from the expression: $Fm = (S - B) / (M - B)$, where B, S and M represent the ¹⁴C/¹²C ratios of the blank, the sample and the modern reference, respectively. Fraction Modern is a measurement of the deviation of the ¹⁴C/¹²C ratio of a sample from "Modern." Modern is defined as 95% of the radiocarbon concentration (in AD 1950) of National Bureau of Standards (NBS) Oxalic Acid I (i.e., standard reference materials (SRM) 4990b) normalized to $\delta^{13}C_{VPDB} = -19$ per mil (Olsson, *The use of Oxalic acid as a Standard*, in, Radiocarbon Variations and Absolute Chronology, Nobel Symposium, 12th Proc., John Wiley & Sons, New York (1970)). Mass spectrometry results, for example, measured by ASM, are calculated using the internationally agreed upon

definition of 0.95 times the specific activity of NBS Oxalic Acid I (SRM 4990b) normalized to $\delta^{13}\text{C}_{\text{VPDB}} = -19$ per mil. This is equivalent to an absolute (AD 1950) $^{14}\text{C}/^{12}\text{C}$ ratio of $1.176 \pm 0.010 \times 10^{-12}$ (Karlen et al., *Arkiv Geofysik*, 4:465-471 (1968)). The standard calculations take into account the differential uptake of one isotope with respect to another, for example, the preferential uptake in biological systems of C^{12} over C^{13} over C^{14} , and these corrections are reflected as a Fm corrected for δ^{13} .

An oxalic acid standard (SRM 4990b or HOx 1) was made from a crop of 1955 sugar beet. Although there were 1000 lbs made, this oxalic acid standard is no longer commercially available. The Oxalic Acid II standard (HOx 2; N.I.S.T designation SRM 4990 C) was made from a crop of 1977 French beet molasses. In the early 1980's, a group of 12 laboratories measured the ratios of the two standards. The ratio of the activity of Oxalic acid II to I is 1.2933 ± 0.001 (the weighted mean). The isotopic ratio of HOx II is -17.8 per mille. ASTM D6866-11 suggests use of the available Oxalic Acid II standard SRM 4990 C (Hox2) for the modern standard (see discussion of original vs. currently available oxalic acid standards in Mann, *Radiocarbon*, 25(2):519-527 (1983)). A Fm = 0% represents the entire lack of carbon-14 atoms in a material, thus indicating a fossil (for example, petroleum based) carbon source. A Fm = 100%, after correction for the post-1950 injection of carbon-14 into the atmosphere from nuclear bomb testing, indicates an entirely modern carbon source. As described herein, such a "modern" source includes biobased sources.

As described in ASTM D6866, the percent modern carbon (pMC) can be greater than 100% because of the continuing but diminishing effects of the 1950s nuclear testing programs, which resulted in a considerable enrichment of carbon-14 in the atmosphere as described in ASTM D6866-11. Because all sample carbon-14 activities are referenced to a "pre-bomb" standard, and because nearly all new biobased products are produced in a post-bomb environment, all pMC values (after correction for isotopic fraction) must be multiplied by 0.95 (as of 2010) to better reflect the true biobased content of the sample. A biobased content that is greater than 103% suggests that either an analytical error has occurred, or that the source of biobased carbon is more than several years old.

ASTM D6866 quantifies the biobased content relative to the material's total organic content and does not consider the inorganic carbon and other non-carbon containing substances present. For example, a product that is 50% starch-based material and 50% water would be considered to have a Biobased Content = 100% (50% organic content that is 100% biobased)

based on ASTM D6866. In another example, a product that is 50% starch-based material, 25% petroleum-based, and 25% water would have a Biobased Content = 66.7% (75% organic content but only 50% of the product is biobased). In another example, a product that is 50% organic carbon and is a petroleum-based product would be considered to have a Biobased Content = 0% (50% organic carbon but from fossil sources). Thus, based on the well known methods and known standards for determining the biobased content of a compound or material, one skilled in the art can readily determine the biobased content and/or prepared downstream products that utilize of the invention having a desired biobased content.

Applications of carbon-14 dating techniques to quantify bio-based content of materials are known in the art (Currie et al., *Nuclear Instruments and Methods in Physics Research B*, 172:281-287 (2000)). For example, carbon-14 dating has been used to quantify bio-based content in terephthalate-containing materials (Colonna et al., *Green Chemistry*, 13:2543-2548 (2011)). Notably, polypropylene terephthalate (PPT) polymers derived from renewable 1,3-propanediol and petroleum-derived terephthalic acid resulted in Fm values near 30% (i.e., since 3/11 of the polymeric carbon derives from renewable 1,3-propanediol and 8/11 from the fossil end member terephthalic acid) (Currie et al., *supra*, 2000). In contrast, polybutylene terephthalate polymer derived from both renewable 1,4-butanediol and renewable terephthalic acid resulted in bio-based content exceeding 90% (Colonna et al., *supra*, 2011).

Accordingly, in some embodiments, the present invention provides primary alcohol, fatty acyl-CoA, fatty ester, wax, acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol or a primary alcohol, fatty acyl-CoA, fatty ester, wax, acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol intermediate that has a carbon-12, carbon-13, and carbon-14 ratio that reflects an atmospheric carbon, also referred to as environmental carbon, uptake source. For example, in some aspects the primary alcohol, fatty acyl-CoA, fatty ester, wax, acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol or a primary alcohol, fatty acyl-CoA, fatty ester, wax, acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol intermediate can have an Fm value of at least 10%, at least 15%, at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 98% or as much as 100%. In some such embodiments, the uptake source is CO₂. In some embodiments, the present invention provides primary alcohol, fatty acyl-CoA, fatty

ester, wax, acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol or a primary alcohol, fatty acyl-CoA, fatty ester, wax, acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol intermediate that has a carbon-12, carbon-13, and carbon-14 ratio that reflects petroleum-based carbon uptake source. In this aspect, the primary alcohol, fatty acyl-CoA, fatty ester, wax, acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol or a primary alcohol, fatty acyl-CoA, fatty ester, wax, acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol intermediate can have an Fm value of less than 95%, less than 90%, less than 85%, less than 80%, less than 75%, less than 70%, less than 65%, less than 60%, less than 55%, less than 50%, less than 45%, less than 40%, less than 35%, less than 30%, less than 25%, less than 20%, less than 15%, less than 10%, less than 5%, less than 2% or less than 1%. In some embodiments, the present invention provides primary alcohol, fatty acyl-CoA, fatty ester, wax, acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol or a primary alcohol, fatty acyl-CoA, fatty ester, wax, acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol intermediate that has a carbon-12, carbon-13, and carbon-14 ratio that is obtained by a combination of an atmospheric carbon uptake source with a petroleum-based uptake source. Using such a combination of uptake sources is one way by which the carbon-12, carbon-13, and carbon-14 ratio can be varied, and the respective ratios would reflect the proportions of the uptake sources.

Further, the present invention relates to the biologically produced primary alcohol, fatty acyl-CoA, fatty ester, wax, acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol or primary alcohol, fatty acyl-CoA, fatty ester, wax, acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol intermediate as disclosed herein, and to the products derived therefrom, wherein the primary alcohol, fatty acyl-CoA, fatty ester, wax, acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol or a primary alcohol, fatty acyl-CoA, fatty ester, wax, acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol intermediate has a carbon-12, carbon-13, and carbon-14 isotope ratio of about the same value as the CO₂ that occurs in the environment. For example, in some aspects the invention provides bioderived primary alcohol, fatty acyl-CoA, fatty ester, wax, acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol or a bioderived primary alcohol, fatty acyl-CoA, fatty ester, wax, acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol intermediate having a carbon-12 versus carbon-13 versus carbon-14 isotope ratio of about the same value as the CO₂ that occurs in the environment, or any of the other ratios disclosed herein. It is understood, as disclosed herein, that a product can have a carbon-12 versus carbon-13 versus carbon-14

isotope ratio of about the same value as the CO₂ that occurs in the environment, or any of the ratios disclosed herein, wherein the product is generated from bioderived primary alcohol, fatty acyl-CoA, fatty ester, wax, acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol or a bioderived primary alcohol, fatty acyl-CoA, fatty ester, wax, acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol intermediate as disclosed herein, wherein the bioderived product is chemically modified to generate a final product. Methods of chemically modifying a bioderived product of primary alcohol, fatty acyl-CoA, fatty ester, wax, acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol, or an intermediate thereof, to generate a desired product are well known to those skilled in the art, as described herein. The invention further provides biofuels, chemicals, polymers, surfactants, soaps, detergents, shampoos, lubricating oil additives, fragrances, flavor materials or acrylates having a carbon-12 versus carbon-13 versus carbon-14 isotope ratio of about the same value as the CO₂ that occurs in the environment, wherein the biofuels, chemicals, polymers, surfactants, soaps, detergents, shampoos, lubricating oil additives, fragrances, flavor materials or acrylates are generated directly from or in combination with bioderived primary alcohol, fatty acyl-CoA, fatty ester, wax, acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol or a bioderived primary alcohol, fatty acyl-CoA, fatty ester, wax, acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol intermediate as disclosed herein.

Primary alcohols, fatty esters, waxes, fatty acids, fatty aldehydes and fatty alcohols are chemicals used in commercial and industrial applications. Non-limiting examples of such applications include production of biofuels, chemicals, polymers, surfactants, soaps, detergents, shampoos, lubricating oil additives, fragrances, flavor materials and acrylates. Accordingly, in some embodiments, the invention provides biobased biofuels, chemicals, polymers, surfactants, soaps, detergents, shampoos, lubricating oil additives, fragrances, flavor materials or acrylates comprising one or more bioderived primary alcohol, fatty ester, wax, fatty acid, fatty aldehyde or fatty alcohol or bioderived primary alcohol, fatty ester, wax, fatty acid, fatty aldehyde or fatty alcohol intermediate produced by a non-naturally occurring microorganism of the invention or produced using a method disclosed herein.

As used herein, the term "bioderived" means derived from or synthesized by a biological organism and can be considered a renewable resource since it can be generated by a biological organism. Such a biological organism, in particular the microbial organisms of the invention disclosed herein, can utilize feedstock or biomass, such as, sugars or carbohydrates

obtained from an agricultural, plant, bacterial, or animal source. Alternatively, the biological organism can utilize atmospheric carbon. As used herein, the term "biobased" means a product as described above that is composed, in whole or in part, of a bioderived compound of the invention. A biobased or bioderived product is in contrast to a petroleum derived product, wherein such a product is derived from or synthesized from petroleum or a petrochemical feedstock.

In some embodiments, the invention provides a biofuel, chemical, polymer, surfactant, soap, detergent, shampoo, lubricating oil additive, fragrance, flavor material or acrylate comprising bioderived primary alcohol, fatty ester, wax, fatty acid, fatty aldehyde or fatty alcohol or bioderived primary alcohol, fatty ester, wax, fatty acid, fatty aldehyde or fatty alcohol intermediate, wherein the bioderived primary alcohol, fatty ester, wax, fatty acid, fatty aldehyde or fatty alcohol or bioderived primary alcohol, fatty ester, wax, fatty acid, fatty aldehyde or fatty alcohol intermediate includes all or part of the primary alcohol, fatty ester, wax, fatty acid, fatty aldehyde or fatty alcohol or primary alcohol, fatty ester, wax, fatty acid, fatty aldehyde or fatty alcohol intermediate used in the production of biofuel, chemical, polymer, surfactant, soap, detergent, shampoo, lubricating oil additive, fragrance, flavor material or acrylate. Thus, in some aspects, the invention provides a biobased biofuel, chemical, polymer, surfactant, soap, detergent, shampoo, lubricating oil additive, fragrance, flavor material or acrylate comprising at least 2%, at least 3%, at least 5%, at least 10%, at least 15%, at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 98% or 100% bioderived primary alcohol, fatty ester, wax, fatty acid, fatty aldehyde or fatty alcohol or bioderived primary alcohol, fatty ester, wax, fatty acid, fatty aldehyde or fatty alcohol intermediate as disclosed herein. Additionally, in some aspects, the invention provides a biobased biofuel, chemical, polymer, surfactant, soap, detergent, shampoo, lubricating oil additive, fragrance, flavor material or acrylate wherein the primary alcohol, fatty ester, wax, fatty acid, fatty aldehyde or fatty alcohol or primary alcohol, fatty ester, wax, fatty acid, fatty aldehyde or fatty alcohol intermediate used in its production is a combination of bioderived and petroleum derived primary alcohol, fatty ester, wax, fatty acid, fatty aldehyde or fatty alcohol or primary alcohol, fatty ester, wax, fatty acid, fatty aldehyde or fatty alcohol intermediate. For example, a biobased biofuel, chemical, polymer, surfactant, soap, detergent, shampoo, lubricating oil additive, fragrance, flavor material or acrylate can be produced using 50% bioderived primary alcohol, fatty ester, wax, fatty acid, fatty aldehyde or

fatty alcohol and 50% petroleum derived primary alcohol, fatty ester, wax, fatty acid, fatty aldehyde or fatty alcohol or other desired ratios such as 60%/40%, 70%/30%, 80%/20%, 90%/10%, 95%/5%, 100%/0%, 40%/60%, 30%/70%, 20%/80%, 10%/90% of

bioderived/petroleum derived precursors, so long as at least a portion of the product

comprises a bioderived product produced by the microbial organisms disclosed herein. It is understood that methods for producing biofuel, chemical, polymer, surfactant, soap, detergent, shampoo, lubricating oil additive, fragrance, flavor material or acrylate using the bioderived primary alcohol, fatty ester, wax, fatty acid, fatty aldehyde or fatty alcohol or bioderived primary alcohol, fatty ester, wax, fatty acid, fatty aldehyde or fatty alcohol

intermediate of the invention are well known in the art.

In an additional embodiment, the invention provides a non-naturally occurring microbial organism having a primary alcohol, a fatty acyl-CoA, a fatty ester, a wax, an acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol pathway, 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 β -

ketoacyl-CoA to β -hydroxyacyl-CoA, β -hydroxyacyl-CoA to trans-2-enoyl-CoA, trans-2-enoyl-CoA to acyl-CoA, acyl-CoA to β -ketoacyl-CoA, acyl-CoA to a fatty aldehyde, a fatty aldehyde to a fatty alcohol, acyl-CoA to a fatty acid, a fatty acid to a fatty aldehyde,

oxaloacetate to malate, malate to fumarate, fumarate to succinate, succinate to succinyl-CoA,

succinyl-CoA to α -ketoglutarate, α -ketoglutarate to D-isocitrate, D-isocitrate to citrate, citrate to acetate, citrate to oxaloacetate, acetate to acetyl-CoA, citrate to acetyl-CoA, acetyl-CoA to malonyl-CoA, acetyl-CoA and malonyl-CoA to acyl-ACP, acyl-ACP to a fatty acid, a fatty acid to acyl-CoA, acyl-CoA to a fatty aldehyde, a fatty aldehyde to a fatty alcohol, a fatty acid to a fatty aldehyde and an acyl-CoA to a fatty alcohol. 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 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 primary alcohol, a fatty acyl-CoA, a fatty ester, a wax, an acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol pathway, such as that shown in Figure 1, 9-22, 25 and 26.

While generally described herein as a microbial organism that contains a primary alcohol, a fatty acyl-CoA, a fatty ester, a wax, an acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol pathway, it is understood that the invention additionally provides a non-naturally occurring microbial organism comprising at least one exogenous nucleic acid

5 encoding a a primary alcohol, a fatty acyl-CoA, a fatty ester, a wax, an acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol pathway enzyme expressed in a sufficient amount to produce an intermediate of a primary alcohol, a fatty acyl-CoA, a fatty ester, a wax, an acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol pathway. For example, as disclosed herein, a primary alcohol, a fatty acyl-CoA, a fatty ester, a wax, an acyl-ACP, fatty
10 acid, acyl-CoA, fatty aldehyde or fatty alcohol pathway is exemplified in Figures 1, 9-22, 25 and 26. Therefore, in addition to a microbial organism containing a primary alcohol, a fatty acyl-CoA, a fatty ester, a wax, an acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol pathway that produces a primary alcohol, a fatty acyl-CoA, a fatty ester, a wax, an acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol, the invention additionally
15 provides a non-naturally occurring microbial organism comprising at least one exogenous nucleic acid encoding a primary alcohol, a fatty acyl-CoA, a fatty ester, a wax, an acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol pathway enzyme, where the microbial organism produces a primary alcohol, a fatty acyl-CoA, a fatty ester, a wax, an acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol pathway intermediate.

20 It is understood that any of the pathways disclosed herein, as described in the Examples and exemplified in the Figures, including the pathways of Figures 1, 9-22, 25 and 26, 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
25 expressing downstream pathway enzymes to produce a desired product. However, it is understood that a non-naturally occurring microbial organism that produces a primary alcohol, a fatty acyl-CoA, a fatty ester, a wax, an acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol pathway intermediate can be utilized to produce the intermediate as a desired product.

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.

As disclosed herein, the product fatty acid, as well as other intermediates, are carboxylic acids, which can occur in various ionized forms, including fully protonated, partially protonated, and fully deprotonated forms. Accordingly, the suffix “-ate,” or the acid form, can be used interchangeably to describe both the free acid form as well as any deprotonated form, in particular since the ionized form is known to depend on the pH in which the compound is found. It is understood that carboxylate products or intermediates includes ester forms of carboxylate products or pathway intermediates, such as O-carboxylate and S-carboxylate esters. O- and S-carboxylates can include lower alkyl, that is C1 to C6, branched or straight chain carboxylates. Some such O- or S-carboxylates include, without limitation, methyl, ethyl, n-propyl, n-butyl, i-propyl, sec-butyl, and tert-butyl, pentyl, hexyl O- or S-carboxylates, any of which can further possess an unsaturation, providing for example, propenyl, butenyl, pentyl, and hexenyl O- or S-carboxylates. O-carboxylates can be the product of a biosynthetic pathway. Exemplary O-carboxylates accessed via biosynthetic pathways can include, without limitation, methyl fatty acid, ethyl fatty acid, and n-propyl fatty acid. Other biosynthetically accessible O-carboxylates can include medium to long chain groups, that is C7-C22, O-carboxylate esters derived from fatty alcohols, such heptyl, octyl, nonyl, decyl, undecyl, lauryl, tridecyl, myristyl, pentadecyl, cetyl, palmitolyl, heptadecyl, stearyl, nonadecyl, arachidyl, heneicosyl, and behenyl alcohols, any one of which can be optionally branched and/or contain unsaturations. O-carboxylate esters can also be accessed via a biochemical or chemical process, such as esterification of a free carboxylic acid product or transesterification of an O- or S-carboxylate. S-carboxylates are exemplified by CoA S-esters, cysteinyl S-esters, alkylthioesters, and various aryl and heteroaryl thioesters.

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 a primary alcohol, a fatty acyl-CoA, a fatty ester, a wax, an acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol biosynthetic pathways.

5 Depending on the host microbial organism chosen for biosynthesis, nucleic acids for some or all of a particular a primary alcohol, a fatty acyl-CoA, a fatty ester, a wax, an acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol 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)
10 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 a primary alcohol, a fatty acyl-CoA, a fatty ester, a wax, an acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol biosynthesis. Thus, a non-naturally occurring
15 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 a primary alcohol, a fatty acyl-CoA, a fatty ester, a wax, an acyl-ACP, fatty acid, acyl-
20 CoA, fatty aldehyde or fatty alcohol .

Depending on the a primary alcohol, a fatty acyl-CoA, a fatty ester, a wax, an acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol 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 a primary alcohol, a fatty acyl-
25 CoA, a fatty ester, a wax, an acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol pathway-encoding nucleic acid and up to all encoding nucleic acids for one or more [a primary alcohol, a fatty acyl-CoA, a fatty ester, a wax, an acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol biosynthetic pathways. For example, a primary alcohol, a fatty acyl-CoA, a fatty ester, a wax, an acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or
30 fatty alcohol biosynthesis can be established in a host deficient in a pathway enzyme or protein through exogenous expression of the corresponding encoding nucleic acid. In a host deficient in all enzymes or proteins of a a primary alcohol, a fatty acyl-CoA, a fatty ester, a wax, an acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol 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 a primary alcohol, a fatty acyl-CoA, a fatty ester, a wax, an acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol can be included, such as a primary alcohol, a fatty acyl-CoA, a fatty ester, a wax, an acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol .

Generally, a host microbial organism is selected such that it produces the precursor of a primary alcohol, a fatty acyl-CoA, a fatty ester, a wax, an acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol 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, a primary alcohol, a fatty acyl-CoA, a fatty ester, a wax, an acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol 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 primary alcohol, a fatty acyl-CoA, a fatty ester, a wax, an acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol pathway.

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 primary alcohol, a fatty acyl-CoA, a fatty ester, a wax, an acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol biosynthetic pathway onto the microbial organism. Alternatively, encoding nucleic acids can be introduced to produce an intermediate microbial organism having the biosynthetic capability to catalyze some of the required reactions to confer primary alcohol, a fatty acyl-CoA, a fatty ester, a wax, an acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol biosynthetic capability. For example, a non-naturally occurring microbial organism having a primary alcohol, a fatty acyl-CoA, a fatty ester, a wax, an acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol biosynthetic pathway can comprise at least two exogenous nucleic acids encoding desired enzymes or proteins, such as the combination of fatty acid synthase and a thioesterase, or alternatively a aconitase and a

aldehyde reductase, or alternatively an alpha-detoglutarate ferredoxin oxidoreductase and a fatty alcohol forming acyl-CoA reductase, and the like. 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 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, for example, a CO dehydrogenase, a fumarase and a enoyl-CoA reductase, or alternatively an acyl-CoA reductase, and alcohol dehydrogenase and a hydrogenase, and so forth, 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. Similarly, any combination of four, a CO dehydrogenase, a ATP citrate lyase, a carboxyl acid reductase and a alcohol dehydrogenase, or alternatively an acetyl-CoA carboxylase, an fatty acid synthase, a thioesterase and a isocitrate dehydrogenase, 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 combination of enzymes and/or proteins of the desired biosynthetic pathway results in production of the corresponding desired product.

In addition to the biosynthesis of primary alcohol, a fatty acyl-CoA, a fatty ester, a wax, an acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol 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 primary alcohol, a fatty acyl-CoA, a fatty ester, a wax, an acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol other than use of the primary alcohol, a fatty acyl-CoA, a fatty ester, a wax, an acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol producers is through addition of another microbial organism capable of converting a primary alcohol, a fatty acyl-CoA, a fatty ester, a wax, an acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol pathway intermediate to primary alcohol, a fatty acyl-CoA, a fatty ester, a wax, an acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol. One such procedure includes, for example, the fermentation of a microbial organism that produces a primary alcohol, a fatty acyl-CoA, a fatty ester, a wax, an acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol pathway intermediate. The primary alcohol, a fatty acyl-CoA, a fatty ester, a wax, an acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol pathway intermediate can then be used as a substrate for a second microbial

organism that converts the primary alcohol, a fatty acyl-CoA, a fatty ester, a wax, an acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol pathway intermediate to primary alcohol, a fatty acyl-CoA, a fatty ester, a wax, an acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol. The primary alcohol, a fatty acyl-CoA, a fatty ester, a wax, an acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol pathway intermediate can be added directly to another culture of the second organism or the original culture of the primary alcohol, a fatty acyl-CoA, a fatty ester, a wax, an acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol 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, primary alcohol, a fatty acyl-CoA, a fatty ester, a wax, an acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol. 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 primary alcohol, a fatty acyl-CoA, a fatty ester, a wax, an acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol 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, primary alcohol, a fatty acyl-CoA, a fatty ester, a wax, an acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol 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 primary alcohol, a fatty acyl-CoA, a fatty ester, a wax, an acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol intermediate and the second microbial organism converts the intermediate to primary alcohol, a fatty acyl-CoA, a fatty ester, a wax, an acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol.

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 primary alcohol, a fatty acyl-CoA, a fatty ester, a wax, an acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol.

Microbial organisms other than *Euglena gracilis* generally lack the capacity to synthesize acyl-CoA through a malonyl-CoA independent FAS pathway. Moreover, organisms having all of the requisite metabolic enzymatic capabilities are not known to produce acyl-CoA from the enzymes described and biochemical pathways exemplified herein. Rather,

microorganisms having the enzymatic constituents of malonyl-CoA independent FAS pathway operate to degrade short, medium, and long chain fatty-acyl-CoA compounds to acetyl-CoA. *E. gracilis*, having a malonyl-CoA independent FAS pathway, utilizes this pathway to produce acylglycerols, trihydric sugar alcohols, phospholipids, wax esters and/or fatty acids. In contrast, the non-naturally occurring microbial organisms of the invention generate acyl-CoA as a product of the malonyl-CoA independent FAS pathway and funnel this product into an acyl-reduction pathway via favorable thermodynamic characteristics. Product biosynthesis of using the non-naturally occurring organisms of the invention is not only particularly useful for the production of primary alcohols, it also allows for the further biosynthesis of compounds using acyl-CoA and/or primary alcohols as an intermediate reactant.

The non-naturally occurring primary alcohol-producing microbial organisms of the invention are generated by ensuring that a host microbial organism includes functional capabilities for the complete biochemical synthesis of a malonyl-CoA independent fatty acid biosynthetic pathway and for an acyl-reduction pathway of the invention. Ensuring complete functional capabilities for both pathways will confer primary alcohol biosynthesis capability onto the host microbial organism. The enzymes participating in a malonyl-CoA independent FAS pathway include ketoacyl-CoA acyltransferase or ketoacyl-CoA thiolase, 3-hydroxyacyl-CoA dehydrogenase, enoyl-CoA hydratase and enoyl-CoA reductase. The enzymes participating in an acyl-reduction pathway include an acyl-CoA reductase and an alcohol dehydrogenase or an enzyme having dual reductase and dehydrogenase activity.

The non-naturally occurring microbial organisms of the invention can be produced by introducing expressible nucleic acids encoding one or more of the enzymes participating in

the malonyl-CoA independent FAS and/or acyl-reduction pathways. Depending on the host microbial organism chosen for biosynthesis, nucleic acids for some or all of these biosynthetic pathways can be expressed. For example, if a chosen host is deficient in all of the enzymes in the malonyl-CoA independent FAS pathway, then expressible nucleic acids for each of the four enzymes ketoacyl-CoA acyltransferase or ketoacyl-CoA thiolase, 3-hydroxyacyl-CoA dehydrogenase, enoyl-CoA hydratase and enoyl-CoA reductase are introduced into the host for subsequent exogenous expression. Alternatively, for example, if the chosen host is deficient less than all four of the above enzymes, then all that is needed is to express nucleic acids encoding the deficient enzymes. For example, if a host is deficient in 3-hydroxyacyl-CoA dehydrogenase and enoyl-CoA hydratase a functionally complete malonyl-CoA independent FAS pathway can be engineered by introduction of nucleic acids encoding these two enzymes.

In like fashion, where endogenous host biosynthetic machinery is complete for an acyl-reduction pathway, then genetic modification is unnecessary. However, if host capabilities are deficient in either or both of the acyl-CoA reductase and/or alcohol dehydrogenase activities, then introduction of the deficient activity by expression of an exogenous encoding nucleic acid is needed. Accordingly, depending on the malonyl-CoA independent FAS and acyl-reduction pathway constituents of a selected host microbial organism, the non-naturally occurring microbial organisms of the invention will include at least one exogenously expressed malonyl-CoA independent FAS pathway-encoding nucleic acid and up to all six malonyl-CoA independent FAS and acyl-reduction pathway encoding nucleic acids.

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 parallel the malonyl-CoA independent FAS and acyl-reduction 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 or six encoding nucleic acids encoding the above enzymes constituting the malonyl-CoA independent FAS pathway, an acyl-reduction pathway or both the malonyl-CoA independent FAS and acyl-reduction biosynthetic pathways. In some embodiments, the non-naturally occurring microbial organisms also can include other genetic modifications that facilitate or optimize acyl-CoA and/or primary alcohol 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 malonyl-CoA independent FAS pathway precursors such as acetyl-CoA, β -ketoacyl-CoA, β -hydroxyacyl-CoA, *trans*-2-enoyl-CoA and/or fatty aldehyde.

In some embodiments, a non-naturally occurring microbial organism of the invention is generated from a host that contains the enzymatic capability to synthesize acyl-CoA through a malonyl-CoA independent FAS pathway, or having the capability to catalyze one or more of the enzymatic steps within the malonyl-CoA independent FAS and/or acyl-reduction pathways. In these specific embodiments it can be useful to increase the synthesis or accumulation of a malonyl-CoA independent FAS pathway product or an acyl-reduction pathway product to, for example, efficiently drive malonyl-CoA independent FAS and/or acyl-reduction pathway reactions toward primary alcohol production. Increased synthesis or accumulation can be accomplished by, for example, overexpression of nucleic acids encoding one or more of the above-described malonyl-CoA independent FAS and/or acyl-reduction pathway enzymes. Overexpression of the desired pathway enzyme or enzymes can occur, for example, through exogenous expression of the endogenous gene or genes, or through exogenous expression of a heterologous gene or genes. Therefore, naturally occurring organisms can readily be generated to be non-naturally primary alcohol producing microbial organisms of the invention through overexpression of one, two, three, four, five or all six nucleic acids encoding a malonyl-CoA independent FAS and/or a acyl-reduction pathway enzymes. 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 malonyl-CoA independent FAS and/or acyl-reduction biosynthetic pathways.

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 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. For example, activation of *fadB*, an *E. coli* gene having malonyl-CoA independent FAS activity corresponding to 3-hydroxyacyl-CoA dehydrogenase and enoyl-CoA hydratase activities can be accomplished by genetically knocking out a negative regulator, *fadR*, and co-expressing a heterologous ketothiolase (*phaA* from *Ralstonia eutropha*; Sato et al., *Journal of Bioscience and Bioengineering* 103:38-44 (2007)). Thus, an endogenous gene having a naturally

occurring inducible promoter can be up-regulated by 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.

Additionally, for example, if an endogenous enzyme or enzymes operate in a reverse direction to the desired malonyl-CoA independent FAS pathway, genetic modifications can be made to attenuate or eliminate such activities. For example, within the malonyl-CoA independent FAS pathway, the ketothiolase, dehydrogenase, and enoyl-CoA hydratase steps are reversible whereas the enoyl-CoA reductase step is primarily oxidative under physiological conditions (Hoffmeister et al., *Journal of Biological Chemistry* 280:4329-4338 (2005); Campbell, J. W. and J. E. Cronan, Jr., *J Bacteriol.* 184:3759-3764 (2002)). To accomplish reduction of a 2-enoyl-CoA intermediate a genetic modification can be introduced to attenuate or eliminate the reverse oxidative reaction.

Sources of encoding nucleic acids for a malonyl-CoA independent FAS and/or acyl-reduction pathway enzyme 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, insect, animal, and mammal, including human. Exemplary species for such sources include, for example, *Escherichia coli*, *Acetobacter aceti*, *Acinetobacter baylyi*, *Acinetobacter calcoaceticus*, *Acinetobacter sp.*, *Acinetobacter sp. Strain M-1*, *Actinobacillus succinogenes*, *Aedes aegypti*, *Aeropyrum pernix*, *Alcanivorax borkumensis SK2*, *Alkaliphilus metalliredigens QYMF*, *Alkaliphilus oremlandii OhILAs*, *Allochromatium vinosum DSM 180*, *Anaerobiospirillum succiniciproducens*, *Anopheles gambiae str. PEST*, *Apis mellifera*, *Aquifex aeolicus*, *Arabidopsis thaliana*, *Archaeoglobus fulgidus*, *Archaeoglobus fulgidus DSM 4304*, *Aromatoleum aromaticum EbN1*, *Aspergillus nidulans*, *Aspergillus niger*, *Azoarcus sp. T*, *Azotobacter vinelandii DJ*, *Bacillus megaterium*, *Bacillus sp. SG-1*, *Bacillus sphaericus*, *Bacillus subtilis*, *Bacillus weihenstephanensis KBAB4*, *Balnearium lithotrophicum*, *Bifidobacterium bifidum*, *Bifidobacterium longum NCC2705*, *Brassica juncea*, *Brevibacterium sp.*, *Burkholderia multivorans ATCC 17616*, *butyrate-producing bacterium L2-50*, *Caenorhabditis briggsae AF16*, *Caenorhabditis*

- elegans*, *Campylobacter curvus* 525.92, *Campylobacter jejuni*, *Candida albicans*, *Candida Antarctica*, *Canis lupus familiaris* (dog), *Carboxydotherrnus hydrogenoformans*, *Chlorobium phaeobacteroides* DSM 266, *Chlorobium limicola*, *Citrobacter youngae* ATCC 29220, *Clostridium acetobutylicum*, *Clostridium carboxidivorans* P7, *Clostridium cellulolyticum*
- 5 *H10*, *Clostridium kluyveri*, *Clostridium kluyveri* DSM 555, *Clostridium pasteurianum*, *Clostridium saccharoperbutylacetonicum*, *Clostridium tetani* E88, *Colwellia psychrerythraea* 34H, *Corynebacterium glutamicum*, *Corynebacterium ulcerans*, *Cryptococcus neoformans* var., *Cyanobium* PCC7001, *Danio rerio*, *Desulfatibacillum alkenivorans* AK-01, *Desulfococcus oleovorans* Hxd3, *Desulfovibrio africanus*, *DesulfoVibrio desulfuricans* G20,
- 10 *Desulfovibrio desulfuricans* subsp. *desulfuricans* str. ATCC 27774, *Desulfovibrio fructosovorans* JJ, *Desulfovibrio vulgaris* str. Hildenborough, *Dictyostelium discoideum*, *Dictyostelium discoideum* AX4, *Escherichia coli* K12, *Escherichia coli* K12 MG1655, *Escherichia coli* str. K-12 substr. MG1655, *Euglena gracilis*, *Flavobacteria bacterium* BAL38, *Fragaria x ananassa*, *Fusarium oxysporum*, *Fusarium proliferatum*, *Geobacter*
- 15 *metallireducens* GS-15, *Geobacter sulfurreducens*, *Haemophilus influenza*, *Haloarcula marismortui*, *Helianthus annuus*, *Helicobacter pylori*, *Helicobacter pylori* 26695, *Homo sapiens*, *Hydrogenobacter thermophilus*, *Klebsiella pneumonia*, *Lactococcus lactis*, *Leuconostoc mesenteroides*, *Macaca mulatta*, *Magnetospirillum magneticum* AMB-1, *Mannheimia succiniciproducens*, *Marinobacter aquaeolei* VT8, *Marinobacter*
- 20 *hydrocarbonoclasticus*, *Methanosarcina thermophila*, *Methanothermobacter thermautotrophicus*, *Methanothermobacter thermautotrophicus*, *Methylobacterium extorquens*, *Monosiga brevicollis* MX1, *Moorella thermoacetica*, *Mus musculus*, *Mycobacterium avium* subsp. *paratuberculosis* K-10, *Mycobacterium bovis* BCG, *Mycobacterium kansasii* ATCC 12478, *Mycobacterium marinum* M, *Mycobacterium*
- 25 *marinum* M, *Mycobacterium smegmatis*, *Mycobacterium smegmatis* MC2 155, *Mycobacterium* sp. MCS, *Mycobacterium tuberculosis* H37Rv, *Myxococcus xanthus* DK 1622, *Nematostella vectensis*, *Nocardia farcinica* IFM 10152, *Nocardia iowensis* (sp. NRRL 5646), *Nostoc* sp. PCC 7120, *Opitutaceae bacterium* TAV2, *Pelobacter carbinolicus* DSM 2380, *Pelotomaculum thermopropionicum*, *Penicillium chrysogenum*, *Photobacterium* sp.
- 30 *SKA34*, *Picea sitchensis*, *Pseudomonas aeruginosa* PA01, *Pseudomonas fluorescens*, *Pseudomonas mendocina*, *Pseudomonas putida*, *Pseudomonas putida* GB-1, *Pseudomonas* sp., *Pseudomonas* sp. CF600, *Pseudomonas stutzeri* A1501, *Pyrobaculum aerophilum* str. IM2, *Ralstonia eutropha*, *Ralstonia eutropha* H16, *Rattus norvegicus*, *Reinekea* sp. MED297, *Rhizobium etli* CFN 42, *Rhizobium leguminosarum*, *Rhodobacter capsulatus*, *Rhodococcus*

erythropolis SK121, Rhodopseudomonas palustris, Rhodopseudomonas palustris CGA009, Rhodospirillum rubrum, Rosa hybrid cultivar, Roseovarius sp. HTCC2601, Saccharomyces cerevisiae, Salmonella enteric, Salmonella enterica subsp. arizonae serva, Salmonella typhimurium, Schizosaccharomyces pombe, Simmondsia chinensis, Simmondsia chinensis,
 5 *Sordaria macrospora, Stenotrophomonas maltophilia, Streptomyces griseus subsp. griseus NBRC 13350, Strongylocentrotus purpuratus, Sulfolobus acidocalarius, Sulfolobus sp. strain 7, Sulfurihydrogenibium subterraneum, Sulfurimonas denitrificans, Synechocystis str. PCC 6803, Synthetic variant, Tetraodon nigroviridis, Thauera aromatic, Thermotoga maritime, Thermus thermophilus, Thiobacillus denitrificans, Thiocapsa roseopersicina, Treponema*
 10 *denticola, Tribolium castaneum, Trichomonas vaginalis G3, Trypanosoma brucei, Trypanosoma brucei, Trypanosoma cruzi strain CL Brener, Tsukamurella paurometabola DSM 20162, Vibrio cholerae V51, Xanthomonas campestris, Xenopus tropicalis, Yarrowia lipolytica, Yersinia intermedia ATCC 29909, and Zea mays,* as well as other exemplary species disclosed herein or available as source organisms for corresponding genes. However,
 15 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 malonyl-CoA independent FAS and/or acyl-reduction biosynthetic activity for one or more genes in related or distant species, including for
 20 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 enabling biosynthesis of primary alcohols of the invention described herein with reference to a particular organism such as *E. coli* can be readily applied to other microorganisms, including prokaryotic and eukaryotic
 25 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 malonyl-CoA independent FAS constituent enzyme or pathway exists in an unrelated species, primary alcohol biosynthesis can be
 30 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. Because certain differences among metabolic networks exist between different organisms, those skilled in the art will understand that the

actual genes 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 microbial organism in a species of interest that will synthesize the primary alcohol products of the invention.

Encoding nucleic acids and species that can be used as sources for conferring malonyl-CoA independent FAS and/or acyl-reduction pathway capability onto a host microbial organism are exemplified further below. In one exemplary embodiment, the genes *fadA* and *fadB* encode a multienzyme complex that exhibits three constituent activities of the malonyl-CoA independent FAS pathway, namely, ketoacyl-CoA thiolase, 3-hydroxyacyl-CoA dehydrogenase, and enoyl-CoA hydratase activities (Nakahigashi, K. and H. Inokuchi, *Nucleic Acids Research* 18:4937 (1990); Yang et al., *Journal of Bacteriology* 173:7405-7406 (1991); Yang et al., *Journal of Biological Chemistry* 265:10424-10429 (1990); Yang et al., *Biochemistry* 30:6788-6795 (1990)). The *fadI* and *fadJ* genes encode similar activities which can substitute for the above malonyl-CoA independent FAS conferring genes *fadA* and *fadB*. These genes are naturally expressed under anaerobic conditions (Campbell and Cronan, *supra*, (2002)). The nucleic acid sequences for each of the above *fad* genes are well known in the art and can be accessed in the public databases such as Genbank using the following accession numbers.

<i>fadA</i>	YP_026272.1	<i>Escherichia coli</i>
<i>fadB</i>	NP_418288.1	<i>Escherichia coli</i>
<i>fadI</i>	NP_416844.1	<i>Escherichia coli</i>
<i>fadJ</i>	NP_416843.1	<i>Escherichia coli</i>
<i>fadR</i>	NP_415705.1	<i>Escherichia coli</i>

Other exemplary genes for the ketothiolase step include *atoB* which can catalyze the reversible condensation of 2 acetyl-CoA molecules (Sato et al., *supra*, 2007), and its homolog *yqeF*. Non-*E. coli* genes that can be used include *phaA* from *R. eutropha* (Jenkins, L. S. and W. D. Nunn. *Journal of Bacteriology* 169:42-52 (1987)), and the two ketothiolases, *thiA* and *thiB*, from *Clostridium acetobutylicum* (Winzer et al., *Journal of Molecular Microbiology and Biotechnology* 2:531-541 (2000)). The sequences for these genes can be found at the following Genbank accession numbers:

<i>atoB</i>	NP_416728.1	<i>Escherichia coli</i>
<i>yqeF</i>	NP_417321.2	<i>Escherichia coli</i>

<i>phaA</i>	YP_725941	<i>Ralstonia eutropha</i>
<i>thiA</i>	NP_349476.1	<i>Clostridium acetobutylicum</i>
<i>thiB</i>	NP_149242.1	<i>Clostridium acetobutylicum</i>

5 An exemplary gene from *E. coli* which can be used for conferring 3-hydroxyacyl-CoA dehydrogenase transformation activity is *paaH* (Ismail et al., *European Journal of Biochemistry* 270:3047-3054 (2003)). Non-*E. coli* genes applicable for conferring this activity include AAO72312.1 from *E. gracilis* (Winkler et al., *Plant Physiology* 131:753-762 (2003)), *paaC* from *Pseudomonas putida* (Olivera et al., *PNAS USA* 95:6419-6424 (1998)),
 10 *paaC* from *Pseudomonas fluorescens* (Di Gennaro et al., *Archives of Microbiology* 188:117-125 (2007)), and *hbd* from *C. acetobutylicum* (Atsumi et al., *Metabolic Engineering* (2007) and Boynton et al., *Journal of Bacteriology* 178:3015-3024 (1996)). The sequences for each of these exemplary genes can be found at the following Genbank accession numbers:

<i>paaH</i>	NP_415913.1	<i>Escherichia coli</i>
	AA072312.1	<i>Euglena gracilis</i>
<i>paaC</i>	NP_745425.1	<i>Pseudomonas putida</i>
<i>paaC</i>	ABF82235.1	<i>Pseudomonas fluorescens</i>
<i>hbd</i>	NP_349314.1	<i>Clostridium acetobutylicum</i>

20 Exemplary genes encoding the enoyl-CoA hydratase step include, for example, *maoC* (Park and Lee, *Journal Bacteriology* 185:5391-5397 (2003)), *paaF* (Ismail et al., *European Journal of Biochemistry* 270:3047-3054 (2003); Park and Lee, *Appl. Biochem. Biotechnol.* 113-116:335-346 (2004) and Park and Yup, *Biotechnol. Bioeng.* 86:681-686 (2004)), and *paaG* (Ismail et al., *European Journal of Biochemistry* 270:3047-3054 (2003); Park and Lee,
 25 *Appl. Biochem. Biotechnol.* 113-116:335-346 (2004) and Park and Yup, *Biotechnol. Bioeng.* 86:681-686 (2004)). Other genes which can be used to produce the gene product catalyzing this step, for example, *paaA*, *paaB*, and *paaN* from *P. putida* (Olivera et al., *PNAS USA* 95:6419-6424 (1998)) and *P. fluorescens* (Di Gennaro et al., *Archives of Microbiology* 188:117-125 (2007)). The gene product of *crt* from *C. acetobutylicum* also can be used
 30 (Atsumi et al., *Metabolic Engineering* (2007) and Boynton et al., *Journal of Bacteriology* 178: 3015-3024 (1996). The sequences for each of these exemplary genes can be found at the following Genbank accession numbers:

<i>maoC</i>	NP_415905.1	<i>Escherichia coli</i>
<i>paaF</i>	NP_415911.1	<i>Escherichia coli</i>
<i>paaG</i>	NP_415912.1	<i>Escherichia coli</i>
<i>paaA</i>	NP_745427.1	<i>Pseudomonas putida</i>
<i>paaA</i>	ABF82233.1	<i>Pseudomonas fluorescens</i>

<i>paaB</i>	NP_745426.1	<i>Pseudomonas putida</i>
<i>paaB</i>	ABF82234.1	<i>Pseudomonas fluorescens</i>
<i>paaN</i>	NP_745413.1	<i>Pseudomonas putida</i>
<i>paaN</i>	ABF82246.1	<i>Pseudomonas fluorescens</i>
<i>crt</i>	NP_349318.1	<i>Clostridium acetobutylicum</i>

An exemplary gene which can be introduced into a non-naturally occurring microbial organism of the invention to confer enoyl-CoA reductase activity is the mitochondrial enoyl-CoA reductase from *E. gracilis* Hoffmeister et al., *supra* (2005)). A construct derived from this sequence following the removal of its mitochondrial targeting leader sequence has been cloned and expressed in *E. coli*. This approach for heterologous expression of membrane targeted polypeptides in a soluble form is well known to those skilled in the art of expressing eukaryotic genes, particularly those with leader sequences that may target the gene product to a specific intracellular compartment, in prokaryotic organisms. A close homolog of this gene, TDE0597, from the prokaryote *Treponema denticola* represents also can be employed to confer enoyl-CoA reductase activity (Tucci and Martin, *FEBS Letters* 581:1561-1566 (2007)). Butyryl-CoA dehydrogenase, encoded by *bcd* from *C. acetobutylicum*, is a further exemplary enzyme that can be used to confer enoyl-CoA reductase activity onto a host microbial organism of the invention (Atsumi et al., *Metabolic Engineering* (2007) and Boynton et al., *Journal of Bacteriology* 178: 3015-3024 (1996)). Alternatively, *E. coli* genes exhibiting this activity can be obtained using methods well known in the art (see, for example, Mizugaki et al., *Chemical & Pharmaceutical Bulletin* 30:206-213 (1982) and Nishimaki et al., *Journal of Biochemistry* 95:1315-1321 (1984)). The sequences for each of the above exemplary genes can be found at the following Genbank accession numbers:

TER	Q5EU90.1	<i>Euglena gracilis</i>
TDE0597	NP_971211.1	<i>Treponema denticola</i>
<i>bcd</i>	NP_349317.1	<i>Clostridium acetobutylicum</i>

At least three mitochondrial enoyl-CoA reductase enzymes exist in *E. gracilis* that similarly are applicable for use in the invention. Each enoyl-CoA reductase enzyme exhibits a unique chain length preference (Inui et al., *European Journal of Biochemistry* 142:121-126 (1984)), which is particularly useful for dictating the chain length of the desired primary alcohol products of the invention. EST's ELL00002199, ELL00002335, and ELL00002648, which are all annotated as mitochondrial trans-2-enoyl-CoA reductases, can be used to isolate these additional enoyl-CoA reductase genes as described further below.

Those skilled in the art also can obtain nucleic acids encoding any or all of the malonyl-CoA independent FAS pathway or acyl-reduction pathway enzymes by cloning using known sequences from available sources. For example, any or all of the encoding nucleic acids for the malonyl-CoA independent FAS pathway can be readily obtained using methods well

known in the art from *E. gracilis* as this pathway has been well characterized in this organism. *E. gracilis* encoding nucleic acids can be isolated from, for example, an *E. gracilis* cDNA library using probes of known sequence. The probes can be designed with whole or partial DNA sequences from the following EST sequences from the publically available sequence database TBestDB (<http://tbestdb.bcm.umontreal.ca>). The nucleic acids generated from this process can be inserted into an appropriate expression vector and transformed into *E. coli* or other microorganisms to generate primary alcohol production organisms of the invention.

ketoacyl-CoA acyltransferase (or ketoacyl-CoA thiolase)

ELL00002550

ELL00002493

ELL00000789

3-hydroxyacyl-CoA dehydrogenase

ELL00000206

ELL00002419

ELL00006286

ELL00006656

enoyl-CoA hydratase

ELL00005926

ELL00001952

ELL00002235

ELL00006206

enoyl-CoA reductase

ELL00002199

ELL00002335

ELL00002648

Alternatively, the above EST sequences can be used to identify homologue polypeptides in GenBank through BLAST search. The resulting homologue polypeptides and their corresponding gene sequences provide additional encoding nucleic acids for transformation into *E. coli* or other microorganisms to generate the primary alcohol producing organisms of the invention. Listed below are exemplary homologue polypeptide and their gene accession numbers in GenBank which are applicable for use in the non-naturally occurring organisms of the invention.

	ketoacyl-CoA acyltransferase (or ketoacyl-CoA thiolase)	
	YP_001530041	<i>Desulfococcus oleovorans</i> Hxd3
	ZP_02133627	<i>Desulfatibacillum alkenivorans</i> AK-01
	ZP_01860900	<i>Bacillus</i> sp. SG-1
5	YP_001511817	<i>Alkaliphilus oremlandii</i> OhILAs
	NP_781017	<i>Clostridium tetani</i> E88
	YP_001646648	<i>Bacillus weihenstephanensis</i> KBAB4
	YP_001322360	<i>Alkaliphilus metalliredigens</i> QYMF
	YP_001397054	<i>Clostridium kluyveri</i> DSM 555
10	NP_070026	<i>Archaeoglobus fulgidus</i> DSM 4304
	YP_001585327	<i>Burkholderia multivorans</i> ATCC 17616
	3-hydroxyacyl-CoA dehydrogenase	
	AA072312	<i>Euglena gracilis</i>
	XP_001655993	<i>Aedes aegypti</i>
15	NP_001011073	<i>Xenopus tropicalis</i>
	NP_001003515	<i>Danio rerio</i>
	XP_973042	<i>Tribolium castaneum</i>
	XP_001638329	<i>Nematostella vectensis</i>
	CAG11476	<i>Tetraodon nigroviridis</i>
20	XP_787188	<i>Strongylocentrotus purpuratus</i>
	XP_001749481	<i>Monosiga brevicollis</i> MX1
	NP_509584	<i>Caenorhabditis elegans</i>
	XP_572875	<i>Cryptococcus neoformans</i> var
	enoyl-CoA hydratase	
25	XP_844077	<i>Trypanosoma brucei</i>
	XP_802711	<i>Trypanosoma cruzi</i> strain CL Brener
	XP_806421	<i>Trypanosoma cruzi</i> strain CL Brener.
	YP_001669856	<i>Pseudomonas putida</i> GB-1
	YP_641317	<i>Mycobacterium</i> sp. MCS
30	YP_959434	<i>Marinobacter aquaeolei</i> VT8
	ABK24445	<i>Picea sitchensis</i>
	XP_640315	<i>Dictyostelium discoideum</i>
	YP_633978	<i>Myxococcus xanthus</i> DK 1622
	YP_467905	<i>Rhizobium etli</i> CFN 42
35	YP_419997	<i>Magnetospirillum magneticum</i> AMB-1
	YP_001172441	<i>Pseudomonas stutzeri</i> A1501
	enoyl-CoA reductase.	
	XP_642118	<i>Dictyostelium discoideum</i> AX4
	XP_001639469	<i>Nematostella vectensis</i>
40	XP_001648220	<i>Aedes aegypti</i>
	XP_974428	<i>Tribolium castaneum</i>
	XP_535334	<i>Canis lupus familiaris</i> (dog)
	NP_001016371	<i>Xenopus tropicalis</i>
	XP_320682	<i>Anopheles gambiae</i> str. PEST
45	ZP_01645699	<i>Stenotrophomonas maltophilia</i>
	XP_001679449	<i>Caenorhabditis briggsae</i> AF16
	ZP_01443601	<i>Roseovarius</i> sp. HTCC2601
	XP_395130	<i>Apis mellifera</i>
	XP_001113746	<i>Macaca mulatta</i>
50	ZP_01485509	<i>Vibrio cholerae</i> V51

	ZP_02012479	<i>Opitutaceae bacterium</i> TAV2
	ZP_01163033	<i>Photobacterium</i> sp. SKA34
	YP_267463	<i>Colwellia psychrerythraea</i> 34H
	ZP_01114282	<i>Reinekea</i> sp. MED297
5	ZP_01732824	<i>Flavobacteria bacterium</i> BAL38

As described previously, after the malonyl-CoA independent elongation cycle, the resulting acyl-CoA can be reduced to produce a primary alcohol by either a single enzyme or pair of enzymes that exhibit acyl-CoA reductase and alcohol dehydrogenase activities. Exemplary genes that encode enzymes for catalyzing the reduction of an acyl-CoA to its corresponding aldehyde include the *Acinetobacter calcoaceticus* *acrI* encoding a fatty acyl-CoA reductase (Reiser and Somerville, *Journal of Bacteriology* 179:2969-2975 (1997)), the *Acinetobacter* sp. M-1 fatty acyl-CoA reductase (Ishige et al., *Appl. Environ. Microbiol.* 68:1192-1195 (2002)), and the *sucD* gene from *Clostridium kluyveri* (Sohling and Gottschalk, *Journal Bacteriology* 178:871-880 (1996)).

	<i>acrI</i>	YP_047869.1	<i>Acinetobacter calcoaceticus</i>
		AAC45217	<i>Acinetobacter baylyi</i>
		BAB85476.1	<i>Acinetobacter</i> sp. Strain M-1
20	<i>sucD</i>	P38947.1	<i>Clostridium kluyveri</i>

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., *Appl. Environ. Microbiol.* 66:5231-5235 (2000)), ADH2 from *Saccharomyces cerevisiae* (Atsumi et al., *Nature* 451:86-89 (2008)), and *yqhD* from *E. coli* which has preference for molecules longer than C3 (Sulzenbacher et al., *Journal of Molecular Biology* 342:489-502 (2004)).

	<i>alrA</i>	BAB12273.1	<i>Acinetobacter</i> sp. Strain M-1
	<i>ADH2</i>	NP_014032.1	<i>Saccharomyces cerevisiae</i>
30	<i>yqhD</i>	NP_417484.1	<i>Escherichia coli</i>

Alternatively, the fatty acyl-CoA can be reduced in one step by a fatty alcohol forming acyl-CoA reductase or any other enzyme with dual acyl-CoA reductase and alcohol dehydrogenase activity. For example, the jojoba (*Simmondsia chinensis*) *FAR* encodes an alcohol-forming fatty acyl-CoA reductase and its overexpression in *E. coli* resulted in *FAR* activity and the accumulation of fatty alcohol (Metz et al., *Plant Physiology* 122:635-644 (2000)). The reductase with narrow substrate chain-length specificities will also function as additional control for product chain-length. Additional gene candidates include the *E. coli*

adhE (Kessler et al., *FEBS Letters* 281:59-63 (2000)) and *C. acetobutylicum bdh I* and *bdh II* (Walter et al., *Journal of Bacteriology* 174:7149-7158 (1992)) which can reduce acetyl-CoA and butyryl-CoA to ethanol and butanol, respectively.

5 FAR AAD38039.1 *Simmondsia chinensis*
 adhE NP_415757.1 *Escherichia coli*
 bdh I NP_349892.1 *Clostridium acetobutylicum*
 bdh II NP_349891.1 *Clostridium acetobutylicum*

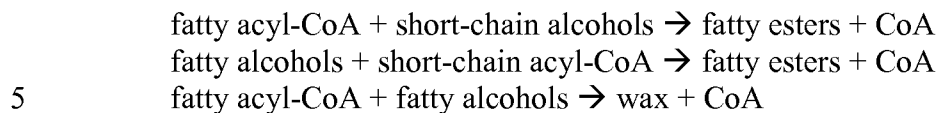
In addition, the *E. gracilis* nucleic acid sequences encoding enzymes for the reduction step can be obtained and transformed into a host as described previously for the malonyl-CoA independent FAS pathway encoding nucleic acids. Isolated from an *E. gracilis* cDNA library using probes, designed with whole or partial DNA sequences from the following EST sequences from TBestDB (<http://tbestdb.bcm.umontreal.ca>) can be performed as described previously.

15 aldehyde dehydrogenase
 ELL00002572
 ELL00002581
 ELL00000108

In addition to the above exemplary encoding nucleic acids, nucleic acids other than those within the malonyl-CoA independent FAS and/or acyl-reduction pathways of the invention also can be introduced into a host organism for further production of primary alcohols. For example, the *Ralstonia eutropha BktB* and *PhbB* genes catalyze the condensation of butyryl-CoA and acetyl-CoA to form β -keto-hexanoyl-CoA and the reduction of β -keto-hexanoyl-CoA to 3-hydroxy-hexanoyl-CoA (Fukui et al., *Biomacromolecules* 3:618-624 (2002)). To improve the production of primary alcohols, exogenous DNA sequences encoding for these specific enzymes can be expressed in the production host of interest. Furthermore, the above described enzymes can be subjected to directed evolution to generate improved versions of these enzymes with high activity and high substrate specificity. A similar approach also can be utilized with any or all other enzymatic steps in the primary alcohol producing pathways of the invention to, for example, improve enzymatic activity and/or specificity and/or to generate long chain alcohols of a predetermined chain length or lengths.

In addition, fatty acyl-CoA and fatty alcohols generated as described above can be applied to produce esters of various lengths. These esters can be formed between: 1) fatty acyl-CoA and short-chain alcohols such as methanol, ethanol, propanol, etc.; 2) fatty alcohols and

short-chain acyl-CoA such as formyl-CoA, acetyl-CoA, and propionyl-CoA, etc.; 3) fatty acyl-CoA and fatty alcohols as shown in the following equations.



The fatty (or long-chain) alcohols can be synthesized intracellularly by the pathways described herein or can be added to the medium and taken up by the engineered microbe. Similarly, short-chain alcohols can be added to the medium or produced endogenously.

10 Ethanol is an exemplary short chain alcohol that is naturally produced by many microorganisms including *Escherichia coli* and *Saccharomyces cerevisiae*. Exemplary fatty esters include, but not limited to, fatty acid methyl esters (FAMES), fatty acid ethyl esters (FAEEs), acetyl esters, and wax. Such molecules have broad applications including in food, personal care, coatings, surfactants, and biodiesel (Gerhard Knothe, Energy & Fuels 2008, 22, 1358–1364). Fatty esters, in this context, are differentiated from wax by the size of the hydrocarbon chain on each side of the ester bond. Waxes have long chain hydrocarbons on each side of the ester bond, whereas fatty esters have one short chain and one long chain hydrocarbon on each side of the ester bond, respectively.

20 Esters can be produced chemically, for example, by heating the acid in the presence of an alcohol or multiple alcohols in the presence of a dehydrating agent such as an acid catalyst. Enzymes with ester-forming activity can also be applied to form esters directly from the acids and alcohols. The ester-forming enzymes can be targeted to the cytosol or another cellular compartment to enable intracellular conversion of alcohols and acids to esters. Alternatively, the ester-forming enzymes can be secreted into the fermentation medium to enable 25 extracellular conversion of alcohols and acids to esters. Another option is to add ester-forming enzymes to a mixture containing acids and alcohols such as a fermentation broth. Several enzymes with ester forming activity are described below.

The amidase from *Brevibacterium* sp. R312 (EC 3.5.1.4) is a likely enzyme with ester-forming activity. This enzyme was shown to hydrolyze ethylacrylate (Thiery et al., J. Gen. Microbiol., 132:2205-8, 1986; Soubrier et al., Gene, 116:99-104, 1992). The microsomal epoxide hydrolase from *Rattus norvegicus* (EC 3.3.2.9) has activity on hydrolyzing glycidyl methacrylate and is another suitable enzyme (Guengerich et al., Rev. Biochem. Toxicol. 4:5-30, 1982). The protein sequences of these genes are provided below.

Gene	GenBank ID	GI Number	Organism
<i>amiE</i>	JC1174	98711	<i>Brevibacterium</i> sp.
Eph-1	P07687.1	123928	<i>Rattus norvegicus</i>

The reactions to produce these esters can be catalyzed by enzymes with acyl-CoA:alcohol transacylase activities. Exemplary enzymes for catalyzing the formation of fatty esters include the acyl-CoA:fatty alcohol acyltransferase (wax ester synthase, WS, EC 2.3.1.75) and acetyl-CoA:alcohol O-acetyltransferase (EC 2.3.1.84). Exemplary genes coding for these enzymes include the *Acinetobacter* sp. ADPl *atfA* encoding a bifunctional enzyme with both wax ester synthase (WS) and acyl-CoA: diacylglycerol acyltransferase (DGAT) activities (Kalscheuer et al. AJ Biol Chem 2003, 278: 8075-8082.); the *Simmondsia chinensis* gene AAD38041 encoding a enzyme required for the accumulation of waxes in jojoba seeds (Lardizabal et al. Plant Physiology 2000, 122: 645-655.); the *Alcanivorax borkumensis atfAl* and *atfA2* encoding bifunctional WS/DGAT enzymes (Kalscheuer et al. J Bacteriol 2007, 189: 918-928.); the *Fragaria x ananassa AAT* encoding an alcohol acetyltransferase (Noichinda et al. FoodSci Technol Res 1999, 5: 239-242.); the *Rosa hybrid cultivar AATl* encoding an alcohol acetyltransferase (Guterman et al. Plant Mol Biol 2006, 60: 555-563.); and the *Saccharomyces cerevisiae ATF1* and *ATF2* encoding alcohol acetyltransferases (Mason et al. Yeast 2000, 16: 1287-1298.); and Ws1 and Ws2 from *Marinobacter hydrocarbonoclasticus* (Holtzapfel, E. and Schmidt-Dannert, C., J. Bacteriol. 189 (10), 3804-3812, 2007). The protein sequences of the enzymes encoded by these genes are provided below.

Gene	GenBank ID	GI Number	Organism
<i>atfA</i>	Q8GGG1	81478805	<i>Acinetobacter</i> sp.
<i>ADPl</i>	AAD38041	5020219	<i>Simmondsia chinensis</i>
<i>atfAl</i>	YP694462	110835603	<i>Alcanivorax borkumensis SK2</i>
<i>atfA2</i>	YP693524	110834665	<i>Alcanivorax borkumensis SK2</i>
<i>AAT</i>	AAG13130	110834665	<i>Fragaria x ananassa</i>
<i>AATl</i>	Q5I6B5	75105208	<i>Rosa hybrid cultivar</i>
<i>ATF1</i>	P40353	2506980	<i>Saccharomyces cerevisiae</i>
<i>ATF2</i>	P53296	1723729	<i>Saccharomyces cerevisiae</i>
<i>Ws2</i>	ABO21021.1	126567232	<i>Marinobacter hydrocarbonoclasticus</i>

<i>Wsl</i>	ABO21020.1	126567230	<i>Marinobacter hydrocarbonoclasticus</i>
------------	------------	-----------	---

The *Homo sapiens* paraoxonase enzymes PON1, *PON1* (G3C9), and PON3 (EC 3.1.8.1) possess both arylesterase and organophosphatase activities and also may possess ester-forming activity. PON1 has a common polymorphic site at residue 192, glutamine (R) or arginine (Q), that results in qualitative differences. For example, the R isozyme has a higher esterase activity on GBL than the S isozyme (Billecke et al., *Drug Metab Dispos.* 28:1335-1342 (2000)). In *H. sapiens* cells, PON1 resides on high-density lipoprotein (HDL) particles, and its activity and stability require this environment. Wild type and recombinant PON1 enzymes have been functionally expressed in other organisms (Rochu et al., *Biochem.Soc.Trans.* 35:1616-1620 (2007); Martin et al., *Appl.Environ.Microbiol.* (2009)). A directed evolution study of PON1 yielded several mutant enzymes with improved solubility and catalytic properties in *E. coli* (nucleotide accession numbers AY499188-AY499199) (Aharoni et al., *Proc.Natl.Acad.Sci.U.S.A* 101:482-487 (2004)). One recombinant variant from this study, G3C9 (Aharoni et al., *Proc.Natl.Acad.Sci.U.S.A* 101:482-487 (2004)), was recently used in an integrated bioprocess for the pH-dependent production of 4-valerolactone from levulinate (Martin et al., *Appl.Environ.Microbiol.* (2009)). Human PON3 is yet another suitable enzyme that may possess ester-forming activity (Draganov et al., *J.Lipid Res.* 46:1239-1247 (2005)).

Gene	Accession No.	GI No.	Organism
<i>PON1</i>	NP_000437.3	19923106	<i>Homo sapiens</i>
<i>PON1</i> (G3C9)	AAR95986.1	40850544	<i>Synthetic variant</i>
<i>PON3</i>	NP_000931.1	29788996	<i>Homo sapiens</i>

Additionally, the *Candida antarctica* lipase B is another suitable candidate enzyme with ester-forming activity (Efe et al., *Biotechnol.Bioeng.* 99:1392-1406 (2008)). The esterase from *Pseudomonas fluorescens*, encoded by *EstF1*, is yet another suitable enzyme (Khalameyzer et al., *Appl.Environ.Microbiol.* 65:477-482 (1999)). Other lipase enzymes from organisms such as *Pseudomonas fluorescens* and *Bacillus subtilis* may also catalyze this transformation. The *B. subtilis* and *P. fluorescens* genes encode triacylglycerol lipase enzymes which have been cloned and characterized in *E. coli* (Dartois et al.,

Biochim.Biophys.Acta 1131:253-260 (1992); Tan et al., *Appl.Environ.Microbiol.* 58:1402-1407 (1992)).

Gene	Accession No.	GI No.	Organism
<i>calB</i>	P41365.1	1170790	<i>Candida Antarctica</i>
<i>EstF1</i>	AAC36352.1	3641341	<i>Pseudomonas fluorescens</i>
<i>lipB</i>	P41773.1	1170792	<i>Pseudomonas fluorescens</i>
<i>estA</i>	P37957.1	7676155	<i>Bacillus subtilis</i>

Formation of esters may also be catalyzed by enzymes in the 3.1.1 family that act on
 5 carboxylic ester bonds molecules for the interconversion between cyclic lactones and the
 open chain hydroxycarboxylic acids. The L-lactonase from *Fusarium proliferatum* ECU2002
 exhibits lactonase and esterase activities on a variety of lactone substrates (Zhang et al.,
Appl.Microbiol.Biotechnol. 75:1087-1094 (2007)). The 1,4-lactone hydroxyacylhydrolase
 (EC 3.1.1.25), also known as 1,4-lactonase or gamma-lactonase, is specific for 1,4-lactones
 10 with 4-8 carbon atoms. The gamma lactonase in human blood and rat liver microsomes was
 purified (Fishbein et al., *J Biol Chem* 241:4835-4841 (1966)) and the lactonase activity was
 activated and stabilized by calcium ions (Fishbein et al., *J Biol Chem* 241:4842-4847 (1966)).
 The optimal lactonase activities were observed at pH 6.0, whereas high pH resulted in
 hydrolytic activities (Fishbein and Bessman, *J Biol Chem* 241:4842-4847 (1966)). Genes
 15 from *Xanthomonas campestris*, *Aspergillus niger* and *Fusarium oxysporum* have been
 annotated as 1,4-lactonase and can be utilized to catalyze the transformation of 4-
 hydroxybutyrate to GBL (Zhang et al., *Appl Microbiol Biotechnol* 75:1087-1094 (2007)).

Gene	Accession No.	GI No.	Organism
<i>EU596535.1:1..1206</i>	ACC61057.1	183238971	<i>Fusarium proliferatum</i>
xccb100_2516	YP_001903921.1	188991911	<i>Xanthomonas campestris</i>
An16g06620	CAK46996.1	134083519	<i>Aspergillus niger</i>
BAA34062	BAA34062.1	3810873	<i>Fusarium oxysporum</i>

Host microbial organisms can be selected from, and the non-naturally occurring microbial
 20 organisms generated in, for example, bacteria, yeast, fungus or any of a variety of other
 microorganisms applicable to fermentation processes. Exemplary bacteria include species

selected from *E. coli*, *Rhodococcus opacus*, *Ralstonia eutropha*, *Klebsiella oxytoca*, *Anaerobiospirillum succiniciproducens*, *Actinobacillus succinogenes*, *Mannheimia succiniciproducens*, *Rhizobium etli*, *Bacillus subtilis*, *Corynebacterium glutamicum*, *Gluconobacter oxydans*, *Zymomonas mobilis*, *Lactococcus lactis*, *Lactobacillus plantarum*,
5 *Streptomyces coelicolor*, *Clostridium acetobutylicum*, *Pseudomonas fluorescens*, *Pseudomonas putida* and *E. gracilis*. Exemplary yeasts or fungi include species selected from *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Kluyveromyces lactis*, *Kluyveromyces marxianus*, *Aspergillus terreus*, *Aspergillus niger* and *Pichia pastoris*.

Methods for constructing and testing the expression levels of a non-naturally occurring
10 primary alcohol-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 Laboratory, New York (2001); Ausubel et al., *Current Protocols in Molecular Biology*, John Wiley and Sons, Baltimore, MD (1999). For example, nucleic acids encoding enzymes in the
15 malonyl-CoA independent FAS and/or acyl-reduction pathway 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, transduction, transfection, and ultrasound transformation. For exogenous expression in *E. coli* or other prokaryotic cells, for example, mitochondrial genes will encode an N-terminal targeting signals, which can be
20 removed before transformation into host cells. For exogenous expression in yeast or other eukaryotic cells, genes can be expressed in the cytosol without the addition of the targeting sequence, or alternatively, can be targeted to mitochondrion with the addition of mitochondrial targeting signal functional in the host organism. Furthermore, genes can be subjected for codon optimization with techniques well known in the art, to achieve optimal
25 expression of the one or more malonyl-CoA independent FAS and/or acyl-reduction pathway gene products.

An expression vector or vectors can be constructed to include one or more malonyl-CoA independent FAS and/or acyl-reduction pathway encoding nucleic acids as exemplified herein operably linked to expression control sequences functional in the host organism.

30 Expression vectors applicable for use in the microbial host organisms of the invention include, for example, plasmids, phage vectors, viral 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 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 nucleic acids encoding 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 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 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 herein.

Primary alcohol production can be detected and/or monitored using methods well known to those skilled in the art. For example, final product of primary alcohol and/or intermediates such as acyl-CoA and organic acids can be analyzed by HPLC, GC-MS and LC-MS. For example, primary alcohols can be separated by HPLC using a Spherisorb 5 ODS1 column and a mobile phase of 70% 10 mM phosphate buffer (pH=7) and 30% methanol, and detected using a UV detector at 215 nm (Hennessy et al. 2004, J. Forensic Sci. 46(6):1-9). The release or secretion of primary alcohol into the culture medium or fermentation broth also can be detected using these procedures. Activities of one or more enzymes in the malonyl-CoA independent FAS and/or acyl-reduction pathway also can be measured using methods well known in the art.

The non-naturally occurring microbial organisms of the invention are constructed using methods well known in the art as exemplified above to exogenously express at least one

nucleic acid encoding a malonyl-CoA independent FAS pathway enzyme in sufficient amounts to produce primary alcohol. Following the teachings and guidance provided herein, the non-naturally occurring microbial organisms of the invention can achieve biosynthesis of greater than that which can be synthesized in naturally occurring organisms. Generally, the intracellular concentration of, for example, octanol is about 54 μ g/L and decanol is about 148 μ g/L.

As described further below, one exemplary growth condition for achieving biosynthesis of primary alcohols includes anaerobic culture or fermentation conditions. In certain embodiments, the non-naturally occurring microbial organisms of the invention can be sustained, cultured or fermented under anaerobic or substantially anaerobic conditions. Briefly, anaerobic conditions refers 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 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 invention further provides a method for the production of primary alcohols. The method includes culturing a non-naturally occurring microbial organism having a malonyl-CoA-independent fatty acid synthesis (FAS) pathway and an acyl-reduction pathway comprising at least one exogenous nucleic acid encoding a malonyl-CoA-independent FAS pathway enzyme expressed in sufficient amounts to produce a primary alcohol under substantially anaerobic conditions for a sufficient period of time to produce said primary alcohol, said malonyl-CoA-independent FAS pathway comprising ketoacyl-CoA acyltransferase or ketoacyl-CoA thiolase, 3-hydroxyacyl-CoA dehydrogenase, enoyl-CoA hydratase and enoyl-CoA reductase, said acyl-reduction pathway comprising an acyl-CoA reductase and an alcohol dehydrogenase.

Any of the non-naturally occurring microbial organisms described previously can be cultured to produce the biosynthetic products of the invention. For example, the primary alcohol producers can be cultured for the biosynthetic production of its engineered target primary

alcohol. The primary alcohol can be isolated or isolated and further utilized in a wide variety of products and procedures.

In one embodiment, provided herein is a method for producing a primary alcohol, comprising culturing a non-naturally occurring microbial organism described herein as capable of

5 producing a primary alcohol under conditions and for a sufficient period of time to produce a primary alcohol. For example, in one embodiment, the invention provides a method for

producing a primary alcohol using a non-naturally occurring microbial organism, comprising a microbial organism having a primary alcohol pathway comprising at least one exogenous nucleic acid encoding a primary alcohol pathway enzyme expressed in a sufficient amount to

10 produce a primary alcohol; said non-naturally occurring microbial organism further comprising: (i) a reductive TCA pathway comprising at least one exogenous nucleic acid encoding a reductive TCA pathway enzyme, wherein said at least one exogenous nucleic acid is selected from an ATP-citrate lyase, a citrate lyase, a citryl-CoA synthetase, a citryl-CoA lyase, a fumarate reductase, and an alpha-ketoglutarate:ferredoxin oxidoreductase; (ii) a

15 reductive TCA pathway comprising at least one exogenous nucleic acid encoding a reductive TCA pathway enzyme, wherein said at least one exogenous nucleic acid is selected from a pyruvate:ferredoxin oxidoreductase, a phosphoenolpyruvate carboxylase, a phosphoenolpyruvate carboxykinase, a CO dehydrogenase, and an H₂ hydrogenase; or (iii)

at least one exogenous nucleic acid encodes an enzyme selected from a CO dehydrogenase, an H₂ hydrogenase, and combinations thereof; and wherein said primary alcohol pathway

20 comprises a malonyl-CoA-independent FAS pathway comprising ketoacyl-CoA acyltransferase or ketoacyl-CoA thiolase, 3-hydroxyacyl-CoA dehydrogenase, enoyl-CoA hydratase and enoyl-CoA reductase; and an acyl-reduction pathway comprising an acyl-CoA reductase and an alcohol dehydrogenase. In a specific embodiment, the invention provides a

25 method for producing a primary alcohol using said microbial organism comprising (i) further comprises an exogenous nucleic acid encoding an enzyme selected from a

pyruvate:ferredoxin oxidoreductase, an aconitase, an isocitrate dehydrogenase, a succinyl-CoA synthetase, a succinyl-CoA transferase, a fumarase, a malate dehydrogenase, an acetate kinase, a phosphotransacetylase, an acetyl-CoA synthetase, an NAD(P)H:ferredoxin

30 oxidoreductase, ferredoxin, and combinations thereof. In another specific embodiment, the invention provides a method for producing a primary alcohol using said microbial organism comprising (ii) further comprises an exogenous nucleic acid encoding an enzyme selected from an aconitase, an isocitrate dehydrogenase, a succinyl-CoA synthetase, a succinyl-CoA

transferase, a fumarase, a malate dehydrogenase, and combinations thereof. In another specific embodiment, said microbial organism comprises two, three, four, five or six exogenous nucleic acids each encoding a primary alcohol pathway enzyme. In another specific embodiment, the invention provides a method for producing a primary alcohol using said microbial organism comprises four exogenous nucleic acids encoding malonyl-CoA-independent FAS pathway enzymes comprising ketoacyl-CoA acyltransferase or ketoacyl-CoA thiolase, 3-hydroxyacyl-CoA dehydrogenase, enoyl-CoA hydratase and enoyl-CoA reductase; and two exogenous nucleic acids encoding acyl-reduction pathway enzymes comprising an acyl-CoA reductase and an alcohol dehydrogenase. In another specific embodiment, the invention provides a method for producing a primary alcohol using said microbial organism comprising (i) comprises two, three or four exogenous nucleic acids each encoding a reductive TCA pathway enzyme. In another specific embodiment, the invention provides a method for producing a primary alcohol using said microbial organism comprising (ii) comprises two, three or four exogenous nucleic acids each encoding a reductive TCA pathway enzyme. In another specific embodiment, said at least one exogenous nucleic acid in said non-naturally occurring microbial organism is a heterologous nucleic acid. In another specific embodiment, said non-naturally occurring microbial organism is in a substantially anaerobic culture medium. In another specific embodiment, said exogenous nucleic acid encoding an acyl-reduction pathway enzyme in said non-naturally occurring microbial organism comprises an enzyme having acyl-CoA reductase and alcohol dehydrogenase activity, e.g., fatty alcohol forming acyl-CoA reductase (FAR). In another embodiment, said non-naturally occurring microbial organism further comprises an acyl-reduction pathway comprising an acyl-CoA hydrolase, an acyl-CoA transferase or an acyl-CoA ligase; a carboxylic acid reductase and an alcohol dehydrogenase. In another specific embodiment, said primary alcohol is produced in amounts at least 10% greater levels compared to a microbial organism lacking said exogenous nucleic acid encoding a malonyl-CoA-independent FAS pathway enzyme. In another specific embodiment, said primary alcohol comprises an alcohol having between 4-24 carbon atoms, e.g., said primary alcohol is selected from butanol, hexanol, heptanol, octanol, nonanol, decanol, dodecanol, tetradecanol or hexadecanol. In a specific embodiment, said non-naturally occurring microbial organism is in a substantially anaerobic culture medium.

In another embodiment, provided herein is a method for producing a fatty acyl-CoA, comprising culturing a non-naturally occurring microbial organism described herein as

capable of producing a fatty acyl-CoA under conditions and for a sufficient period of time to produce a fatty acyl-CoA. For example, in one embodiment, the invention provides a method for producing a fatty acyl-CoA using a non-naturally occurring microbial organism, comprising a microbial organism having a fatty acyl-CoA pathway comprising at least one

5 exogenous nucleic acid encoding a fatty acyl-CoA pathway enzyme expressed in a sufficient amount to produce a fatty acyl-CoA; said non-naturally occurring microbial organism further comprising: (i) a reductive TCA pathway comprising at least one exogenous nucleic acid encoding a reductive TCA pathway enzyme, wherein said at least one exogenous nucleic acid is selected from an ATP-citrate lyase, a citrate lyase, a citryl-CoA synthetase, a citryl-CoA

10 lyase, a fumarate reductase, and an alpha-ketoglutarate:ferredoxin oxidoreductase; (ii) a reductive TCA pathway comprising at least one exogenous nucleic acid encoding a reductive TCA pathway enzyme, wherein said at least one exogenous nucleic acid is selected from a pyruvate:ferredoxin oxidoreductase, a phosphoenolpyruvate carboxylase, a phosphoenolpyruvate carboxykinase, a CO dehydrogenase, and an H₂ hydrogenase; or (iii)

15 at least one exogenous nucleic acid encodes an enzyme selected from a CO dehydrogenase, an H₂ hydrogenase, and combinations thereof; and wherein said fatty acyl-CoA pathway comprises a malonyl-CoA-independent FAS pathway comprising ketoacyl-CoA acyltransferase or ketoacyl-CoA thiolase, 3-hydroxyacyl-CoA dehydrogenase, enoyl-CoA hydratase and enoyl-CoA reductase. In a specific embodiment, the invention provides a

20 method for producing a fatty acyl-CoA using said microbial organism comprising (i) further comprises an exogenous nucleic acid encoding an enzyme selected from a pyruvate:ferredoxin oxidoreductase, an aconitase, an isocitrate dehydrogenase, a succinyl-CoA synthetase, a succinyl-CoA transferase, a fumarase, a malate dehydrogenase, an acetate kinase, a phosphotransacetylase, an acetyl-CoA synthetase, an NAD(P)H:ferredoxin

25 oxidoreductase, ferredoxin, and combinations thereof. In another specific embodiment, the invention provides a method for producing a fatty acyl-CoA using said microbial organism comprising (ii) further comprises an exogenous nucleic acid encoding an enzyme selected from an aconitase, an isocitrate dehydrogenase, a succinyl-CoA synthetase, a succinyl-CoA transferase, a fumarase, a malate dehydrogenase, and combinations thereof. In another

30 specific embodiment, said microbial organism comprises two, three, four, five or six exogenous nucleic acids each encoding a fatty acyl-CoA pathway enzyme. In another specific embodiment, the invention provides a method for producing a fatty acyl-CoA using said microbial organism comprises four exogenous nucleic acids encoding malonyl-CoA-independent FAS pathway enzymes comprising ketoacyl-CoA acyltransferase or ketoacyl-

CoA thiolase, 3-hydroxyacyl-CoA dehydrogenase, enoyl-CoA hydratase and enoyl-CoA reductase. In another specific embodiment, the invention provides a method for producing a fatty acyl-CoA using said microbial organism comprising (i) comprises two, three or four exogenous nucleic acids each encoding a reductive TCA pathway enzyme. In another specific embodiment, said microbial organism comprising (ii) comprises two, three or four exogenous nucleic acids each encoding a reductive TCA pathway enzyme. In another specific embodiment, said at least one exogenous nucleic acid in said non-naturally occurring microbial organism is a heterologous nucleic acid. In another specific embodiment, said non-naturally occurring microbial organism is in a substantially anaerobic culture medium. In a specific embodiment, said non-naturally occurring microbial organism is in a substantially anaerobic culture medium.

In another embodiment, provided herein is a method for producing a fatty ester, comprising culturing a non-naturally occurring microbial organism described herein as capable of producing a fatty ester under conditions and for a sufficient period of time to produce a fatty ester. For example in one embodiment, the invention provides a method for producing a fatty ester using a non-naturally occurring microbial organism, comprising a microbial organism having a fatty ester pathway comprising at least one exogenous nucleic acid encoding a fatty ester pathway enzyme expressed in a sufficient amount to produce a fatty ester; said non-naturally occurring microbial organism further comprising: (i) a reductive TCA pathway comprising at least one exogenous nucleic acid encoding a reductive TCA pathway enzyme, wherein said at least one exogenous nucleic acid is selected from an ATP-citrate lyase, a citrate lyase, a citryl-CoA synthetase, a citryl-CoA lyase, a fumarate reductase, and an alpha-ketoglutarate:ferredoxin oxidoreductase; (ii) a reductive TCA pathway comprising at least one exogenous nucleic acid encoding a reductive TCA pathway enzyme, wherein said at least one exogenous nucleic acid is selected from a pyruvate:ferredoxin oxidoreductase, a phosphoenolpyruvate carboxylase, a phosphoenolpyruvate carboxykinase, a CO dehydrogenase, and an H₂ hydrogenase; or (iii) at least one exogenous nucleic acid encodes an enzyme selected from a CO dehydrogenase, an H₂ hydrogenase, and combinations thereof; and wherein said fatty ester pathway comprises a pathway selected from: (a) a malonyl-CoA-independent fatty acid synthesis (FAS) pathway and a wax ester synthase, wherein said malonyl-CoA-independent FAS pathway comprises ketoacyl-CoA acyltransferase or ketoacyl-CoA thiolase, 3-hydroxyacyl-CoA dehydrogenase, enoyl-CoA hydratase and enoyl-CoA reductase; and (b) a malonyl-CoA-independent fatty acid synthesis

(FAS) pathway and an alcohol acetyltransferase, wherein said malonyl-CoA-independent FAS pathway comprises ketoacyl-CoA acyltransferase or ketoacyl-CoA thiolase, 3-hydroxyacyl-CoA dehydrogenase, enoyl-CoA hydratase and enoyl-CoA reductase. In a specific embodiment, the invention provides a method for producing a fatty ester using said microbial organism comprising (i) further comprises an exogenous nucleic acid encoding an enzyme selected from a pyruvate:ferredoxin oxidoreductase, an aconitase, an isocitrate dehydrogenase, a succinyl-CoA synthetase, a succinyl-CoA transferase, a fumarase, a malate dehydrogenase, an acetate kinase, a phosphotransacetylase, an acetyl-CoA synthetase, an NAD(P)H:ferredoxin oxidoreductase, ferredoxin, and combinations thereof. In another specific embodiment, the invention provides a method for producing a fatty ester using said microbial organism comprising (ii) further comprises an exogenous nucleic acid encoding an enzyme selected from an aconitase, an isocitrate dehydrogenase, a succinyl-CoA synthetase, a succinyl-CoA transferase, a fumarase, a malate dehydrogenase, and combinations thereof. In another specific embodiment, the invention provides a method for producing a fatty ester using said microbial organism comprises two, three, four, five or six exogenous nucleic acids each encoding a fatty ester enzyme. In another specific embodiment, the invention provides a method for producing a fatty ester using said microbial organism comprises (A) four exogenous nucleic acids encoding malonyl-CoA-independent fatty acid synthesis (FAS) pathway enzymes and one exogenous nucleic acid encoding a wax ester synthase, wherein said malonyl-CoA-independent FAS pathway enzymes comprise a ketoacyl-CoA acyltransferase or a ketoacyl-CoA thiolase, a 3-hydroxyacyl-CoA dehydrogenase, an enoyl-CoA hydratase and an enoyl-CoA reductase; or (B) four exogenous nucleic acids encoding malonyl-CoA-independent fatty acid synthesis (FAS) pathway enzymes and one exogenous nucleic acid encoding an alcohol acetyltransferase, wherein said malonyl-CoA-independent FAS pathway enzymes comprise a ketoacyl-CoA acyltransferase or a ketoacyl-CoA thiolase, a 3-hydroxyacyl-CoA dehydrogenase, an enoyl-CoA hydratase and an enoyl-CoA reductase. In another specific embodiment, the invention provides a method for producing a fatty ester using said microbial organism comprising (i) comprises two, three or four exogenous nucleic acids each encoding a reductive TCA pathway enzyme. In another specific embodiment, the invention provides a method for producing a fatty ester using said microbial organism comprising (ii) comprises two, three or four exogenous nucleic acids each encoding a reductive TCA pathway enzyme. In another specific embodiment, said at least one exogenous nucleic acid in said non-naturally occurring microbial organism is a heterologous nucleic acid. In another specific embodiment, said non-naturally occurring microbial

organism is in a substantially anaerobic culture medium. In a specific embodiment, said non-naturally occurring microbial organism is in a substantially anaerobic culture medium.

In another embodiment, provided herein is a method for producing a wax, comprising culturing a non-naturally occurring microbial organism described herein as capable of

5 producing a wax under conditions and for a sufficient period of time to produce a wax. For example, In one embodiment, the invention provides a method for producing a wax using a non-naturally occurring microbial organism, comprising a microbial organism having a wax pathway comprising at least one exogenous nucleic acid encoding a wax pathway enzyme expressed in a sufficient amount to produce a wax; said non-naturally occurring microbial
10 organism further comprising: (i) a reductive TCA pathway comprising at least one exogenous nucleic acid encoding a reductive TCA pathway enzyme, wherein said at least one exogenous nucleic acid is selected from an ATP-citrate lyase, a citrate lyase, a citryl-CoA synthetase, a citryl-CoA lyase, a fumarate reductase, and an alpha-ketoglutarate:ferredoxin oxidoreductase; (ii) a reductive TCA pathway comprising at least one exogenous nucleic
15 acid encoding a reductive TCA pathway enzyme, wherein said at least one exogenous nucleic acid is selected from a pyruvate:ferredoxin oxidoreductase, a phosphoenolpyruvate carboxylase, a phosphoenolpyruvate carboxykinase, a CO dehydrogenase, and an H₂ hydrogenase; or (iii) at least one exogenous nucleic acid encodes an enzyme selected from a CO dehydrogenase, an H₂ hydrogenase, and combinations thereof; and wherein said wax
20 pathway comprises a pathway selected from: (a) a malonyl-CoA-independent fatty acid synthesis (FAS) pathway and a wax ester synthase, wherein said malonyl-CoA-independent FAS pathway comprises ketoacyl-CoA acyltransferase or ketoacyl-CoA thiolase, 3-hydroxyacyl-CoA dehydrogenase, enoyl-CoA hydratase and enoyl-CoA reductase; and (b) a malonyl-CoA-independent fatty acid synthesis (FAS) pathway and an alcohol
25 acetyltransferase, wherein said malonyl-CoA-independent FAS pathway comprises ketoacyl-CoA acyltransferase or ketoacyl-CoA thiolase, 3-hydroxyacyl-CoA dehydrogenase, enoyl-CoA hydratase and enoyl-CoA reductase. In a specific embodiment, the invention provides a method for producing a wax using said microbial organism comprising (i) further comprises an exogenous nucleic acid encoding an enzyme selected from a pyruvate:ferredoxin
30 oxidoreductase, an aconitase, an isocitrate dehydrogenase, a succinyl-CoA synthetase, a succinyl-CoA transferase, a fumarase, a malate dehydrogenase, an acetate kinase, a phosphotransacetylase, an acetyl-CoA synthetase, an NAD(P)H:ferredoxin oxidoreductase, ferredoxin, and combinations thereof. In another specific embodiment, the invention

provides a method for producing a wax using said microbial organism comprising (ii) further comprises an exogenous nucleic acid encoding an enzyme selected from an aconitase, an isocitrate dehydrogenase, a succinyl-CoA synthetase, a succinyl-CoA transferase, a fumarase, a malate dehydrogenase, and combinations thereof. In another specific embodiment, the invention provides a method for producing a wax using said microbial organism comprises two, three, four, five or six exogenous nucleic acids each encoding a primary alcohol, a fatty acyl-CoA, a fatty ester or a wax pathway enzyme. In another specific embodiment, the invention provides a method for producing a wax using said microbial organism comprises (A) four exogenous nucleic acids encoding malonyl-CoA-independent fatty acid synthesis (FAS) pathway enzymes and one nucleic acid encoding a wax ester synthase, wherein said malonyl-CoA-independent FAS pathway enzymes comprise a ketoacyl-CoA acyltransferase or a ketoacyl-CoA thiolase, a 3-hydroxyacyl-CoA dehydrogenase, an enoyl-CoA hydratase and an enoyl-CoA reductase; or (B) four exogenous nucleic acids encoding malonyl-CoA-independent fatty acid synthesis (FAS) pathway enzymes and one nucleic acid encoding an alcohol acetyltransferase, wherein said malonyl-CoA-independent FAS pathway enzymes comprise a ketoacyl-CoA acyltransferase or a ketoacyl-CoA thiolase, a 3-hydroxyacyl-CoA dehydrogenase, an enoyl-CoA hydratase and an enoyl-CoA reductase. In another specific embodiment, the invention provides a method for producing a wax using said microbial organism comprising (i) comprises two, three or four exogenous nucleic acids each encoding a reductive TCA pathway enzyme. In another specific embodiment, the invention provides a method for producing a wax using said microbial organism comprising (ii) comprises two, three or four exogenous nucleic acids each encoding a reductive TCA pathway enzyme. In another specific embodiment, said at least one exogenous nucleic acid in said non-naturally occurring microbial organism is a heterologous nucleic acid. In another specific embodiment, said non-naturally occurring microbial organism is in a substantially anaerobic culture medium. In a specific embodiment, said non-naturally occurring microbial organism is in a substantially anaerobic culture medium.

In another embodiment, provided herein is a method for producing acyl-ACP, comprising culturing a non-naturally occurring microbial organism described herein as capable of producing acyl-ACP under conditions and for a sufficient period of time to produce acyl-ACP. For example, in one embodiment, the invention provides a method for producing acyl-ACP using a non-naturally occurring microbial organism, comprising a microbial organism having an acyl-ACP pathway comprising at least one exogenous nucleic acid encoding a

acyl-ACP pathway enzyme expressed in a sufficient amount to produce acyl-ACP; said non-naturally occurring microbial organism further comprising: (i) a reductive TCA pathway comprising at least one exogenous nucleic acid encoding a reductive TCA pathway enzyme, wherein said at least one exogenous nucleic acid is selected from an ATP-citrate lyase, a citrate lyase, a citryl-CoA synthetase, a citryl-CoA lyase, a fumarate reductase, and an alpha-ketoglutarate:ferredoxin oxidoreductase; (ii) a reductive TCA pathway comprising at least one exogenous nucleic acid encoding a reductive TCA pathway enzyme, wherein said at least one exogenous nucleic acid is selected from a pyruvate:ferredoxin oxidoreductase, a phosphoenolpyruvate carboxylase, a phosphoenolpyruvate carboxykinase, a CO dehydrogenase, and an H₂ hydrogenase; or (iii) at least one exogenous nucleic acid encodes an enzyme selected from a CO dehydrogenase, an H₂ hydrogenase, and combinations thereof; and wherein said acyl-ACP pathway comprises a pathway selected from: (A) acetyl-CoA carboxylase and a fatty acid synthase; and (B) a fatty acid synthase. In another specific embodiment, the invention provides a method for producing acyl-ACP using said microbial organism comprising (i) further comprises an exogenous nucleic acid encoding an enzyme selected from a pyruvate:ferredoxin oxidoreductase, an aconitase, an isocitrate dehydrogenase, a succinyl-CoA synthetase, a succinyl-CoA transferase, a fumarase, a malate dehydrogenase, an acetate kinase, a phosphotransacetylase, an acetyl-CoA synthetase, an NAD(P)H:ferredoxin oxidoreductase, ferredoxin, and combinations thereof. In another specific embodiment, the invention provides a method for producing acyl-ACP using said microbial organism comprising (ii) further comprises an exogenous nucleic acid encoding an enzyme selected from an aconitase, an isocitrate dehydrogenase, a succinyl-CoA synthetase, a succinyl-CoA transferase, a fumarase, a malate dehydrogenase, and combinations thereof. In another specific embodiment, the invention provides a method for producing acyl-ACP using said microbial organism comprises two, three, four, five, or six exogenous nucleic acids each encoding a acyl-ACP pathway enzyme. In another specific embodiment, the invention provides a method for producing acyl-ACP using said microbial organism comprising (i) comprises two, three or four exogenous nucleic acids each encoding a reductive TCA pathway enzyme. In another specific embodiment, the invention provides a method for producing acyl-ACP using said microbial organism comprising (ii) comprises two, three or four exogenous nucleic acids each encoding a reductive TCA pathway enzyme. In another specific embodiment, said at least one exogenous nucleic acid is a heterologous nucleic acid. In a specific embodiment, said non-naturally occurring microbial organism is in a substantially anaerobic culture medium.

In another embodiment, provided herein is a method for producing a fatty acid, comprising culturing a non-naturally occurring microbial organism described herein as capable of producing a fatty acid under conditions and for a sufficient period of time to produce a fatty acid. For example, in one embodiment, the invention provides a method for producing a fatty acid using a non-naturally occurring microbial organism having an acyl-ACP pathway as disclosed herein further comprises a fatty acid pathway comprising an exogenous nucleic acid encoding a thioesterase expressed in a sufficient amount to produce fatty acid. In another specific embodiment, the invention provides a method for producing a fatty acid using said microbial organism comprising (i) further comprises an exogenous nucleic acid encoding an enzyme selected from a pyruvate:ferredoxin oxidoreductase, an aconitase, an isocitrate dehydrogenase, a succinyl-CoA synthetase, a succinyl-CoA transferase, a fumarase, a malate dehydrogenase, an acetate kinase, a phosphotransacetylase, an acetyl-CoA synthetase, an NAD(P)H:ferredoxin oxidoreductase, ferredoxin, and combinations thereof. In another specific embodiment, the invention provides a method for producing a fatty acid using said microbial organism comprising (ii) further comprises an exogenous nucleic acid encoding an enzyme selected from an aconitase, an isocitrate dehydrogenase, a succinyl-CoA synthetase, a succinyl-CoA transferase, a fumarase, a malate dehydrogenase, and combinations thereof. In another specific embodiment, the invention provides a method for producing a fatty acid using said microbial organism comprises two, three, four, five, or six exogenous nucleic acids each encoding a acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol pathway enzyme. In another specific embodiment, the invention provides a method for producing a fatty acid using said microbial organism comprising (i) comprises two, three or four exogenous nucleic acids each encoding a reductive TCA pathway enzyme. In another specific embodiment, the invention provides a method for producing a fatty acid using said microbial organism comprising (ii) comprises two, three or four exogenous nucleic acids each encoding a reductive TCA pathway enzyme. In another specific embodiment, said at least one exogenous nucleic acid is a heterologous nucleic acid. In a specific embodiment, said non-naturally occurring microbial organism is in a substantially anaerobic culture medium.

In another embodiment, provided herein is a method for producing acyl-CoA, comprising culturing a non-naturally occurring microbial organism described herein as capable of producing acyl-CoA under conditions and for a sufficient period of time to produce acyl-CoA. For example, in one embodiment, the invention provides a method for producing a

fatty acid using said non-naturally occurring microbial organism having a fatty acid pathway as disclosed herein further comprises an acyl-CoA pathway comprising an exogenous nucleic acid encoding an acyl-CoA synthetase or an acyl-CoA ligase expressed in a sufficient amount to produce acyl-CoA. In another specific embodiment, said microbial organism comprising

5 (i) further comprises an exogenous nucleic acid encoding an enzyme selected from a pyruvate:ferredoxin oxidoreductase, an aconitase, an isocitrate dehydrogenase, a succinyl-CoA synthetase, a succinyl-CoA transferase, a fumarase, a malate dehydrogenase, an acetate kinase, a phosphotransacetylase, an acetyl-CoA synthetase, an NAD(P)H:ferredoxin oxidoreductase, ferredoxin, and combinations thereof. In another specific embodiment, the
10 invention provides a method for producing acyl-CoA using said microbial organism comprising (ii) further comprises an exogenous nucleic acid encoding an enzyme selected from an aconitase, an isocitrate dehydrogenase, a succinyl-CoA synthetase, a succinyl-CoA transferase, a fumarase, a malate dehydrogenase, and combinations thereof. In another specific embodiment, the invention provides a method for producing acyl-CoA using said
15 microbial organism comprises two, three, four, five, or six exogenous nucleic acids each encoding acyl-CoA pathway enzyme. In another specific embodiment, the invention provides a method for producing acyl-CoA using said microbial organism comprising (i) comprises two, three or four exogenous nucleic acids each encoding a reductive TCA pathway enzyme. In another specific embodiment, said microbial organism comprising (ii)
20 comprises two, three or four exogenous nucleic acids each encoding a reductive TCA pathway enzyme. In another specific embodiment, said said at least one exogenous nucleic acid is a heterologous nucleic acid. In a specific embodiment, said non-naturally occurring microbial organism is in a substantially anaerobic culture medium.

In another embodiment, provided herein is a method for producing a fatty aldehyde,
25 comprising culturing a non-naturally occurring microbial organism described herein as capable of producing fatty aldehyde under conditions and for a sufficient period of time to produce fatty aldehyde. For example, in one embodiment, the invention provides a method for producing a fatty aldehyde using said non-naturally occurring microbial organism having an acyl-CoA pathway as disclosed herein further comprises a fatty aldehyde pathway
30 comprising an exogenous nucleic acid encoding an acyl-CoA reductase (aldehyde forming) or carboxylic acid reductase expressed in a sufficient amount to produce a fatty aldehyde. In another specific embodiment, the invention provides a method for producing a fatty aldehyde using said microbial organism comprising (i) further comprises an exogenous nucleic acid

encoding an enzyme selected from a pyruvate:ferredoxin oxidoreductase, an aconitase, an isocitrate dehydrogenase, a succinyl-CoA synthetase, a succinyl-CoA transferase, a fumarase, a malate dehydrogenase, an acetate kinase, a phosphotransacetylase, an acetyl-CoA synthetase, an NAD(P)H:ferredoxin oxidoreductase, ferredoxin, and combinations thereof.

- 5 In another specific embodiment, the invention provides a method for producing a fatty aldehyde using said microbial organism comprising (ii) further comprises an exogenous nucleic acid encoding an enzyme selected from an aconitase, an isocitrate dehydrogenase, a succinyl-CoA synthetase, a succinyl-CoA transferase, a fumarase, a malate dehydrogenase, and combinations thereof. In another specific embodiment, the invention provides a method
- 10 for producing a fatty aldehyde using said microbial organism comprises two, three, four, five, or six exogenous nucleic acids each encoding a fatty aldehyde pathway enzyme. In another specific embodiment, the invention provides a method for producing a fatty aldehyde using said microbial organism comprising (i) comprises two, three or four exogenous nucleic acids each encoding a reductive TCA pathway enzyme. In another specific embodiment, the
- 15 invention provides a method for producing a fatty aldehyde using said microbial organism comprising (ii) comprises two, three or four exogenous nucleic acids each encoding a reductive TCA pathway enzyme. In another specific embodiment, said at least one exogenous nucleic acid is a heterologous nucleic acid. In a specific embodiment, said non-naturally occurring microbial organism is in a substantially anaerobic culture medium.
- 20 In another embodiment, provided herein is a method for producing a fatty alcohol, comprising culturing a non-naturally occurring microbial organism described herein as capable of producing fatty alcohol under conditions and for a sufficient period of time to produce fatty alcohol. For example, in one embodiment, the invention provides a method for producing a fatty alcohol using said non-naturally occurring microbial organism having a
- 25 fatty aldehyde pathway as disclosed herein further comprises a fatty alcohol pathway comprising an exogenous nucleic acid encoding an aldehyde reductase expressed in a sufficient amount to produce a fatty alcohol. In another embodiment, the invention provides a method for producing a fatty alcohol using said non-naturally occurring microbial organism an acyl-CoA pathway as disclosed herein further comprising a fatty alcohol pathway
- 30 comprising an exogenous nucleic acid encoding an acyl-CoA reductase (alcohol forming) expressed in a sufficient amount to produce a fatty alcohol. In another specific embodiment, the invention provides a method for producing a fatty alcohol using said microbial organism comprising (i) further comprises an exogenous nucleic acid encoding an enzyme selected

from a pyruvate:ferredoxin oxidoreductase, an aconitase, an isocitrate dehydrogenase, a succinyl-CoA synthetase, a succinyl-CoA transferase, a fumarase, a malate dehydrogenase, an acetate kinase, a phosphotransacetylase, an acetyl-CoA synthetase, an

NAD(P)H:ferredoxin oxidoreductase, ferredoxin, and combinations thereof. In another

5 specific embodiment, the invention provides a method for producing a fatty alcohol using said microbial organism comprising (ii) further comprises an exogenous nucleic acid encoding an enzyme selected from an aconitase, an isocitrate dehydrogenase, a succinyl-CoA synthetase, a succinyl-CoA transferase, a fumarase, a malate dehydrogenase, and

combinations thereof. In another specific embodiment, the invention provides a method for

10 producing a fatty alcohol using said microbial organism comprises two, three, four, five, or six exogenous nucleic acids each encoding a fatty alcohol pathway enzyme. In another

specific embodiment, the invention provides a method for producing a fatty alcohol using said microbial organism comprising (i) comprises two, three or four exogenous nucleic acids each encoding a reductive TCA pathway enzyme. In another specific embodiment, said

15 microbial organism comprising (ii) comprises two, three or four exogenous nucleic acids each encoding a reductive TCA pathway enzyme. In another specific embodiment, said at least one exogenous nucleic acid is a heterologous nucleic acid. In a specific embodiment, said non-naturally occurring microbial organism is in a substantially anaerobic culture medium.

20 The primary alcohol, a fatty acyl-CoA, a fatty ester, a wax, an acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol 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,

25 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.

In some embodiments, culture conditions include anaerobic or substantially anaerobic growth or maintenance conditions. Exemplary anaerobic conditions have been described previously

30 and 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. Exemplary anaerobic conditions for fermentation processes are described

below and are well known in the art. 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.

In addition to the culturing and fermentation conditions disclosed herein, growth condition
5 for achieving biosynthesis of primary alcohol, a fatty acyl-CoA, a fatty ester, a wax, an acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol 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
10 compound that acts as an osmolyte and helps a microbial organism as described herein survive osmotic stress. Osmoprotectants include, but are not limited to, betaines, amino acids, and the sugar trehalose. Non-limiting examples of such are glycine betaine, praline betaine, dimethylthetin, dimethylsulfoniopropionate, 3-dimethylsulfonio-2-methylpropionate, pipercolic acid, dimethylsulfonioacetate, choline, L-carnitine and ectoine.

15 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 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
20 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 than about 7.0 mM, no more than about 10mM, no more than about 50mM, no more than about 100mM or no more than about 500mM.

The culture conditions can include, for example, liquid culture procedures as well as fermentation and other large scale culture procedures. As described further below in the
25 Examples, particularly useful yields of the biosynthetic products of the invention can be obtained under anaerobic or substantially anaerobic culture conditions. 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.

30 Fermentation procedures are particularly useful for the biosynthetic production of commercial quantities of primary alcohols. Generally, and as with non-continuous culture procedures, the continuous and/or near-continuous production of primary alcohols will

include culturing a non-naturally occurring primary alcohol 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, 1 day, 2, 3, 4, 5, 6 or 7 days or more. Additionally, continuous culture can include 1 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.

Fermentation procedures are well known in the art. Briefly, fermentation for the biosynthetic production of primary alcohol products of the invention 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 well known in the art are exemplified further below in the Examples.

In a further embodiment, the primary alcohol producing microbial organisms of the invention utilize renewable feedstocks and carbon-containing gas as carbon sources for growth.

Employing these alternative materials as a feedstock is particularly useful because they are beneficial from an environmental standpoint and lower production costs of bioprocess-derived products such as the primary alcohols of the invention.

Renewable feedstocks useful for growth of the primary alcohol producing organisms of the invention, including fermentation processes with the modified organisms of the invention, can include any regenerative raw material which can be used by the cell as a supply a carbon or other energy source. In general, renewable feedstock are derived from living organisms or their metabolic byproducts including material derived from biomass, often consisting of underutilized components like chaff. Agricultural products specifically grown for use as renewable feedstocks and useful in the methods of the invention include, for example, corn, soybeans and cotton; flaxseed and rapeseed; sugar cane and palm oil. Renewable feedstocks that can be used therefore include an array of carbohydrates, fats and proteins derived from agricultural and/or animal matter which can be harnessed by the primary alcohol producing organisms of the invention as a source for carbon.

Plant-derived biomass which is available as an energy source on a sustainable basis includes, for example, herbaceous and woody energy crops, agricultural food and feed crops,

agricultural crop wastes and residues, wood wastes and residues, aquatic plants, and other waste materials including some municipal wastes (see, for example, the URL 1.eere.energy.gov/biomass/information_resources.html, which includes a database describing more than 150 exemplary kinds of biomass sources). Exemplary types of biomasses that can be used as feedstocks in the methods of the invention include cellulosic biomass, hemicellulosic biomass and lignin feedstocks or portions of feedstocks. Such biomass feedstocks contain, for example, carbohydrate substrates useful as carbon sources such as glucose, xylose, arabinose, galactose, mannose, fructose and starch. Given the teachings and guidance provided herein, those skilled in the art will understand that renewable feedstocks and biomass other than those exemplified above also can be used for culturing the microbial organisms of the invention for the production of a wide variety of primary alcohols.

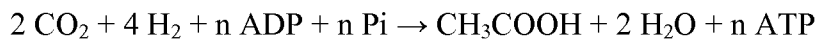
In addition to renewable feedstocks such as those exemplified above, the primary alcohol producing microbial organisms of the invention also can be modified for growth on syngas as its source of carbon. In this specific embodiment, one or more proteins or enzymes are expressed in the primary alcohol producing organisms to provide a metabolic pathway for utilization of syngas or other gaseous carbon source.

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 H_2 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 H_2 and CO , syngas can also include CO_2 and other gases in smaller quantities. Thus, synthesis gas provides a cost effective source of gaseous carbon such as CO and, additionally, CO_2 .

The Wood-Ljungdahl pathway catalyzes the conversion of CO and H_2 to acetyl-CoA and other products such as acetate. Organisms capable of utilizing CO and syngas also generally have the capability of utilizing CO_2 and CO_2/H_2 mixtures through the same basic set of enzymes and transformations encompassed by the Wood-Ljungdahl pathway. H_2 -dependent conversion of CO_2 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 CO_2 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
5 can utilize CO₂ and H₂ mixtures as well for the production of acetyl-CoA and other desired products.

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
10 branch converts methyl-THF to acetyl-CoA. The reactions in the methyl branch are catalyzed in order by the following enzymes: ferredoxin oxidoreductase, formate dehydrogenase, formyltetrahydrofolate synthetase, methenyltetrahydrofolate cyclodehydratase, methylenetetrahydrofolate dehydrogenase and methylenetetrahydrofolate reductase. The reactions in the carbonyl branch are catalyzed in order by the following
15 enzymes: cobalamide corrinoid/iron-sulfur protein, methyltransferase, carbon monoxide dehydrogenase, acetyl-CoA synthase, acetyl-CoA synthase disulfide reductase and hydrogenase. Following the teachings and guidance provided above for introducing a sufficient number of encoding nucleic acids to complete the either or both the malonyl-CoA independent FAS and/or the acyl-reduction pathway, those skilled in the art will understand
20 that the same engineering design also can be performed with respect to introducing at least the nucleic acids encoding the Wood-Ljungdahl enzymes absent 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 the complete Wood-Ljungdahl pathway will confer syngas utilization ability.

25 The invention is also directed, in part, to the design and creation of cells and organisms having growth-coupled production of LCA. In one embodiment, the invention utilizes optimization-based approaches based on *in silico* stoichiometric model of *Escherichia coli* metabolism that identify metabolic designs for optimal production of LCA. A bilevel programming framework, OptKnock, is applied within an iterative algorithm to predict
30 multiple sets of gene disruptions, that collectively result in the growth-coupled production of LCA. The results described herein indicate that combinations of strategically placed gene deletions or functional disruptions of genes significantly improve the LCA production

capabilities of *Escherichia coli* and other cells or organisms. The strain design strategies are equally applicable if an organism other than *E. coli* is chosen as the production host, even if the organism naturally lacks the activity or exhibits low activity of a subset of the gene products marked for disruption. In those cases, disruptions must only be introduced to
5 eliminate or lessen the enzymatic activities of the gene products that are naturally present in the chosen production host. Growth-coupled production of LCA for the *in silico* designs are confirmed by construction of strains having the designed metabolic genotype. These metabolically engineered cells or organisms also can be subjected to adaptive evolution to further augment growth-coupled product production.

10 The invention is also directed, in part, to the design and creation of cells and organisms that produce long chain alcohols, LCAs based on *in silico* stoichiometric model of *Saccharomyces cerevisiae* metabolism. One skilled in the art will recognize the ability to also produce LCAs by non-growth-coupled production by providing a non-producing growth phase, followed by a non-growth production phase, for example. The results described herein
15 indicate that combinations of gene deletions or functional disruptions of genes significantly improve the LCA production capabilities of *Saccharomyces cerevisiae* and other cells of eukaryotic organisms and eukaryotic microbial organisms. The strain design pathways are equally applicable if a eukaryotic microbial organism other than *S. cerevisiae* is chosen as the production host, even if the organism naturally lacks the activity or exhibits low activity of a
20 subset of the gene products marked for disruption. In the latter case, disruptions can be introduced to eliminate or lessen the enzymatic activities of the gene products that are naturally present in the chosen production host. In some embodiments, growth-coupled production of LCA for the *in silico* determined metabolic pathways is confirmed by construction of strains having the designed metabolic genotype. These metabolically
25 engineered cells or organisms can also be subjected to adaptive evolution to further augment growth-coupled product production. In some embodiments, the engineered cells or organisms can also incorporate additional copies of beneficial genes to increase flux through a particular metabolic pathway. Alternatively, exogenous gene insertions from another organism can be used to install functionality that is not present in the host organism.

30 In some embodiments, the designed LCA production pathway utilizes a malonyl-CoA-independent fatty acid synthesis pathway coupled with reduction of the fatty acid to form primary alcohol as shown in Figure 1. The malonyl-CoA independent LCA production

pathway (MI-LCA pathway) comprises the malonyl-CoA-independent fatty acid synthesis steps and the acyl-CoA reduction steps. An engineered microorganism possessing the MI-LCA pathway will convert low cost renewable feedstocks, such as glucose and sucrose, to acetyl-CoA through glycolysis. Acetyl-CoA then is used as both primer and extension units

5 in an elongation cycle that involves the ketoacyl-CoA thiolase, 3-hydroxyacyl-CoA dehydrogenase, enoyl-CoA hydratase, and enoyl-CoA reductase. At the end of each elongation cycle, an acyl-CoA is formed that is one C₂ unit longer than the acyl-CoA entering the elongation cycle. The acyl-CoA with a desired chain-length is then reduced through the combination of acyl-CoA reductase and alcohol dehydrogenase or the fatty alcohol forming

10 acyl-CoA reductase to form the desired primary alcohol. The carbon chain-length of the LCA can be controlled by chain-length specific enoyl-CoA reductase, ketoacyl-CoA thiolase, and/or acyl-CoA reductase.

The MI-LCA pathway has the advantage of better product and ATP yields than that through the typical energy-intensive fatty acid synthesis pathways for LCA production. For example,

15 the maximum theoretical yield for dodecanol (C₁₂) using the MI-LCA pathway is 0.333 mol per mol of glucose consumed under both aerobic and anaerobic conditions:



Additionally, the energy and redox characteristics of the MI-LCA pathway make it suited for

20 the creation of strains that couple LCA production to growth using OptKnock algorithms (Burgard, A.P., P. Pharkya, and C.D. Maranas, *Optknock: a bilevel programming framework for identifying gene knockout strategies for microbial strain optimization*. Biotechnol Bioeng, 2003. **84**(6): p. 647-57; Pharkya, P., A.P. Burgard, and C.D. Maranas, *Exploring the overproduction of amino acids using the bilevel optimization framework OptKnock*.

25 Biotechnol Bioeng, 2003. **84**(7): p. 887-99; Pharkya, P., A.P. Burgard, and C.D. Maranas, *OptStrain: a computational framework for redesign of microbial production systems*. Genome Res, 2004. **14**(11): p. 2367-76.). The resulting growth-coupled production strains will be inherently stable, self-optimizing, and suited for batch, fed-batch, and continuous process designs.

30 In some embodiments, the invention is directed to an integrated computational and engineering platform for developing metabolically altered microorganism strains having enhanced LCA producing characteristics. Strains identified via the computational component

of the platform are put into actual production by genetically engineering the predicted metabolic alterations which lead to the enhanced production of LCA. Production of the desired product is coupled to optimal growth of the microorganism to optimize yields of this product during fermentation. In yet another embodiment, strains exhibiting growth-coupled production of LCA are further subjected to adaptive evolution to further augment product biosynthesis. The levels of growth-coupled product production following adaptive evolution also can be predicted by the computational component of the system where, in this specific embodiment, the elevated product levels are realized only following evolution.

In some embodiments, the invention provides a non-naturally occurring microbial organism, that includes one or more gene disruptions. The disruptions occur in genes encoding an enzyme that couples LCA production to growth of the organism when the gene disruption reduces the activity of the enzyme, such that the gene disruptions confer stable growth-coupled production of LCA onto the non-naturally occurring organism.

In particular embodiments, the invention provides a non-naturally occurring eukaryotic organism, that includes one or more gene disruptions. The one or more gene disruptions occur in genes that encode enzymes that include, for example a cytosolic pyruvate decarboxylase, a mitochondrial pyruvate dehydrogenase, a cytosolic ethanol-specific alcohol dehydrogenase or a mitochondrial ethanol-specific alcohol dehydrogenase. These gene disruptions confer production of long chain alcohols in the cytosol or mitochondrion (*vide infra*) of the organism.

Further, the present invention provides methods of producing such non-naturally microbial organisms having stable growth-coupled production of LCA. For LCA production, for example, the method includes: (a) identifying *in silico* a set of metabolic modifications requiring LCA production during cell growth, and (b) genetically modifying a microorganism to contain the set of metabolic modifications requiring LCA production.

One consideration for bioprocessing is whether to use a batch or continuous fermentation scheme. One difference between the two schemes that will influence the amount of product produced is the presence of a preparation, lag, and stationary phase for the batch scheme in addition to the exponential growth phase. In contrast, continuous processes are kept in a state of constant exponential growth and, if properly operated, can run for many months at a time. For growth-associated and mixed-growth-associated product formation, continuous processes

provide much higher productivities (i.e., dilution rate times cell mass) due to the elimination of the preparation, lag, and stationary phases. For example, given the following reasonable assumptions:

Monod kinetics (i.e., $\mu = \mu_m \cdot S / (K_s + S)$)

$\mu_m = 1.0 \text{ hr}^{-1}$

final cell concentration/initial cell concentration = 20

$t_{prep} + t_{lag} + t_{stat} = 5 \text{ hr}$

feed concentration of limiting nutrient $\gg K_s$

increased productivity from a continuous process has been estimated at 8-fold, Shuler et al,

Prentice Hall, Inc.: Upper Saddle River, NJ., 245-247.

Despite advantages in productivity, many more batch processes are in operation than continuous processes for a number of reasons. First, for non-growth associated product formation (e.g., penicillin), the productivity of a batch system may significantly exceed that of a continuous process because the latter would have to operate at very low dilution rates.

Next, production strains generally have undergone modifications to their genetic material to improve their biochemical or protein production capabilities. These specialized strains are likely to grow less rapidly than their parental complements whereas continuous processes such as those employing chemostats (fermenters operated in continuous mode) impose large selection pressures for the fastest growing cells. Cells containing recombinant DNA or

carrying point mutations leading to the desired overproduction phenotype are susceptible to back-mutation into the original less productive parental strain. It also is possible for strains having single gene deletions to develop compensatory mutations that will tend to restore the wild-type growth phenotype. The faster growing cells usually out-compete their more productive counterparts for limiting nutrients, drastically reducing productivity. Batch

processes, on the other hand, limit the number of generations available by not reusing cells at the end of each cycle, thus decreasing the probability of the production strain reverting back to its wild-type phenotype. Finally, continuous processes are more difficult to operate long-term due to potential engineering obstacles such as equipment failure and foreign organism contamination. The consequences of such failures also are much more considerable for a

continuous process than with a batch culture.

For small-volume production of specialty chemicals and/or proteins, the productivity increases of continuous processes rarely outweigh the risks associated with strain stability and reliability. However, for the production of large-volume, growth-associated products

such as LCA, the increases in productivity for a continuous process can result in significant economic gains when compared to a batch process. Although the engineering obstacles associated with continuous bioprocess operation would always be present, the strain stability concerns can be overcome through metabolic engineering strategies that reroute metabolic pathways to reduce or avoid negative selective pressures and favor production of the target product during the exponential growth phase.

One computational method for identifying and designing metabolic alterations favoring growth-coupled production of a product is the OptKnock computational framework, Burgard et al., *Biotechnol Bioeng*, 84: 647-57 (2003). OptKnock is a metabolic modeling and simulation program that suggests gene disruption strategies that result in genetically stable microorganisms which overproduce the target product. Specifically, the framework examines the complete metabolic and/or biochemical network of a microorganism in order to suggest genetic manipulations that force the desired biochemical to become a 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 performance as a result of the compulsory growth-coupled biochemical production.

The concept of growth-coupled biochemical production can be visualized in the context of the biochemical production envelopes of a typical metabolic network calculated using an *in silico* model. These limits are obtained by fixing the uptake rate(s) of the limiting substrate(s) to their experimentally measured value(s) and calculating the maximum and minimum rates of biochemical production at each attainable level of growth. Although exceptions exist, typically the production of a desired biochemical is in direct competition with biomass formation for intracellular resources. Thus, enhanced rates of biochemical production will necessarily result in sub-maximal growth rates. The disruptions suggested by OptKnock are designed to restrict the allowable solution boundaries forcing a change in metabolic behavior from the wild-type strain as depicted in Figure 2. Although the actual solution boundaries for a given strain will expand or contract as the substrate uptake rate(s) increase or decrease, each experimental point should lie within its calculated solution boundary. Plots such as these enable one to visualize how close strains are to their performance limits or, in other words, how much room is available for improvement. The OptKnock framework has already been able to identify promising gene disruption strategies

for biochemical overproduction, (Burgard, A.P., P. Pharkya, and C.D. Maranas, *Biotechnol Bioeng*, 84(6):647-657 (2003); Pharkya, P., A.P. Burgard, and C.D. Maranas, *Biotechnol Bioeng*, 84(7):887-899 (2003)) and establishes a systematic framework that will naturally encompass future improvements in metabolic and regulatory modeling frameworks.

- 5 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.

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
10 methods that incorporate particular constraints into flux balance analysis (FBA) models. These constraints include, for example, qualitative kinetic information, qualitative regulatory 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
15 networks in the presence of gene additions or disruptions. OptKnock computational framework allows the construction of model formulations that enable an effective query of the performance limits of 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. Patent
20 Application Serial No. 10/043,440, filed January 10, 2002, and in International Patent No. PCT/US02/00660, filed January 10, 2002.

Another computational method for identifying and designing metabolic alterations favoring growth-coupled production of a product is metabolic modeling and simulation system termed SimPheny[®]. This computational method and system is described in, for example, U.S. Patent
25 Application Serial No. 10/173,547, 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 the chemical reactions of a biological system to define a solution space that contains any and all possible functionalities
30 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 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. Analysis methods such as convex analysis, linear programming and the calculation of extreme pathways as described, for example, in Schilling et al., *J. Theor. Biol.* 203:229-248 (2000); Schilling et al., *Biotech. Bioeng.* 71:286-306 (2000) and Schilling et al., *Biotech. Prog.* 15:288-295 (1999), can be used to determine such phenotypic capabilities.

As described above, one constraints-based method used in the computational programs applicable to the invention is flux balance analysis. Flux balance analysis is based on flux balancing in a steady state condition and can be performed as described in, for example, Varma and Palsson, *Biotech. Bioeng.* 12:994-998 (1994). Flux balance approaches have been applied to reaction networks to simulate or predict systemic properties of, for example, adipocyte metabolism as described in Fell and Small, *J. Biochem.* 138:781-786 (1986), acetate secretion from *E. coli* under ATP maximization conditions as described in Majewski and Domach, *Biotech. Bioeng.* 35:732-738 (1990) or ethanol secretion by yeast as described in Vanrolleghem et al., *Biotech. Prog.* 12:434-448 (1996). Additionally, this approach can be used to predict or simulate the growth of *S. cerevisiae* on a variety of single-carbon sources as well as the metabolism of *H. influenzae* as described in Edwards and Palsson, *Proc. Natl. Acad. Sci.* 97:5528-5533 (2000), Edwards and Palsson, *J. Bio. Chem.* 274:17410-17416 (1999) and Edwards et al., *Nature Biotech.* 19:125-130 (2001).

Once the solution space has been defined, it can be analyzed to determine possible solutions under various conditions. This computational approach is consistent with biological realities because biological 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 will be able to apply various computational frameworks for metabolic modeling and simulation to design and implement growth-coupled production of a biochemical product. Such metabolic modeling and simulation methods include, for example, the computational systems

5 exemplified above as SimPheny[®] and OptKnock. For simplicity in illustrating the invention, the methods and strains will be 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
10 well known in the art.

The ability of a cell or organism to couple growth to the production of a biochemical product can be illustrated in the context of the biochemical production limits of a typical metabolic network calculated using an *in silico* model. These limits are obtained by fixing the uptake rate(s) of the limiting substrate(s) to their experimentally measured value(s) and calculating
15 the maximum and minimum rates of biochemical production at each attainable level of growth. As shown in Figure 2, the production of a desired biochemical generally is in direct competition with biomass formation for intracellular resources. Under these circumstances, enhanced rates of biochemical production will necessarily result in sub-maximal growth rates. The disruptions suggested by the above metabolic modeling and simulation programs
20 such as OptKnock are designed to restrict the allowable solution boundaries forcing a change in metabolic behavior from the wild-type strain as depicted in Figure 2. Although the actual solution boundaries for a given strain will expand or contract as the substrate uptake rate(s) increase or decrease, each experimental point will lie within its calculated solution boundary. Plots such as these enable accurate predictions of how close the designed strains are to their
25 performance limits which also indicates how much room is available for improvement.

The OptKnock mathematical framework is exemplified herein for pinpointing gene disruptions leading to growth-coupled biochemical production as illustrated in Figure 2. The procedure builds upon constraint-based metabolic modeling which narrows the range of possible phenotypes that a cellular system can display through the successive imposition of
30 governing physico-chemical constraints, Price et al., *Nat Rev Microbiol*, 2: 886-97 (2004). As described above, constraint-based models and simulations are well known in the art and

generally invoke the optimization of a particular cellular objective, subject to network stoichiometry, to suggest a likely flux distribution.

Briefly, the maximization of a cellular objective quantified as an aggregate reaction flux for a steady state metabolic network comprising a set $N = \{1, \dots, N\}$ of metabolites and a set $M =$

5 $\{1, \dots, M\}$ of metabolic reactions is expressed mathematically as follows:

$$\begin{aligned} &\text{maximize} && v_{\text{cellular objective}} \\ &\text{subject to} && \sum_{j=1}^M S_{ij} v_j = 0, && \forall i \in N \end{aligned}$$

$$\begin{aligned} &v_{\text{substrate}} = v_{\text{substrate_uptake}} \text{ mmol/gDW}\cdot\text{hr} \\ &\forall i \in \{\text{limiting substrate(s)}\} \end{aligned}$$

$$\begin{aligned} &10 \quad v_{\text{atp}} \geq v_{\text{atp_main}} \text{ mmol/gDW}\cdot\text{hr} \\ &\quad v_j \geq 0, && \forall j \in \{\text{irrev. reactions}\} \end{aligned}$$

where S_{ij} is the stoichiometric coefficient of metabolite i in reaction j , v_j is the flux of reaction j , $v_{\text{substrate_uptake}}$ represents the assumed or measured uptake rate(s) of the limiting substrate(s),

15 and $v_{\text{atp_main}}$ is the non-growth associated ATP maintenance requirement. The vector v includes both internal and external fluxes. In this study, the cellular objective is often assumed to be a drain of biosynthetic precursors in the ratios required for biomass formation, Neidhardt, F.C. et al., 2nd ed. 1996, Washington, D.C.: ASM Press. 2 v. (xx, 2822, lxxvi). The fluxes are generally reported per 1 *gDW*·hr (gram of dry weight times hour) such that

20 biomass formation is expressed as *g biomass produced/gDW*·hr or *1/hr*.

The modeling of gene deletions, and thus reaction elimination, first employs the incorporation of binary variables into the constraint-based approach framework, Burgard et al., *Biotechnol Bioeng*, 74: 364-375 (2001), Burgard et al., *Biotechnol Prog*, 17: 791-797 (2001). These binary variables,

$$25 \quad y_j = \begin{cases} 1, & \text{if reaction flux } v_j \text{ is active} \\ 0, & \text{if reaction flux } v_j \text{ is not active} \end{cases}, \forall j \in M$$

assume a value of 1 if reaction j is active and a value of 0 if it is inactive. The following constraint,

$$v_j^{\min} \cdot y_j \leq v_j \leq v_j^{\max} \cdot y_j, \quad \forall j \in M$$

ensures that reaction flux v_j is set to zero only if variable y_j is equal to zero. Alternatively, when y_j is equal to one, v_j is free to assume any value between a lower $v_j^{\min}$ and an upper $v_j^{\max}$ bound. Here, $v_j^{\min}$ and $v_j^{\max}$ are identified by minimizing and maximizing, respectively, every reaction flux subject to the network constraints described above, Mahadevan et al., *Metab Eng*, 5: 264-76 (2003).

Optimal gene/reaction disruptions are identified by solving a bilevel optimization problem that chooses the set of active reactions ($y_j = 1$) such that an optimal growth solution for the resulting network overproduces the chemical of interest. Schematically, this bilevel optimization problem is illustrated in Figure 2. Mathematically, this bilevel optimization problem is expressed as the following bilevel mixed-integer optimization problem:

$$\begin{array}{ll}
 \text{maximize} & v_{\text{chemical}} \\
 y_j & \\
 \text{subject to} & \text{maximize } v_{\text{biomass}} \\
 & \text{subject to} \quad \sum_{j=1}^M S_{ij} v_j = 0, \quad \forall i \in N \\
 & v_{\text{substrate}} = v_{\text{substrate_uptake}} \\
 & \forall i \in \{\text{limiting substrate(s)}\} \\
 & v_{\text{atp}} \geq v_{\text{atp_main}} \\
 & v_{\text{biomass}} \geq v_{\text{biomass}}^{\text{target}}
 \end{array} \quad \text{(OptKnock)}$$

$$\begin{array}{ll}
 v_j^{\min} \cdot y_j \leq v_j \leq v_j^{\max} \cdot y_j, & \forall j \in M \\
 \sum_{j \in M_{\text{forward}}} (1 - y_j) = K \\
 y_j \in \{0, 1\}, & \forall j \in M
 \end{array}$$

where v_{chemical} is the production of the desired target product, for example LCA or other biochemical product, and K is the number of allowable knockouts. Note that setting K equal to zero returns the maximum biomass solution of the complete network, while setting K equal to one identifies the single gene/reaction knockout ($y_j = 0$) such that the resulting network involves the maximum overproduction given its maximum biomass yield. The final

constraint ensures that the resulting network meets a minimum biomass yield. Burgard et al., *Biotechnol Bioeng*, 84: 647-57 (2003), provide a more detailed description of the model formulation and solution procedure. Problems containing hundreds of binary variables can be solved in the order of minutes to hours using CPLEX 8.0, *GAMS: The Solver Manuals*. 2003: 5 GAMS Development Corporation, accessed via the GAMS, Brooke et al., *GAMS Development Corporation* (1998), modeling environment on an IBM RS6000-270 workstation. The OptKnock framework has already been able to identify promising gene disruption strategies for biochemical overproduction, Burgard et al., *Biotechnol Bioeng*, 84: 647-57 (2003), Pharkya et al., *Biotechnol Bioeng*, 84: 887-899 (2003), and establishes a 10 systematic framework that will naturally encompass future improvements in metabolic and regulatory modeling frameworks.

Any solution of the above described bilevel OptKnock problem will provide one set of metabolic reactions to disrupt. Elimination of each reaction within the set or metabolic modification can result in LCA as a product during the growth phase of the organism.

15 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 20 database having a relationship between enzymes and encoding genes.

Once identified, the set of reactions that are to be disrupted in order to achieve growth-coupled LCA production are implemented in the target cell or organism by functional disruption of at least one gene encoding each metabolic reaction within the set. As described previously, one particularly useful means to achieve functional disruption of the reaction set 25 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 30 assessments of the product coupling 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 growth-coupled production of LCA, or other biochemical products, 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 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: $y_1 + y_2 + y_3 \geq 1$. The integer cut method is well known in the art and can be found described in, for example, reference, 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.

Constraints of the above form preclude identification of larger reaction sets that include previously identified sets. For example, employing the integer cut optimization method above in a further iteration would preclude identifying a quadruple reaction set that specified reactions 1, 2, and 3 for disruption since these reactions had been previously identified. To ensure identification of all possible reaction sets leading to growth-coupled production of a product, a modification of the integer cut method was employed.

Briefly, the modified integer cut procedure begins with iteration 'zero' which calculates the maximum production of the desired biochemical at optimal growth for a wild-type network. This calculation corresponds to an OptKnock solution with K equaling 0. Next, single disruptions are considered and the two parameter sets, $objstore_{iter}$ and $y_{store_{iter,j}}$, are introduced to store the objective function ($v_{chemical}$) and reaction on-off information (y_j), respectively, at each iteration, $iter$. The following constraints are then successively added to the OptKnock formulation at each iteration.

$$v_{chemical} \geq objstore_{iter} + \varepsilon - M \cdot \sum_{j \in y_{store_{iter,j}}=0} y_j$$

In the above equation, ε and M are a small and a large numbers, respectively. In general, ε can be set at about 0.01 and M can be set at about 1000. However, numbers smaller and/or larger than these numbers also can be used. M ensures that the constraint can be binding only for previously identified disruption strategies, while ε ensures that adding disruptions to a previously identified strategy must lead to an increase of at least ε in biochemical production at optimal growth. The approach moves onto double disruptions whenever a single disruption strategy fails to improve upon the wild-type strain. Triple disruptions are then considered when no double disruption strategy improves upon the wild-type strain, and so on. The end result is a ranked list, represented as desired biochemical production at optimal growth, of distinct disruption strategies that differ from each other by at least one disruption. This optimization procedure as well as the identification of a wide variety of reaction sets that, when disrupted, lead to the growth-coupled production of a biochemical product are exemplified in detail further below. Given the teachings and guidance provided herein, those skilled in the art will understand that the methods and metabolic engineering designs exemplified herein are applicable to the coupling of cell or microorganism growth to any biochemical product.

Employing the methods exemplified above, the methods of the invention enable the construction of cells and organisms that couple the production of a target biochemical product to growth of the cell or organism engineered to harbor the identified genetic alterations. In this regard, metabolic alterations have been identified that obligatorily couple the production of LCA to organism growth. Microbial organism strains constructed with the identified metabolic alterations produce elevated levels of LCA during the exponential growth phase. These strains can be beneficially used for the commercial production of LCA in continuous fermentation process without being subjected to the negative selective pressures described previously.

Therefore, the methods of the invention provide a set of metabolic modifications that are identified by an *in silico* method selected from OptKnock. The set of metabolic modifications can include functional disruption of one or more metabolic reactions including, for example, disruption by gene deletion. For LCA production metabolic modifications can be selected from the set of metabolic modifications listed in Table 1.

Also provided is a method of producing a non-naturally occurring microbial organism having stable growth-coupled production of LCA. The method includes: (a) identifying *in silico* a

set of metabolic modifications requiring LCA production during exponential growth; (b) genetically modifying an organism to contain the set of metabolic modifications requiring product production, and culturing the genetically modified organism. Culturing can include adaptively evolving the genetically modified organism under conditions requiring product

5 production. The methods of the invention are applicable to bacterium, yeast and fungus as well as a variety of other cells and microorganism. Exemplary bacteria include species selected from *E. coli*, *Anaerobiospirillum succiniciproducens*, *Actinobacillus succinogenes*, *Mannheimia succiniciproducens*, *Rhizobium etli*, *Bacillus subtilis*, *Corynebacterium glutamicum*, *Gluconobacter oxydans*, *Zymomonas mobilis*, *Lactococcus lactis*, *Lactobacillus*

10 *plantarum*, *Streptomyces coelicolor*, *Clostridium acetobutylicum*, *Pseudomonas fluorescens*, and *Pseudomonas putida*. Exemplary eukaryotic organisms include species selected from *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Kluyveromyces lactis*, *Kluyveromyces marxianus*, *Aspergillus terreus*, *Aspergillus niger*, *Rhizopus arrhizus*, *Rhizopus oryzae*, and *Pichia pastoris*.

15 A microbial organism produced by the methods of the invention is further provided. Additionally, the invention provides a non-naturally occurring microbial organism comprising one or more gene disruptions encoding an enzyme associated with growth-coupled production of LCA and exhibiting stable growth-coupled production of these products. The non-naturally occurring microbial organism of the invention includes one or

20 more gene disruptions occurring in genes encoding an enzyme obligatorily coupling LCA production to growth of the microbial organism when the gene disruption reduces an activity of the enzyme, whereby the one or more gene disruptions confers stable growth-coupled production of LCA onto the non-naturally occurring microbial organism.

The non-naturally occurring microbial organism can have one or more gene disruptions

25 included in a metabolic modification listed in Table 1. The one or more gene disruptions can be a deletion. The non-naturally occurring microbial organism of the invention can be selected from a group of microbial organism having a metabolic modification listed in Tables 1. Non-naturally occurring microbial organisms of the invention include bacteria, yeast, fungus, or any of a variety of other microorganisms applicable to fermentation processes.

30 Exemplary bacteria include species selected from *E. coli*, *A. succiniciproducens*, *A. succinogenes*, *M. succiniciproducens*, *R. etli*, *Bacillus subtilis*, *C. glutamicum*, *G. oxydans*, *Z. mobilis*, *L. lactis*, *L. plantarum*, *S. coelicolor*, *C. acetobutylicum*, *P. fluorescens*, and *P.*

putida. Exemplary eukaryotic organisms include species selected from *S. cerevisiae*, *S. pombe*, *K. lactis*, *K. marxianus*, *A. terreus*, *A. niger*, *R. arrhizus*, *R. oryzae*, and *P. pastoris*.

The microbial organisms having growth-coupled LCA production are exemplified herein with reference to an *Escherichia coli* genetic background. However, with the complete

5 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 an alternate species homolog for one or more genes, including for example, orthologs, paralogs and nonorthologous gene displacements, and the interchange of genetic alterations between
10 organisms is routine and well known in the art. Accordingly, the metabolic alterations enabling growth-coupled production of LCA described herein with reference to a particular organism such as *Escherichia coli* can be readily applied to other microorganisms. 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.

15 As described previously, homologues can include othologs and/or nonorthologous gene displacements. In some instances, such as when a substitute metabolic pathway exists in the species of interest, functional disruption can be accomplished by, for example, deletion of a paralog that catalyzes a similar, yet non-identical metabolic reaction which replaces the referenced reaction. Because certain differences among metabolic networks between
20 different organisms, those skilled in the art will understand that the actual genes disrupted between different organisms may differ. However, the given the teachings and guidance provided herein, those skilled in the art also will understand that the methods of the invention can be applied to all microorganisms to identify the cognate metabolic alterations between organisms and to construct an organism in a species of interest that will enhance the coupling
25 of LCA biosynthesis to growth.

The invention will be described herein with general reference to the metabolic reaction, reactant or product thereof, or with specific reference to one or more genes 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
30 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 that reference to any of these metabolic constitutes also references the gene or genes encoding the enzymes

that catalyze the referenced reaction, reactant or product. Likewise, given the well known fields of metabolic biochemistry, enzymology and genomics, reference herein to a gene also constitutes a reference to the corresponding encoded enzyme and the reaction it catalyzes as well as the reactants and products of the reaction. As described previously and further below, 5 exemplary reactions, reaction nomenclature, reactants, products, cofactors and genes encoding enzymes catalyzing a reaction involved in the growth-coupled production of LCA are set forth in Tables 2 and 3.

The invention provides non naturally occurring microbial organisms having growth-coupled production of LCA. Product production is obligatorily linked to the exponential growth 10 phase of the microorganism by genetically altering the metabolic pathways of the cell. The genetic alterations make the desired product a product during the growth phase. Sets of metabolic alterations or transformations that result in elevated levels of LCA biosynthesis are exemplified in Table 1, respectively. Each alteration within a set corresponds to the requisite metabolic reaction that should be functionally disrupted. Functional disruption of all 15 reactions within each set results in the production of LCA by the engineered strain during the growth phase. The corresponding reactions to the referenced alterations and the gene or genes that potentially encode them in *Escherichia coli*, are set forth in Table 2. The various metabolites, their abbreviations and location are set forth in Table 3.

For example, for each strain exemplified in Table 1, the metabolic alterations that can be 20 generated for growth coupled LCA production are shown in each row. These alterations include the functional disruption of from one to six or more reactions. In particular, 995 strains are exemplified in Table 1 that have non-naturally occurring metabolic genotypes. Each of these non-naturally occurring alterations result in an enhanced level of LCA production during the exponential growth phase of the microbial organism compared to a 25 wild-type strain, under appropriate culture conditions. Appropriate conditions include, for example, those exemplified further below in the Example I such as particular carbon sources or reactant availabilities and/or adaptive evolution.

Given the teachings and guidance provided herein, those skilled in the art will understand that to disrupt an enzymatic reaction it is necessary to disrupt the catalytic activity of the one or 30 more enzymes involved in the reaction. Disruption can occur by a variety of means including, for example, deletion of an encoding gene or incorporation of a genetic alteration in one or more of the encoding gene sequences. The encoding genes targeted for disruption

can be one, some, or all of the genes encoding enzymes involved in the catalytic activity. For example, where a single enzyme is involved in a targeted catalytic activity disruption can occur by a genetic alteration that reduces or destroys the catalytic activity of the encoded gene product. Similarly, where the single enzyme is multimeric, including heteromeric, disruption can occur by a genetic alteration that reduces or destroys the function of one or all subunits of the encoded gene products. Destruction of activity can be accomplished by loss of the binding activity of one or more subunits in order to form an active complex, by destruction of the catalytic subunit of the multimeric complex or by both. Other functions of multimeric protein association and activity also can be targeted in order to disrupt a metabolic reaction of the invention. Such other functions are well known to those skilled in the art. Further, some or all of the functions of a single polypeptide or multimeric complex can be disrupted according to the invention in order to reduce or abolish the catalytic activity of one or more enzymes involved in a reaction or metabolic modification of the invention. Similarly, some or all of enzymes involved in a reaction or metabolic modification of the invention can be disrupted so long as the targeted reaction is reduced or destroyed.

Given the teachings and guidance provided herein, those skilled in the art also will understand that an enzymatic reaction can be disrupted by reducing or eliminating reactions encoded by a common gene and/or by one or more orthologs of that gene exhibiting similar or substantially the same activity. Reduction of both the common gene and all orthologs can lead to complete abolishment of any catalytic activity of a targeted reaction. However, disruption of either the common gene or one or more orthologs can lead to a reduction in the catalytic activity of the targeted reaction sufficient to promote coupling of growth to product biosynthesis. Exemplified herein are both the common genes encoding catalytic activities for a variety of metabolic modifications as well as their orthologs. Those skilled in the art will understand that disruption of some or all of the genes encoding a enzyme of a targeted metabolic reaction can be practiced in the methods of the invention and incorporated into the non-naturally occurring microbial organisms of the invention in order to achieve the growth-coupled product production.

Herein below are described the designs identified for increasing LCA production in *Escherichia coli*. The OptKnock algorithm identified designs based on a stoichiometric model of *Escherichia coli* metabolism. Assumptions include (i) a glucose uptake rate of 10 mmol/gdw/hr; (ii) anaerobic or microaerobic conditions; and (iii) a minimum non-growth

associated maintenance requirement of 3 mmol/gdw/hr. Dodecanol, a C₁₂ molecule, was chosen as an exemplary long chain alcohol whose production can be coupled to growth following the teachings of this invention. Although glucose was assumed to be the growth substrate, it is understood that the strategies are applicable to any substrate including glucose, sucrose, xylose, arabinose, or glycerol. The complete set of growth-coupled LCA productions designs are listed in Table 1. The enzyme names, their abbreviations, and the corresponding reaction stoichiometries are listed in Table 2. Finally, metabolites names corresponding to the abbreviations in the reaction equations are listed in Table 3. Although the designs were identified using a metabolic model of *E. coli* metabolism, and the gene names listed in Table 2 are specific to *E. coli*, the method of choosing the metabolic engineering strategies and also the designs themselves are applicable to any LCA-producing organism. Thus the designs are essentially lists of enzymatic transformations whose activity must be either eliminated, attenuated, or initially absent from a microorganism to enable the growth coupled production of long chain alcohols.

One criterion for prioritizing the final selection of designs was the growth-coupled yield of dodecanol. To examine this, production cones were constructed for each strategy by first maximizing and, subsequently minimizing the dodecanol yields at different rates of biomass formation (as described in the previous section). If the rightmost boundary of all possible phenotypes of the mutant network is a single point, it implies that there is a unique optimum yield of the product at the maximum biomass formation rate possible in the network. In other cases, the rightmost boundary of the feasible phenotypes is a vertical line, indicating that at the point of maximum biomass the network can make any amount of the dodecanol in the calculated range, including the lowest amount at the bottommost point of the vertical line. Such designs were given a lower priority. A short list of the highest priority OptKnock designs is provided here in Table I which represents a subset of the designs of Table 1.

TABLE I

Design	Enzyme activity	Abbreviation	Other notes
I	Acetaldehyde-CoA dehydrogenase D-lactate dehydrogenase	ADH _{ER} LDH_D	Design I + PFL
II	Acetaldehyde-CoA dehydrogenase D-lactate dehydrogenase Pyruvate formate lyase	ADH _{ER} LDH_D PFLi	
III	Acetaldehyde-CoA dehydrogenase D-lactate dehydrogenase Pyruvate formate lyase Fumarate reductase	ADH _{ER} LDH_D PFLi FRD2	Design II + FRD2
IV	Acetaldehyde-CoA dehydrogenase D-lactate dehydrogenase Pyruvate formate lyase Fumarase	ADH _{ER} LDH_D PFLi FUM	Design II + FUM
V	Acetaldehyde-CoA dehydrogenase D-lactate dehydrogenase Pyruvate formate lyase Malate dehydrogenase	ADH _{ER} LDH_D PFLi MDH	Design II + MDH
VI	Acetaldehyde-CoA dehydrogenase D-lactate dehydrogenase Pyruvate formate lyase Fumarate reductase Glutamate dehydrogenase	ADH _{ER} LDH_D PFLi FRD2 GLUDy	Design III + GLUDy
VII	Acetaldehyde-CoA dehydrogenase D-lactate dehydrogenase Pyruvate formate lyase Fumarase	ADH _{ER} LDH_D PFLi FUM	Design IV + GLUDy
VIII	Glutamate dehydrogenase Acetaldehyde-CoA dehydrogenase D-lactate dehydrogenase Pyruvate formate lyase Malate dehydrogenase Glutamate dehydrogenase	GLUDy ADH _{ER} LDH_D PFLi MDH	Design V + GLUDy
IX	Acetaldehyde-CoA dehydrogenase D-lactate dehydrogenase Pyruvate formate lyase Fumarate reductase NAD(P) transhydrogenase	ADH _{ER} LDH_D PFLi FRD2 THD2	Design III + THD2
X	Acetaldehyde-CoA dehydrogenase D-lactate dehydrogenase Pyruvate formate lyase Fumarase NAD(P) transhydrogenase	ADH _{ER} LDH_D PFLi FUM THD2	Design IV + THD2

Design	Enzyme activity	Abbreviation	Other notes
XI	Acetaldehyde-CoA dehydrogenase D-lactate dehydrogenase Pyruvate formate lyase Malate dehydrogenase NAD(P) transhydrogenase	ADHER LDH_D PFLi MDH THD2	Design V + THD2
XII	Acetaldehyde-CoA dehydrogenase D-lactate dehydrogenase Phosphotransacetylase and/or Acetate kinase	ADHER LDH_D PTAr and/or ACKr	Design I + PTAr and/or ACKr
XIII	Acetaldehyde-CoA dehydrogenase D-lactate dehydrogenase Phosphotransacetylase and/or Acetate kinase Fumarate reductase	ADHER LDH_D PTAr and/or ACKr FRD2	Design XII + FRD2
XIV	Acetaldehyde-CoA dehydrogenase D-lactate dehydrogenase Phosphotransacetylase and/or Acetate kinase Fumarase	ADHER LDH_D PTAr and/or ACKr FUM	Design XII + FUM
XV	Acetaldehyde-CoA dehydrogenase D-lactate dehydrogenase Phosphotransacetylase and/or Acetate kinase Malate dehydrogenase	ADHER LDH_D PTAr and/or ACKr MDH	Design XII + MDH
XVI	Acetaldehyde-CoA dehydrogenase D-lactate dehydrogenase Fumarate reductase	ADHER LDH_D FRD2	Design I + FRD
XVII	Acetaldehyde-CoA dehydrogenase D-lactate dehydrogenase Fumarase	ADHER LDH_D FUM	Design I + FUM
XVIII	Acetaldehyde-CoA dehydrogenase D-lactate dehydrogenase Malate dehydrogenase	ADHER LDH_D MDH	Design I + MDH
XIX	Acetaldehyde-CoA dehydrogenase D-lactate dehydrogenase Fumarate reductase ATP synthase	ADHER LDH_D FRD2 ATPS4r	Design XVI + ATPS4r
XX	Acetaldehyde-CoA dehydrogenase D-lactate dehydrogenase Fumarase ATP synthase	ADHER LDH_D FUM ATPS4r	Design XVII + ATPS4r
XXI	Acetaldehyde-CoA dehydrogenase D-lactate dehydrogenase Fumarate reductase ATP synthase	ADHER LDH_D MDH ATPS4r	Design XVIII + ATPS4r

All growth coupled designs in this document build upon Design I which calls for the joint disruption of acetylaldehyde-CoA dehydrogenase (ADHEr) and lactate dehydrogenase (LDH_D) activities to reduce the formation of ethanol and lactate, respectively. A dodecanol yield of 0.14 mol/mol glucose is predicted to be attained upon achieving a maximum growth rate of 0.20 1/hr (Design I, Figure 3). Design II specifies the removal, attenuation, or absence of ADHEr, LDH_D, and pyruvate formate lyase (PFLi) and is predicted to result in a dodecanol yield of 0.28 mol/mol glucose at maximum growth as shown in Figure 4. A tighter coupling of LCA production to growth is attained by the further disruption of fumarate reductase (FRD2), fumarase (FUM), or malate dehydrogenase (MDH) activity as indicated by the solution boundary of Designs III-V in Figure 4. An even tighter coupling of production to growth is attained by the further disruption of glutamate dehydrogenase (GLUDy) or NADP transhydrogenase (THD2) activity as shown in solution boundary of Designs VI – XI in Figure 4. Designs VI – XI actually require a non-insignificant yield of LCA, specifically, 0.05 mol dodecanol/mol glucose, to enable a minimal amount of cell growth.

Design XII calls for the disruption of phosphotransacetylase (PTAr) and/or acetate kinase (ACKr) activity in addition to ADHEr and LDH_D to prevent or lessen the production of acetate, ethanol, and lactate, respectively. A dodecanol yield of 0.28 mol/mol is required to attain a maximum growth rate of 0.16 1/hr assuming a glucose uptake rate of 10 mmol/gDW/hr as shown in Figure 5. A tighter coupling of LCA production to growth is attained by the further disruption of FRD2, FUM, or MDH as indicated by the solution boundary of Designs XIII - XV. Designs XVI – XVIII specify that the disruption of FRD2, FUM, or MDH activity in addition to ADHEr and LDH_D results in a tighter coupling of dodecanol production to cell growth as compared to Design I as shown in Figure 6. Further disrupting ATP synthase activity in designs XIX – XXI is predicted to result in a dodecanol yield of 0.30 mol/mol at a maximum growth rate of 0.13 1/hr as shown in Figure 6. The disruption of this activity forces the organism to rely on the MI-LCA pathway for energy generation. Accordingly, a minimum dodecanol yield of 0.05 mol/mol is required for any growth to be attained assuming that the organism lacks the activities listed in Designs XIX – XXI.

It is understood that the disruption of certain activities in addition to those listed by Designs I – XXI can lead to even higher production yields as illustrated in the following examples. Design

V_A involves disruption of Acetaldehyde-CoA dehydrogenase (ADHEr), lactate dehydrogenase (LDH_D), malate dehydrogenase (MDH), pyruvate formate lyase (PFLi), L-aspartase (ASPT), pyruvate kinase (PYK), glucose 6-phosphate dehydrogenase (G6PDHy), and dihydroxyacetone phosphotransferase (DHAPT). Upon addition of the MI-LCA pathway, an engineered strain containing disruptions in these activities is predicted to have a growth-coupled dodecanol yield of 0.327 mol/mol glucose at the maximum growth rate of 0.02 1/hr (Figure 7, point A). This corresponds to 98% of the maximum theoretical yield of 0.333 mol dodecanol/mol glucose. The maximum growth rate of such a strain is predicted to be approximately 10% of the wide type strain while a minimum dodecanol yield of 0.09 mol/mol is required for growth (Figure 7, point B). A recombinant strain containing reduced activity of these functionalities can be constructed in a single step or in subsequent steps by, for example, disrupting 2-3 activities each step. For example, one can engineer *E. coli* for growth coupled LCA production by first removing genes encoding ADHEr and LDH_D activities resulting in Design I. Design V is then constructed by further deleting genes responsible for MDH and PFLi activities. Design V_A is then constructed by deleting genes encoding ASPT, PYK, G6PDHy, and DHAPT activities. Finally, note that several activities (i.e., 6-phosphogluconolactonase (PGL), phosphogluconate dehydratase (PGDHY), or 2-dehydro-3-deoxy-phosphogluconate aldolase (EDA)) can replace G6PDHy for disruption and yield the same characteristics as Design V_A.

Design XII_A involves disruption of Acetaldehyde-CoA dehydrogenase (ADHEr), lactate dehydrogenase (LDH_D), acetate kinase (ACKr) and/or phosphotransacetylase (PTAr), glutamate dehydrogenase (NADP) (GLUDy), phosphogluconate dehydrogenase (PGDH), and glucose-6-phosphate isomerase (PGI). Design XII_B involves disruption of Acetaldehyde-CoA dehydrogenase (ADHEr), lactate dehydrogenase (LDH_D), acetate kinase (ACKr) and/or phosphotransacetylase (PTAr), glutamate dehydrogenase (NADP) (GLUDy), phosphogluconate dehydrogenase (PGDH), glucose-6-phosphate isomerase (PGI), and D-glucose transport via PEP:Pyr PTS (GLCpts). Upon addition of the MI-LCA pathway, an engineered strain lacking the activities specified by Design XII_B is predicted to have a growth-coupled dodecanol yield of 0.322 mol/mol glucose at the maximum growth rate of 0.04 1/hr (Figure 8, point A). This corresponds to 97% of the maximum theoretical yield of 0.333 mol dodecanol/mol glucose. The maximum growth rate of such a strain is predicted to be approximately 20% of the wild type strain while a minimum dodecanol yield of 0.05 mol/mol is required for growth (Figure 8, point

B). A recombinant strain containing reduced activity of these functionalities can be constructed in a single step or in subsequent steps by, for example, removing additional activities each step. For example, one can engineer *E. coli* for growth coupled LCA production by first removing genes encoding ADHER and LDH_D activities resulting in Design I. Design XII is then constructed by further deleting genes encoding PTAr and/or ACKr activities. Design XII_A is then constructed by deleting the genes responsible for GLUDy, PGDH, and PGI activities. Finally, Design XII_B is constructed by further deleting a gene essential for GLCpts activity.

Accordingly, the invention also provides a non-naturally occurring microbial organism having a set of metabolic modifications coupling LCA production to growth of the organism, the set of metabolic modifications includes disruption of one or more genes selected from the set of genes encoding proteins that include an acetylaldehyde-CoA dehydrogenase and a lactate dehydrogenase.

The present invention also provides a strain lacking the activities listed for Design I above that further lack at least one of the following activities: pyruvate formate lyase (PFLi), phosphotransacetylase (PTAr), acetate kinase (ACKr), fumarate reductase (FRD2), fumarase (FUM), or malate dehydrogenase (MDH) as exemplified by Designs II, XII, XVI, XVII, and XVIII.

In further embodiments, the invention provides a strain lacking the activities listed for Design II above and further lacks at least one of the following activities: fumarate reductase (FRD2), fumarase (FUM), or malate dehydrogenase (MDH) as exemplified by Designs III, IV, and V.

In still further embodiments, the invention provides strains lacking the activities listed for Designs III, IV, or V, above and further lack glutamate dehydrogenase (GLUDy) activity as exemplified by Designs VI, VII, and VIII.

The invention also provides strains lacking the activities listed for designs III, IV, or V, above and further lack NAD(P) transhydrogenase (THD2) activity as exemplified by Designs IX, X, and XI.

In yet further embodiments, the invention provides a strain lacking the activities listed for Design XII above and further lack at least one of the following activities: fumarate reductase (FRD2),

fumarase (FUM), or malate dehydrogenase (MDH) as exemplified by Designs XIII, XIV, and XV.

Finally, the invention provides strains lacking the activities listed for designs XVI, XVII, and XVIII, above and further lack ATP synthase (ATPS4r) activity as exemplified by Designs XIX, XX, and XXI.

Herein below are described the pathways identified for increasing LCA production in *S. cerevisiae*. The OptKnock algorithm, described herein further below, identified designs based on a stoichiometric model of *Saccharomyces cerevisiae* metabolism. Assumptions include (i) a glucose uptake rate of 10 mmol/gdw/hr; (ii) anaerobic or microaerobic conditions; and (iii) a minimum non-growth associated maintenance requirement of 3 mmol/gdw/hr. Dodecanol, a C₁₂ molecule, was chosen as an exemplary long chain alcohol whose production can be coupled to growth following the teachings of this invention. Although glucose was assumed to be the growth substrate, it is understood that the methods are applicable to any substrate including glucose, sucrose, xylose, arabinose, or glycerol. Although the designs were identified using a metabolic model of *S. cerevisiae* metabolism the method of choosing the metabolic engineering pathways and also the designs themselves are applicable to any LCA-producing eukaryotic organism. Thus, the designs are essentially lists of enzymatic transformations whose activity must be either eliminated, attenuated, or initially absent from a microorganism to enable the production of long chain alcohols.

One criterion for prioritizing the final selection of pathways was the yield of dodecanol. To examine this, production cones were constructed for each set of pathways by first maximizing and, subsequently minimizing the dodecanol yields at different rates of biomass formation. If the rightmost boundary of all possible phenotypes of the mutant network is a single point, it implies that there is a unique optimum yield of the product at the maximum biomass formation rate possible in the network. In other cases, the rightmost boundary of the feasible phenotypes is a vertical line, indicating that at the point of maximum biomass the network can make any amount of the dodecanol in the calculated range, including the lowest amount at the bottommost point of the vertical line. Such designs were given a lower priority.

The organisms of the present invention can be cultured in a substantially anaerobic culture medium or a microaerobic culture medium as detailed herein below further. Such organisms have one or more gene disruptions which may include complete deletion in some embodiments, or disruption by removal or changes in functional portions encoded by fragments of the entire gene.

In some embodiments, the present invention provides non-naturally occurring eukaryotic microbial organisms that produce LCAs in the cytosol. Note that cytosol herein refers to any compartment outside the mitochondrion. In such embodiments, one or more gene disruptions in the eukaryotic organism encoding an enzyme include, for example, a cytosolic pyruvate decarboxylase, a mitochondrial pyruvate dehydrogenase, a cytosolic ethanol-specific alcohol dehydrogenase and a mitochondrial ethanol-specific alcohol dehydrogenase. Exemplary genes encoding these enzymes include, for example, YLR044C, YLR134W, YGR087C, PDC3, YNL071W, YER178W, YBR221C, YGR193C, YFL018C, YBR145W, YGL256W, YOL086C, YMR303, YMR083W, YPL088W, YAL061W, YMR318C, YCR105W, and YDL168W.

Other gene disruptions encoding an enzyme include, for example, a cytosolic malate dehydrogenase, a glycerol-3-phosphate dehydrogenase shuttle, an external NADH dehydrogenase, and an internal mitochondrial NADH dehydrogenase can also be effected. Exemplary genes of the later include, for example, YOL126C, YDL022W, YOL059W, YIL155C, YMR145C, YDL085W, and YML120C.

These organisms can also include an exogenous nucleic acid encoding an enzyme in the cytosol including, for example, an acetyl-CoA synthetase (AMP-forming), an ADP-dependent acetate-CoA ligase, an acylating acetaldehyde dehydrogenase, a pyruvate dehydrogenase, a pyruvate:NADP oxidoreductase, and a pyruvate formate lyase, or their corresponding gene regulatory regions. An exogenous nucleic acid encoding a cytosolic transhydrogenase or its gene regulatory region can also be incorporated. In some embodiments these gene products may be natively expressed in the cytosol, while in other embodiments, they may be overexpressed by, for example, adding copies of the gene from the same source or from other organisms, or by introducing or changing gene regulatory regions. Such gene regulatory regions include, for example, alternate promoters, inducible promoters, variant promoters or enhancers to enhance

gene expression. Functional disruption of negative regulatory elements such as repressors and/or silencers also can be employed to enhance gene expression. Similar modifications can be made to translation regulatory regions to enhance polypeptide synthesis and include, for example, substitution of a ribosome binding site with an optimal or consensus sequence and/or removing secondary structures.

These organisms maximize the availability of acetyl CoA, ATP and reducing equivalents (NADH) for dodecanol production. Acetyl CoA is the primary carbon precursor for the production of LCA via the proposed MI-LCA route. All the reactions enabling the formation of dodecanol via the malonyl-CoA independent pathway are operational in the cytosol. Specifically, ketoacyl-CoA thiolase, 3-hydroxyacyl-CoA dehydrogenase, enoyl-CoA hydratase, and enoyl-CoA reductase function in the appropriate direction to form acyl CoA which is then reduced to fatty aldehyde and dodecanol via acyl CoA reductase and alcohol dehydrogenase.

Introduction of the MI-LCA pathway in the cytosol prevented any flux through the native pyruvate dehydrogenase *in silico*. Under anaerobic conditions and in conditions where glucose concentrations are high in the medium, the capacity of this mitochondrial enzyme is very limited and there is no significant flux through it. However, in some embodiments, this enzyme can be deleted or attenuated to increase LCA production.

In one embodiment, LCA production in the cytosol uses the AMP-forming acetyl-CoA synthetase. Dodecanol production in the cytosol relies on the native cell machinery to provide the precursors needed in LCA production. A majority of the pyruvate flux generated by glycolysis is channeled into the formation of acetyl CoA via the pyruvate dehydrogenase bypass comprised of the pyruvate decarboxylase, the acetaldehyde dehydrogenase and the AMP-forming acetyl-CoA synthetase (Figure 9a). This bypass is reported to have significant flux through it even under aerobic conditions at high concentrations of glucose (Pronk et al., *Yeast* 12:1607-1633 (1996)).

The last step of the bypass that converts acetate into acetyl-CoA is catalyzed by acetyl-CoA synthetase, encoded by the *ACS1* and *ACS2* genes. Since *ACS2* is constitutively expressed on glucose and is present in cytosol among other compartments, in some embodiments the non-naturally occurring eukaryotic organism is engineered to overexpress *ACS2*. In other

embodiments the *ACS2* gene is replaced with a mutant *ACS* from *Salmonella enterica* (Genbank id NP_807785.1) that is not subject to post-translational modification and has higher activity in *S. cerevisiae* as compared to *ACS1* or *ACS2* (Shiba et al., *Metab Eng.* 9:160-168 (2007)).

The AMP-generating acetyl CoA synthetase uses two ATP equivalents for the conversion of each molecule of acetate into acetyl CoA ($\text{CoA} + \text{acetate} + \text{ATP} \rightarrow \text{acetyl-CoA} + \text{PPi} + \text{AMP}$). Under anaerobic conditions, when energy is available only through substrate-level phosphorylation, the production of dodecanol via the AMP-forming acetyl CoA synthetase is not energetically favorable. Therefore, a small amount of oxygen is made available to the cell to fulfill its energetic requirements, simultaneously increasing the conversion of acetate into acetyl CoA.

The production of dodecanol can be improved by disruption of ethanol-specific alcohol dehydrogenases to prevent acetyl-CoA and NADH from being used for ethanol production. Additionally, the production of LCA benefits from preventing NADH from being used in the respiratory electron-transport chain. Thus, disruptions in the internal mitochondrial NADH dehydrogenase, the glycerol-3-phosphate dehydrogenase shuttle (consisting of cytosolic NADH-linked glycerol-3-phosphate dehydrogenase and a membrane-bound glycerol-3-phosphate:ubiquinone oxidoreductase) (Bakker et al., *FEMS Microbiol. Rev.* 25:15-37 (2001)) and the external NADH dehydrogenase are introduced in some embodiments. Further, cytosolic malate dehydrogenase that can potentially draw NADH away from dodecanol production is also disrupted. A growth-coupled production envelope after imposing these disruptions is shown in dark gray in Figure 9b and compared with the dodecanol production characteristics under aerobic conditions.

In some embodiments, the non-naturally occurring eukaryotic organism incorporates an exogenous gene encoding an ADP-forming acetate CoA ligase. In this embodiment, the AMP-forming acetyl CoA synthetase in the cytosol is replaced by the ADP-forming acetate CoA ligase ($\text{CoA} + \text{acetate} + \text{ATP} \rightarrow \text{acetyl-CoA} + \text{Pi} + \text{ADP}$) (Figure 10a). Exogenous genes to introduce acetate CoA ligase include, for example, *acdA* and *acdB* from *Pyrococcus furiosus* (Glasemacher et al., *Eur. J. Biochem.* 244:561-567 (1997)) (Mai and Adams, *J. Bacteriol.* 178: 5897-5903 (1996)). The introduction of this enzyme that uses one equivalent of ATP for formation of each

molecule of acetyl CoA (as opposed to 2 ATP equivalents) allows the production of dodecanol to be energetically neutral. In this embodiment, a small amount of oxygen or other electron acceptor respiration is used to generate energy to support growth. Such small amounts of oxygen are referred to as microaerobic conditions, as described further below. In some embodiments, the ethanol-specific alcohol dehydrogenase is disrupted to prevent ethanol formation. In embodiments incorporating CoA ligase, one or more of the following knockouts can be introduced for LCA production: cytosolic malate dehydrogenase, glycerol-3-phosphate dehydrogenase shuttle, the external NADH dehydrogenase, and the internal mitochondrial NADH dehydrogenase. The growth-coupled production after imposition of these disruptions is shown in Figure 10b in dark gray. The black curve shows the production envelope for the wild-type strain under aerobic conditions and the light gray curve shows the envelope when the network is augmented with acetate-CoA ligase. Note the increase in the maximum theoretical yield of dodecanol after introduction of this enzyme.

In some embodiments, the non-naturally occurring eukaryotic organism incorporates an exogenous gene encoding an acylating acetaldehyde dehydrogenase. Improvement in the energetics of the dodecanol process can be accomplished by using the acylating acetaldehyde dehydrogenase (acetaldehyde + CoA + NAD $\rightarrow$ acetyl-CoA + NADH) for the conversion of acetaldehyde into acetyl CoA (Figure 11a). The benefits of using this enzyme are that (i) no energy is expended for production of acetyl CoA, and (ii) one molecule of NADH is formed for every molecule of acetyl CoA formed. Thus, the reducing equivalents needed for the production of acetyl CoA can also be generated. The introduction of this enzyme allows production of LCA under anaerobic conditions.

Acylating acetaldehyde dehydrogenase has been reported in several bacteria, including *Acetobacterium woodii* (Mai and Adams, *J. Bacteriol.* 178:5897-5903 (1996)), *Clostridium kluyveri* (Seedorf et al., *Proc. Natl. Acad. Sci. U. S. A* 105:2128-2133 (2008); Smith and Kaplan, *Arch. Biochem. Biophys.* 203:663-675(1980)), *Clostridium beijerinckii* (Yan et al., *Appl. Environ. Microbiol* 56:2591-2599 (1990)), and in species of *Pseudomonas* such as strain CF600 (Lei et al., *Biochemistry* 47:6870-6882 (2008); Manjasetty et al., *Acta Crystallogr. D. Biol. Crystallogr.* 57:582-585 (2001)). The Genbank ids of genes are shown in Table 5 below.

Table 5

<i>Ald</i>	YP_001394464.1	<i>Clostridium kluyveri</i>
<i>dmpF</i>	CAA43226.1	<i>Pseudomonas sp. CF600</i>
<i>bphG</i>	BAA03892.1	<i>Pseudomonas sp</i>
hpD	NP_4148 84.1	Escherichia coli K12 MG1655

In some embodiments each of the strains above can be supplemented with additional disruptions. Alternatively, some other enzymes not known to possess significant activity under the growth conditions can become active due to adaptive evolution or random mutagenesis and can also be disrupted.

The anaerobic growth-coupled production of dodecanol (or any LCA) can be accomplished by disrupting ethanol-specific alcohol dehydrogenase activity. The introduction of an acylating acetaldehyde dehydrogenase, with its favorable energetics, prevents or reduces carbon flux through the native acetaldehyde dehydrogenase and the acetyl-CoA synthetase. The production envelope is shown in Figure 11b. The wild-type *S. cerevisiae* (black) network can form only small amounts of dodecanol as a byproduct of growth under anaerobic conditions. When the network is augmented with acylating dehydrogenase, there is an increase in the theoretical maximum yield in the network, but no growth-coupling is observed (dotted light gray curve). However, disruption of ethanol-specific alcohol dehydrogenase from the augmented network shows that dodecanol production is coupled to growth at the maximum feasible biomass in the network (dark gray curve).

In some embodiments, the non-naturally occurring eukaryotic organism uses a cytosolic pyruvate dehydrogenase for dodecanol production. Cytosolic pyruvate dehydrogenase for generating the precursors for the MI-LCA pathway are shown in Figure 12. In such embodiments, (i) pyruvate is directly converted into acetyl CoA in the cytosol without the expenditure of energy, and (ii) more reducing equivalents are available to the cell.

In some embodiments, the non-naturally occurring eukaryotic organism is engineered to retarget the native mitochondrial pyruvate dehydrogenase to the cytosol. In other embodiments, a heterologous cytosolic enzyme is introduced into the organism. The retargeting of an enzyme to

a different compartment can be accomplished by changing the targeting sequence of the protein (van Loon and Young, *EMBO J.* 5:161-165 (1986)). Disruption of the native pyruvate decarboxylase enables a majority of the carbon flux to be introduced into the cytosol for processing by cytosolic pyruvate dehydrogenase. This also allows the production of dodecanol under anaerobic conditions. The growth-coupled production envelope is similar to that depicted in Figure 11b. Note that pyruvate decarboxylase is disrupted instead of alcohol dehydrogenase to achieve growth-coupling in the network.

In some embodiments, the non-naturally occurring eukaryotic organism uses a cytosolic pyruvate:NADP oxidoreductase. Pyruvate: NADP oxidoreductase allows for the production of acetyl CoA and reducing equivalents in the cytosol as shown in Figure 13. The addition of this enzyme allows for the production of acetyl CoA without expending energy that would otherwise have been required by acetyl CoA synthetase. The enzyme has been purified from the mitochondrion of *Euglena gracilis* and is oxygen-sensitive (Inui et al., *Journal of Biochemistry* 96:931-934 (1984); Inui et al., *Archives of Biochemistry and Biophysics* 237:423-429 (1985); Inui et al., *Archives of Biochemistry and Biophysics* 274:434-442 (1989); Inui et al., *Archives of Biochemistry and Biophysics* 280:292-298 (1990)). It is used for generating acetyl CoA from pyruvate, simultaneously producing NADPH. The corresponding gene is *pno* and its Genbank id is: CAC37628.1. It can be targeted to the cytosol by removing the mitochondrial targeting sequence. In some embodiments, a transhydrogenase is also added. This enzyme can be introduced as an exogenous gene from an organism such as *E. coli* to convert the generated NADPH into NADH (Nissen et al., *Yeast* 18:19-32 (2001)).

With its low ATP requirements, the pathway is energetically favorable even under anaerobic conditions. To prevent or reduce the utilization of NADH and pyruvate for ethanol production, pyruvate decarboxylase activity can be disrupted. This leads to a growth-coupled production of dodecanol similar to that shown in Figure 11b.

In some embodiments, a non-naturally occurring eukaryotic organism uses a pyruvate formate lyase. In such embodiments, a heterologous cytosolic pyruvate formate lyase (*pfl*) is used to generate both acetyl CoA and NADH as shown in Figure 14. This enzyme is active typically under anaerobic conditions in organisms such as *E. coli*. The lack of energy requirement for

conversion of pyruvate into acetyl CoA makes the production of dodecanol feasible under anaerobic conditions.

The conversion of pyruvate into acetyl CoA is accompanied by the production of formate. This is metabolized by the native formate dehydrogenase, leading to additional generation of reducing equivalents in stoichiometric quantities. In some embodiments that use this strain design, one or more of the three pyruvate decarboxylases, *PDC1*, *PDC5* and *PDC6*, can be disrupted. The Genbank ids of exemplary genes encoding pyruvate formate lyase are shown in Table 6 below.

Table 6

<i>pflB</i>	NP_415423.1	<i>Escherichia coli</i>
<i>pfl</i>	YP_001588758.	<i>Lactococcus lactis</i>
<i>pfl</i>	YP_001394497.1	<i>Clostridium kluyveri</i>

The disruption of pyruvate decarboxylase along with the introduction of a heterologous pyruvate formate lyase in the network leads to a growth-coupled production of dodecanol. The production curve is similar to what is shown in Figure 11b.

While the non-naturally occurring eukaryotic organisms described above produce LCAs in the cytosol, it is also possible to produce LCAs in the mitochondrion. Exemplary designs for the distribution of the carbon flux towards dodecanol production are detailed herein below.

Organisms that produce LCAs in the mitochondrion include one or more disruptions in genes that encode enzymes such as a cytosolic pyruvate decarboxylase, a cytosolic ethanol-specific alcohol dehydrogenase, and a mitochondrial ethanol-specific alcohol dehydrogenase. Exemplary genes encoding these enzymes include, for example, YLR044C, YLR134W, YGR087C, PDC3, YBR145W, YGL256W, YOL086C, YMR303, YMR083W, YPL088W, YAL061W, YMR318C, YCR105W, and YDL168W.

Other genes disruptions include those encoding an enzyme such as a cytosolic malate dehydrogenase, glycerol-3-phosphate dehydrogenase shuttle, catalyzed by, the external NADH dehydrogenase, and internal NADH dehydrogenase. Exemplary genes of the latter include, for example, YOL126C, YDL022W, YOL059W, YIL155C, YMR145C, YDL085W, and YML120C.

Organisms that produce LCAs in the mitochondrion can also include an exogenous nucleic acid encoding an enzyme such as a pyruvate dehydrogenase, a pyruvate: NADP oxidoreductase, a pyruvate formate lyase, an acylating acetaldehyde dehydrogenase, an acetate CoA ligase, and an AMP-forming acetyl CoA synthetase or their corresponding gene regulatory regions as described above. Additionally, such organisms benefit from enhanced NADH transporting shuttle systems for transport of NADH from the cytosol into the mitochondrion. Other exogenous nucleic acids encoding an enzyme that can be inserted in such organisms include a transhydrogenase, formate dehydrogenase, a pyruvate decarboxylase, and a pyruvate oxidase, all in the mitochondrion, or their corresponding gene regulatory regions.

In one embodiment a mitochondrial pyruvate dehydrogenase is used in the non-naturally occurring eukaryotic organism. This can be the native pyruvate dehydrogenase which produces both acetyl CoA and NADH as shown in Figure 15a. Since, there is no energy requirement for the conversion of pyruvate to acetyl CoA via this route; the production of dodecanol, for example, is energetically favorable even under anaerobic conditions.

The mitochondrial pyruvate dehydrogenase is known to be active in both aerobic and anaerobic conditions in *S. cerevisiae* (Pronk et al., *Yeast* 12:1607-1633 (1996)). In some embodiments the enzyme is overexpressed in its native or a heterologous form. The native enzyme can be overexpressed by using a stronger promoter. Additionally, mutations can be introduced aimed at increasing its activity under anaerobic conditions (Kim et al., *J. Bacteriol.* 190:3851-3858 (2008)). Reducing equivalents generated in the cytosol are made available in the mitochondrion for dodecanol production by using the redox shuttles present in *S. cerevisiae*. Note that these shuttles transport NADH into the mitochondrion for energy generation under respiratory conditions (Overkamp et al., *J. Bacteriol.* 182:2823-2830 (2000)). For growth-coupled production, pyruvate decarboxylase activity is disrupted to allow for pyruvate flux to be directed towards pyruvate dehydrogenase and to inhibit ethanol formation. The production curve for the mutant network is shown in Figure 15b.

In some embodiments, a non-naturally occurring eukaryotic organism uses a heterologous pyruvate:NADP-oxidoreductase. The production of dodecanol in the mitochondrion can be achieved by introduction of the pyruvate:NADP oxidoreductase in the mitochondrion as shown

in Figure 16. This enzyme is purified from *E. gracilis*. Since the enzyme is naturally present in mitochondrion and is active under anaerobic conditions, it is possible to get high activity of the enzyme under anaerobic conditions. The introduction of this enzyme provides the precursor acetyl CoA for dodecanol production and also reducing equivalents. The NADPH generated by the enzyme is converted into NADH by a transhydrogenase, which can be introduced into the mitochondrion. For additional reducing equivalents, the redox shuttles need to transport NADH from the cytosol to the mitochondrion. The growth-coupled production of LCA using this enzyme can be obtained by disruption of pyruvate decarboxylase. The production curve of the mutant strain is very similar to the one shown in Figure 15b.

In some embodiments, a non-naturally occurring eukaryotic organism uses a heterologous pyruvate formate lyase. The production of dodecanol using a pyruvate formate lyase in mitochondrion is shown in Figure 17. These genes have been outlined herein above. In such embodiments, the native formate dehydrogenase is retargeted to the mitochondrion to allow for further metabolizing formate and generating more reducing equivalents. This strain can be adopted to carry sufficient flux to sustain high yield and productivity of LCA production in the mitochondrion in the absence of oxygen.

Anaerobic growth conditions are feasible for the production of dodecanol using this strain design. Redox shuttles can be overexpressed to transport NADH generated in the cytosol to the mitochondrion. Production in this scenario is possible by disrupting the cytosolic pyruvate decarboxylase activity. The production characteristics of the mutant strain are similar to that shown in Figure 15b.

In some embodiments, a non-naturally occurring eukaryotic organism uses a heterologous acetaldehyde dehydrogenase (acylating). In such embodiments, an acylating acetaldehyde dehydrogenase is introduced into the mitochondrion to provide both acetyl-CoA and NADH for LCA production as shown in Figure 18. A pyruvate decarboxylase isozyme is retargeted to the mitochondrion to convert pyruvate into acetaldehyde in some embodiments. The expression of these two activities in the mitochondrion is equivalent to the activity of pyruvate dehydrogenase. The growth-coupled production curve is the same as that shown in Figure 15b. The growth-coupled production strain has the native mitochondrial acetaldehyde dehydrogenase (Pronk et al.,

Yeast 12:1607-1633 (1996)) and the cytosolic pyruvate decarboxylase disrupted in some embodiments. In other embodiments, the mitochondrial ethanol-specific alcohol dehydrogenase is also disrupted to prevent the conversion of acetaldehyde into ethanol.

In some embodiments, a non-naturally occurring eukaryotic organism uses a mitochondrial acetyl CoA synthetase (AMP-forming). As discussed above, the expression of this enzyme requires oxygen for favorable energetics. ACS1, an isozyme of acetyl CoA synthetase is expressed in *S. cerevisiae* in the mitochondrion under aerobic conditions but is repressed by glucose. This enzyme can be mutated to eliminate the repression or a heterologous enzyme that is expressed under the conditions of interest can be introduced. Additionally, pyruvate decarboxylase also can be expressed in the mitochondrion to form acetate. *S. cerevisiae*, for example, already possesses a mitochondrial acetaldehyde dehydrogenase (Pronk et al., *Yeast* 12:1607-1633 (1996)). Alternatively, enzymes such as pyruvate oxidase can be heterologously expressed to convert pyruvate into acetate. One such enzyme candidate is pyruvate oxidase from *E. coli* (Genbank id: NP_451392.1). This enzyme is naturally expressed in the presence of oxygen.

The production of LCA using this strain design benefits from one or more of the following disrupted enzymes: cytosolic malate dehydrogenase, the glycerol-3-phosphate dehydrogenase shuttle, the external NADH dehydrogenase, and the internal mitochondrial NADH dehydrogenase. The glycerol-3-phosphate shuttle is comprised of the cytosolic glycerol-3-phosphate dehydrogenase and the membrane-bound glycerol-3-phosphate:ubiquinone oxidoreductase, with the latter also functioning as the mitochondrial glycerol-3-phosphate dehydrogenase. In some embodiments, the mitochondrial ethanol-specific alcohol dehydrogenase is also disrupted to prevent or reduce the conversion of acetaldehyde into ethanol. The production curve for the wild type strain with a mitochondrial pyruvate decarboxylase added to the network is shown in black in Figure 19b. This curve is shown for aerobic conditions. The production characteristics when the aforementioned disruptions are imposed on the network are shown in light gray. The downregulation of the oxidative part of the pentose phosphate pathway, especially the committing step, glucose-6-phosphate dehydrogenase, further improves the LCA production characteristics of the network.

In some embodiments, a non-naturally occurring eukaryotic organism uses a mitochondrial acetate CoA ligase (ADP-forming). Mitochondrial LCA production can also be accomplished using an acetate-CoA ligase to convert acetate into acetyl-CoA as shown in Figure 20. As described above, the use of this enzyme is energetically favorable and LCA production is energetically neutral unless oxygen is supplied. The mitochondrial expression of pyruvate decarboxylase is used in such embodiments. LCA production is obtained by imposing disruptions in cytosolic malate dehydrogenase, the glycerol-3-phosphate dehydrogenase shuttle, the external NADH dehydrogenase, and the internal NADH dehydrogenase. The down-regulation of the oxidative part of the pentose phosphate pathway further improves the growth-coupled production characteristics to yield a production curve similar to the one shown in Figure 19b. In some embodiments, the mitochondrial ethanol-specific alcohol dehydrogenase is also disrupted to prevent or reduce the conversion of acetaldehyde into ethanol.

The design strategies described herein are useful not only for enhancing growth-coupled production, but they are also well-suited for enhancing non-growth coupled production because they link the production of long chain alcohols to energy generation and/or redox balance. Exemplary non-growth coupled production methods include implementing an aerobic growth phase followed by an anaerobic production phase. For example, Vemuri *et al. J. Ind. Microbiol. Biotechnol.* (6):325-332, (2002) describe a dual-phase process for the production of succinate in *E. Coli*. Okino *et al. Appl. Microbiol. Biotechnol.* Sep 6. (2008) [Currently available in online edition]. describe a similar non-growth couple production process in a strain of *Corynebacterium glutamicum* strain.

Another such method involves withholding an essential nutrient from a propagated cell culture, thereby limiting growth, but not precluding production as described in Durner *et al. Appl. Environ. Microbiol.* (8):3408-3414(2000). Yet another strategy aimed at decoupling growth from production involves replacing the growth substrate with another compound that is more slowly metabolizable as described in Altamirano *et al. Biotechnol. Bioeng.* 76:351-360 (2001). Growth decoupled-product formation can also be brought about by specific genetic modifications as described in Blombach *et al. Appl. Microbiol. Biotechnol.* 79:471-9 (2008).

Some microbial organisms capable of LCA production are exemplified herein with reference to an *Saccharomyces cerevisiae* genetic background. However, with the complete genome sequence available now for 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 an alternate species homolog for one or more genes, including for example, orthologs, paralogs and nonorthologous gene displacements, and the interchange of genetic alterations between eukaryotic organisms is routine and well known in the art. Accordingly, the metabolic alterations enabling production of LCA described herein with reference to a particular organism such as *Saccharomyces cerevisiae* can be readily applied to other microorganisms. 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.

The methods of the invention are applicable to various eukaryotic organisms such as yeast and fungus. The yeast can include *S. cerevisiae* and *Rhizopus arrhizus*, for example. Exemplary eukaryotic species include those selected from *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Kluyveromyces lactis*, *Kluyveromyces marxianus*, *Aspergillus terreus*, *Aspergillus niger*, *Rhizopus arrhizus*, *Rhizopus oryzae*, *Candida albicans*, *Candida boidinii* and *Pichia pastoris*. Additionally, select cells from larger eukaryotic organisms are also applicable to methods of the present invention.

Genes can be inserted into *S. cerevisiae*, using several methods; some of these are plasmid-based whereas others allow for the incorporation of the gene in the chromosome. The latter approach employs an integrative promoter based expression vector, for example, the pGAPZ or the pGAPZ α vector based on the GAP promoter. The expression vector constitutes the GAP promoter, the HIS4 wild-type allele for integration and the 3' AOX transcription termination region of *P. pastoris* in addition to a KanMX cassette, flanked by loxP sites enabling removal and recycling of the resistance marker. The vectors are commercially available from Invitrogen. The details of which are elaborated in the Example below.

The engineered strains are characterized by measuring the growth rate, the substrate uptake rate, and the product/byproduct secretion rate. Cultures are grown overnight and used as inoculum for

a fresh batch culture for which measurements are taken during exponential growth. The growth rate is determined by measuring optical density using a spectrophotometer (A600).

Concentrations of glucose, alcohols, and other organic acid byproducts in the culture supernatant will be determined by analytical methods including HPLC using an HPX-87H column (BioRad), or GC-MS, and used to calculate uptake and secretion rates. All experiments are performed with triplicate cultures.

The invention also provides a method for producing long chain alcohols by culturing the non-naturally occurring eukaryotic organism described herein above. The one or more gene disruptions occur in genes encoding an enzyme to coupling long chain alcohol production to growth of the organism when the gene disruption reduces an activity of the enzyme. The one or more gene disruptions confers stable growth-coupled production of long chain alcohols onto the organism. In alternate embodiments the gene disruptions can enhance LCA production in a non-growth dependent manner.

Each of the strains presented herein may be supplemented with additional disruptions if it is determined that the predicted strain designs do not sufficiently couple the formation of LCAs with biomass formation. However, the list of gene disruption sets provided here serves as an excellent starting point for the construction of high-yielding growth-coupled LCA production strains.

Each of the proposed strains can be supplemented with additional disruptions if it is determined that the predicted strain designs do not sufficiently couple the formation of the product with biomass formation. Alternatively, some other enzymes not known to possess significant activity under the growth conditions can become active due to adaptive evolution or random mutagenesis and can also be disrupted. However, the list of gene disruption sets provided here serves as a starting point for construction of high-yielding growth-coupled LCA production strains.

The non-naturally occurring microbial organisms of the invention can be employed in the growth-coupled production of LCA. Essentially any quantity, including commercial quantities, can be synthesized using the growth-coupled LCA producers of the invention. Because the organisms of the invention obligatorily couple LCA to continuous growth or near-continuous growth processes are particularly useful for biosynthetic production of LCA. Such continuous

and/or near continuous growth processes are described above and exemplified below in the Example I. Continuous and/or near-continuous microorganism growth processes also are well known in the art. Briefly, continuous and/or near-continuous growth processes involve maintaining the microorganism in an exponential growth or logarithmic phase. Procedures include using apparatuses such as the EvolugatorTM evolution machine (Evolugate LLC, Gainesville, FL), fermentors and the like. Additionally, shake flask fermentation and grown under microaerobic conditions also can be employed. Given the teachings and guidance provided herein those skilled in the art will understand that the growth-coupled LCA producing microorganisms can be employed in a variety of different settings under a variety of different conditions using a variety of different processes and/or apparatuses well known in the art.

Generally, the continuous and/or near-continuous production of LCA will include culturing a non-naturally occurring growth-coupled LCA 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 be grown, for example, for a day, 2, 3, 4, 5, 6 or 7 days or more. Additionally, continuous cultures can include time durations of 1 week, 2, 3, 4 or 5 or more weeks and up to several months. 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. In particular embodiments, culturing is conducted in a substantially anaerobic culture medium.

LCA can be harvested or isolated at any time point during the continuous and/or near-continuous culture period exemplified above. As exemplified below, the longer the microorganisms are maintained in a continuous and/or near-continuous growth phase, the proportionally greater amount of LCA can be produced.

Therefore, the invention provides a method for producing LCA that includes culturing a non-naturally occurring microbial organism that includes one or more gene disruptions. The disruptions can occur in genes encoding an enzyme to coupling LCA production to growth of the microorganism when the gene disruption reduces an activity of the enzyme, such that the disruptions confer stable growth-coupled production of LCA onto the non-naturally microbial organism.

In some embodiments, the gene disruption can include a complete gene deletion. In some embodiments other means to disrupt a gene include, for example, frameshifting by omission or addition of oligonucleotides or by mutations that render the gene inoperable. One skilled in the art will recognize the advantages of gene deletions, however, because of the stability it may confer to the non-naturally occurring organism from reverting to its wild-type. In particular, the gene disruptions are selected from the gene set that includes genes detailed herein above.

The metabolic engineering strategies listed in this disclosure assume that the organism can produce long chain alcohols via the malonyl-CoA independent pathway. The construction of a recombinant host organism capable of producing long chain alcohols via the malonyl-CoA independent pathway involves engineering a non-naturally occurring microbial organism having a malonyl-CoA-independent fatty acid synthesis (FAS) pathway and an acyl-reduction pathway having at least one exogenous nucleic acid encoding a malonyl-CoA-independent FAS pathway enzyme expressed in sufficient amounts to produce a primary alcohol. Such a malonyl-CoA-independent FAS pathway includes a ketoacyl-CoA acyltransferase or ketoacyl-CoA thiolase, 3-hydroxyacyl-CoA dehydrogenase, enoyl-CoA hydratase and enoyl-CoA reductase. The acyl-reduction pathway includes an acyl-CoA reductase and an alcohol dehydrogenase.

In order to validate the computational predictions presented herein, the strains must be constructed, evolved, and tested. *Escherichia coli* K-12 MG1655 housing the MI-LCA pathway will serve as the strain into which the disruptions will be introduced. The disruptions will be constructed by incorporating in-frame deletions using homologous recombination via the λ Red recombinase system of Datsenko and Wanner (Datsenko, K.A. and B.L. Wanner, *One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products*. Proc Natl Acad Sci U S A, 2000. **97**(12): p. 6640-5.). The approach involves replacing a chromosomal sequence (i.e., the gene targeted for removal) with a selectable antibiotic resistance gene, which itself is later removed. Knockouts are integrated one by one into the recipient strain. No antibiotic resistance markers remain after each deletion allowing accumulation of multiple mutations in each target strain. The deletion technology completely removes the gene targeted for removal so as to substantially reduce the possibility of the constructed mutants reverting back to the wild-type.

As intermediate strains are being constructed, strain performance will be quantified by performing shake flask fermentations. Anaerobic conditions will be obtained by sealing the flasks with a rubber septum and then sparging the medium with nitrogen. For strains where growth is not observed under strict anaerobic conditions, microaerobic conditions can be applied by covering the flask with foil and poking a small hole for limited aeration. All experiments are performed using M9 minimal medium supplemented with glucose unless otherwise stated. Pre-cultures are grown overnight and used as inoculum for a fresh batch culture for which measurements are taken during exponential growth. The growth rate is determined by measuring optical density using a spectrophotometer (600 nm), and the glucose uptake rate by monitoring carbon source depletion over time. LCAs, ethanol, and organic acids are analyzed by GC-MS or HPLC using routine procedures. Triplicate cultures are grown for each strain.

The performance of select strains is tested in anaerobic, pH-controlled batch fermentations. This enables reliable quantification of the growth, glucose uptake, and formation rates of all products, as well as ensuring that the accumulation of acidic fermentation products will not limit cell growth. In addition, it allows accurate determination of LCA volumetric productivity and yield, two important parameters in benchmarking strain performance. Fermentations are carried out in 1-L bioreactors with 600 mL working volume, equipped with temperature and pH control. The reactor is continuously sparged with N₂ at approximately 0.5 L/min to ensure that DO levels remain below detection levels. The culture medium is the same as described above, except that the glucose concentration is increased in accordance with the higher cell density achievable in a fermentation vessel.

Chemostat experiments will be conducted to obtain a direct measure of how the switch in fermentation mode from batch to continuous affects LCA yield and volumetric productivity. The bioreactors described above using batch mode are operated in chemostat mode through continuous supply of medium and removal of spent culture. The inlet flow rate is set to maintain a constant dilution rate of 80% of the maximum growth rate observed for each strain in batch, and the outlet flow is controlled to maintain level. Glucose is the limiting nutrient in the medium, and set to achieve the desired optical density in the vessel.

The recombinant strains are initially expected to exhibit suboptimal growth rates until their metabolic networks have adjusted to their missing functionalities. To enable this adjustment, the strains are adaptively evolved. By subjecting the strains to adaptive evolution, cellular growth rate becomes the primary selection pressure and the mutant cells are compelled to reallocate their metabolic fluxes in order to enhance their rates of growth. This reprogramming of metabolism has been recently demonstrated for several *E. coli* mutants that had been adaptively evolved on various substrates to reach the growth rates predicted *a priori* by an *in silico* model (Fong, S.S. and B.O. Palsson, *Metabolic gene-deletion strains of Escherichia coli evolve to computationally predicted growth phenotypes*. Nat Genet, 2004. **36**(10): p. 1056-8.). The OptKnock-generated strains are adaptively evolved in triplicate (running in parallel) due to differences in the evolutionary patterns witnessed previously in *E. coli* (Fong, S.S. and B.O. Palsson, *Metabolic gene-deletion strains of Escherichia coli evolve to computationally predicted growth phenotypes*. Nat Genet, 2004. **36**(10): p. 1056-8; Fong, S.S., J.Y. Marciniak, and B.O. Palsson, *Description and interpretation of adaptive evolution of Escherichia coli K-12 MG1655 by using a genome-scale in silico metabolic model*. J Bacteriol, 2003. **185**(21): p. 6400-8; Ibarra, R.U., J.S. Edwards, and B.O. Palsson, *Escherichia coli K-12 undergoes adaptive evolution to achieve in silico predicted optimal growth*. Nature, 2002. **420**(6912): p. 186-189.) that could potentially result in one strain having superior production qualities over the others. Evolutions are run for a period of 2-6 weeks, depending upon the rate of growth improvement attained. In general, evolutions are stopped once a stable phenotype is obtained. The growth-coupled biochemical production concept behind the OptKnock approach results in the generation of genetically stable overproducers.

The engineered strains can be characterized by measuring the growth rate, the substrate uptake rate, and the product/byproduct secretion rate. Cultures are grown overnight and used as inoculum for a fresh batch culture for which measurements are taken during exponential growth. The growth rate can be determined by measuring optical density using a spectrophotometer (A600). Concentrations of glucose and other organic acid byproducts in the culture supernatant are determined by HPLC using an HPX-87H column (BioRad), and used to calculate uptake and secretion rates. All experiments are performed with triplicate cultures.

Following the adaptive evolution process, the new strains are characterized again by measuring the growth rate, the substrate uptake rate, and the product/byproduct secretion rate. These results will be compared to the OptKnock predictions by plotting actual growth and production yields along side the production envelopes in the above figures. The most successful OptKnock design/evolution combinations are chosen to pursue further, and are characterized in lab-scale batch and continuous fermentations. The growth-coupled biochemical production concept behind the OptKnock approach should also result in the generation of genetically stable overproducers. Thus, the cultures are maintained in continuous mode for one month to evaluate long-term stability. Periodic samples are taken to ensure that yield and productivity are maintained throughout the experiment.

As previously mentioned, 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)). The framework examines the complete metabolic and/or biochemical network of a microorganism in order to suggest genetic manipulations that force the desired biochemical to become a 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 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 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 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 disruptions. OptKnock computational framework allows the construction of model formulations that enable an effective query of the performance limits of 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. patent application serial No. 11/891,602, 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. 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 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 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 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 will 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 a 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 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 disruption combinations that force a coupling between biochemical production and cell growth based on network stoichiometry. The identification of optimal gene/reaction disruptions 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 *Escherichia coli* metabolism can be employed to identify essential genes for metabolic pathways as exemplified previously and described in, 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 in U.S. Patent No. 7,127,379. As disclosed herein, the OptKnock mathematical framework can be applied to pinpoint gene disruptions leading to the growth-coupled production of a desired product. Further, the solution of the bilevel OptKnock problem provides only one set of disruptions. To enumerate all meaningful solutions, that is, all sets of disruptions 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.

Adaptive evolution is a powerful experimental technique that can be used to increase growth rates of mutant or engineered microbial strains, or of wild-type strains growing under unnatural environmental conditions. It is especially useful for strains designed via the OptKnock formalism, which results in growth-coupled product formation. Therefore, evolution toward optimal growing strains will indirectly optimize production as well. Unique strains of *E. coli* K-12 MG1655 were created through gene knockouts and adaptive evolution. (Fong, S. S. and B. O. Palsson, *Nat. Genet.* 36:1056-1058 (2004).) In this work, all adaptive evolutionary cultures were maintained in prolonged exponential growth by serial passage of batch cultures into fresh medium before the stationary phase was reached, thus rendering growth rate as the primary selection pressure. The genes that were selected for this knockout study were *ackA*, *frdA*, *pckA*, *ppc*, *tpiA*, and *zwf*. Knockout strains were constructed and evolved on minimal medium supplemented with different carbon substrates (four for each knockout strain). Evolution cultures were carried out in duplicate or triplicate, giving a total of 50 evolution knockout strains. The evolution cultures were maintained in exponential growth until a stable growth rate was reached. The computational predictions were accurate (i.e., within 10%) at predicting the post-evolution growth rate of the knockout strains in 38 out of the 50 cases examined. Furthermore, a combination of OptKnock design with adaptive evolution has led to improved lactic acid

production strains. (Fong, S. S., A. P. Burgard, C. D. Herring, E. M. Knight, F. R. Blattner, C. D. Maranas, and B. O. Palsson, *Biotechnol Bioeng* 91:643-648 (2005).) The guidance of these teachings relevant to *E. coli* can be applied to other organisms.

There are a number of developed technologies for carrying out adaptive evolution. Exemplary methods are provided herein below. In some embodiments, optimization of a non-naturally occurring organism of the present invention includes subject the use of any of the these adaptive evolution techniques.

Serial culture involves repetitive transfer of a small volume of grown culture to a much larger vessel containing fresh growth medium. When the cultured organisms have grown to saturation in the new vessel, the process is repeated. This method has been used to achieve the longest demonstrations of sustained culture in the literature, (Lenski, R. E. and M. Travisano, *Proc Natl Acad Sci U S.A.* 91:6808-6814 (1994).) in experiments which clearly demonstrated consistent improvement in reproductive rate over period of years. In the experiments performed in the Palsson lab described above, transfer is usually performed during exponential phase, so each day the transfer volume is precisely calculated to maintain exponential growth through the next 24 hour period. This process is usually done manually, with considerable labor investment, and is subject to contamination through exposure to the outside environment. Furthermore, since such small volumes are transferred each time, the evolution is inefficient and many beneficial mutations are lost. On the positive side, serial dilution is inexpensive and easy to parallelize.

In continuous culture the growth of cells in a chemostat represents an extreme case of dilution in which a very high fraction of the cell population remains. As a culture grows and becomes saturated, a small proportion of the grown culture is replaced with fresh media, allowing the culture to continually grow at close to its maximum population size. Chemostats have been used to demonstrate short periods of rapid improvement in reproductive rate. (Dykhuizen, D. E., *Methods Enzymol.* 613-631 (1993).) The potential power of these devices was recognized, but traditional chemostats were unable to sustain long periods of selection for increased reproduction rate, due to the unintended selection of dilution-resistant (static) variants. These variants are able to resist dilution by adhering to the surface of the chemostat, and by doing so, outcompete less sticky individuals including those that have higher reproductive rates, thus obviating the intended

purpose of the device. (Chao, L. and G. Ramsdell *J.Gen.Microbiol* 20:132-138 (1985).) One possible way to overcome this drawback is the implementation of a device with two growth chambers, which periodically undergo transient phases of sterilization, as described in the patent by the Pasteur Institute (Marliere and Mutzel, US Patent 6,686,194, filed 1999).

EvolugatorTM is a continuous culture device developed by Evolugate, LLC (Gainesville, FL) exhibits significant time and effort savings over traditional evolution techniques. (de Crecy, E., Metzgar, D., Allen, C., Penicaud, M., Lyons, B., Hansen, C.J., de Crecy-Lagard, V. *Appl. Microbiol. Biotechnol.* 77:489-496 (2007).) The cells are maintained in prolonged exponential growth by the serial passage of batch cultures into fresh medium before the stationary phase is attained. By automating optical density measurement and liquid handling, the Evolugator can perform serial transfer at high rates using large culture volumes, thus approaching the efficiency of a chemostat in evolution of cell fitness. For example, a mutant of *Acinetobacter* sp ADP1 deficient in a component of the translation apparatus, and having severely hampered growth, was evolved in 200 generations to 80% of the wild-type growth rate. However, in contrast to the chemostat which maintains cells in a single vessel, the machine operates by moving from one "reactor" to the next in subdivided regions of a spool of tubing, thus eliminating any selection for wall-growth. The transfer volume is adjustable, and normally set to about 50%.

As disclosed herein, a nucleic acid encoding a desired activity of a primary alcohol, a fatty acyl-CoA, a fatty ester, a wax, an acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol pathway can be introduced into a host organism. In some cases, it can be desirable to modify an activity of a primary alcohol, a fatty acyl-CoA, a fatty ester, a wax, an acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol pathway enzyme or protein to increase production of primary alcohol, a fatty acyl-CoA, a fatty ester, a wax, an acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol. 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, 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 the automated screening of many enzyme variants (for example, >104). 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 (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 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 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 primary alcohol, a fatty acyl-CoA, a fatty ester, a wax, an acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol 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

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 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 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 Chimeragenesis 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 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 (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 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 (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 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 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 Verenium 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)).

It is understood that modifications which do not substantially affect the activity of the various 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

Primary Alcohol Biosynthesis

This Example describes the generation of a microbial organism capable of producing primary alcohols using a malonyl-CoA independent FAS metabolic pathway and acyl-reduction metabolic pathways.

Escherichia coli is used as a target organism to engineer a malonyl-CoA-independent FAS and acyl-reduction pathway as shown in Figure 1. *E. coli* provides a good host for generating a non-naturally occurring microorganism capable of producing primary alcohol, such as octanol. *E. coli* is amenable to genetic manipulation and is known to be capable of producing various products, like ethanol, effectively under anaerobic conditions.

To generate an *E. coli* strain engineered to produce primary alcohol, nucleic acids encoding the enzymes utilized in the malonyl-CoA-independent FAS and acyl-reduction pathway as described previously, are expressed in *E. coli* using well known molecular biology techniques (see, for example, Sambrook, *supra*, 2001; Ausubel *supra*, 1999; Roberts et al., *supra*, 1989). In particular, the *fadI/fadJ* genes (NP_416844.1 and NP_416843.1), encoding the multienzyme complex with ketoacyl-CoA thiolase, 3-hydroxyacyl-CoA dehydrogenase, and enoyl-CoA hydratase activities under anaerobic conditions, and the TDE0597 (NP_971211.1), encoding enoyl-CoA reductase, are cloned into the pZE13 vector (Expressys, Ruelzheim, Germany) under the PA1/lacO promoter. The *acrI* gene (YP_047869.1), encoding acyl-CoA reductase, and the *alrA* gene (BAB12273.1), encoding alcohol dehydrogenase, are cloned into the pZA33 vector (Expressys, Ruelzheim, Germany) under the PA1/lacO promoter. The two sets of plasmids are transformed into *E. coli* strain MG1655 to express the proteins and enzymes required for the malonyl-CoA-independent FAS and acyl-reduction pathway.

The resulting genetically engineered organism is cultured in glucose containing medium following procedures well known in the art (see, for example, Sambrook et al., *supra*, 2001). The expression of malonyl-CoA-independent FAS and acyl-reduction pathway genes is corroborated using methods well known in the art for determining polypeptide expression or enzymatic activity, including for example, Northern blots, PCR amplification of mRNA, immunoblotting. Enzymatic activities of the expressed enzymes are confirmed using assays specific for the individually activities (see, for example, Tucci, *supra*, 2007; Hoffmeister et al., 2005; Inui et al., *supra*, 1984; Winkler, 2003; Tani, 2000; Reiser, 1997; Ishige, 2000). The ability of the engineered *E. coli* strain to produce primary alcohol, such as octanol is confirmed using HPLC, gas chromatography-mass spectrometry (GCMS) or liquid chromatography-mass spectrometry (LCMS).

Microbial strains engineered to have a functional malonyl-CoA-independent FAS and acyl-reduction pathway is further augmented by optimization for efficient utilization of the pathway. Briefly, the engineered strain is assessed to determine whether any of the exogenous genes are expressed at a rate limiting level. Expression is increased for any enzymes expressed at low levels that can limit the flux through the pathway by, for example, introduction of additional gene copy numbers.

To generate better producers, metabolic modeling is utilized to optimize growth conditions. Modeling is also 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 in 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 primary alcohols. One modeling method is the bilevel optimization approach, OptKnock (Burgard et al., *Biotechnol. Bioengineer.* 84:647-657 (2003)), which is applied to select gene knockouts that collectively result in better production of primary alcohols. Adaptive evolution also can be used to generate better producers of, for example, the acetyl-CoA intermediate or the primary alcohol product. Adaptive evolution is performed to improve both growth and production characteristics (Fong and Palsson, *Nat. Genet.* 36:1056-1058 (2004); Alper et al., *Science* 314:1565-1568 (2006)). Based on the results, subsequent rounds of modeling, genetic engineering and adaptive evolution can be applied to the primary alcohol producer to further increase their production.

For large-scale production of primary alcohols, the above malonyl-CoA independent FAS pathway-containing organism is cultured in a fermenter using a medium known in the art to support growth of the organism under anaerobic conditions. Fermentations are performed in either a batch, fed-batch or continuous manner. Anaerobic conditions are maintained by first sparging the medium with nitrogen and then sealing culture vessel (e.g., flasks can be sealed with a septum and crimp-cap). Microaerobic conditions also can be utilized by providing a small hole for limited aeration. The pH of the medium is maintained at a pH of 7 by addition of an acid, such as H₂SO₄. The growth rate is determined by measuring optical density using a spectrophotometer (600 nm), and the glucose uptake rate by monitoring carbon source depletion over time. Byproducts such as undesirable alcohols, organic acids, and residual glucose can be

quantified by HPLC (Shimadzu) with an HPX-087 column (BioRad), using a refractive index detector for glucose and alcohols, and a UV detector for organic acids, Lin et al., *Biotechnol. Bioeng.*, 775-779 (2005).

Isolation of the product primary alcohol is performed based their insolubility in water. In particular, a two-phase fermentation process is used for separation of these product alcohols where they can either form a separate phase or be readily extracted in an organic phase from the fermentation broth. Residual cells and any other insoluble impurities are removed by filtration, allowing a continuous or semi-continuous fermentation process.

EXAMPLE II

Microorganisms Having Growth-coupled Production of LCA

This Example describes the construction *in silico* designed strains for the growth-coupled production of LCA.

E. coli K-12 MG1655 serves as the wild-type strain into which the disruptions are introduced. The disruptions are constructed by incorporating in-frame deletions using homologous recombination via the λ Red recombinase system of Datsenko and Wanner. (Datsenko, K.A. and B.L. Wanner, *Proc Natl Acad Sci U S A.*, 97(12):6640-5 (2000).) The approach involves replacing a chromosomal sequence (i.e., the gene targeted for removal) with a selectable antibiotic resistance gene, which itself is later removed. Knockouts are integrated one by one into the recipient strain. No antibiotic resistance markers will remain after each deletion allowing accumulation of multiple mutations in each target strain. The deletion technology completely removes the gene targeted for removal so as to substantially reduce the possibility of the constructed mutants reverting back to the wild-type.

As described further below, one exemplary growth condition for achieving biosynthesis of LCA includes anaerobic culture or fermentation conditions. In certain embodiments, the non-naturally occurring microbial organism of the invention can be sustained, cultured or fermented under anaerobic or substantially anaerobic conditions. Briefly, anaerobic conditions refers 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 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 engineered strains are characterized by measuring the growth rate, the substrate uptake rate, and the product/byproduct secretion rate. Cultures are grown overnight and used as inoculum for a fresh batch culture for which measurements are taken during exponential growth. The growth rate is determined by measuring optical density using a spectrophotometer (A600).

Concentrations of glucose, LCA, and other organic acid byproducts in the culture supernatant are determined by HPLC using an HPX-87H column (BioRad), and are used to calculate uptake and secretion rates. All experiments are performed with triplicate cultures.

The recombinant strains can exhibit suboptimal growth rates until their metabolic networks have adjusted to their missing functionalities. To enable this adjustment, the strains are adaptively evolved. By subjecting the strains to adaptive evolution, cellular growth rate becomes the primary selection pressure and the mutant cells are compelled to reallocate their metabolic fluxes in order to enhance their rates of growth. This reprogramming of metabolism has been recently demonstrated for several *E. coli* mutants that had been adaptively evolved on various substrates to reach the growth rates predicted *a priori* by an *in silico* model. (Fong, S.S. and B.O. Palsson, *Nat Genet*, 36(10):1056-8 (2004).) These teachings can be applied to *Escherichia coli*.

Should the OptKnock predictions prove successful; the growth improvements brought about by adaptive evolution will be accompanied by enhanced rates of LCA production. The OptKnock-generated strains are adaptively evolved in triplicate (running in parallel) due to differences in the evolutionary patterns witnessed previously in *E. coli* (Fong, S.S. and B.O. Palsson, *Nat Genet*, 36(10):1056-8 (2004); Fong, S.S., J.Y. Marciniak, and B.O. Palsson, *J Bacteriol*, 185(21):6400-8 (2003); Ibarra, R.U., J.S. Edwards, and B.O. Palsson, *Nature*, 420(6912):186-189 (2002)) that could potentially result in one strain having superior production qualities over the others. Evolutions are run for a period of 2-6 weeks, depending upon the rate of growth improvement attained. In general, evolutions are stopped once a stable phenotype is obtained.

The adaptive evolution procedure involves maintaining the cells in prolonged exponential growth by the serial passage of batch cultures into fresh medium before the stationary phase is attained. Briefly, one procedure allows cells to reach mid-exponential growth ($A_{600}=0.5$) before being diluted and passed to fresh medium (i.e., M9 minimal media with 2 g/L carbon source). This process is repeated, allowing for about 500 generations for each culture. Culture samples are taken, frozen with liquid nitrogen, and the optical culture density recorded for each day throughout the course of the evolutions. The evolutions are performed in triplicate due to differences in the evolutionary patterns witnessed previously Donnelly et al., *Appl Biochem Biotechnol* 70-72: 187-98 (1998); Vemuri et al., *Appl Environ Microbiol* 68:1715-27 (2002), that could potentially result in one strain having superior production qualities over the others. The adaptive evolution step can take up to about two months or more. The adaptive evolution step also can be less than two months depending on the strain design, for example.

Another process can evolve cells using automation technology and is commercially available by Evolugate, LLC (Gainesville, FL) under a service contract. The procedure employs the EvolugatorTM evolution machine which results in significant time and effort savings over non-automated evolution techniques. Cells are maintained in prolonged exponential growth by the serial passage of batch cultures into fresh medium before the stationary phase is attained. By automating optical density measurement and liquid handling, the Evolugator can perform serial transfer at high rates using large culture volumes, thus approaching the efficiency of a chemostat for evolution of cell fitness. For example, a mutant of *Acinetobacter* sp ADP1 deficient in a component of the translation apparatus, and having severely hampered growth, was evolved in 200 generations to 80% of the wild-type growth rate. However, in contrast to the chemostat which maintains cells in a single vessel, the machine operates by moving from one “reactor” to the next in subdivided regions of a spool of tubing, thus eliminating any selection for wall-growth. The transfer volume is adjustable, and normally set to about 50%.

In contrast to a chemostat, which maintains cells in a single vessel, the machine operates by moving from one “reactor” to the next in subdivided regions of a spool of tubing, thus eliminating any selection for wall-growth. Culture samples are taken, frozen with liquid nitrogen, and the optical culture density recorded each day throughout the course of the evolutions. The Evolugator is used for each strain until a stable growth rate is achieved. Growth

rate improvements of nearly 50% have been observed in two weeks using this device. The above-described strains are adaptively evolved in triplicate (running in parallel). At ten day intervals, culture samples are taken from the Evolugator, purified on agar plates, and cultured in triplicate as discussed above to assess strain physiology. EvolugatorTM is a continuous culture device that exhibits significant time and effort savings over traditional evolution techniques. (de Crecy et al., *Appl. Microbiol. Biotechnol.* 77:489-496 (2007)).

Following the adaptive evolution process, the new strains are again characterized by measuring the growth rate, the substrate uptake rate, and the product/byproduct secretion rate. These results are compared to the OptKnock predictions by plotting actual growth and production yields alongside the production envelopes. The most successful OptKnock design/evolution combinations are chosen to pursue further, and is characterized in lab-scale batch and continuous fermentations. The growth-coupled biochemical production concept behind the OptKnock approach should also result in the generation of genetically stable overproducers. Thus, the cultures can be maintained in continuous mode for one month to evaluate long-term stability. Periodic samples will be taken to ensure that yield and productivity are maintained throughout the experiment.

EXAMPLE III

Microorganisms Having Growth-coupled Production of LCA

This Example describes the construction *in silico* designed strains for the growth-coupled production of LCA.

Gene deletions are introduced into *S. cerevisiae* by homologous recombination of the gene interrupted by the KanMX cassette, flanked by loxP sites enabling removal and recycling of the resistance marker (e.g. URA3) (Wach, A., et al., *PCR-based gene targeting in Saccharomyces cerevisiae*, in *Yeast Gene Analysis*, M.F. Tuite, Editor. 1998, Academic Press: San Diego). Starting with a loxP-kanMX-loxP sequence on a plasmid, an artificial construct with this sequence flanked by fragments of the gene of interest are created by PCR using primers containing both 45-50 bp target sequence followed by a region homologous to the above cassette. This linear DNA is transformed into wild-type *S. cerevisiae*, and recombinants are selected by geneticin resistance (Wach, A., *et al. supra*]. Colonies are purified and tested for

correct double crossover by PCR. To remove the KanMX marker, a plasmid containing the Cre recombinase and bleomycin resistance are introduced, promoting recombination between the loxP sites (Gueldener, U., *et al.*, *Nucleic Acids Res.* e23 (2002)). Finally, the resulting strain is cured of the Cre plasmid by successive culturing on media without any antibiotic present. The final strain will have a markerless gene deletion, and thus the same method can be used to introduce multiple deletions in the same strain.

The strains are constructed, evolved, and tested by methods disclosed herein. Genes can be inserted into *S. cerevisiae*, for example, using several methods. These methods can be plasmid-based whereas others allow for the incorporation of the gene in the chromosome. The latter approach employs an integrative promoter based expression vector, for example, the pGAPZ or the pGAPZ α vector based on the GAP promoter. The expression vector constitutes the GAP promoter, the HIS4 wild-type allele for integration and the 3' AOX transcription termination region of *P. pastoris* in addition to a KanMX cassette, flanked by loxP sites enabling removal and recycling of the resistance marker. Both of these vectors are commercially available from Invitrogen (Carlsbad, CA).

The method entails the synthesis and amplification of the gene of interest with suitable primers, followed by the digestion of the gene at a unique restriction site, such as that created by the *EcoRI/XhoI* enzymes (Vellanki *et al.*, *Biotechnol. Lett.* 29:313-318 (2007)). The gene is inserted at the *EcoRI* and *XhoI* sites in the expression vector, downstream of the GAP promoter. The gene insertion is verified by PCR and/or DNA sequence analysis. The recombinant plasmid is then linearized with *NarI* for histidine integration, purified and integrated into the chromosomal DNA of *S. cerevisiae* using an appropriate transformation method. The cells are plated on the YPD medium with the appropriate selection marker (e.g., kanamycin) and incubated for 2-3 days. The transformants will then be analyzed for the requisite gene insert by colony PCR.

To remove the antibiotic marker, a plasmid containing the Cre recombinase is introduced, promoting recombination between the loxP sites (Gueldener *et al.*, *supra*). Finally, the resulting strain is cured of the Cre plasmid by successive culturing on media without any antibiotic present. The final strain will have a markerless gene deletion, and thus the same method can be used to introduce multiple insertions in the same strain.

The engineered strains are characterized by measuring the growth rate, the substrate uptake rate, and the product/byproduct secretion rate. Cultures are grown overnight and used as inoculum for a fresh batch culture for which measurements are taken during exponential growth. The growth rate is determined by measuring optical density using a spectrophotometer (A₆₀₀).

Concentrations of glucose, alcohols, and other organic acid byproducts in the culture supernatant are determined by analytical methods including HPLC using an HPX-87H column (BioRad), or GC-MS, and used to calculate uptake and secretion rates. All experiments are performed with triplicate cultures.

The knockout strains are initially anticipated to exhibit suboptimal growth rates until their metabolic networks have adjusted to their missing functionalities. To enable this adjustment, the strains will be adaptively evolved. By subjecting the strains to adaptive evolution, cellular growth rate becomes the primary selection pressure and the mutant cells will be compelled to reallocate their metabolic fluxes in order to enhance their rates of growth. This reprogramming of metabolism has been recently demonstrated for several *E. coli* mutants that had been adaptively evolved on various substrates to reach the growth rates predicted *a priori* by an *in silico* model. The growth improvements brought about by adaptive evolution can be accompanied by enhanced rates of LCA production. The OptKnock-generated strains can be adaptively evolved in triplicate (running in parallel) due to differences in the evolutionary patterns witnessed previously in *E. coli* that could potentially result in one strain having superior production qualities over the others. Evolutions can be run for a period of 2-6 weeks, or longer depending upon the rate of growth improvement attained. In general, evolutions can be stopped once a stable phenotype is obtained.

The adaptive evolution procedure involves maintaining the cells in prolonged exponential growth by the serial passage of batch cultures into fresh medium before the stationary phase is attained. Briefly, one procedure allows cells to reach mid-exponential growth (A₆₀₀=0.5) before being diluted and passed to fresh medium (i.e., M9 minimal media with 2 g/L carbon source). This process is repeated, allowing for about 500 generations for each culture. Culture samples are taken, frozen with liquid nitrogen, and the optical culture density recorded for each day throughout the course of the evolutions. The evolutions are performed in triplicate due to differences in the evolutionary patterns witnessed previously Donnelly et al., *Appl Biochem*

Biotechnol. 70-72: 187-98 (1998); Vemuri et al., *Appl Environ Microbiol.* 68:1715-27 (2002), that could potentially result in one strain having superior production qualities over the others. The adaptive evolution step can take up to about two months or more. The adaptive evolution step also can be less than two months depending on the strain design, for example.

Another process can evolve cells using automation technology and is commercially available by Evolugate, LLC (Gainesville, FL) under a service contract. The procedure employs the EvolugatorTM evolution machine which results in significant time and effort savings over non-automated evolution techniques. Cells are maintained in prolonged exponential growth by the serial passage of batch cultures into fresh medium before the stationary phase is attained. By automating optical density measurement and liquid handling, the Evolugator can perform serial transfer at high rates using large culture volumes, thus approaching the efficiency of a chemostat for evolution of cell fitness. In contrast to a chemostat, which maintains cells in a single vessel, the machine operates by moving from one “reactor” to the next in subdivided regions of a spool of tubing, thus eliminating any selection for wall-growth. Culture samples are taken, frozen with liquid nitrogen, and the optical culture density recorded each day throughout the course of the evolutions. The Evolugator is used for each strain until a stable growth rate is achieved. Growth rate improvements of nearly 50% have been observed in two weeks using this device. The above-described strains are adaptively evolved in triplicate (running in parallel). At ten day intervals, culture samples are taken from the Evolugator, purified on agar plates, and cultured in triplicate as discussed above to assess strain physiology.

Following the adaptive evolution process, the new strains are again characterized by measuring the growth rate, the substrate uptake rate, and the product/byproduct secretion rate. These results are compared to the OptKnock predictions by plotting actual growth and production yields alongside the production envelopes. The most successful OptKnock design/evolution combinations are chosen to pursue further, and is characterized in lab-scale batch and continuous fermentations. The growth-coupled biochemical production concept behind the OptKnock approach should also result in the generation of genetically stable overproducers. Thus, the cultures can be maintained in continuous mode for one month to evaluate long-term stability. Periodic samples will be taken to ensure that yield and productivity are maintained throughout the experiment.

Described herein above, is the application of the OptKnock methodology for generating useful gene disruption targets. Multiple disruption strategies were enumerated for establishing the coupling between LCA production and *Escherichia coli* growth. This methodology is applicable to a wide variety of cells and microorganisms other than *Escherichia coli* and also can utilize metabolic modeling and simulation systems other than OptKnock.

The combined computational and engineering platform described herein is generally applicable to any stoichiometric model organism and the teachings and guidance provided herein will allow those skilled in the art to design and implement sets of genetic manipulations for metabolic alterations that lead to the growth-coupled production of any biochemical product.

The present disclosure provides gene disruption strategies for growth-coupled production of LCA in *Escherichia coli* under anaerobic conditions. The suggested strategies can increase product yields significantly over the reported yields for each of these products. A comprehensive list of the strategies is listed in Table 1 for LCA production. The associated genes and stoichiometries for each of the reaction disruption are catalogued in Table 2. Table 3 lists metabolite abbreviations and their corresponding names along with their location.

Table 1: The list of all disruption strategies identified by OptKnock that are most likely to provide growth-coupled LCA production.

1	ADHEr	LDH_D	ASPT	MDH	PFLi	PGDHY	PYK	DHAPT
2	ADHEr	LDH_D	ASPT	MDH	PFLi	PGL	PYK	DHAPT
3	ADHEr	LDH_D	ASPT	G6PDHy	MDH	PFLi	PYK	DHAPT
4	ADHEr	LDH_D	ASPT	EDA	MDH	PFLi	PYK	DHAPT
5	ADHEr	LDH_D	GLCpts	GLUDy	PGDH	PGI	PTAr	
6	ADHEr	LDH_D	ACKr	GLCpts	GLUDy	PGDH	PGI	
7	ADHEr	LDH_D	GLCpts	GLUDy	PGI	PTAr	TAL	
8	ADHEr	LDH_D	GLCpts	GLUDy	PGI	PTAr	TKT1	
9	ADHEr	LDH_D	ACKr	GLCpts	GLUDy	PGI	TKT1	
10	ADHEr	LDH_D	ACKr	GLCpts	GLUDy	PGI	TAL	
11	ADHEr	LDH_D	FBA	GLCpts	GLUDy	PTAr		
12	ADHEr	LDH_D	GLCpts	GLUDy	PTAr	TPI		
13	ADHEr	LDH_D	ACKr	GLCpts	GLUDy	TPI		
14	ADHEr	LDH_D	GLCpts	GLUDy	PFK	PTAr		
15	ADHEr	LDH_D	ACKr	FBA	GLCpts	GLUDy		
16	ADHEr	LDH_D	ACKr	GLCpts	GLUDy	PFK		

17	ADHEr	LDH_D	ACKr	GLCpts	GLUDy	PGI	RPE
18	ADHEr	LDH_D	GLCpts	GLUDy	PGI	PTAr	RPE
19	ADHEr	LDH_D	GLCpts	GLUDy	PGI	PTAr	TKT2
20	ADHEr	LDH_D	ACKr	GLCpts	GLUDy	PGI	TKT2
21	ADHEr	LDH_D	ACKr	GLCpts	PGDH	PGI	
22	ADHEr	LDH_D	GLCpts	PGDH	PGI	PTAr	
23	ADHEr	LDH_D	ACKr	GLCpts	PGI	TKT1	
24	ADHEr	LDH_D	ACKr	GLCpts	PGI	TAL	
25	ADHEr	LDH_D	GLCpts	PGI	PTAr	TAL	
26	ADHEr	LDH_D	GLCpts	PGI	PTAr	TKT1	
27	ADHEr	LDH_D	ACKr	GLCpts	PFK		
28	ADHEr	LDH_D	ACKr	GLCpts	TPI		
29	ADHEr	LDH_D	ACKr	FBA	GLCpts		
30	ADHEr	LDH_D	FBA	GLCpts	PTAr		
31	ADHEr	LDH_D	GLCpts	PFK	PTAr		
32	ADHEr	LDH_D	GLCpts	PTAr	TPI		
33	ADHEr	LDH_D	ACKr	GLCpts	PGI	RPE	
34	ADHEr	LDH_D	GLCpts	PGI	PTAr	RPE	
35	ADHEr	LDH_D	FRD2	GLCpts	GLUDy	PFLi	PGI
36	ADHEr	LDH_D	ACKr	GLCpts	PGI	TKT2	
37	ADHEr	LDH_D	GLCpts	PGI	PTAr	TKT2	
38	ADHEr	LDH_D	FRD2	GLCpts	GLUDy	PFLi	TPI
39	ADHEr	LDH_D	FBA	FRD2	GLCpts	GLUDy	PFLi
40	ADHEr	LDH_D	FRD2	GLCpts	GLUDy	PFK	PFLi
41	ADHEr	LDH_D	ASPT	ATPS4r	FUM	NADH6	PGI
42	ADHEr	LDH_D	ASPT	ATPS4r	MDH	NADH6	PGI
43	ADHEr	LDH_D	ASPT	ATPS4r	MDH	NADH6	PFK
44	ADHEr	LDH_D	ASPT	ATPS4r	FBA	FUM	NADH6
45	ADHEr	LDH_D	ASPT	ATPS4r	FUM	NADH6	TPI
46	ADHEr	LDH_D	ASPT	ATPS4r	MDH	NADH6	TPI
47	ADHEr	LDH_D	ASPT	ATPS4r	FBA	MDH	NADH6
48	ADHEr	LDH_D	ASPT	ATPS4r	FUM	NADH6	PFK
49	ADHEr	LDH_D	FUM	GLCpts	GLUDy	PFLi	PGI
50	ADHEr	LDH_D	GLCpts	GLUDy	MDH	PFLi	PGI
51	ADHEr	LDH_D	GLCpts	PFLi	PGI	SUCD4	
52	ADHEr	LDH_D	GLCpts	NADH6	PFLi	PGI	
53	ADHEr	LDH_D	FRD2	GLCpts	PFLi	PGI	
54	ADHEr	LDH_D	ACKr	GLUDy	HEX1	PGDH	PGI
55	ADHEr	LDH_D	GLUDy	HEX1	PGDH	PGI	PTAr
56	ADHEr	LDH_D	FUM	GLCpts	GLUDy	PFK	PFLi
57	ADHEr	LDH_D	FBA	FUM	GLCpts	GLUDy	PFLi
58	ADHEr	LDH_D	GLCpts	GLUDy	MDH	PFK	PFLi
59	ADHEr	LDH_D	FUM	GLCpts	GLUDy	PFLi	TPI
60	ADHEr	LDH_D	FBA	GLCpts	GLUDy	MDH	PFLi
61	ADHEr	LDH_D	GLCpts	GLUDy	MDH	PFLi	TPI
62	ADHEr	LDH_D	GLCpts	NADH6	PFLi	TPI	

63	ADHEr	LDH_D	FRD2	GLCpts	PFLi	TPI	
64	ADHEr	LDH_D	FBA	FRD2	GLCpts	PFLi	
65	ADHEr	LDH_D	FBA	GLCpts	NADH6	PFLi	
66	ADHEr	LDH_D	FBA	GLCpts	PFLi	SUCD4	
67	ADHEr	LDH_D	FRD2	GLCpts	PFK	PFLi	
68	ADHEr	LDH_D	GLCpts	PFLi	SUCD4	TPI	
69	ADHEr	LDH_D	GLCpts	PFK	PFLi	SUCD4	
70	ADHEr	LDH_D	GLCpts	NADH6	PFK	PFLi	
71	ADHEr	LDH_D	ASPT	GLCpts	MDH	PFLi	PGI
72	ADHEr	LDH_D	ASPT	FUM	GLCpts	PFLi	PGI
73	ADHEr	LDH_D	ASPT	ATPS4r	MDH	PGI	PPS
74	ADHEr	LDH_D	ASPT	ATPS4r	FUM	PGI	PPS
75	ADHEr	LDH_D	GLUDy	HEX1	PGI	PTAr	TAL
76	ADHEr	LDH_D	ACKr	GLUDy	HEX1	PGI	TAL
77	ADHEr	LDH_D	ACKr	GLUDy	HEX1	PGI	TKT1
78	ADHEr	LDH_D	GLUDy	HEX1	PGI	PTAr	TKT1
79	ADHEr	LDH_D	ACKr	GLUDy	HEX1	TPI	
80	ADHEr	LDH_D	ACKr	GLUDy	HEX1	PFK	
81	ADHEr	LDH_D	GLUDy	HEX1	PTAr	TPI	
82	ADHEr	LDH_D	GLUDy	HEX1	PFK	PTAr	
83	ADHEr	LDH_D	ACKr	FBA	GLUDy	HEX1	
84	ADHEr	LDH_D	FBA	GLUDy	HEX1	PTAr	
85	ADHEr	LDH_D	ASPT	GLCpts	MDH	PFLi	TPI
86	ADHEr	LDH_D	ASPT	FBA	GLCpts	MDH	PFLi
87	ADHEr	LDH_D	ASPT	GLCpts	MDH	PFK	PFLi
88	ADHEr	LDH_D	ASPT	FUM	GLCpts	PFK	PFLi
89	ADHEr	LDH_D	ASPT	FUM	GLCpts	PFLi	TPI
90	ADHEr	LDH_D	ASPT	FBA	FUM	GLCpts	PFLi
91	ADHEr	LDH_D	GLUDy	HEX1	PGI	PTAr	RPE
92	ADHEr	LDH_D	ACKr	GLUDy	HEX1	PGI	RPE
93	ADHEr	LDH_D	ASPT	ATPS4r	FUM	GLUDy	PGI
94	ADHEr	LDH_D	ASPT	ATPS4r	GLUDy	MDH	PGI
95	ADHEr	LDH_D	ASPT	ATPS4r	FBA	MDH	PPS
96	ADHEr	LDH_D	ASPT	ATPS4r	FUM	PFK	PPS
97	ADHEr	LDH_D	ASPT	ATPS4r	MDH	PFK	PPS
98	ADHEr	LDH_D	ASPT	ATPS4r	MDH	PPS	TPI
99	ADHEr	LDH_D	ASPT	ATPS4r	FUM	PPS	TPI
100	ADHEr	LDH_D	ASPT	ATPS4r	FBA	FUM	PPS
101	ADHEr	LDH_D	ACKr	GLUDy	HEX1	PGI	TKT2
102	ADHEr	LDH_D	GLUDy	HEX1	PGI	PTAr	TKT2
103	ADHEr	LDH_D	ASPT	ATPS4r	FBA	FUM	GLUDy
104	ADHEr	LDH_D	ASPT	ATPS4r	GLUDy	MDH	PFK
105	ADHEr	LDH_D	ASPT	ATPS4r	FBA	GLUDy	MDH
106	ADHEr	LDH_D	ASPT	ATPS4r	FUM	GLUDy	TPI
107	ADHEr	LDH_D	ASPT	ATPS4r	FUM	GLUDy	PFK
108	ADHEr	LDH_D	ASPT	ATPS4r	GLUDy	MDH	TPI

109	ADHEr	LDH_D	ACKr	GLUDy	PGDH	PGI	
110	ADHEr	LDH_D	GLUDy	PGDH	PGI	PTAr	
111	ADHEr	LDH_D	ACKr	GLUDy	PGI	TAL	
112	ADHEr	LDH_D	GLUDy	PGI	PTAr	TKT1	
113	ADHEr	LDH_D	ACKr	GLUDy	PGI	TKT1	
114	ADHEr	LDH_D	GLUDy	PGI	PTAr	TAL	
115	ADHEr	LDH_D	ACKr	GLUDy	TPI		
116	ADHEr	LDH_D	GLUDy	PFK	PTAr		
117	ADHEr	LDH_D	FBA	GLUDy	PTAr		
118	ADHEr	LDH_D	ACKr	FBA	GLUDy		
119	ADHEr	LDH_D	ACKr	GLUDy	PFK		
120	ADHEr	LDH_D	GLUDy	PTAr	TPI		
121	ADHEr	LDH_D	ACKr	GLUDy	PGI	RPE	
122	ADHEr	LDH_D	GLUDy	PGI	PTAr	RPE	
123	ADHEr	LDH_D	GLUDy	PGI	PTAr	TKT2	
124	ADHEr	LDH_D	ACKr	GLUDy	PGI	TKT2	
125	ADHEr	LDH_D	HEX1	PGDH	PGI	PTAr	
126	ADHEr	LDH_D	ACKr	HEX1	PGDH	PGI	
127	ADHEr	LDH_D	ASPT	ATPS4r	CBMK2	FUM	PGI
128	ADHEr	LDH_D	ASPT	ATPS4r	CBMK2	MDH	PGI
129	ADHEr	LDH_D	HEX1	PGI	PTAr	TAL	
130	ADHEr	LDH_D	HEX1	PGI	PTAr	TKT1	
131	ADHEr	LDH_D	ACKr	HEX1	PGI	TKT1	
132	ADHEr	LDH_D	ACKr	HEX1	PGI	TAL	
133	ADHEr	LDH_D	GLUDy	HEX1	PFLi	PGI	SUCD4
134	ADHEr	LDH_D	FRD2	GLUDy	HEX1	PFLi	PGI
135	ADHEr	LDH_D	GLUDy	HEX1	NADH6	PFLi	PGI
136	ADHEr	LDH_D	ACKr	FBA	HEX1		
137	ADHEr	LDH_D	FBA	HEX1	PTAr		
138	ADHEr	LDH_D	HEX1	PFK	PTAr		
139	ADHEr	LDH_D	ACKr	HEX1	PFK		
140	ADHEr	LDH_D	ACKr	HEX1	TPI		
141	ADHEr	LDH_D	HEX1	PTAr	TPI		
142	ADHEr	LDH_D	HEX1	PGI	PTAr	RPE	
143	ADHEr	LDH_D	ACKr	HEX1	PGI	RPE	
144	ADHEr	LDH_D	ASPT	ATPS4r	CBMK2	FBA	FUM
145	ADHEr	LDH_D	ASPT	ATPS4r	CBMK2	FBA	MDH
146	ADHEr	LDH_D	ASPT	ATPS4r	CBMK2	MDH	TPI
147	ADHEr	LDH_D	ASPT	ATPS4r	CBMK2	FUM	PFK
148	ADHEr	LDH_D	ASPT	ATPS4r	CBMK2	FUM	TPI
149	ADHEr	LDH_D	ASPT	ATPS4r	CBMK2	MDH	PFK
150	ADHEr	LDH_D	FBA	GLUDy	HEX1	NADH6	PFLi
151	ADHEr	LDH_D	GLUDy	HEX1	NADH6	PFK	PFLi
152	ADHEr	LDH_D	FBA	GLUDy	HEX1	PFLi	SUCD4
153	ADHEr	LDH_D	FRD2	GLUDy	HEX1	PFK	PFLi
154	ADHEr	LDH_D	GLUDy	HEX1	PFK	PFLi	SUCD4

155	ADHEr	LDH_D	GLUDy	HEX1	NADH6	PFLi	TPI
156	ADHEr	LDH_D	FBA	FRD2	GLUDy	HEX1	PFLi
157	ADHEr	LDH_D	GLUDy	HEX1	PFLi	SUCD4	TPI
158	ADHEr	LDH_D	FRD2	GLUDy	HEX1	PFLi	TPI
159	ADHEr	LDH_D	GLUDy	HEX1	MDH	PFLi	PGI
160	ADHEr	LDH_D	FUM	GLUDy	HEX1	PFLi	PGI
161	ADHEr	LDH_D	HEX1	PGI	PTAr	TKT2	
162	ADHEr	LDH_D	ACKr	HEX1	PGI	TKT2	
163	ADHEr	LDH_D	ATPS4r	GLUDy	HEX1	MDH	PFK
164	ADHEr	LDH_D	ATPS4r	FBA	GLUDy	HEX1	MDH
165	ADHEr	LDH_D	ATPS4r	GLUDy	HEX1	MDH	TPI
166	ADHEr	LDH_D	ATPS4r	FBA	FUM	GLUDy	HEX1
167	ADHEr	LDH_D	ATPS4r	FUM	GLUDy	HEX1	PFK
168	ADHEr	LDH_D	ATPS4r	FUM	GLUDy	HEX1	TPI
169	ADHEr	LDH_D	FBA	FUM	GLUDy	HEX1	PFLi
170	ADHEr	LDH_D	FUM	GLUDy	HEX1	PFLi	TPI
171	ADHEr	LDH_D	GLUDy	HEX1	MDH	PFLi	TPI
172	ADHEr	LDH_D	GLUDy	HEX1	MDH	PFK	PFLi
173	ADHEr	LDH_D	FBA	GLUDy	HEX1	MDH	PFLi
174	ADHEr	LDH_D	FUM	GLUDy	HEX1	PFK	PFLi
175	ADHEr	LDH_D	ATPS4r	FUM	GLUDy	HEX1	PGI
176	ADHEr	LDH_D	ATPS4r	GLUDy	HEX1	MDH	PGI
177	ADHEr	LDH_D	ASPT	ATPS4r	MDH	PGI	
178	ADHEr	LDH_D	ASPT	ATPS4r	FUM	PGI	
179	ADHEr	LDH_D	ATPS4r	GLUDy	MDH	NADH6	PGI
180	ADHEr	LDH_D	ATPS4r	FUM	GLUDy	NADH6	PGI
181	ADHEr	LDH_D	ATPS4r	GLUDy	HEX1	PGDH	PGI
182	ADHEr	LDH_D	PGDH	PGI	PTAr		
183	ADHEr	LDH_D	ACKr	PGDH	PGI		
184	ADHEr	LDH_D	ATPS4r	GLUDy	HEX1	PFLi	PGI
185	ADHEr	LDH_D	ASPT	ATPS4r	MDH	TPI	
186	ADHEr	LDH_D	ASPT	ATPS4r	FUM	TPI	
187	ADHEr	LDH_D	ASPT	ATPS4r	MDH	PFK	
188	ADHEr	LDH_D	ASPT	ATPS4r	FBA	FUM	
189	ADHEr	LDH_D	ASPT	ATPS4r	FBA	MDH	
190	ADHEr	LDH_D	ASPT	ATPS4r	FUM	PFK	
191	ADHEr	LDH_D	ACKr	PGI	TKT1		
192	ADHEr	LDH_D	PGI	PTAr	TAL		
193	ADHEr	LDH_D	PGI	PTAr	TKT1		
194	ADHEr	LDH_D	ACKr	PGI	TAL		
195	ADHEr	LDH_D	ATPS4r	FBA	GLUDy	MDH	NADH6
196	ADHEr	LDH_D	ATPS4r	GLUDy	MDH	NADH6	PFK
197	ADHEr	LDH_D	ATPS4r	GLUDy	MDH	NADH6	TPI
198	ADHEr	LDH_D	ATPS4r	FUM	GLUDy	NADH6	TPI
199	ADHEr	LDH_D	ATPS4r	FUM	GLUDy	NADH6	PFK
200	ADHEr	LDH_D	ATPS4r	FBA	FUM	GLUDy	NADH6

201	ADHEr	LDH_D	ATPS4r	GLUDy	HEX1	PGI	TAL
202	ADHEr	LDH_D	ATPS4r	GLUDy	HEX1	PGI	TKT1
203	ADHEr	LDH_D	ATPS4r	GLUDy	HEX1	PFK	
204	ADHEr	LDH_D	ATPS4r	GLUDy	HEX1	TPI	
205	ADHEr	LDH_D	ATPS4r	FBA	GLUDy	HEX1	
206	ADHEr	LDH_D	GLUDy	PTAr	PYK	SUCD4	
207	ADHEr	LDH_D	ACKr	GLUDy	PYK	SUCD4	
208	ADHEr	LDH_D	FRD2	GLUDy	PTAr	PYK	
209	ADHEr	LDH_D	ACKr	FRD2	GLUDy	PYK	
210	ADHEr	LDH_D	FDH2	GLUDy	NADH6	PTAr	PYK
211	ADHEr	LDH_D	ACKr	FDH2	GLUDy	NADH6	PYK
212	ADHEr	LDH_D	PFK	PTAr			
213	ADHEr	LDH_D	ACKr	TPI			
214	ADHEr	LDH_D	ACKr	FBA			
215	ADHEr	LDH_D	PTAr	TPI			
216	ADHEr	LDH_D	FBA	PTAr			
217	ADHEr	LDH_D	ACKr	PFK			
218	ADHEr	LDH_D	FRD2	GLUDy	PFLi	PGI	
219	ADHEr	LDH_D	GLUDy	PFLi	PGI	PRO1z	SUCD4
220	ADHEr	LDH_D	ACKr	PGI	RPE		
221	ADHEr	LDH_D	PGI	PTAr	RPE		
222	ADHEr	LDH_D	ACKr	PGI	TKT2		
223	ADHEr	LDH_D	PGI	PTAr	TKT2		
224	ADHEr	LDH_D	ATPS4r	GLUDy	HEX1	PGI	RPE
225	ADHEr	LDH_D	FRD2	GLUDy	PFLi	TPI	
226	ADHEr	LDH_D	FRD2	GLUDy	PFK	PFLi	
227	ADHEr	LDH_D	FBA	FRD2	GLUDy	PFLi	
228	ADHEr	LDH_D	GLUDy	PFK	PFLi	PRO1z	SUCD4
229	ADHEr	LDH_D	GLUDy	PFLi	PRO1z	SUCD4	TPI
230	ADHEr	LDH_D	FBA	GLUDy	PFLi	PRO1z	SUCD4
231	ADHEr	LDH_D	GLUDy	MDH	PFLi	PGI	SUCD4
232	ADHEr	LDH_D	FUM	GLUDy	NADH6	PFLi	PGI
233	ADHEr	LDH_D	GLUDy	MDH	NADH6	PFLi	PGI
234	ADHEr	LDH_D	FUM	GLUDy	PFLi	PGI	SUCD4
235	ADHEr	LDH_D	ASPT	GLUDy	MDH	PFLi	PGI
236	ADHEr	LDH_D	ASPT	FUM	GLUDy	PFLi	PGI
237	ADHEr	LDH_D	ATPS4r	GLUDy	HEX1	PGI	TKT2
238	ADHEr	LDH_D	FUM	GLUDy	PFK	PFLi	SUCD4
239	ADHEr	LDH_D	GLUDy	MDH	NADH6	PFK	PFLi
240	ADHEr	LDH_D	FUM	GLUDy	PFLi	SUCD4	TPI
241	ADHEr	LDH_D	FUM	GLUDy	NADH6	PFK	PFLi
242	ADHEr	LDH_D	FBA	FUM	GLUDy	PFLi	SUCD4
243	ADHEr	LDH_D	GLUDy	MDH	PFLi	SUCD4	TPI
244	ADHEr	LDH_D	GLUDy	MDH	PFK	PFLi	SUCD4
245	ADHEr	LDH_D	FBA	FUM	GLUDy	NADH6	PFLi
246	ADHEr	LDH_D	FBA	GLUDy	MDH	PFLi	SUCD4

247	ADHEr	LDH_D	FBA	GLUDy	MDH	NADH6	PFLi
248	ADHEr	LDH_D	GLUDy	MDH	NADH6	PFLi	TPI
249	ADHEr	LDH_D	FUM	GLUDy	NADH6	PFLi	TPI
250	ADHEr	LDH_D	ASPT	ATPS4r	FUM	NADH6	PYK
251	ADHEr	LDH_D	ASPT	ATPS4r	MDH	NADH6	PYK
252	ADHEr	LDH_D	GLCpts	GLUDy	PFLi	PGI	PTAr
253	ADHEr	LDH_D	ACKr	GLCpts	GLUDy	PFLi	PGI
254	ADHEr	LDH_D	ASPT	FBA	GLUDy	MDH	PFLi
255	ADHEr	LDH_D	ASPT	GLUDy	MDH	PFK	PFLi
256	ADHEr	LDH_D	ASPT	FBA	FUM	GLUDy	PFLi
257	ADHEr	LDH_D	ASPT	GLUDy	MDH	PFLi	TPI
258	ADHEr	LDH_D	ASPT	FUM	GLUDy	PFLi	TPI
259	ADHEr	LDH_D	ASPT	FUM	GLUDy	PFK	PFLi
260	ADHEr	LDH_D	ME2	PGL	PTAr	PYK	SUCD4
261	ADHEr	LDH_D	FRD2	G6PDHy	ME2	PTAr	PYK
262	ADHEr	LDH_D	ACKr	ME2	PGL	PYK	SUCD4
263	ADHEr	LDH_D	ACKr	FRD2	ME2	PGL	PYK
264	ADHEr	LDH_D	FRD2	ME2	PGL	PTAr	PYK
265	ADHEr	LDH_D	G6PDHy	ME2	PTAr	PYK	SUCD4
266	ADHEr	LDH_D	ACKr	FRD2	G6PDHy	ME2	PYK
267	ADHEr	LDH_D	ACKr	G6PDHy	ME2	PYK	SUCD4
268	ADHEr	LDH_D	G6PDHy	MDH	PTAr	PYK	SUCD4
269	ADHEr	LDH_D	ACKr	G6PDHy	MDH	NADH6	PYK
270	ADHEr	LDH_D	FRD2	G6PDHy	MDH	PTAr	PYK
271	ADHEr	LDH_D	FRD2	MDH	PGL	PTAr	PYK
272	ADHEr	LDH_D	ACKr	G6PDHy	MDH	PYK	SUCD4
273	ADHEr	LDH_D	ACKr	MDH	PGL	PYK	SUCD4
274	ADHEr	LDH_D	MDH	NADH6	PGL	PTAr	PYK
275	ADHEr	LDH_D	ACKr	MDH	NADH6	PGL	PYK
276	ADHEr	LDH_D	ACKr	FRD2	G6PDHy	MDH	PYK
277	ADHEr	LDH_D	MDH	PGL	PTAr	PYK	SUCD4
278	ADHEr	LDH_D	ACKr	FRD2	MDH	PGL	PYK
279	ADHEr	LDH_D	G6PDHy	MDH	NADH6	PTAr	PYK
280	ADHEr	LDH_D	ATPS4r	GLUDy	NADH6	PGI	
281	ADHEr	LDH_D	FUM	GLUDy	PTAr	PYK	
282	ADHEr	LDH_D	ACKr	GLUDy	MDH	PYK	
283	ADHEr	LDH_D	ACKr	FUM	GLUDy	PYK	
284	ADHEr	LDH_D	GLUDy	MDH	PTAr	PYK	
285	ADHEr	LDH_D	ATPS4r	HEX1	PGDH	PGI	
286	ADHEr	LDH_D	ATPS4r	GLUDy	NADH6	TPI	
287	ADHEr	LDH_D	ATPS4r	GLUDy	NADH6	PFK	
288	ADHEr	LDH_D	ATPS4r	FBA	GLUDy	NADH6	
289	ADHEr	LDH_D	HEX1	PFLi	PGI		
290	ADHEr	LDH_D	ASPT	ATPS4r	GLUDy	MDH	PYK
291	ADHEr	LDH_D	ASPT	ATPS4r	FUM	GLUDy	PYK
292	ADHEr	LDH_D	ATPS4r	HEX1	PGI	TKT1	

293	ADHEr	LDH_D	ATPS4r	HEX1	PGI	TAL	
294	ADHEr	LDH_D	ATPS4r	HEX1	PFK		
295	ADHEr	LDH_D	ATPS4r	FBA	HEX1		
296	ADHEr	LDH_D	ATPS4r	HEX1	TPI		
297	ADHEr	LDH_D	HEX1	PFLi	TPI		
298	ADHEr	LDH_D	HEX1	PFK	PFLi		
299	ADHEr	LDH_D	FBA	HEX1	PFLi		
300	ADHEr	LDH_D	ATPS4r	HEX1	PGI	RPE	
301	ADHEr	LDH_D	ACKr	GLUDy	NADH6	PGI	PYK
302	ADHEr	LDH_D	GLUDy	NADH6	PGI	PTAr	PYK
303	ADHEr	LDH_D	ATPS4r	HEX1	PGI	TKT2	
304	ADHEr	LDH_D	ACKr	FRD2	PYK		
305	ADHEr	LDH_D	ACKr	PYK	SUCD4		
306	ADHEr	LDH_D	FRD2	PTAr	PYK		
307	ADHEr	LDH_D	PTAr	PYK	SUCD4		
308	ADHEr	LDH_D	ACKr	FDH2	NADH6	PYK	
309	ADHEr	LDH_D	FDH2	NADH6	PTAr	PYK	
310	ADHEr	LDH_D	ATPS4r	NADH6	PGI		
311	ADHEr	LDH_D	ACKr	GLCpts	PFLi	PGI	
312	ADHEr	LDH_D	GLCpts	PFLi	PGI	PTAr	
313	ADHEr	LDH_D	FRD2	GLUDy	PFLi	PYK	
314	ADHEr	LDH_D	ATPS4r	FUM	GLUDy	PGDH	PGI
315	ADHEr	LDH_D	ATPS4r	GLUDy	MDH	PGDH	PGI
316	ADHEr	LDH_D	FUM	GLUDy	PFLi	PGI	
317	ADHEr	LDH_D	GLUDy	MDH	PFLi	PGI	
318	ADHEr	LDH_D	ATPS4r	FBA	NADH6		
319	ADHEr	LDH_D	ATPS4r	NADH6	PFK		
320	ADHEr	LDH_D	ATPS4r	NADH6	TPI		
321	ADHEr	LDH_D	ATPS4r	FBA	FUM	GLUDy	
322	ADHEr	LDH_D	ATPS4r	FUM	GLUDy	PFK	
323	ADHEr	LDH_D	ATPS4r	FBA	GLUDy	MDH	
324	ADHEr	LDH_D	ATPS4r	GLUDy	MDH	TPI	
325	ADHEr	LDH_D	ATPS4r	FUM	GLUDy	TPI	
326	ADHEr	LDH_D	ATPS4r	GLUDy	MDH	PFK	
327	ADHEr	LDH_D	FRD2	G6PDHy	ME2	PFLi	PYK
328	ADHEr	LDH_D	FRD2	ME2	PFLi	PGL	PYK
329	ADHEr	LDH_D	EDA	FRD2	ME2	PFLi	PYK
330	ADHEr	LDH_D	FRD2	ME2	PFLi	PGDHY	PYK
331	ADHEr	LDH_D	GLUDy	MDH	PFK	PFLi	
332	ADHEr	LDH_D	FBA	GLUDy	MDH	PFLi	
333	ADHEr	LDH_D	GLUDy	MDH	PFLi	TPI	
334	ADHEr	LDH_D	FBA	FUM	GLUDy	PFLi	
335	ADHEr	LDH_D	FUM	GLUDy	PFLi	TPI	
336	ADHEr	LDH_D	FUM	GLUDy	PFK	PFLi	
337	ADHEr	LDH_D	PFLi	PGI	SUCD4		
338	ADHEr	LDH_D	FRD2	PFLi	PGI		

339	ADHEr	LDH_D	NADH6	PFLi	PGI		
340	ADHEr	LDH_D	FRD2	MDH	PFLi	PGL	PYK
341	ADHEr	LDH_D	FRD2	G6PDHy	MDH	PFLi	PYK
342	ADHEr	LDH_D	FRD2	MDH	PFLi	PGDHY	PYK
343	ADHEr	LDH_D	EDA	FRD2	MDH	PFLi	PYK
344	ADHEr	LDH_D	ACKr	ASPT	MDH	PYK	
345	ADHEr	LDH_D	ASPT	MDH	PTAr	PYK	
346	ADHEr	LDH_D	ACKr	ASPT	FUM	PYK	
347	ADHEr	LDH_D	ASPT	FUM	PTAr	PYK	
348	ADHEr	LDH_D	ATPS4r	GLUDy	MDH	PGI	
349	ADHEr	LDH_D	ATPS4r	FUM	GLUDy	PGI	
350	ADHEr	LDH_D	FBA	PFLi	SUCD4		
351	ADHEr	LDH_D	FRD2	PFK	PFLi		
352	ADHEr	LDH_D	PFLi	SUCD4	TPI		
353	ADHEr	LDH_D	FBA	FRD2	PFLi		
354	ADHEr	LDH_D	PFK	PFLi	SUCD4		
355	ADHEr	LDH_D	FRD2	PFLi	TPI		
356	ADHEr	LDH_D	NADH6	PFLi	TPI		
357	ADHEr	LDH_D	FBA	NADH6	PFLi		
358	ADHEr	LDH_D	NADH6	PFK	PFLi		
359	ADHEr	LDH_D	ASPT	MDH	PFLi	PGI	
360	ADHEr	LDH_D	ASPT	FUM	PFLi	PGI	
361	ADHEr	LDH_D	ASPT	GLUDy	MDH	PFLi	PYK
362	ADHEr	LDH_D	ASPT	FUM	GLUDy	PFLi	PYK
363	ADHEr	LDH_D	ASPT	ATPS4r	CBMK2	FUM	PYK
364	ADHEr	LDH_D	ASPT	MDH	PFLi	TPI	
365	ADHEr	LDH_D	ASPT	FUM	PFLi	TPI	
366	ADHEr	LDH_D	ASPT	FBA	MDH	PFLi	
367	ADHEr	LDH_D	ASPT	FBA	FUM	PFLi	
368	ADHEr	LDH_D	ASPT	MDH	PFK	PFLi	
369	ADHEr	LDH_D	ASPT	FUM	PFK	PFLi	
370	ADHEr	LDH_D	ACKr	NADH6	PGI	PYK	
371	ADHEr	LDH_D	NADH6	PGI	PTAr	PYK	
372	ADHEr	LDH_D	ASPT	ATPS4r	FUM	PYK	
373	ADHEr	LDH_D	ASPT	ATPS4r	MALS	MDH	PYK
374	ADHEr	LDH_D	ASPT	ATPS4r	ICL	MDH	PYK
375	ADHEr	LDH_D	GLUDy	PFLi	PGDH	PGI	
376	ADHEr	LDH_D	ATPS4r	GLUDy	PFLi	PGI	
377	ADHEr	LDH_D	FBA	GLUDy	PFLi		
378	ADHEr	LDH_D	GLUDy	PFLi	TPI		
379	ADHEr	LDH_D	GLUDy	PFK	PFLi		
380	ADHEr	LDH_D	GLUDy	PFLi	PGI	TAL	
381	ADHEr	LDH_D	GLUDy	PFLi	PGI	TKT1	
382	ADHEr	LDH_D	GLUDy	PFLi	PRO1z	PYK	SUCD4
383	ADHEr	LDH_D	GLUDy	MDH	NADH6	PFLi	PYK
384	ADHEr	LDH_D	GLUDy	MDH	PFLi	PYK	SUCD4

385	ADHEr	LDH_D	FUM	GLUDy	PFLi	PYK	SUCD4
386	ADHEr	LDH_D	FUM	GLUDy	NADH6	PFLi	PYK
387	ADHEr	LDH_D	GLUDy	PFLi	PGI		
388	ADHEr	LDH_D	EDA	MDH	PFLi	PYK	SUCD4
389	ADHEr	LDH_D	MDH	PFLi	PGDHY	PYK	SUCD4
390	ADHEr	LDH_D	MDH	PFLi	PGL	PYK	SUCD4
391	ADHEr	LDH_D	G6PDHy	MDH	PFLi	PYK	SUCD4
392	ADHEr	LDH_D	ATPS4r	GLUDy	MDH	NADH6	PYK
393	ADHEr	LDH_D	ATPS4r	FUM	GLUDy	NADH6	PYK
394	ADHEr	LDH_D	ACKr	AKGD	ATPS4r	GLUDy	PYK
395	ADHEr	LDH_D	AKGD	ATPS4r	GLUDy	PTAr	PYK
396	ADHEr	LDH_D	FRD2	PFLi	PYK		
397	ADHEr	LDH_D	ALAR	PFLi	PRO1z	PYK	SUCD4
398	ADHEr	LDH_D	DAAD	PFLi	PRO1z	PYK	SUCD4
399	ADHEr	LDH_D	PFLi	PGDH	PGI		
400	ADHEr	LDH_D	ATPS4r	PFLi	PGI		
401	ADHEr	LDH_D	ATPS4r	FUM	GLUDy	PFLi	PYK
402	ADHEr	LDH_D	ATPS4r	GLUDy	MDH	PFLi	PYK
403	ADHEr	LDH_D	PFLi	TPI			
404	ADHEr	LDH_D	FBA	PFLi			
405	ADHEr	LDH_D	PFK	PFLi			
406	ADHEr	LDH_D	ASPT	FUM	PFLi	PYK	
407	ADHEr	LDH_D	ASPT	MDH	PFLi	PYK	
408	ADHEr	LDH_D	PFLi	PGI	TKT1		
409	ADHEr	LDH_D	PFLi	PGI	TAL		
410	ADHEr	LDH_D	ASPT	ATPS4r	FUM	GLUDy	NADH6
411	ADHEr	LDH_D	ASPT	ATPS4r	GLUDy	MDH	NADH6
412	ADHEr	LDH_D	G6PDHy	ME2	PFLi	PYK	SUCD4
413	ADHEr	LDH_D	EDA	ME2	PFLi	PYK	SUCD4
414	ADHEr	LDH_D	ME2	PFLi	PGDHY	PYK	SUCD4
415	ADHEr	LDH_D	ME2	PFLi	PGL	PYK	SUCD4
416	ADHEr	LDH_D	MDH	NADH6	PFLi	PGDHY	PYK
417	ADHEr	LDH_D	G6PDHy	MDH	NADH6	PFLi	PYK
418	ADHEr	LDH_D	EDA	MDH	NADH6	PFLi	PYK
419	ADHEr	LDH_D	MDH	NADH6	PFLi	PGL	PYK
420	ADHEr	LDH_D	ASPT	ATPS4r	CBMK2	MDH	NADH6
421	ADHEr	LDH_D	ASPT	ATPS4r	CBMK2	FUM	NADH6
422	ADHEr	LDH_D	CBMK2	PFLi	PGI	RPE	
423	ADHEr	LDH_D	ASNS2	GLU5K	PFLi	PGI	RPE
424	ADHEr	LDH_D	ASNS2	G5SD	PFLi	PGI	RPE
425	ADHEr	LDH_D	ASPT	ATPS4r	GLUDy	MDH	PTAr
426	ADHEr	LDH_D	ASPT	ATPS4r	FUM	GLUDy	PTAr
427	ADHEr	LDH_D	PFLi	PGI			
428	ADHEr	LDH_D	ASPT	ATPS4r	FUM	GLUDy	
429	ADHEr	LDH_D	ASPT	ATPS4r	GLUDy	MDH	
430	ADHEr	LDH_D	ACKr	AKGD	ATPS4r	PYK	

431	ADHEr	LDH_D	AKGD	ATPS4r	PTAr	PYK
432	ADHEr	LDH_D	ASPT	ATPS4r	MDH	NADH6
433	ADHEr	LDH_D	ASPT	ATPS4r	FUM	NADH6
434	ADHEr	LDH_D	G6PDHy	GLCpts	GLUDy	PTAr
435	ADHEr	LDH_D	ACKr	GLCpts	GLUDy	PGL
436	ADHEr	LDH_D	GLCpts	GLUDy	PGDH	PTAr
437	ADHEr	LDH_D	GLCpts	GLUDy	PGL	PTAr
438	ADHEr	LDH_D	ACKr	G6PDHy	GLCpts	GLUDy
439	ADHEr	LDH_D	ACKr	GLCpts	GLUDy	PGDH
440	ADHEr	LDH_D	GLCpts	GLUDy	PTAr	TKT1
441	ADHEr	LDH_D	GLCpts	GLUDy	PTAr	TAL
442	ADHEr	LDH_D	ACKr	GLCpts	GLUDy	TKT1
443	ADHEr	LDH_D	ACKr	GLCpts	GLUDy	TAL
444	ADHEr	LDH_D	ACKr	GLCpts	GLUDy	RPE
445	ADHEr	LDH_D	GLCpts	GLUDy	PTAr	RPE
446	ADHEr	LDH_D	ACKr	GLCpts	GLUDy	TKT2
447	ADHEr	LDH_D	GLCpts	GLUDy	PTAr	TKT2
448	ADHEr	LDH_D	GLCpts	PGDH	PTAr	THD2
449	ADHEr	LDH_D	G6PDHy	GLCpts	PTAr	THD2
450	ADHEr	LDH_D	ACKr	G6PDHy	GLCpts	THD2
451	ADHEr	LDH_D	ACKr	GLCpts	PGL	THD2
452	ADHEr	LDH_D	ACKr	GLCpts	PGDH	THD2
453	ADHEr	LDH_D	GLCpts	PGL	PTAr	THD2
454	ADHEr	LDH_D	ACKr	GLCpts	THD2	TKT1
455	ADHEr	LDH_D	ACKr	GLCpts	TAL	THD2
456	ADHEr	LDH_D	GLCpts	PTAr	TAL	THD2
457	ADHEr	LDH_D	GLCpts	PTAr	THD2	TKT1
458	ADHEr	LDH_D	ASPT	ATPS4r	MDH	
459	ADHEr	LDH_D	ASPT	ATPS4r	FUM	
460	ADHEr	LDH_D	GLCpts	PTAr	RPE	THD2
461	ADHEr	LDH_D	ACKr	GLCpts	RPE	THD2
462	ADHEr	LDH_D	ACKr	ATPS4r	PYK	SUCOAS
463	ADHEr	LDH_D	ATPS4r	PTAr	PYK	SUCOAS
464	ADHEr	LDH_D	FRD2	GLCpts	GLUDy	PFLi
465	ADHEr	LDH_D	GLCpts	PTAr	THD2	TKT2
466	ADHEr	LDH_D	ACKr	GLCpts	THD2	TKT2
467	ADHEr	LDH_D	FRD2	GLCpts	PFLi	THD2
468	ADHEr	LDH_D	ACKr	GLUDy	PGDH	THD2
469	ADHEr	LDH_D	GLUDy	PGL	PTAr	THD2
470	ADHEr	LDH_D	G6PDHy	GLUDy	PTAr	THD2
471	ADHEr	LDH_D	GLUDy	PGDH	PTAr	THD2
472	ADHEr	LDH_D	ACKr	GLUDy	PGL	THD2
473	ADHEr	LDH_D	ACKr	G6PDHy	GLUDy	THD2
474	ADHEr	LDH_D	FRD2	GLUDy	PFLi	THD2
475	ADHEr	LDH_D	GLUDy	PTAr	THD2	TKT1
476	ADHEr	LDH_D	GLUDy	PTAr	TAL	THD2

477	ADHEr	LDH_D	ACKr	GLUDy	TAL	THD2
478	ADHEr	LDH_D	ACKr	GLUDy	THD2	TKT1
479	ADHEr	LDH_D	ACKr	GLCpts	PGDH	
480	ADHEr	LDH_D	ACKr	GLCpts	PGL	
481	ADHEr	LDH_D	GLCpts	PGDH	PTAr	
482	ADHEr	LDH_D	GLCpts	PGL	PTAr	
483	ADHEr	LDH_D	ACKr	G6PDHy	GLCpts	
484	ADHEr	LDH_D	G6PDHy	GLCpts	PTAr	
485	ADHEr	LDH_D	GLUDy	PTAr	RPE	THD2
486	ADHEr	LDH_D	ACKr	GLUDy	RPE	THD2
487	ADHEr	LDH_D	GLCpts	GLUDy	PTAr	
488	ADHEr	LDH_D	ACKr	GLCpts	GLUDy	
489	ADHEr	LDH_D	GLCpts	PTAr	TKT1	
490	ADHEr	LDH_D	GLCpts	PTAr	TAL	
491	ADHEr	LDH_D	ACKr	GLCpts	TAL	
492	ADHEr	LDH_D	ACKr	GLCpts	TKT1	
493	ADHEr	LDH_D	NADH6	PFLi	PTAr	PYK
494	ADHEr	LDH_D	ACKr	NADH6	PFLi	PYK
495	ADHEr	LDH_D	ACKr	GLUDy	THD2	TKT2
496	ADHEr	LDH_D	GLUDy	PTAr	THD2	TKT2
497	ADHEr	LDH_D	ACKr	GLCpts	RPE	
498	ADHEr	LDH_D	GLCpts	PTAr	RPE	
499	ADHEr	LDH_D	ACKr	GLCpts	TKT2	
500	ADHEr	LDH_D	GLCpts	PTAr	TKT2	
501	ADHEr	LDH_D	ACKr	GLUDy	PGDH	
502	ADHEr	LDH_D	GLUDy	PGL	PTAr	
503	ADHEr	LDH_D	ACKr	GLUDy	PGL	
504	ADHEr	LDH_D	ACKr	G6PDHy	GLUDy	
505	ADHEr	LDH_D	GLUDy	PGDH	PTAr	
506	ADHEr	LDH_D	G6PDHy	GLUDy	PTAr	
507	ADHEr	LDH_D	GLUDy	PTAr	TKT1	
508	ADHEr	LDH_D	ACKr	GLUDy	TKT1	
509	ADHEr	LDH_D	ACKr	GLUDy	TAL	
510	ADHEr	LDH_D	GLUDy	PTAr	TAL	
511	ADHEr	LDH_D	GLUDy	PTAr	RPE	
512	ADHEr	LDH_D	ACKr	GLUDy	RPE	
513	ADHEr	LDH_D	GLUDy	PTAr	TKT2	
514	ADHEr	LDH_D	ACKr	GLUDy	TKT2	
515	ADHEr	LDH_D	PGDH	PTAr	THD2	
516	ADHEr	LDH_D	ACKr	PGDH	THD2	
517	ADHEr	LDH_D	G6PDHy	PTAr	THD2	
518	ADHEr	LDH_D	PGL	PTAr	THD2	
519	ADHEr	LDH_D	ACKr	PGL	THD2	
520	ADHEr	LDH_D	ACKr	G6PDHy	THD2	
521	ADHEr	LDH_D	PTAr	TAL	THD2	
522	ADHEr	LDH_D	ACKr	THD2	TKT1	

523	ADHEr	LDH_D	ACKr	TAL	THD2	
524	ADHEr	LDH_D	PTAr	THD2	TKT1	
525	ADHEr	LDH_D	PTAr	RPE	THD2	
526	ADHEr	LDH_D	ACKr	RPE	THD2	
527	ADHEr	LDH_D	FRD2	GLUDy	PFLi	
528	ADHEr	LDH_D	GLUDy	PFLi	PRO1z	SUCD4
529	ADHEr	LDH_D	FRD2	GLCpts	PFLi	
530	ADHEr	LDH_D	PTAr	THD2	TKT2	
531	ADHEr	LDH_D	ACKr	THD2	TKT2	
532	ADHEr	LDH_D	ACKr	GLCpts		
533	ADHEr	LDH_D	GLCpts	PTAr		
534	ADHEr	LDH_D	FRD2	PFLi	THD2	
535	ADHEr	LDH_D	ATPS4r	FUM	GLUDy	
536	ADHEr	LDH_D	ATPS4r	GLUDy	MDH	
537	ADHEr	LDH_D	FUM	GLCpts	PFLi	SUCD4
538	ADHEr	LDH_D	GLCpts	MDH	PFLi	SUCD4
539	ADHEr	LDH_D	FUM	GLUDy	PFLi	SUCD4
540	ADHEr	LDH_D	GLUDy	MDH	PFLi	SUCD4
541	ADHEr	LDH_D	GLUDy	MDH	NADH6	PFLi
542	ADHEr	LDH_D	FUM	GLUDy	NADH6	PFLi
543	ADHEr	LDH_D	MDH	PFLi	SUCD4	THD2
544	ADHEr	LDH_D	FUM	PFLi	SUCD4	THD2
545	ADHEr	LDH_D	ASPT	FUM	GLCpts	PFLi
546	ADHEr	LDH_D	ASPT	GLCpts	MDH	PFLi
547	ADHEr	LDH_D	ASPT	FUM	GLUDy	PFLi
548	ADHEr	LDH_D	ASPT	GLUDy	MDH	PFLi
549	ADHEr	LDH_D	GLCpts	PFLi	SUCD4	THD2
550	ADHEr	LDH_D	PGDH	PTAr		
551	ADHEr	LDH_D	PGL	PTAr		
552	ADHEr	LDH_D	ACKr	PGL		
553	ADHEr	LDH_D	G6PDHy	PTAr		
554	ADHEr	LDH_D	ACKr	G6PDHy		
555	ADHEr	LDH_D	ACKr	PGDH		
556	ADHEr	LDH_D	ASPT	FUM	PFLi	THD2
557	ADHEr	LDH_D	ASPT	MDH	PFLi	THD2
558	ADHEr	LDH_D	ACKr	GLUDy		
559	ADHEr	LDH_D	GLUDy	PTAr		
560	ADHEr	LDH_D	PTAr	TAL		
561	ADHEr	LDH_D	ACKr	TAL		
562	ADHEr	LDH_D	ACKr	TKT1		
563	ADHEr	LDH_D	PTAr	TKT1		
564	ADHEr	LDH_D	ACKr	RPE		
565	ADHEr	LDH_D	PTAr	RPE		
566	ADHEr	LDH_D	GLCpts	GLUDy	PFLi	SUCD4
567	ADHEr	LDH_D	FUM	GLCpts	GLUDy	PFLi
568	ADHEr	LDH_D	GLCpts	GLUDy	MDH	PFLi

569	ADHEr	LDH_D	ACKr	TKT2		
570	ADHEr	LDH_D	PTAr	TKT2		
571	ADHEr	LDH_D	GLUDy	PFLi	SUCD4	THD2
572	ADHEr	LDH_D	FUM	GLUDy	PFLi	THD2
573	ADHEr	LDH_D	GLUDy	MDH	PFLi	THD2
574	ADHEr	LDH_D	GLCpts	GLUDy	NADH6	PFLi
575	ADHEr	LDH_D	ATPS4r	GLUDy	NADH6	PFLi
576	ADHEr	LDH_D	GLCpts	MDH	PFLi	THD2
577	ADHEr	LDH_D	FUM	GLCpts	PFLi	THD2
578	ADHEr	LDH_D	ACKr	CBMK2	FRD2	PFLi
579	ADHEr	LDH_D	CBMK2	FRD2	PFLi	PTAr
580	ADHEr	LDH_D	MDH	PTAr	SUCD4	
581	ADHEr	LDH_D	FRD2	MDH	PTAr	
582	ADHEr	LDH_D	ACKr	MDH	SUCD4	
583	ADHEr	LDH_D	ACKr	FRD2	MDH	
584	ADHEr	LDH_D	FDH2	MDH	NADH6	PTAr
585	ADHEr	LDH_D	ACKr	FDH2	MDH	NADH6
586	ADHEr	LDH_D	GLCpts	NADH6	PFLi	THD2
587	ADHEr	LDH_D	GLCpts	PFLi	SUCD4	
588	ADHEr	LDH_D	GLCpts	NADH12	NADH6	PFLi
589	ADHEr	LDH_D	ATPS4r	FUM	PGL	
590	ADHEr	LDH_D	ATPS4r	MDH	PGDH	
591	ADHEr	LDH_D	ATPS4r	FUM	PGDH	
592	ADHEr	LDH_D	ATPS4r	FUM	G6PDHy	
593	ADHEr	LDH_D	GLCpts	MDH	NADH6	PFLi
594	ADHEr	LDH_D	FUM	GLCpts	NADH6	PFLi
595	ADHEr	LDH_D	FRD2	PFLi		
596	ADHEr	LDH_D	ALAR	PFLi	PRO1z	SUCD4
597	ADHEr	LDH_D	DAAD	PFLi	PRO1z	SUCD4
598	ADHEr	LDH_D	ACKr			
599	ADHEr	LDH_D	PTAr			
600	ADHEr	LDH_D	FUM	PFLi	SUCD4	
601	ADHEr	LDH_D	MDH	PFLi	SUCD4	
602	ADHEr	LDH_D	FUM	NADH12	NADH6	PFLi
603	ADHEr	LDH_D	MDH	NADH12	NADH6	PFLi
604	ADHEr	LDH_D	ATPS4r	MDH	TKT1	
605	ADHEr	LDH_D	ATPS4r	FUM	TKT1	
606	ADHEr	LDH_D	ATPS4r	MDH	TAL	
607	ADHEr	LDH_D	ATPS4r	FUM	TAL	
608	ADHEr	LDH_D	ATPS4r	NADH6	PFLi	PYK
609	ADHEr	LDH_D	ASPT	FUM	PFLi	
610	ADHEr	LDH_D	ASPT	MDH	PFLi	
611	ADHEr	LDH_D	ATPS4r	MDH	RPE	
612	ADHEr	LDH_D	ATPS4r	FUM	RPE	
613	ADHEr	LDH_D	PFLi	SUCD4	THD2	
614	ADHEr	LDH_D	NADH12	NADH6	PFLi	THD2

615	ADHEr	LDH_D	FUM	NADH6	PFLi	THD2
616	ADHEr	LDH_D	MDH	NADH6	PFLi	THD2
617	ADHEr	LDH_D	ATPS4r	MDH	TKT2	
618	ADHEr	LDH_D	ATPS4r	FUM	TKT2	
619	ADHEr	LDH_D	GLCpts	NADH6	PFLi	
620	ADHEr	LDH_D	GLUDy	NADH6	PFLi	THD2
621	ADHEr	LDH_D	GLUDy	PFLi	SUCD4	
622	ADHEr	LDH_D	GLUDy	NADH12	NADH6	PFLi
623	ADHEr	LDH_D	FUM	GLUDy	PFLi	
624	ADHEr	LDH_D	GLUDy	MDH	PFLi	
625	ADHEr	LDH_D	ATPS4r	FUM	NADH6	
626	ADHEr	LDH_D	ATPS4r	MDH	NADH6	
627	ADHEr	LDH_D	ATPS4r	G6PDHy	GLUDy	NADH6
628	ADHEr	LDH_D	ATPS4r	GLUDy	NADH6	PGDH
629	ADHEr	LDH_D	ATPS4r	GLUDy	NADH6	PGL
630	ADHEr	LDH_D	ATPS4r	MDH	PFLi	THD2
631	ADHEr	LDH_D	ATPS4r	FUM	PFLi	THD2
632	ADHEr	LDH_D	ATPS4r	GLUDy	NADH6	TKT1
633	ADHEr	LDH_D	ATPS4r	GLUDy	NADH6	TAL
634	ADHEr	LDH_D	ATPS4r	GLUDy	PFLi	THD2
635	ADHEr	LDH_D	GLCpts	MDH	PFLi	
636	ADHEr	LDH_D	FUM	GLCpts	PFLi	
637	ADHEr	LDH_D	GLUDy	NADH6	PFLi	
638	ADHEr	LDH_D	ATPS4r	GLUDy	NADH6	RPE
639	ADHEr	LDH_D	ATPS4r	GLUDy	NADH6	TKT2
640	ADHEr	LDH_D	FUM	PFLi	THD2	
641	ADHEr	LDH_D	MDH	PFLi	THD2	
642	ADHEr	LDH_D	NADH6	PFLi	THD2	
643	ADHEr	LDH_D	PFLi	SUCD4		
644	ADHEr	LDH_D	NADH12	NADH6	PFLi	
645	ADHEr	LDH_D	ATPS4r	NADH6	PFLi	
646	ADHEr	LDH_D	FUM	NADH6	PFLi	
647	ADHEr	LDH_D	MDH	NADH6	PFLi	
648	ADHEr	LDH_D	ATPS4r	NADH6	PGL	
649	ADHEr	LDH_D	ATPS4r	NADH6	PGDH	
650	ADHEr	LDH_D	ATPS4r	G6PDHy	NADH6	
651	ADHEr	LDH_D	ATPS4r	NADH6	TAL	
652	ADHEr	LDH_D	ATPS4r	NADH6	TKT1	
653	ADHEr	LDH_D	CBMK2	GLU5K	NADH6	PFLi
654	ADHEr	LDH_D	CBMK2	G5SD	NADH6	PFLi
655	ADHEr	LDH_D	ASNS2	CBMK2	NADH6	PFLi
656	ADHEr	LDH_D	ATPS4r	PFLi	THD2	
657	ADHEr	LDH_D	NADH6	PFLi		
658	ADHEr	LDH_D	ATPS4r	NADH6	RPE	
659	ADHEr	LDH_D	ATPS4r	NADH6	TKT2	
660	ADHEr	LDH_D	CBMK2	FUM	G5SD	PFLi

661	ADHEr	LDH_D	CBMK2	GLU5K	MDH	PFLi
662	ADHEr	LDH_D	CBMK2	FUM	GLU5K	PFLi
663	ADHEr	LDH_D	CBMK2	G5SD	MDH	PFLi
664	ADHEr	LDH_D	ASNS2	CBMK2	FUM	PFLi
665	ADHEr	LDH_D	ASNS2	CBMK2	MDH	PFLi
666	ADHEr	LDH_D	MDH	PFLi		
667	ADHEr	LDH_D	FUM	PFLi		
668	ADHEr	LDH_D	ATPS4r	GLUDy	PFLi	RPE
669	ADHEr	LDH_D	ATPS4r	GLUDy	PFLi	TAL
670	ADHEr	LDH_D	ATPS4r	GLUDy	PFLi	TKT1
671	ADHEr	LDH_D	ATPS4r	GLUDy	PFLi	TKT2
672	ADHEr	LDH_D	ATPS4r	GLUDy	PFLi	
673	ADHEr	LDH_D	ATPS4r	GLUDy	NADH6	
674	ADHEr	LDH_D	ATPS4r	PFLi	RPE	
675	ADHEr	LDH_D	ATPS4r	PFLi	TAL	
676	ADHEr	LDH_D	ATPS4r	PFLi	TKT1	
677	ADHEr	LDH_D	ATPS4r	PFLi	TKT2	
678	ADHEr	LDH_D	ATPS4r	CBMK2	PFLi	
679	ADHEr	LDH_D	ATPS4r	PFLi		
680	ADHEr	LDH_D	ASPT	MDH	PGDHY	PYK
681	ADHEr	LDH_D	ASPT	EDA	MDH	PYK
682	ADHEr	LDH_D	ATPS4r	CBMK2	NADH6	
683	ADHEr	LDH_D	ATPS4r	NADH6		
684	ADHEr	LDH_D	ATPS4r	HEX1	PGI	PPS
685	ADHEr	LDH_D	G6PDHy	ME2	THD2	
686	ADHEr	LDH_D	ME2	PGL	THD2	
687	ADHEr	LDH_D	ME2	PGDH	PGDHY	THD2
688	ADHEr	LDH_D	EDA	ME2	PGDH	THD2
689	ADHEr	LDH_D	EDA	ME2	TAL	THD2
690	ADHEr	LDH_D	ME2	PGDHY	TAL	THD2
691	ADHEr	LDH_D	ME2	PGDHY	THD2	TKT1
692	ADHEr	LDH_D	EDA	ME2	THD2	TKT1
693	ADHEr	LDH_D	ME2	PGDHY	RPE	THD2
694	ADHEr	LDH_D	EDA	ME2	RPE	THD2
695	ADHEr	LDH_D	MDH	PGL	THD2	
696	ADHEr	LDH_D	G6PDHy	MDH	THD2	
697	ADHEr	LDH_D	EDA	MDH	PGDH	THD2
698	ADHEr	LDH_D	MDH	PGDH	PGDHY	THD2
699	ADHEr	LDH_D	ME2	PGDHY	THD2	TKT2
700	ADHEr	LDH_D	EDA	ME2	THD2	TKT2
701	ADHEr	LDH_D	MDH	PGDHY	THD2	TKT1
702	ADHEr	LDH_D	EDA	MDH	THD2	TKT1
703	ADHEr	LDH_D	MDH	PGDHY	TAL	THD2
704	ADHEr	LDH_D	EDA	MDH	TAL	THD2
705	ADHEr	LDH_D	ATPS4r	GLUDy	HEX1	PGI
706	ADHEr	LDH_D	MDH	PGDHY	RPE	THD2

707	ADHEr	LDH_D	EDA	MDH	RPE	THD2
708	ADHEr	LDH_D	MDH	PGDHY	THD2	TKT2
709	ADHEr	LDH_D	EDA	MDH	THD2	TKT2
710	ADHEr	LDH_D	ATPS4r	HEX1	PGI	
711	ADHEr	LDH_D	FRD2	HEX1	MDH	PGI
712	ADHEr	LDH_D	HEX1	MDH	PGI	SUCD4
713	ADHEr	LDH_D	HEX1	PGI	SUCOAS	
714	ADHEr	LDH_D	HEX1	MDH	NADH6	PGI
715	ADHEr	LDH_D	FUM	HEX1	NADH6	PGI
716	ADHEr	LDH_D	FRD2	FUM	HEX1	PGI
717	ADHEr	LDH_D	HEX1	PGI		
718	ADHEr	LDH_D	SUCOAS	THD2		
719	ADHEr	LDH_D	THD2			
720	ADHEr	LDH_D	GLCpts	SUCOAS	TKT2	TPI
721	ADHEr	LDH_D	GLCpts	PFK	SUCOAS	TKT2
722	ADHEr	LDH_D	FBA	GLCpts	SUCOAS	TKT2
723	ADHEr	LDH_D	GLCpts	GLUDy	TKT2	TPI
724	ADHEr	LDH_D	FBA	GLCpts	GLUDy	TKT2
725	ADHEr	LDH_D	GLCpts	GLUDy	PFK	TKT2
726	ADHEr	LDH_D	GLCpts	PGI	SUCOAS	
727	ADHEr	LDH_D	GLCpts	GLUDy	PGI	
728	ADHEr	LDH_D	GLCpts	PFK	RPE	SUCOAS
729	ADHEr	LDH_D	GLCpts	RPE	SUCOAS	TPI
730	ADHEr	LDH_D	FBA	GLCpts	RPE	SUCOAS
731	ADHEr	LDH_D	GLCpts	GLUDy	RPE	TPI
732	ADHEr	LDH_D	FBA	GLCpts	GLUDy	RPE
733	ADHEr	LDH_D	GLCpts	GLUDy	PFK	RPE
734	ADHEr	LDH_D	FBA	GLUDy	SUCOAS	TKT2
735	ADHEr	LDH_D	GLUDy	PFK	SUCOAS	TKT2
736	ADHEr	LDH_D	GLUDy	SUCOAS	TKT2	TPI
737	ADHEr	LDH_D	GLCpts	GLUDy	PFK	SUCOAS
738	ADHEr	LDH_D	GLCpts	GLUDy	SUCOAS	TPI
739	ADHEr	LDH_D	FBA	GLCpts	GLUDy	SUCOAS
740	ADHEr	LDH_D	GLCpts	PFK	TKT2	
741	ADHEr	LDH_D	FBA	GLCpts	TKT2	
742	ADHEr	LDH_D	GLCpts	TKT2	TPI	
743	ADHEr	LDH_D	GLUDy	PGI	SUCOAS	
744	ADHEr	LDH_D	PGDHY	PGI		
745	ADHEr	LDH_D	EDA	PGI		
746	ADHEr	LDH_D	GLCpts	PGI		
747	ADHEr	LDH_D	GLUDy	PFK	RPE	SUCOAS
748	ADHEr	LDH_D	GLUDy	RPE	SUCOAS	TPI
749	ADHEr	LDH_D	FBA	GLUDy	RPE	SUCOAS
750	ADHEr	LDH_D	GLCpts	RPE	TPI	
751	ADHEr	LDH_D	GLCpts	PFK	RPE	
752	ADHEr	LDH_D	FBA	GLCpts	RPE	

753	ADHEr	LDH_D	PFK	SUCOAS	TKT2	
754	ADHEr	LDH_D	FBA	SUCOAS	TKT2	
755	ADHEr	LDH_D	SUCOAS	TKT2	TPI	
756	ADHEr	LDH_D	GLCpts	SUCOAS	TPI	
757	ADHEr	LDH_D	GLCpts	PFK	SUCOAS	
758	ADHEr	LDH_D	FBA	GLCpts	SUCOAS	
759	ADHEr	LDH_D	FBA	GLCpts	GLUDy	
760	ADHEr	LDH_D	GLCpts	GLUDy	TPI	
761	ADHEr	LDH_D	GLCpts	GLUDy	PFK	
762	ADHEr	LDH_D	GLUDy	PFK	TKT2	
763	ADHEr	LDH_D	FBA	GLUDy	TKT2	
764	ADHEr	LDH_D	GLUDy	TKT2	TPI	
765	ADHEr	LDH_D	PGI	SUCOAS		
766	ADHEr	LDH_D	GLUDy	PGI		
767	ADHEr	LDH_D	ASPT	G6PDHy	MDH	PYK
768	ADHEr	LDH_D	ASPT	MDH	PGL	PYK
769	ADHEr	LDH_D	FBA	RPE	SUCOAS	
770	ADHEr	LDH_D	PFK	RPE	SUCOAS	
771	ADHEr	LDH_D	RPE	SUCOAS	TPI	
772	ADHEr	LDH_D	HEX1	PFK	SUCOAS	TKT1
773	ADHEr	LDH_D	FBA	HEX1	SUCOAS	TAL
774	ADHEr	LDH_D	HEX1	PFK	SUCOAS	TAL
775	ADHEr	LDH_D	HEX1	SUCOAS	TKT1	TPI
776	ADHEr	LDH_D	FBA	HEX1	SUCOAS	TKT1
777	ADHEr	LDH_D	HEX1	SUCOAS	TAL	TPI
778	ADHEr	LDH_D	GLUDy	RPE	TPI	
779	ADHEr	LDH_D	FBA	GLUDy	RPE	
780	ADHEr	LDH_D	GLUDy	PFK	RPE	
781	ADHEr	LDH_D	GLUDy	HEX1	TKT1	TPI
782	ADHEr	LDH_D	GLUDy	HEX1	PFK	TKT1
783	ADHEr	LDH_D	FBA	GLUDy	HEX1	TKT1
784	ADHEr	LDH_D	GLUDy	HEX1	TAL	TPI
785	ADHEr	LDH_D	FBA	GLUDy	HEX1	TAL
786	ADHEr	LDH_D	GLUDy	HEX1	PFK	TAL
787	ADHEr	LDH_D	GLUDy	SUCOAS	TPI	
788	ADHEr	LDH_D	GLUDy	PFK	SUCOAS	
789	ADHEr	LDH_D	FBA	GLUDy	SUCOAS	
790	ADHEr	LDH_D	FRD2	PYK	SUCOAS	TKT2
791	ADHEr	LDH_D	PYK	SUCD4	SUCOAS	TKT2
792	ADHEr	LDH_D	GLCpts	TPI		
793	ADHEr	LDH_D	GLCpts	PFK		
794	ADHEr	LDH_D	FBA	GLCpts		
795	ADHEr	LDH_D	FRD2	GLUDy	PYK	TKT2
796	ADHEr	LDH_D	GLUDy	PYK	SUCD4	TKT2
797	ADHEr	LDH_D	PFK	TKT2		
798	ADHEr	LDH_D	FBA	TKT2		

799	ADHEr	LDH_D	TKT2	TPI		
800	ADHEr	LDH_D	CBMK2	SUCOAS	TAL	TPI
801	ADHEr	LDH_D	CBMK2	FBA	SUCOAS	TAL
802	ADHEr	LDH_D	CBMK2	FBA	SUCOAS	TKT1
803	ADHEr	LDH_D	CBMK2	PFK	SUCOAS	TAL
804	ADHEr	LDH_D	CBMK2	PFK	SUCOAS	TKT1
805	ADHEr	LDH_D	CBMK2	SUCOAS	TKT1	TPI
806	ADHEr	LDH_D	CBMK2	FBA	HEX1	SUCOAS
807	ADHEr	LDH_D	CBMK2	HEX1	SUCOAS	TPI
808	ADHEr	LDH_D	CBMK2	HEX1	PFK	SUCOAS
809	ADHEr	LDH_D	PGI			
810	ADHEr	LDH_D	HEX1	PFK	TAL	
811	ADHEr	LDH_D	HEX1	TAL	TPI	
812	ADHEr	LDH_D	FBA	HEX1	TAL	
813	ADHEr	LDH_D	HEX1	PFK	TKT1	
814	ADHEr	LDH_D	HEX1	TKT1	TPI	
815	ADHEr	LDH_D	FBA	HEX1	TKT1	
816	ADHEr	LDH_D	PYK	RPE	SUCD4	SUCOAS
817	ADHEr	LDH_D	FRD2	PYK	RPE	SUCOAS
818	ADHEr	LDH_D	FRD2	GLUDy	PYK	RPE
819	ADHEr	LDH_D	GLUDy	PYK	RPE	SUCD4
820	ADHEr	LDH_D	RPE	TPI		
821	ADHEr	LDH_D	PFK	RPE		
822	ADHEr	LDH_D	FBA	RPE		
823	ADHEr	LDH_D	SUCOAS	TPI		
824	ADHEr	LDH_D	PFK	SUCOAS		
825	ADHEr	LDH_D	FBA	SUCOAS		
826	ADHEr	LDH_D	GLUDy	TPI		
827	ADHEr	LDH_D	FBA	GLUDy		
828	ADHEr	LDH_D	GLUDy	PFK		
829	ADHEr	LDH_D	FRD2	GLUDy	PYK	SUCOAS
830	ADHEr	LDH_D	GLUDy	PYK	SUCD4	SUCOAS
831	ADHEr	LDH_D	HEX1	MDH	PFK	SUCD4
832	ADHEr	LDH_D	HEX1	MDH	SUCD4	TPI
833	ADHEr	LDH_D	FBA	HEX1	MDH	SUCD4
834	ADHEr	LDH_D	FRD2	HEX1	MDH	TPI
835	ADHEr	LDH_D	FBA	FRD2	HEX1	MDH
836	ADHEr	LDH_D	FRD2	HEX1	MDH	PFK
837	ADHEr	LDH_D	FRD2	MDH	TKT1	TPI
838	ADHEr	LDH_D	FRD2	MDH	TAL	TPI
839	ADHEr	LDH_D	MDH	PFK	SUCD4	TKT1
840	ADHEr	LDH_D	MDH	PFK	SUCD4	TAL
841	ADHEr	LDH_D	FBA	MDH	SUCD4	TKT1
842	ADHEr	LDH_D	FBA	MDH	SUCD4	TAL
843	ADHEr	LDH_D	MDH	SUCD4	TAL	TPI
844	ADHEr	LDH_D	FRD2	MDH	PFK	TKT1

845	ADHEr	LDH_D	FRD2	MDH	PFK	TAL
846	ADHEr	LDH_D	FBA	FRD2	MDH	TAL
847	ADHEr	LDH_D	MDH	SUCD4	TKT1	TPI
848	ADHEr	LDH_D	FBA	FRD2	MDH	TKT1
849	ADHEr	LDH_D	PYK	SUCD4	TKT2	
850	ADHEr	LDH_D	FRD2	PYK	TKT2	
851	ADHEr	LDH_D	FDH2	NADH6	PYK	TKT2
852	ADHEr	LDH_D	CBMK2	PFK	TAL	
853	ADHEr	LDH_D	CBMK2	TAL	TPI	
854	ADHEr	LDH_D	CBMK2	FBA	TKT1	
855	ADHEr	LDH_D	CBMK2	TKT1	TPI	
856	ADHEr	LDH_D	CBMK2	FBA	TAL	
857	ADHEr	LDH_D	CBMK2	PFK	TKT1	
858	ADHEr	LDH_D	CBMK2	HEX1	PFK	
859	ADHEr	LDH_D	CBMK2	HEX1	TPI	
860	ADHEr	LDH_D	CBMK2	FBA	HEX1	
861	ADHEr	LDH_D	GLU5K	TAL	TPI	
862	ADHEr	LDH_D	G5SD	TAL	TPI	
863	ADHEr	LDH_D	FBA	GLU5K	TKT1	
864	ADHEr	LDH_D	G5SD	TKT1	TPI	
865	ADHEr	LDH_D	G5SD	PFK	TKT1	
866	ADHEr	LDH_D	GLU5K	PFK	TAL	
867	ADHEr	LDH_D	FBA	G5SD	TAL	
868	ADHEr	LDH_D	FBA	G5SD	TKT1	
869	ADHEr	LDH_D	G5SD	PFK	TAL	
870	ADHEr	LDH_D	GLU5K	TKT1	TPI	
871	ADHEr	LDH_D	GLU5K	PFK	TKT1	
872	ADHEr	LDH_D	FBA	GLU5K	TAL	
873	ADHEr	LDH_D	GLU5K	HEX1	TPI	
874	ADHEr	LDH_D	GLU5K	HEX1	PFK	
875	ADHEr	LDH_D	G5SD	HEX1	PFK	
876	ADHEr	LDH_D	FBA	G5SD	HEX1	
877	ADHEr	LDH_D	FBA	GLU5K	HEX1	
878	ADHEr	LDH_D	G5SD	HEX1	TPI	
879	ADHEr	LDH_D	ASNS2	PFK	TKT1	
880	ADHEr	LDH_D	ASNS2	TKT1	TPI	
881	ADHEr	LDH_D	ASNS2	PFK	TAL	
882	ADHEr	LDH_D	ASNS2	FBA	TKT1	
883	ADHEr	LDH_D	ASNS2	FBA	TAL	
884	ADHEr	LDH_D	ASNS2	TAL	TPI	
885	ADHEr	LDH_D	ASNS2	HEX1	PFK	
886	ADHEr	LDH_D	ASNS2	FBA	HEX1	
887	ADHEr	LDH_D	ASNS2	HEX1	TPI	
888	ADHEr	LDH_D	PYK	SUCD4	SUCOAS	TKT1
889	ADHEr	LDH_D	FRD2	PYK	SUCOAS	TAL
890	ADHEr	LDH_D	PYK	SUCD4	SUCOAS	TAL

891	ADHEr	LDH_D	FRD2	PYK	SUCOAS	TKT1
892	ADHEr	LDH_D	PYK	RPE	SUCD4	
893	ADHEr	LDH_D	FRD2	PYK	RPE	
894	ADHEr	LDH_D	FDH2	NADH6	PYK	RPE
895	ADHEr	LDH_D	GLUDy	MDH	PYK	TKT2
896	ADHEr	LDH_D	FUM	GLUDy	PYK	TKT2
897	ADHEr	LDH_D	GLCpts	GLUDy	SUCOAS	TKT2
898	ADHEr	LDH_D	GLUDy	PYK	SUCD4	
899	ADHEr	LDH_D	FRD2	GLUDy	PYK	
900	ADHEr	LDH_D	FDH2	GLUDy	NADH6	PYK
901	ADHEr	LDH_D	FBA			
902	ADHEr	LDH_D	TPI			
903	ADHEr	LDH_D	PFK			
904	ADHEr	LDH_D	PYK	SUCD4	SUCOAS	
905	ADHEr	LDH_D	FRD2	PYK	SUCOAS	
906	ADHEr	LDH_D	FDH2	NADH6	PYK	SUCOAS
907	ADHEr	LDH_D	FRD2	ME2	PGDHY	PYK
908	ADHEr	LDH_D	EDA	FRD2	ME2	PYK
909	ADHEr	LDH_D	FRD2	ME2	PGL	PYK
910	ADHEr	LDH_D	EDA	ME2	PYK	SUCD4
911	ADHEr	LDH_D	ME2	PGDHY	PYK	SUCD4
912	ADHEr	LDH_D	ME2	PGL	PYK	SUCD4
913	ADHEr	LDH_D	FRD2	G6PDHy	ME2	PYK
914	ADHEr	LDH_D	G6PDHy	ME2	PYK	SUCD4
915	ADHEr	LDH_D	MDH	NADH6	PGDHY	PYK
916	ADHEr	LDH_D	MDH	PGL	PYK	SUCD4
917	ADHEr	LDH_D	FRD2	MDH	PGL	PYK
918	ADHEr	LDH_D	FRD2	MDH	PGDHY	PYK
919	ADHEr	LDH_D	G6PDHy	MDH	PYK	SUCD4
920	ADHEr	LDH_D	MDH	NADH6	PGL	PYK
921	ADHEr	LDH_D	EDA	FRD2	MDH	PYK
922	ADHEr	LDH_D	EDA	MDH	PYK	SUCD4
923	ADHEr	LDH_D	MDH	PGDHY	PYK	SUCD4
924	ADHEr	LDH_D	EDA	MDH	NADH6	PYK
925	ADHEr	LDH_D	FRD2	G6PDHy	MDH	PYK
926	ADHEr	LDH_D	G6PDHy	MDH	NADH6	PYK
927	ADHEr	LDH_D	GLUDy	MDH	PYK	RPE
928	ADHEr	LDH_D	FUM	GLUDy	PYK	RPE
929	ADHEr	LDH_D	FRD2	PYK	TAL	
930	ADHEr	LDH_D	PYK	SUCD4	TKT1	
931	ADHEr	LDH_D	PYK	SUCD4	TAL	
932	ADHEr	LDH_D	FRD2	PYK	TKT1	
933	ADHEr	LDH_D	FDH2	NADH6	PYK	TAL
934	ADHEr	LDH_D	FDH2	NADH6	PYK	TKT1
935	ADHEr	LDH_D	GLCpts	GLUDy	RPE	SUCOAS
936	ADHEr	LDH_D	GLUDy	MDH	PYK	SUCOAS

937	ADHEr	LDH_D	FUM	GLUDy	PYK	SUCOAS
938	ADHEr	LDH_D	FUM	GLUDy	NADH6	PYK
939	ADHEr	LDH_D	GLUDy	MDH	NADH6	PYK
940	ADHEr	LDH_D	GLCpts	SUCOAS	TKT2	
941	ADHEr	LDH_D	GLUDy	SUCOAS	TKT2	
942	ADHEr	LDH_D	ASPT	MDH	PYK	TKT2
943	ADHEr	LDH_D	ASPT	FUM	PYK	TKT2
944	ADHEr	LDH_D	FRD2	PYK		
945	ADHEr	LDH_D	PYK	SUCD4		
946	ADHEr	LDH_D	FDH2	NADH6	PYK	
947	ADHEr	LDH_D	GLCpts	GLUDy	TKT2	
948	ADHEr	LDH_D	GLCpts	GLUDy	SUCOAS	TAL
949	ADHEr	LDH_D	GLCpts	GLUDy	SUCOAS	TKT1
950	ADHEr	LDH_D	FUM	GLUDy	PYK	
951	ADHEr	LDH_D	GLUDy	MDH	PYK	
952	ADHEr	LDH_D	GLCpts	RPE	SUCOAS	
953	ADHEr	LDH_D	ASPT	FUM	PYK	RPE
954	ADHEr	LDH_D	ASPT	MDH	PYK	RPE
955	ADHEr	LDH_D	GLUDy	RPE	SUCOAS	
956	ADHEr	LDH_D	GLCpts	GLUDy	RPE	
957	ADHEr	LDH_D	ASPT	FUM	PYK	SUCOAS
958	ADHEr	LDH_D	ASPT	MDH	PYK	SUCOAS
959	ADHEr	LDH_D	GLCpts	GLUDy	SUCOAS	
960	ADHEr	LDH_D	ASPT	FUM	NADH6	PYK
961	ADHEr	LDH_D	ASPT	MDH	NADH6	PYK
962	ADHEr	LDH_D	SUCOAS	TKT2		
963	ADHEr	LDH_D	GLCpts	TKT2		
964	ADHEr	LDH_D	ASPT	MDH	PYK	TKT1
965	ADHEr	LDH_D	ASPT	FUM	PYK	TAL
966	ADHEr	LDH_D	ASPT	MDH	PYK	TAL
967	ADHEr	LDH_D	ASPT	FUM	PYK	TKT1
968	ADHEr	LDH_D	GLCpts	SUCOAS	TAL	
969	ADHEr	LDH_D	GLCpts	SUCOAS	TKT1	
970	ADHEr	LDH_D	GLUDy	TKT2		
971	ADHEr	LDH_D	GLCpts	GLUDy	TKT1	
972	ADHEr	LDH_D	GLCpts	GLUDy	TAL	
973	ADHEr	LDH_D	GLUDy	SUCOAS	TKT1	
974	ADHEr	LDH_D	GLUDy	SUCOAS	TAL	
975	ADHEr	LDH_D	ASPT	MDH	PYK	
976	ADHEr	LDH_D	ASPT	FUM	PYK	
977	ADHEr	LDH_D	RPE	SUCOAS		
978	ADHEr	LDH_D	GLCpts	RPE		
979	ADHEr	LDH_D	GLCpts	SUCOAS		
980	ADHEr	LDH_D	GLUDy	RPE		
981	ADHEr	LDH_D	GLCpts	GLUDy		
982	ADHEr	LDH_D	GLUDy	SUCOAS		

983 ADHEr LDH_D TKT2
 984 ADHEr LDH_D GLCpts TAL
 985 ADHEr LDH_D GLCpts TKT1
 986 ADHEr LDH_D SUCOAS TAL
 987 ADHEr LDH_D SUCOAS TKT1
 988 ADHEr LDH_D GLUDy TKT1
 989 ADHEr LDH_D GLUDy TAL
 990 ADHEr LDH_D RPE
 991 ADHEr LDH_D GLCpts
 992 ADHEr LDH_D SUCOAS
 993 ADHEr LDH_D GLUDy
 994 ADHEr LDH_D TAL
 995 ADHEr LDH_D TKT1

Table 2: A list of all the reaction stoichiometries and the associated genes known to be associated with the reactions identified for disruption in the strategies listed in Tables 1.

Reaction Abbreviation	Reaction Name	Reaction Stoichiometry	Assigned Genes
ACKr	acetate kinase	[c] : ac + atp <==> actp + adp	b2296, b3115
ADHEr	acetaldehyde-CoA dehydrogenase	[c] : accoa + (2) h + (2) nadh <==> coa + etoh + (2) nad	b1241
AKGD	2-oxoglutarate dehydrogenase	[c] : akg + coa + nad --> co2 + nadh + succoa	b0727, b0726, b0116
ALAR	alanine racemase	[c] : ala-L <==> ala-D	b4053
ASNS2	asparagine synthetase	[c] : asp-L + atp + nh4 --> amp + asn-L + h + ppi	b3744
ASPT	L-aspartase	[c] : asp-L --> fum + nh4	b4139
ATPS4r	ATP synthase (four protons for one ATP)	adp[c] + (4) h[e] + pi[c] <==> atp[c] + (3) h[c] + h2o[c]	b3738+b3736+b3737, b3739, b3734+b3732+b3735+b3733+b3731
CBMK2	Carbamate kinase	[c] : atp + co2 + nh4 --> adp + cbp + (2) h	b0323, b0521, b2874
DAAD	D-Amino acid dehydrogenase	[c] : ala-D + fad + h2o --> fadh2 + nh4 + pyr	b1189
EDA	2-dehydro-3-deoxy-phosphogluconate aldolase	[c] : 2ddg6p --> g3p + pyr	b1850
FBA	fructose-bisphosphate aldolase	[c] : fdp <==> dhap + g3p	b1773, b2097, b2925
FDH2	formate dehydrogenase (quinone-8: 2 protons)	for[c] + (3) h[c] + ubq8[c] --> co2[c] + (2) h[e] + ubq8h2[c]	b3893+b3894+b3892, b1476+b1475+b1474, b4079

FRD2	fumarate reductase	[c] : fum + mql8 --> mqn8 + succ	b4153+b4152+b4151+b4154
FUM	fumarase	[c] : fum + h2o <==> mal-L	b1612, b4122, b1611
G5SD	glutamate-5-semialdehyde dehydrogenase	[c] : glu5p + h + nadph --> glu5sa + nadp + pi	b0243
G6PDHy	glucose 6-phosphate dehydrogenase	[c] : g6p + nadp <==> 6pgl + h + nadph	b1852
GLCpts	D-glucose transport via PEP:Pyr PTS	glc-D[e] + pep[c] --> g6p[c] + pyr[c]	b1817, b1818, b2417, b1621, b2416, b1819, b1101, b2415
GLU5K	glutamate 5-kinase	[c] : atp + glu-L --> adp + glu5p	b0242
GLUDy	glutamate dehydrogenase (NADP)	[c] : glu-L + h2o + nadp <==> akg + h + nadph + nh4	b1761
HEX1	hexokinase (D-glucose:ATP)	[c] : atp + glc-D --> adp + g6p + h	b2388
ICL	Isocitrate lyase	[c] : icit --> glx + succ	b4015
LDH_D	D-lactate dehydrogenase	[c] : lac-D + nad <==> h + nadh + pyr	b2133, b1380
MALS	malate synthase	[c] : accoa + glx + h2o --> coa + h + mal-L	b4014, b2976
MDH	malate dehydrogenase	[c] : mal-L + nad <==> h + nadh + oaa	b3236
ME2	malic enzyme (NADP)	[c] : mal-L + nadp --> co2 + nadph + pyr	b2463
NADH12	NADH dehydrogenase (ubiquinone-8)	[c] : h + nadh + ubq8 --> nad + ubq8h2	b1109
NADH6	NADH dehydrogenase (ubiquinone-8 & 3.5 protons)	(4.5) h[c] + nadh[c] + ubq8[c] --> (3.5) h[e] + nad[c] + ubq8h2[c]	b2288+b2277+b2285+b2278+b2276+b2286+b2287+b2279+b2280+b2284+b2283+b2282+b2281
PFK	phosphofructokinase	[c] : atp + f6p --> adp + fdp + h	b3916, b1723
PFLi	pyruvate formate lyase	[c] : coa + pyr --> accoa + for	b3114, b3951+b3952, b0902+b2579+b0903
PGDH	phosphogluconate dehydrogenase	[c] : 6pgc + nadp --> co2 + nadph + ru5p-D	b2029
PGDHY	phosphogluconate dehydratase	[c] : 6pgc --> 2ddg6p + h2o	b1851
PGI	glucose-6-phosphate isomerase	[c] : g6p <==> f6p	b4025
PGL	6-phosphogluconolactonase	[c] : 6pgl + h2o --> 6pgc + h	b0767

PPS	phosphoenolpyruvate synthase	[c] : atp + h2o + pyr --> amp + (2) h + pep + pi	b1702
PRO1z	proline oxidase	[c] : fad + pro-L --> 1pyr5c + fadh2 + h	b1014
PTAr	phosphotransacetylase	[c] : accoa + pi <==> actp + coa	b2297
PYK	pyruvate kinase	[c] : adp + h + pep --> atp + pyr	b1854, b1676
RPE	ribulose 5-phosphate 3-epimerase	[c] : ru5p-D <==> xu5p-D	b4301, b3386
SUCD4	succinate dehydrogenase	[c] : fadh2 + ubq8 <==> fad + ubq8h2	b0723+b0721+b0724+b0722
SUCOAS	succinyl-CoA synthetase (ADP-forming)	[c] : atp + coa + succ <==> adp + pi + succoa	b0729+b0728
TAL	transaldolase	[c] : g3p + s7p <==> e4p + f6p	b2464, b0008
THD2	NAD(P) transhydrogenase	(2) h[e] + nadh[c] + nadp[c] --> (2) h[c] + nad[c] + nadph[c]	b1602+b1603
TKT1	transketolase	[c] : r5p + xu5p-D <==> g3p + s7p	b2935, b2465
TKT2	transketolase	[c] : e4p + xu5p-D <==> f6p + g3p	b2935, b2465
TPI	triose-phosphate isomerase	[c] : dhap <==> g3p	b3919

Table 3: List of the metabolite abbreviations, the corresponding names and locations of all the metabolites that participate in the reactions listed in Table 2.

Metabolite Abbreviation	Compartment	Metabolite Name
1pyr5c	Cytosol	1-Pyrroline-5-carboxylate
2ddg6p	Cytosol	2-Dehydro-3-deoxy-D-gluconate 6-phosphate
6pgc	Cytosol	6-Phospho-D-gluconate
6pgl	Cytosol	6-phospho-D-glucono-1,5-lactone
ac	Cytosol	Acetate
accoa	Cytosol	Acetyl-CoA
actp	Cytosol	Acetyl phosphate
adp	Cytosol	ADP
akg	Cytosol	2-Oxoglutarate
ala-D	Cytosol	D-Alanine
ala-L	Cytosol	L-Alanine
amp	Cytosol	AMP
asn-L	Cytosol	L-Asparagine
asp-L	Cytosol	L-Aspartate
atp	Cytosol	ATP
cbp	Cytosol	Carbamoyl phosphate
cit	Cytosol	Citrate
co2	Cytosol	CO2
coa	Cytosol	Coenzyme A
ctp	Cytosol	CTP
dha	Cytosol	Dihydroxyacetone
dhap	Cytosol	Dihydroxyacetone phosphate
e4p	Cytosol	D-Erythrose 4-phosphate
etoh	Cytosol	Ethanol
f6p	Cytosol	D-Fructose 6-phosphate
fad	Cytosol	FAD
fadh2	Cytosol	FADH2
fdp	Cytosol	D-Fructose 1,6-bisphosphate
for	Cytosol	Formate
fum	Cytosol	Fumarate
g3p	Cytosol	Glyceraldehyde 3-phosphate
g6p	Cytosol	D-Glucose 6-phosphate
glc-D	Cytosol	D-Glucose
glc-D[e]	Extra-organism	D-Glucose
glu5p	Cytosol	L-Glutamate 5-phosphate
glu5sa	Cytosol	L-Glutamate 5-semialdehyde
glu-L	Cytosol	L-Glutamate
glx	Cytosol	Glyoxylate
h	Cytosol	H+
h[e]	Extra-organism	H+
h2	Cytosol	H2
h2o	Cytosol	H2O
icit	Cytosol	Isocitrate

k	Cytosol	K+
lac-D	Cytosol	D-Lactate
mal-L	Cytosol	L-Malate
mgl8	Cytosol	Menaquinol 8
mqn8	Cytosol	Menaquinone 8
nad	Cytosol	Nicotinamide adenine dinucleotide
nadh	Cytosol	Nicotinamide adenine dinucleotide - reduced
nadp	Cytosol	Nicotinamide adenine dinucleotide phosphate
nadph	Cytosol	Nicotinamide adenine dinucleotide phosphate - reduced
nh4	Cytosol	Ammonium
o2	Cytosol	O2
oaa	Cytosol	Oxaloacetate
pep	Cytosol	Phosphoenolpyruvate
pi	Cytosol	Phosphate
ppi	Cytosol	Diphosphate
pro-L	Cytosol	L-Proline
pyr	Cytosol	Pyruvate
r5p	Cytosol	alpha-D-Ribose 5-phosphate
ru5p-D	Cytosol	D-Ribulose 5-phosphate
s7p	Cytosol	Sedoheptulose 7-phosphate
succ	Cytosol	Succinate
succoa	Cytosol	Succinyl-CoA
ubq8	Cytosol	Ubiquinone-8
ubq8h2	Cytosol	Ubiquinol-8
xu5p-D	Cytosol	D-Xylulose 5-phosphate

Example IV

Exemplary Hydrogenase and CO Dehydrogenase Enzymes for Extracting Reducing Equivalents from Syngas and Exemplary Reductive TCA Cycle Enzymes

Enzymes of the reductive TCA cycle useful in the non-naturally occurring microbial organisms of the present invention include one or more of ATP-citrate lyase and three CO₂-fixing enzymes: isocitrate dehydrogenase, alpha-ketoglutarate:ferredoxin oxidoreductase, pyruvate:ferredoxin oxidoreductase. The presence of ATP-citrate lyase or citrate lyase and alpha-ketoglutarate:ferredoxin oxidoreductase indicates the presence of an active reductive TCA cycle in an organism. Enzymes for each step of the reductive TCA cycle are shown below.

ATP-citrate lyase (ACL, EC 2.3.3.8), also called ATP citrate synthase, catalyzes the ATP-dependent cleavage of citrate to oxaloacetate and acetyl-CoA. ACL is an enzyme of the RTCA cycle that has been studied in green sulfur bacteria *Chlorobium limicola* and *Chlorobium tepidum*. The alpha(4)beta(4) heteromeric enzyme from *Chlorobium limicola* was cloned and

characterized in *E. coli* (Kanao et al., *Eur. J. Biochem.* 269:3409-3416 (2002)). The *C. limicola* enzyme, encoded by *aclAB*, is irreversible and activity of the enzyme is regulated by the ratio of ADP/ATP. A recombinant ACL from *Chlorobium tepidum* was also expressed in *E. coli* and the holoenzyme was reconstituted *in vitro*, in a study elucidating the role of the alpha and beta subunits in the catalytic mechanism (Kim and Tabita, *J. Bacteriol.* 188:6544-6552 (2006)). ACL enzymes have also been identified in *Balnearium lithotrophicum*, *Sulfurihydrogenibium subterraneum* and other members of the bacterial phylum *Aquificae* (Hugler et al., *Environ. Microbiol.* 9:81-92 (2007)). This activity has been reported in some fungi as well. Exemplary organisms include *Sordaria macrospora* (Nowrousian et al., *Curr. Genet.* 37:189-93 (2000)), *Aspergillus nidulans*, *Yarrowia lipolytica* (Hynes and Murray, *Eukaryotic Cell*, July: 1039-1048, (2010) and *Aspergillus niger* (Meijer et al. *J. Ind. Microbiol. Biotechnol.* 36:1275-1280 (2009)). Other candidates can be found based on sequence homology. Information related to these enzymes is tabulated below:

<u>Protein</u>	<u>GenBank ID</u>	<u>GI Number</u>	<u>Organism</u>
<i>aclA</i>	BAB21376.1	12407237	<i>Chlorobium limicola</i>
<i>aclB</i>	BAB21375.1	12407235	<i>Chlorobium limicola</i>
<i>aclA</i>	AAM72321.1	21647054	<i>Chlorobium tepidum</i>
<i>aclB</i>	AAM72322.1	21647055	<i>Chlorobium tepidum</i>
<i>aclA</i>	ABI50076.1	114054981	<i>Balnearium lithotrophicum</i>
<i>aclB</i>	ABI50075.1	114054980	<i>Balnearium lithotrophicum</i>
<i>aclA</i>	ABI50085.1	114055040	<i>Sulfurihydrogenibium subterraneum</i>
<i>aclB</i>	ABI50084.1	114055039	<i>Sulfurihydrogenibium subterraneum</i>
<i>aclA</i>	AAX76834.1	62199504	<i>Sulfurimonas denitrificans</i>
<i>aclB</i>	AAX76835.1	62199506	<i>Sulfurimonas denitrificans</i>
<i>acl1</i>	XP_504787.1	50554757	<i>Yarrowia lipolytica</i>
<i>acl2</i>	XP_503231.1	50551515	<i>Yarrowia lipolytica</i>
<i>SPBC1703.07</i>	NP_596202.1	19112994	<i>Schizosaccharomyces pombe</i>
<i>SPAC22A12.16</i>	NP_593246.1	19114158	<i>Schizosaccharomyces pombe</i>
<i>acl1</i>	CAB76165.1	7160185	<i>Sordaria macrospora</i>
<i>acl2</i>	CAB76164.1	7160184	<i>Sordaria macrospora</i>

<u>Protein</u>	<u>GenBank ID</u>	<u>GI Number</u>	<u>Organism</u>
<i>aclA</i>	CBF86850.1	259487849	<i>Aspergillus nidulans</i>
<i>aclB</i>	CBF86848	259487848	<i>Aspergillus nidulans</i>

In some organisms the conversion of citrate to oxaloacetate and acetyl-CoA proceeds through a citryl-CoA intermediate and is catalyzed by two separate enzymes, citryl-CoA synthetase (EC 6.2.1.18) and citryl-CoA lyase (EC 4.1.3.34) (Aoshima, M., *Appl. Microbiol. Biotechnol.* 75:249-255 (2007)). Citryl-CoA synthetase catalyzes the activation of citrate to citryl-CoA. The *Hydrogenobacter thermophilus* enzyme is composed of large and small subunits encoded by *ccsA* and *ccsB*, respectively (Aoshima et al., *Mol. Microbiol.* 52:751-761 (2004)). The citryl-CoA synthetase of *Aquifex aeolicus* is composed of alpha and beta subunits encoded by *sucC1* and *sucD1* (Hugler et al., *Environ. Microbiol.* 9:81-92 (2007)). Citryl-CoA lyase splits citryl-CoA into oxaloacetate and acetyl-CoA. This enzyme is a homotrimer encoded by *ccl* in *Hydrogenobacter thermophilus* (Aoshima et al., *Mol. Microbiol.* 52:763-770 (2004)) and *aq_150* in *Aquifex aeolicus* (Hugler et al., *supra* (2007)). The genes for this mechanism of converting citrate to oxaloacetate and citryl-CoA have also been reported recently in *Chlorobium tepidum* (Eisen et al., *PNAS* 99(14): 9509-14 (2002)).

<u>Protein</u>	<u>GenBank ID</u>	<u>GI Number</u>	<u>Organism</u>
<i>ccsA</i>	BAD17844.1	46849514	<i>Hydrogenobacter thermophilus</i>
<i>ccsB</i>	BAD17846.1	46849517	<i>Hydrogenobacter thermophilus</i>
<i>sucC1</i>	AAC07285	2983723	<i>Aquifex aeolicus</i>
<i>sucD1</i>	AAC07686	2984152	<i>Aquifex aeolicus</i>
<i>ccl</i>	BAD17841.1	46849510	<i>Hydrogenobacter thermophilus</i>
<i>aq_150</i>	AAC06486	2982866	<i>Aquifex aeolicus</i>
<i>CT0380</i>	NP_661284	21673219	<i>Chlorobium tepidum</i>
<i>CT0269</i>	NP_661173.1	21673108	<i>Chlorobium tepidum</i>
<i>CT1834</i>	AAM73055.1	21647851	<i>Chlorobium tepidum</i>

Oxaloacetate is converted into malate by malate dehydrogenase (EC 1.1.1.37), an enzyme which functions in both the forward and reverse direction. *S. cerevisiae* possesses three copies of malate

dehydrogenase, *MDH1* (McAlister-Henn and Thompson, *J. Bacteriol.* 169:5157-5166 (1987), *MDH2* (Minard and McAlister-Henn, *Mol. Cell. Biol.* 11:370-380 (1991); Gibson and McAlister-Henn, *J. Biol. Chem.* 278:25628-25636 (2003)), and *MDH3* (Steffan and McAlister-Henn, *J. Biol. Chem.* 267:24708-24715 (1992)), which localize to the mitochondrion, cytosol, and peroxisome, respectively. *E. coli* is known to have an active malate dehydrogenase encoded by *mdh*.

<u>Protein</u>	<u>GenBank ID</u>	<u>GI Number</u>	<u>Organism</u>
<i>MDH1</i>	NP_012838	6322765	<i>Saccharomyces cerevisiae</i>
<i>MDH2</i>	NP_014515	116006499	<i>Saccharomyces cerevisiae</i>
<i>MDH3</i>	NP_010205	6320125	<i>Saccharomyces cerevisiae</i>
<i>Mdh</i>	NP_417703.1	16131126	<i>Escherichia coli</i>

Fumarate hydratase (EC 4.2.1.2) catalyzes the reversible hydration of fumarate to malate. The three fumarases of *E. coli*, encoded by *fumA*, *fumB* and *fumC*, are regulated under different conditions of oxygen availability. FumB is oxygen sensitive and is active under anaerobic conditions. FumA is active under microanaerobic conditions, and FumC is active under aerobic growth conditions (Tseng et al., *J. Bacteriol.* 183:461-467 (2001); Woods et al., *Biochim. Biophys. Acta* 954:14-26 (1988); Guest et al., *J. Gen. Microbiol.* 131:2971-2984 (1985)). *S. cerevisiae* contains one copy of a fumarase-encoding gene, *FUM1*, whose product localizes to both the cytosol and mitochondrion (Sass et al., *J. Biol. Chem.* 278:45109-45116 (2003)). Additional fumarase enzymes are found in *Campylobacter jejuni* (Smith et al., *Int. J. Biochem. Cell. Biol.* 31:961-975 (1999)), *Thermus thermophilus* (Mizobata et al., *Arch. Biochem. Biophys.* 355:49-55 (1998)) and *Rattus norvegicus* (Kobayashi et al., *J. Biochem.* 89:1923-1931 (1981)). Similar enzymes with high sequence homology include *fumI* from *Arabidopsis thaliana* and *fumC* from *Corynebacterium glutamicum*. The *MmcBC* fumarase from *Pelotomaculum thermopropionicum* is another class of fumarase with two subunits (Shimoyama et al., *FEMS Microbiol. Lett.* 270:207-213 (2007)).

<u>Protein</u>	<u>GenBank ID</u>	<u>GI Number</u>	<u>Organism</u>
<i>fumA</i>	NP_416129.1	16129570	<i>Escherichia coli</i>

<u>Protein</u>	<u>GenBank ID</u>	<u>GI Number</u>	<u>Organism</u>
<i>fumB</i>	NP_418546.1	16131948	<i>Escherichia coli</i>
<i>fumC</i>	NP_416128.1	16129569	<i>Escherichia coli</i>
<i>FUM1</i>	NP_015061	6324993	<i>Saccharomyces cerevisiae</i>
<i>fumC</i>	Q8NRN8.1	39931596	<i>Corynebacterium glutamicum</i>
<i>fumC</i>	O69294.1	9789756	<i>Campylobacter jejuni</i>
<i>fumC</i>	P84127	75427690	<i>Thermus thermophilus</i>
<i>fumH</i>	P14408.1	120605	<i>Rattus norvegicus</i>
<i>MmcB</i>	YP_001211906	147677691	<i>Pelotomaculum thermopropionicum</i>
<i>MmcC</i>	YP_001211907	147677692	<i>Pelotomaculum thermopropionicum</i>

Fumarate reductase catalyzes the reduction of fumarate to succinate. The fumarate reductase of *E. coli*, composed of four subunits encoded by *frdABCD*, is membrane-bound and active under anaerobic conditions. The electron donor for this reaction is menaquinone and the two protons produced in this reaction do not contribute to the proton gradient (Iverson et al., *Science* 284:1961-1966 (1999)). The yeast genome encodes two soluble fumarate reductase isozymes encoded by *FRDS1* (Enomoto et al., *DNA Res.* 3:263-267 (1996)) and *FRDS2* (Muratsubaki et al., *Arch. Biochem. Biophys.* 352:175-181 (1998)), which localize to the cytosol and promitochondrion, respectively, and are used during anaerobic growth on glucose (Arikawa et al., *FEMS Microbiol. Lett.* 165:111-116 (1998)).

<u>Protein</u>	<u>GenBank ID</u>	<u>GI Number</u>	<u>Organism</u>
<i>FRDS1</i>	P32614	418423	<i>Saccharomyces cerevisiae</i>
<i>FRDS2</i>	NP_012585	6322511	<i>Saccharomyces cerevisiae</i>
<i>frdA</i>	NP_418578.1	16131979	<i>Escherichia coli</i>
<i>frdB</i>	NP_418577.1	16131978	<i>Escherichia coli</i>
<i>frdC</i>	NP_418576.1	16131977	<i>Escherichia coli</i>
<i>frdD</i>	NP_418475.1	16131877	<i>Escherichia coli</i>

The ATP-dependent acylation of succinate to succinyl-CoA is catalyzed by succinyl-CoA synthetase (EC 6.2.1.5). 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* (Buck et al., *Biochemistry* 24:6245-6252 (1985)). These proteins are identified below:

<u>Protein</u>	<u>GenBank ID</u>	<u>GI Number</u>	<u>Organism</u>
<i>LSC1</i>	NP_014785	6324716	<i>Saccharomyces cerevisiae</i>
<i>LSC2</i>	NP_011760	6321683	<i>Saccharomyces cerevisiae</i>
<i>sucC</i>	NP_415256.1	16128703	<i>Escherichia coli</i>
<i>sucD</i>	AAC73823.1	1786949	<i>Escherichia coli</i>

Alpha-ketoglutarate:ferredoxin oxidoreductase (EC 1.2.7.3), also known as 2-oxoglutarate synthase or 2-oxoglutarate:ferredoxin oxidoreductase (OFOR), forms alpha-ketoglutarate from CO₂ and succinyl-CoA with concurrent consumption of two reduced ferredoxin equivalents. OFOR and pyruvate:ferredoxin oxidoreductase (PFOR) are members of a diverse family of 2-oxoacid:ferredoxin (flavodoxin) oxidoreductases which utilize thiamine pyrophosphate, CoA and iron-sulfur clusters as cofactors and ferredoxin, flavodoxin and FAD as electron carriers (*Adams et al., Archaea. Adv. Protein Chem.* 48:101-180 (1996)). Enzymes in this class are reversible and function in the carboxylation direction in organisms that fix carbon by the RTCA cycle such as *Hydrogenobacter thermophilus*, *Desulfobacter hydrogenophilus* and *Chlorobium species* (Shiba et al. 1985; Evans et al., *Proc. Natl. Acad. Sci. U.S.A.* 55:92934 (1966); Buchanan, 1971). The two-subunit enzyme from *H. thermophilus*, encoded by *korAB*, has been cloned and expressed in *E. coli* (Yun et al., *Biochem. Biophys. Res. Commun.* 282:589-594 (2001)). A five subunit OFOR from the same organism with strict substrate specificity for succinyl-CoA, encoded by *forDABGE*, was recently identified and expressed in *E. coli* (Yun et al. 2002). The kinetics of CO₂ fixation of both *H. thermophilus* OFOR enzymes have been characterized (Yamamoto et al., *Extremophiles* 14:79-85 (2010)). A CO₂-fixing OFOR from *Chlorobium thiosulfatophilum* has been purified and characterized but the genes encoding this enzyme have not been identified to date. Enzyme candidates in *Chlorobium* species can be inferred by sequence similarity to the *H. thermophilus* genes. For example, the *Chlorobium limicola* genome encodes two similar proteins. Acetogenic bacteria such as *Moorella thermoacetica* are predicted

to encode two OFOR enzymes. The enzyme encoded by *Moth_0034* is predicted to function in the CO₂-assimilating direction. The genes associated with this enzyme, *Moth_0034* have not been experimentally validated to date but can be inferred by sequence similarity to known OFOR enzymes.

OFOR enzymes that function in the decarboxylation direction under physiological conditions can also catalyze the reverse reaction. The OFOR from the thermoacidophilic archaeon *Sulfolobus* sp. strain 7, encoded by ST2300, has been extensively studied (Zhang et al. 1996. A plasmid-based expression system has been developed for efficiently expressing this protein in *E. coli* (Fukuda et al., Eur. J. Biochem. 268:5639-5646 (2001)) and residues involved in substrate specificity were determined (Fukuda and Wakagi, Biochim. Biophys. Acta 1597:74-80 (2002)). The OFOR encoded by *Ape1472/Ape1473* from *Aeropyrum pernix* str. K1 was recently cloned into *E. coli*, characterized, and found to react with 2-oxoglutarate and a broad range of 2-oxoacids (Nishizawa et al., FEBS Lett. 579:2319-2322 (2005)). Another exemplary OFOR is encoded by *oorDABC* in *Helicobacter pylori* (Hughes et al. 1998). An enzyme specific to alpha-ketoglutarate has been reported in *Thauera aromatica* (Dorner and Boll, J. Bacteriol. 184 (14), 3975-83 (2002). A similar enzyme can be found in *Rhodospirillum rubrum* by sequence homology. A two subunit enzyme has also been identified in *Chlorobium tepidum* (Eisen et al., PNAS 99(14): 9509-14 (2002)).

<u>Protein</u>	<u>GenBank ID</u>	<u>GI Number</u>	<u>Organism</u>
<i>korA</i>	BAB21494	12583691	<i>Hydrogenobacter thermophilus</i>
<i>korB</i>	BAB21495	12583692	<i>Hydrogenobacter thermophilus</i>
<i>forD</i>	BAB62132.1	14970994	<i>Hydrogenobacter thermophilus</i>
<i>forA</i>	BAB62133.1	14970995	<i>Hydrogenobacter thermophilus</i>
<i>forB</i>	BAB62134.1	14970996	<i>Hydrogenobacter thermophilus</i>
<i>forG</i>	BAB62135.1	14970997	<i>Hydrogenobacter thermophilus</i>
<i>forE</i>	BAB62136.1	14970998	<i>Hydrogenobacter thermophilus</i>
<i>Clim_0204</i>	ACD89303.1	189339900	<i>Chlorobium limicola</i>
<i>Clim_0205</i>	ACD89302.1	189339899	<i>Chlorobium limicola</i>
<i>Clim_1123</i>	ACD90192.1	189340789	<i>Chlorobium limicola</i>
<i>Clim_1124</i>	ACD90193.1	189340790	<i>Chlorobium limicola</i>

<i>Moth_1984</i>	YP_430825.1	83590816	<i>Moorella thermoacetica</i>
<i>Moth_1985</i>	YP_430826.1	83590817	<i>Moorella thermoacetica</i>
<i>Moth_0034</i>	YP_428917.1	83588908	<i>Moorella thermoacetica</i>
<i>ST2300</i>	NP_378302.1	15922633	<i>Sulfolobus</i> sp. strain 7
<i>Ape1472</i>	BAA80470.1	5105156	<i>Aeropyrum pernix</i>
<i>Ape1473</i>	BAA80471.2	116062794	<i>Aeropyrum pernix</i>
<i>oorD</i>	NP_207383.1	15645213	<i>Helicobacter pylori</i>
<i>oorA</i>	NP_207384.1	15645214	<i>Helicobacter pylori</i>
<i>oorB</i>	NP_207385.1	15645215	<i>Helicobacter pylori</i>
<i>oorC</i>	NP_207386.1	15645216	<i>Helicobacter pylori</i>
<i>CT0163</i>	NP_661069.1	21673004	<i>Chlorobium tepidum</i>
<i>CT0162</i>	NP_661068.1	21673003	<i>Chlorobium tepidum</i>
<i>korA</i>	CAA12243.2	19571179	<i>Thauera aromatica</i>
<i>korB</i>	CAD27440.1	19571178	<i>Thauera aromatica</i>
<i>Rru_A2721</i>	YP_427805.1	83594053	<i>Rhodospirillum rubrum</i>
<i>Rru_A2722</i>	YP_427806.1	83594054	<i>Rhodospirillum rubrum</i>

Isocitrate dehydrogenase catalyzes the reversible decarboxylation of isocitrate to 2-oxoglutarate coupled to the reduction of NAD(P)⁺. IDH enzymes in *Saccharomyces cerevisiae* and *Escherichia coli* are encoded by *IDP1* and *icd*, respectively (Haselbeck and McAlister-Henn, *J. Biol. Chem.* 266:2339-2345 (1991); Nimmo, H.G., *Biochem. J.* 234:317-2332 (1986)). The reverse reaction in the reductive TCA cycle, the reductive carboxylation of 2-oxoglutarate to isocitrate, is favored by the NADPH-dependent CO₂-fixing IDH from *Chlorobium limicola* and was functionally expressed in *E. coli* (Kanao et al., *Eur. J. Biochem.* 269:1926-1931 (2002)). A similar enzyme with 95% sequence identity is found in the *C. tepidum* genome in addition to some other candidates listed below.

<u>Protein</u>	<u>GenBank ID</u>	<u>GI Number</u>	<u>Organism</u>
<i>Icd</i>	ACI84720.1	209772816	<i>Escherichia coli</i>
<i>IDP1</i>	AAA34703.1	171749	<i>Saccharomyces cerevisiae</i>
<i>Idh</i>	BAC00856.1	21396513	<i>Chlorobium limicola</i>

<u>Protein</u>	<u>GenBank ID</u>	<u>GI Number</u>	<u>Organism</u>
<i>Icd</i>	AAM71597.1	21646271	<i>Chlorobium tepidum</i>
<i>icd</i>	NP_952516.1	39996565	<i>Geobacter sulfurreducens</i>
<i>icd</i>	YP_393560.	78777245	<i>Sulfurimonas denitrificans</i>

In *H. thermophilus* the reductive carboxylation of 2-oxoglutarate to isocitrate is catalyzed by two enzymes: 2-oxoglutarate carboxylase and oxalosuccinate reductase. 2-Oxoglutarate carboxylase (EC 6.4.1.7) catalyzes the ATP-dependent carboxylation of alpha-ketoglutarate to oxalosuccinate (Aoshima and Igarashi, *Mol. Microbiol.* 62:748-759 (2006)). This enzyme is a large complex composed of two subunits. Biotinylation of the large (A) subunit is required for enzyme function (Aoshima et al., *Mol. Microbiol.* 51:791-798 (2004)). Oxalosuccinate reductase (EC 1.1.1.-) catalyzes the NAD-dependent conversion of oxalosuccinate to D-threo-isocitrate. The enzyme is a homodimer encoded by *icd* in *H. thermophilus*. The kinetic parameters of this enzyme indicate that the enzyme only operates in the reductive carboxylation direction *in vivo*, in contrast to isocitrate dehydrogenase enzymes in other organisms (Aoshima and Igarashi, *J. Bacteriol.* 190:2050-2055 (2008)). Based on sequence homology, gene candidates have also been found in *Thiobacillus denitrificans* and *Thermocrinis albus*.

<u>Protein</u>	<u>GenBank ID</u>	<u>GI Number</u>	<u>Organism</u>
<i>cfiA</i>	BAF34932.1	116234991	<i>Hydrogenobacter thermophilus</i>
<i>cifB</i>	BAF34931.1	116234990	<i>Hydrogenobacter thermophilus</i>
<i>Icd</i>	BAD02487.1	38602676	<i>Hydrogenobacter thermophilus</i>
<i>Tbd_1556</i>	YP_315314	74317574	<i>Thiobacillus denitrificans</i>
<i>Tbd_1555</i>	YP_315313	74317573	<i>Thiobacillus denitrificans</i>
<i>Tbd_0854</i>	YP_314612	74316872	<i>Thiobacillus denitrificans</i>
<i>Thal_0268</i>	YP_003473030	289548042	<i>Thermocrinis albus</i>
<i>Thal_0267</i>	YP_003473029	289548041	<i>Thermocrinis albus</i>
<i>Thal_0646</i>	YP_003473406	289548418	<i>Thermocrinis albus</i>

Aconitase (EC 4.2.1.3) is an iron-sulfur-containing protein catalyzing the reversible isomerization of citrate and iso-citrate via the intermediate *cis*-aconitate. Two aconitase enzymes

are encoded in the *E. coli* genome by *acnA* and *acnB*. AcnB is the main catabolic enzyme, while AcnA is more stable and appears to be active under conditions of oxidative or acid stress (Cunningham et al., *Microbiology* 143 (Pt 12):3795-3805 (1997)). Two isozymes of aconitase in *Salmonella typhimurium* are encoded by *acnA* and *acnB* (Horswill and Escalante-Semerena, *Biochemistry* 40:4703-4713 (2001)). The *S. cerevisiae* aconitase, encoded by *ACO1*, is localized to the mitochondria where it participates in the TCA cycle (Gangloff et al., *Mol. Cell. Biol.* 10:3551-3561 (1990)) and the cytosol where it participates in the glyoxylate shunt (Regev-Rudzki et al., *Mol. Biol. Cell.* 16:4163-4171 (2005)).

<u>Protein</u>	<u>GenBank ID</u>	<u>GI Number</u>	<u>Organism</u>
<i>acnA</i>	AAC7438.1	1787531	<i>Escherichia coli</i>
<i>acnB</i>	AAC73229.1	2367097	<i>Escherichia coli</i>
<i>HP0779</i>	NP_207572.1	15645398	<i>Helicobacter pylori</i> 26695
<i>H16_B0568</i>	CAJ95365.1	113529018	<i>Ralstonia eutropha</i>
<i>DesfrDRAFT_3783</i>	ZP_07335307.1	303249064	<i>Desulfovibrio fructosovorans JJ</i>
<i>Suden_1040 (acnB)</i>	ABB44318.1	78497778	<i>Sulfurimonas denitrificans</i>
<i>Hydth_0755</i>	ADO45152.1	308751669	<i>Hydrogenobacter thermophilus</i>
<i>CT0543 (acn)</i>	AAM71785.1	21646475	<i>Chlorobium tepidum</i>
<i>Clim_2436</i>	YP_001944436.1	189347907	<i>Chlorobium limicola</i>
<i>Clim_0515</i>	ACD89607.1	189340204	<i>Chlorobium limicola</i>
<i>acnA</i>	NP_460671.1	16765056	<i>Salmonella typhimurium</i>
<i>acnB</i>	NP_459163.1	16763548	<i>Salmonella typhimurium</i>
<i>ACO1</i>	AAA34389.1	170982	<i>Saccharomyces cerevisiae</i>

Pyruvate:ferredoxin oxidoreductase (PFOR) catalyzes the reversible oxidation of pyruvate to form 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.* 179:5684-5692 (1997)). Oxygen stability is relatively uncommon in PFORs and is believed to be conferred by a 60 residue extension in the polypeptide chain of the *D. africanus* enzyme. Two cysteine residues in this enzyme form a disulfide bond that protects it against inactivation in the form of oxygen. This disulfide bond and

the stability in the presence of oxygen has been found in other *Desulfovibrio* species also (Vita et al., *Biochemistry*, 47: 957-64 (2008)). The *M. thermoacetica* PFOR is also well characterized (Menon and Ragsdale, *Biochemistry* 36:8484-8494 (1997)) and was shown to have high activity in the direction of pyruvate synthesis during autotrophic growth (Furdui and Ragsdale, *J. Biol. Chem.* 275:28494-28499 (2000)). Further, *E. coli* possesses an uncharacterized open reading frame, *ydbK*, encoding a protein that is 51% identical to the *M. thermoacetica* PFOR. Evidence for pyruvate oxidoreductase activity in *E. coli* has been described (Blaschkowski et al., *Eur. J. Biochem.* 123:563-569 (1982)). PFORs have also been described in other organisms, including *Rhodobacter capsulatus* (Yakunin and Hallenbeck, *Biochimica et Biophysica Acta* 1409 (1998) 39-49 (1998)) and *Choloboum tepidum* (Eisen et al., *PNAS* 99(14): 9509-14 (2002)). The five subunit PFOR from *H. thermophilus*, encoded by *porEDABG*, was cloned into *E. coli* and shown to function in both the decarboxylating and CO₂-assimilating directions (Ikeda et al. 2006; Yamamoto et al., *Extremophiles* 14:79-85 (2010)). Homologs also exist in *C. carboxidivorans* P7. Several additional PFOR enzymes are described in the following review (Ragsdale, S.W., *Chem. Rev.* 103:2333-2346 (2003)). Finally, flavodoxin reductases (e.g., *fqrB* from *Helicobacter pylori* or *Campylobacter jejuni*) (St Maurice et al., *J. Bacteriol.* 189:4764-4773 (2007)) or Rnf-type proteins (Seedorf et al., *Proc. Natl. Acad. Sci. U.S.A.* 105:2128-2133 (2008); and Herrmann, *J. Bacteriol* 190:784-791 (2008)) provide a means to generate NADH or NADPH from the reduced ferredoxin generated by PFOR. These proteins are identified below.

<u>Protein</u>	<u>GenBank ID</u>	<u>GI Number</u>	<u>Organism</u>
<i>DesfrDRAFT_0121</i>	ZP_07331646.1	303245362	<i>Desulfovibrio fructosovorans</i> <i>JJ</i>
Por	CAA70873.1	1770208	<i>Desulfovibrio africanus</i>
por	YP_012236.1	46581428	<i>Desulfovibrio vulgaris</i> str. <i>Hildenborough</i>
Dde_3237	ABB40031.1	78220682	<i>DesulfoVibrio desulfuricans</i> <i>G20</i>
Ddes_0298	YP_002478891.1	220903579	<i>Desulfovibrio desulfuricans</i> subsp. <i>desulfuricans</i> str. <i>ATCC 27774</i>
Por	YP_428946.1	83588937	<i>Moorella thermoacetica</i>

YdbK	NP_415896.1	16129339	<i>Escherichia coli</i>
nifJ (CT1628)	NP_662511.1	21674446	<i>Chlorobium tepidum</i>
CJE1649	YP_179630.1	57238499	<i>Campylobacter jejuni</i>
nifJ	ADE85473.1	294476085	<i>Rhodobacter capsulatus</i>
<i>porE</i>	BAA95603.1	7768912	<i>Hydrogenobacter thermophilus</i>
<i>porD</i>	BAA95604.1	7768913	<i>Hydrogenobacter thermophilus</i>
<i>porA</i>	BAA95605.1	7768914	<i>Hydrogenobacter thermophilus</i>
<i>porB</i>	BAA95606.1	776891	<i>Hydrogenobacter thermophilus</i>
<i>porG</i>	BAA95607.1	7768916	<i>Hydrogenobacter thermophilus</i>
FqrB	YP_001482096.1	157414840	<i>Campylobacter jejuni</i>
HP1164	NP_207955.1	15645778	<i>Helicobacter pylori</i>
RnfC	EDK33306.1	146346770	<i>Clostridium kluyveri</i>
RnfD	EDK33307.1	146346771	<i>Clostridium kluyveri</i>
RnfG	EDK33308.1	146346772	<i>Clostridium kluyveri</i>
RnfE	EDK33309.1	146346773	<i>Clostridium kluyveri</i>
RnfA	EDK33310.1	146346774	<i>Clostridium kluyveri</i>
RnfB	EDK33311.1	146346775	<i>Clostridium kluyveri</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 (5). Crystal structures of the enzyme complex from bovine kidney (18) and the E2 catalytic domain from *Azotobacter vinelandii* are available (4). 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 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 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 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*.

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 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.

For enzymes that use reducing equivalents in the form of NADH or NADPH, these reduced carriers can be generated by transferring electrons from reduced ferredoxin. Two enzymes catalyze the reversible transfer of electrons from reduced ferredoxins to NAD(P)⁺, ferredoxin:NAD⁺ oxidoreductase (EC 1.18.1.3) and ferredoxin:NADP⁺ oxidoreductase (FNR, EC 1.18.1.2). Ferredoxin:NADP⁺ oxidoreductase (FNR, EC 1.18.1.2) has a noncovalently bound FAD cofactor that facilitates the reversible transfer of electrons from NADPH to low-potential acceptors such as ferredoxins or flavodoxins (Blaschkowski et al., Eur. J. Biochem. 123:563-569 (1982); Fujii et al., 1977). The *Helicobacter pylori* FNR, encoded by *HP1164* (*fqrB*), is coupled to the activity of pyruvate:ferredoxin oxidoreductase (PFOR) resulting in the pyruvate-dependent production of NADPH (St et al. 2007). An analogous enzyme is found in *Campylobacter jejuni* (St et al. 2007). A ferredoxin:NADP⁺ oxidoreductase enzyme is encoded in the *E. coli* genome by *fpr* (Bianchi et al. 1993). Ferredoxin:NAD⁺ oxidoreductase utilizes

reduced ferredoxin to generate NADH from NAD⁺. In several organisms, including *E. coli*, this enzyme is a component of multifunctional dioxygenase enzyme complexes. The ferredoxin:NAD⁺ oxidoreductase of *E. coli*, encoded by *hcaD*, is a component of the 3-phenylproppionate dioxygenase system involved in involved in aromatic acid utilization (Diaz et al. 1998). NADH:ferredoxin reductase activity was detected in cell extracts of *Hydrogenobacter thermophilus* strain TK-6, although a gene with this activity has not yet been indicated (Yoon et al. 2006). Finally, the energy-conserving membrane-associated Rnf-type proteins (Seedorf et al., Proc. Natl. Acad. Sci. U.S.A. 105:2128-2133 (2008); Herrmann et al., *J. Bacteriol.* 190:784-791 (2008)) provide a means to generate NADH or NADPH from reduced ferredoxin. Additional ferredoxin:NAD(P)⁺ oxidoreductases have been annotated in *Clostridium carboxydivorans* P7.

<u>Protein</u>	<u>GenBank ID</u>	<u>GI Number</u>	<u>Organism</u>
<i>HP1164</i>	NP_207955.1	15645778	<i>Helicobacter pylori</i>
<i>RPA3954</i>	CAE29395.1	39650872	<i>Rhodopseudomonas palustris</i>
<i>fpr</i>	BAH29712.1	225320633	<i>Hydrogenobacter thermophilus</i>
<i>yumC</i>	NP_391091.2	255767736	<i>Bacillus subtilis</i>
<i>CJE0663</i>	AAW35824.1	57167045	<i>Campylobacter jejuni</i>
<i>fpr</i>	P28861.4	399486	<i>Escherichia coli</i>
<i>hcaD</i>	AAC75595.1	1788892	<i>Escherichia coli</i>
<i>LOC100282643</i>	NP_001149023.1	226497434	<i>Zea mays</i>
<i>RnfC</i>	EDK33306.1	146346770	<i>Clostridium kluyveri</i>
<i>RnfD</i>	EDK33307.1	146346771	<i>Clostridium kluyveri</i>
<i>RnfG</i>	EDK33308.1	146346772	<i>Clostridium kluyveri</i>
<i>RnfE</i>	EDK33309.1	146346773	<i>Clostridium kluyveri</i>
<i>RnfA</i>	EDK33310.1	146346774	<i>Clostridium kluyveri</i>
<i>RnfB</i>	EDK33311.1	146346775	<i>Clostridium kluyveri</i>
<i>CcarbDRAFT_2639</i>	ZP_05392639.1	255525707	<i>Clostridium carboxidivorans</i> P7
<i>CcarbDRAFT_2638</i>	ZP_05392638.1	255525706	<i>Clostridium carboxidivorans</i> P7

<i>CcarbDRAFT_2636</i>	ZP_05392636.1	255525704	<i>Clostridium carboxidivorans P7</i>
<i>CcarbDRAFT_5060</i>	ZP_05395060.1	255528241	<i>Clostridium carboxidivorans P7</i>
<i>CcarbDRAFT_2450</i>	ZP_05392450.1	255525514	<i>Clostridium carboxidivorans P7</i>
<i>CcarbDRAFT_1084</i>	ZP_05391084.1	255524124	<i>Clostridium carboxidivorans P7</i>

Ferredoxins are small acidic proteins containing one or more iron-sulfur clusters that function as intracellular electron carriers with a low reduction potential. Reduced ferredoxins donate electrons to Fe-dependent enzymes such as ferredoxin-NADP⁺ oxidoreductase, pyruvate:ferredoxin oxidoreductase (PFOR) and 2-oxoglutarate:ferredoxin oxidoreductase (OFOR). The *H. thermophilus* gene *fdx1* encodes a [4Fe-4S]-type ferredoxin that is required for the reversible carboxylation of 2-oxoglutarate and pyruvate by OFOR and PFOR, respectively (Yamamoto et al., *Extremophiles* 14:79-85 (2010)). The ferredoxin associated with the *Sulfolobus solfataricus* 2-oxoacid:ferredoxin reductase is a monomeric dicluster [3Fe-4S][4Fe-4S] type ferredoxin (Park et al. 2006). While the gene associated with this protein has not been fully sequenced, the N-terminal domain shares 93% homology with the *zfx* ferredoxin from *S. acidocaldarius*. The *E. coli* genome encodes a soluble ferredoxin of unknown physiological function, *fdx*. Some evidence indicates that this protein can function in iron-sulfur cluster assembly (Takahashi and Nakamura, 1999). Additional ferredoxin proteins have been characterized in *Helicobacter pylori* (Mukhopadhyay et al. 2003) and *Campylobacter jejuni* (van Vliet et al. 2001). A 2Fe-2S ferredoxin from *Clostridium pasteurianum* has been cloned and expressed in *E. coli* (Fujinaga and Meyer, Biochemical and Biophysical Research Communications, 192(3): (1993)). Acetogenic bacteria such as *Moorella thermoacetica*, *Clostridium carboxidivorans P7* and *Rhodospirillum rubrum* are predicted to encode several ferredoxins, listed in the table below.

<u>Protein</u>	<u>GenBank ID</u>	<u>GI Number</u>	<u>Organism</u>
<i>fdx1</i>	BAE02673.1	68163284	<i>Hydrogenobacter thermophilus</i>

M11214.1	AAA83524.1	144806	<i>Clostridium pasteurianum</i>
Zfx	AAAY79867.1	68566938	<i>Sulfolobus acidocalarius</i>
Fdx	AAC75578.1	1788874	<i>Escherichia coli</i>
hp_0277	AAD07340.1	2313367	<i>Helicobacter pylori</i>
fdxA	CAL34484.1	112359698	<i>Campylobacter jejuni</i>
Moth_0061	ABC18400.1	83571848	<i>Moorella thermoacetica</i>
Moth_1200	ABC19514.1	83572962	<i>Moorella thermoacetica</i>
Moth_1888	ABC20188.1	83573636	<i>Moorella thermoacetica</i>
Moth_2112	ABC20404.1	83573852	<i>Moorella thermoacetica</i>
Moth_1037	ABC19351.1	83572799	<i>Moorella thermoacetica</i>
CcarbDRAFT_4383	ZP_05394383.1	255527515	<i>Clostridium carboxidivorans</i> P7
CcarbDRAFT_2958	ZP_05392958.1	255526034	<i>Clostridium carboxidivorans</i> P7
CcarbDRAFT_2281	ZP_05392281.1	255525342	<i>Clostridium carboxidivorans</i> P7
CcarbDRAFT_5296	ZP_05395295.1	255528511	<i>Clostridium carboxidivorans</i> P7
CcarbDRAFT_1615	ZP_05391615.1	255524662	<i>Clostridium carboxidivorans</i> P7
CcarbDRAFT_1304	ZP_05391304.1	255524347	<i>Clostridium carboxidivorans</i> P7
cooF	AAG29808.1	11095245	<i>Carboxydotherrnus</i> <i>hydrogenoformans</i>
fdxN	CAA35699.1	46143	<i>Rhodobacter capsulatus</i>
Rru_A2264	ABC23064.1	83576513	<i>Rhodospirillum rubrum</i>
Rru_A1916	ABC22716.1	83576165	<i>Rhodospirillum rubrum</i>
Rru_A2026	ABC22826.1	83576275	<i>Rhodospirillum rubrum</i>
cooF	AAC45122.1	1498747	<i>Rhodospirillum rubrum</i>
fdxN	AAA26460.1	152605	<i>Rhodospirillum rubrum</i>
Alvin_2884	ADC63789.1	288897953	<i>Allochrotrattum vinosum</i> DSM 180
fdx	YP_002801146.1	226946073	<i>Azotobacter vinelandii</i> DJ

CKL_3790	YP_001397146.1	153956381	<i>Clostridium kluyveri</i> DSM 555
fer1	NP_949965.1	39937689	<i>Rhodopseudomonas palustris</i> CGA009
fdx	CAA12251.1	3724172	<i>Thauera aromatica</i>
CHY_2405	YP_361202.1	78044690	<i>Carboxydotherrnus hydrogenoformans</i>
fer	YP_359966.1	78045103	<i>Carboxydotherrnus hydrogenoformans</i>
fer	AAC83945.1	1146198	<i>Bacillus subtilis</i>
fdx1	NP_249053.1	15595559	<i>Pseudomonas aeruginosa</i> PA01
yfhL	AP_003148.1	89109368	<i>Escherichia coli</i> K-12

Succinyl-CoA transferase catalyzes the conversion of succinyl-CoA to succinate while transferring the CoA moiety to a CoA acceptor molecule. Many transferases have broad specificity and can utilize CoA acceptors as diverse as acetate, succinate, propionate, butyrate, 2-methylacetoacetate, 3-ketohexanoate, 3-ketopentanoate, valerate, crotonate, 3-mercaptopropionate, propionate, vinylacetate, and butyrate, among others.

The conversion of succinate to succinyl-CoA can be carried by a transferase which does not require the direct consumption of an ATP or GTP. This type of reaction is common in a number of organisms. The conversion of succinate to succinyl-CoA can also be catalyzed by succinyl-CoA:Acetyl-CoA transferase. The gene product of *catI* of *Clostridium kluyveri* has been shown to exhibit succinyl-CoA: acetyl-CoA transferase activity (Sohling and Gottschalk, J. Bacteriol. 178:871-880 (1996)). In addition, the activity is present in *Trichomonas vaginalis* (van Grinsven et al. 2008) and *Trypanosoma brucei* (Riviere et al. 2004). The succinyl-CoA:acetate CoA-transferase from *Acetobacter aceti*, encoded by *aarC*, replaces succinyl-CoA synthetase in a variant TCA cycle (Mullins et al. 2008). Similar succinyl-CoA transferase activities are also present in *Trichomonas vaginalis* (van Grinsven et al. 2008), *Trypanosoma brucei* (Riviere et al. 2004) and *Clostridium kluyveri* (Sohling and Gottschalk, 1996c). The beta-ketoadipate:succinyl-CoA transferase encoded by *pcaI* and *pcaJ* in *Pseudomonas putida* is yet another candidate (Kaschabek et al. 2002). The aforementioned proteins are identified below.

<u>Protein</u>	<u>GenBank ID</u>	<u>GI Number</u>	<u>Organism</u>
<i>catI</i>	P38946.1	729048	<i>Clostridium kluyveri</i>
<i>TVAG_395550</i>	XP_001330176	123975034	<i>Trichomonas vaginalis</i> G3
<i>Tb11.02.0290</i>	XP_828352	71754875	<i>Trypanosoma brucei</i>
<i>pcaI</i>	AAN69545.1	24985644	<i>Pseudomonas putida</i>
<i>pcaJ</i>	NP_746082.1	26990657	<i>Pseudomonas putida</i>
<i>aarC</i>	ACD85596.1	189233555	<i>Acetobacter aceti</i>

An additional exemplary transferase that converts succinate to succinyl-CoA while converting a 3-ketoacyl-CoA to a 3-ketoacid is succinyl-CoA:3:ketoacid-CoA transferase (EC 2.8.3.5).

Exemplary succinyl-CoA:3:ketoacid-CoA transferases are present in *Helicobacter pylori* (Corthesy-Theulaz et al. 1997), *Bacillus subtilis*, and *Homo sapiens* (Fukao et al. 2000; Tanaka et al. 2002). The aforementioned proteins are identified below.

<u>Protein</u>	<u>GenBank ID</u>	<u>GI Number</u>	<u>Organism</u>
<i>HPAG1_0676</i>	YP_627417	108563101	<i>Helicobacter pylori</i>
<i>HPAG1_0677</i>	YP_627418	108563102	<i>Helicobacter pylori</i>
<i>ScoA</i>	NP_391778	16080950	<i>Bacillus subtilis</i>
<i>ScoB</i>	NP_391777	16080949	<i>Bacillus subtilis</i>
<i>OXCT</i> 1	NP_000427	4557817	<i>Homo sapiens</i>
<i>OXCT</i> 2	NP_071403	11545841	<i>Homo sapiens</i>

Converting succinate to succinyl-CoA by succinyl-CoA:3:ketoacid-CoA transferase requires the simultaneous conversion of a 3-ketoacyl-CoA such as acetoacetyl-CoA to a 3-ketoacid such as acetoacetate. Conversion of a 3-ketoacid back to a 3-ketoacyl-CoA can be catalyzed by an acetoacetyl-CoA:acetate:CoA transferase. Acetoacetyl-CoA:acetate:CoA transferase converts acetoacetyl-CoA and acetate to acetoacetate and acetyl-CoA, or vice versa. 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)) are shown below.

<u>Protein</u>	<u>GenBank ID</u>	<u>GI Number</u>	<u>Organism</u>
<i>AtoA</i>	NP_416726.1	2492994	<i>Escherichia coli</i>
<i>AtoD</i>	NP_416725.1	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>

Yet another possible CoA acceptor is benzylsuccinate. Succinyl-CoA:(R)-Benzylsuccinate CoA-Transferase functions as part of an anaerobic degradation pathway for toluene in organisms such as *Thauera aromatica* (Leutwein and Heider, *J. Bact.* 183(14) 4288-4295 (2001)). Homologs can be found in *Azoarcus* sp. T, *Aromatoleum aromaticum* EbN1, and *Geobacter metallireducens* GS-15. The aforementioned proteins are identified below.

<u>Protein</u>	<u>GenBank ID</u>	<u>GI Number</u>	<u>Organism</u>
<i>bbsE</i>	AAF89840	9622535	<i>Thauera aromatica</i>
<i>Bbsf</i>	AAF89841	9622536	<i>Thauera aromatica</i>
<i>bbsE</i>	AAU45405.1	52421824	<i>Azoarcus</i> sp. T
<i>bbsF</i>	AAU45406.1	52421825	<i>Azoarcus</i> sp. T
<i>bbsE</i>	YP_158075.1	56476486	<i>Aromatoleum aromaticum</i> EbN1
<i>bbsF</i>	YP_158074.1	56476485	<i>Aromatoleum aromaticum</i> EbN1
<i>Gmet_1521</i>	YP_384480.1	78222733	<i>Geobacter metallireducens</i> GS-15
<i>Gmet_1522</i>	YP_384481.1	78222734	<i>Geobacter metallireducens</i> GS-15

Additionally, *ygfH* encodes a propionyl CoA:succinate CoA transferase in *E. coli* (Haller et al., *Biochemistry*, 39(16) 4622-4629). Close homologs can be found in, for example, *Citrobacter youngae* ATCC 29220, *Salmonella enterica* subsp. *arizonae* serovar, and *Yersinia intermedia* ATCC 29909. The aforementioned proteins are identified below.

<u>Protein</u>	<u>GenBank ID</u>	<u>GI Number</u>	<u>Organism</u>
<i>ygfH</i>	NP_417395.1	16130821	<i>Escherichia coli str. K-12 substr. MG1655</i>
<i>CIT292_04485</i>	ZP_03838384.1	227334728	<i>Citrobacter youngae</i> ATCC 29220
<i>SARI_04582</i>	YP_001573497.1	161506385	<i>Salmonella enterica subsp. arizonae serovar</i>
<i>yinte0001_14430</i>	ZP_04635364.1	238791727	<i>Yersinia intermedia</i> ATCC 29909

Citrate lyase (EC 4.1.3.6) catalyzes a series of reactions resulting in the cleavage of citrate to acetate and oxaloacetate. The enzyme is active under anaerobic conditions and is composed of three subunits: an acyl-carrier protein (ACP, gamma), an ACP transferase (alpha), and a acyl lyase (beta). Enzyme activation uses covalent binding and acetylation of an unusual prosthetic group, 2'-(5''-phosphoribosyl)-3'-dephospho-CoA, which is similar in structure to acetyl-CoA. Acylation is catalyzed by CitC, a citrate lyase synthetase. Two additional proteins, CitG and CitX, are used to convert the apo enzyme into the active holo enzyme (Schneider et al., *Biochemistry* 39:9438-9450 (2000)). Wild type *E. coli* does not have citrate lyase activity; however, mutants deficient in molybdenum cofactor synthesis have an active citrate lyase (Clark, *FEMS Microbiol. Lett.* 55:245-249 (1990)). The *E. coli* enzyme is encoded by *citEFD* and the citrate lyase synthetase is encoded by *citC* (Nilekani and SivaRaman, *Biochemistry* 22:4657-4663 (1983)). The *Leuconostoc mesenteroides* citrate lyase has been cloned, characterized and expressed in *E. coli* (Bekal et al., *J. Bacteriol.* 180:647-654 (1998)). Citrate lyase enzymes have also been identified in enterobacteria that utilize citrate as a carbon and energy source, including *Salmonella typhimurium* and *Klebsiella pneumoniae* (Bott, *Arch. Microbiol.* 167: 78-88 (1997); Bott and Dimroth, *Mol. Microbiol.* 14:347-356 (1994)). The aforementioned proteins are tabulated below.

<u>Protein</u>	<u>GenBank ID</u>	<u>GI Number</u>	<u>Organism</u>
<i>citF</i>	AAC73716.1	1786832	<i>Escherichia coli</i>
<i>Cite</i>	AAC73717.2	87081764	<i>Escherichia coli</i>
<i>citD</i>	AAC73718.1	1786834	<i>Escherichia coli</i>
<i>citC</i>	AAC73719.2	87081765	<i>Escherichia coli</i>

<u>Protein</u>	<u>GenBank ID</u>	<u>GI Number</u>	<u>Organism</u>
<i>citG</i>	AAC73714.1	1786830	<i>Escherichia coli</i>
<i>citX</i>	AAC73715.1	1786831	<i>Escherichia coli</i>
<i>citF</i>	CAA71633.1	2842397	<i>Leuconostoc mesenteroides</i>
<i>Cite</i>	CAA71632.1	2842396	<i>Leuconostoc mesenteroides</i>
<i>citD</i>	CAA71635.1	2842395	<i>Leuconostoc mesenteroides</i>
<i>citC</i>	CAA71636.1	3413797	<i>Leuconostoc mesenteroides</i>
<i>citG</i>	CAA71634.1	2842398	<i>Leuconostoc mesenteroides</i>
<i>citX</i>	CAA71634.1	2842398	<i>Leuconostoc mesenteroides</i>
<i>citF</i>	NP_459613.1	16763998	<i>Salmonella typhimurium</i>
<i>cite</i>	AAL19573.1	16419133	<i>Salmonella typhimurium</i>
<i>citD</i>	NP_459064.1	16763449	<i>Salmonella typhimurium</i>
<i>citC</i>	NP_459616.1	16764001	<i>Salmonella typhimurium</i>
<i>citG</i>	NP_459611.1	16763996	<i>Salmonella typhimurium</i>
<i>citX</i>	NP_459612.1	16763997	<i>Salmonella typhimurium</i>
<i>citF</i>	CAA56217.1	565619	<i>Klebsiella pneumoniae</i>
<i>cite</i>	CAA56216.1	565618	<i>Klebsiella pneumoniae</i>
<i>citD</i>	CAA56215.1	565617	<i>Klebsiella pneumoniae</i>
<i>citC</i>	BAH66541.1	238774045	<i>Klebsiella pneumoniae</i>
<i>citG</i>	CAA56218.1	565620	<i>Klebsiella pneumoniae</i>
<i>citX</i>	AAL60463.1	18140907	<i>Klebsiella pneumoniae</i>

Acetate kinase (EC 2.7.2.1) catalyzes the reversible ATP-dependent phosphorylation of acetate to acetylphosphate. Exemplary acetate kinase enzymes have been characterized in many organisms including *E. coli*, *Clostridium acetobutylicum* and *Methanosarcina thermophila*

(Ingram-Smith et al., *J. Bacteriol.* 187:2386-2394 (2005); Fox and Roseman, *J. Biol. Chem.* 261:13487-13497 (1986); Winzer et al., *Microbiology* 143 (Pt 10):3279-3286 (1997)). Acetate kinase activity has also been demonstrated in the gene product of *E. coli purT* (Marolewski et al., *Biochemistry* 33:2531-2537 (1994)). Some butyrate kinase enzymes (EC 2.7.2.7), for example *buk1* and *buk2* from *Clostridium acetobutylicum*, also accept acetate as a substrate (Hartmanis, M.G., *J. Biol. Chem.* 262:617-621 (1987)).

<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>Ack</i>	AAB18301.1	1491790	<i>Clostridium acetobutylicum</i>
<i>Ack</i>	AAA72042.1	349834	<i>Methanosarcina thermophila</i>
<i>purT</i>	AAC74919.1	1788155	<i>Escherichia coli</i>
<i>buk1</i>	NP_349675	15896326	<i>Clostridium acetobutylicum</i>
<i>buk2</i>	Q97III	20137415	<i>Clostridium acetobutylicum</i>

The formation of acetyl-CoA from acetylphosphate is catalyzed by phosphotransacetylase (EC 2.3.1.8). The *pta* gene from *E. coli* encodes an enzyme that reversibly converts acetyl-CoA into acetyl-phosphate (Suzuki, T., *Biochim. Biophys. Acta* 191:559-569 (1969)). Additional acetyltransferase enzymes have been characterized in *Bacillus subtilis* (Rado and Hoch, *Biochim. Biophys. Acta* 321:114-125 (1973)), *Clostridium kluyveri* (Stadtman, E., *Methods Enzymol.* 1:5896-599 (1955)), and *Thermotoga maritima* (Bock et al., *J. Bacteriol.* 181:1861-1867 (1999)). This reaction is also catalyzed by some phosphotranbutyrylase enzymes (EC 2.3.1.19) including the *ptb* gene products from *Clostridium acetobutylicum* (Wiesenborn et al., *App. Environ. Microbiol.* 55:317-322 (1989); Walter et al., *Gene* 134:107-111 (1993)). Additional *ptb* genes are 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)).

<u>Protein</u>	<u>GenBank ID</u>	<u>GI Number</u>	<u>Organism</u>
<i>Pta</i>	NP_416800.1	71152910	<i>Escherichia coli</i>
<i>Pta</i>	P39646	730415	<i>Bacillus subtilis</i>
<i>Pta</i>	A5N801	146346896	<i>Clostridium kluyveri</i>
<i>Pta</i>	Q9X0L4	6685776	<i>Thermotoga maritima</i>

<u>Protein</u>	<u>GenBank ID</u>	<u>GI Number</u>	<u>Organism</u>
<i>Ptb</i>	NP_349676	34540484	<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>

The acylation of acetate to acetyl-CoA is catalyzed by enzymes with acetyl-CoA synthetase activity. Two enzymes that catalyze this reaction are AMP-forming acetyl-CoA synthetase (EC 6.2.1.1) and ADP-forming acetyl-CoA synthetase (EC 6.2.1.13). AMP-forming acetyl-CoA synthetase (ACS) is the predominant enzyme for activation of acetate to acetyl-CoA. Exemplary ACS enzymes are found in *E. coli* (Brown et al., *J. Gen. Microbiol.* 102:327-336 (1977)), *Ralstonia eutropha* (Priefert and Steinbuchel, *J. Bacteriol.* 174:6590-6599 (1992)), *Methanothermobacter thermautotrophicus* (Ingram-Smith and Smith, *Archaea* 2:95-107 (2007)), *Salmonella enterica* (Gulick et al., *Biochemistry* 42:2866-2873 (2003)) and *Saccharomyces cerevisiae* (Jogl and Tong, *Biochemistry* 43:1425-1431 (2004)). ADP-forming acetyl-CoA synthetases are reversible enzymes with a generally broad substrate range (Musfeldt and Schonheit, *J. Bacteriol.* 184:636-644 (2002)). Two isozymes of ADP-forming acetyl-CoA synthetases are encoded in the *Archaeoglobus fulgidus* genome by are encoded by AF1211 and AF1983 (Musfeldt and Schonheit, *supra* (2002)). The enzyme from *Haloarcula marismortui* (annotated as a succinyl-CoA synthetase) also accepts acetate as a substrate and reversibility of the enzyme was demonstrated (Brasen and Schonheit, *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 acetate, isobutyryl-CoA (preferred substrate) and phenylacetyl-CoA (Brasen and Schonheit, *supra* (2004)).

Directed evolution or engineering can be used to modify this enzyme to operate at the physiological temperature of the host organism. The enzymes from *A. fulgidus*, *H. marismortui* and *P. aerophilum* have all been cloned, functionally expressed, and characterized in *E. coli* (Brasen and Schonheit, *supra* (2004); Musfeldt and Schonheit, *supra* (2002)). Additional candidates include the succinyl-CoA synthetase encoded by *sucCD* in *E. coli* (Buck et al., *Biochemistry* 24:6245-6252 (1985)) and the acyl-CoA ligase from *Pseudomonas putida* (Fernandez-Valverde et al., *Appl. Environ. Microbiol.* 59:1149-1154 (1993)). The aforementioned proteins are tabulated below.

<u>Protein</u>	<u>GenBank ID</u>	<u>GI Number</u>	<u>Organism</u>
<i>acs</i>	AAC77039.1	1790505	<i>Escherichia coli</i>
<i>acoE</i>	AAA21945.1	141890	<i>Ralstonia eutropha</i>
<i>acsI</i>	ABC87079.1	86169671	<i>Methanothermobacter thermautotrophicus</i>
<i>acsI</i>	AAL23099.1	16422835	<i>Salmonella enterica</i>
<i>ACSI</i>	Q01574.2	257050994	<i>Saccharomyces cerevisiae</i>
<i>AF1211</i>	NP_070039.1	11498810	<i>Archaeoglobus fulgidus</i>
<i>AF1983</i>	NP_070807.1	11499565	<i>Archaeoglobus fulgidus</i>
<i>scs</i>	YP_135572.1	55377722	<i>Haloarcula marismortui</i>
<i>PAE3250</i>	NP_560604.1	18313937	<i>Pyrobaculum aerophilum</i> str. IM2
<i>sucC</i>	NP_415256.1	16128703	<i>Escherichia coli</i>
<i>sucD</i>	AAC73823.1	1786949	<i>Escherichia coli</i>
<i>paaF</i>	AAC24333.2	22711873	<i>Pseudomonas putida</i>

The product yields per C-mol of substrate of microbial cells synthesizing reduced fermentation products such as a primary alcohol, a fatty acyl-CoA, a fatty ester, or a wax, are limited by insufficient reducing equivalents in the carbohydrate feedstock. Reducing equivalents, or electrons, can be extracted from synthesis gas components such as CO and H₂ using carbon monoxide dehydrogenase (CODH) and hydrogenase enzymes, respectively. The reducing equivalents are then passed to acceptors such as oxidized ferredoxins, oxidized quinones, oxidized cytochromes, NAD(P)⁺, water, or hydrogen peroxide to form reduced ferredoxin, reduced quinones, reduced cytochromes, NAD(P)H, H₂, or water, respectively. Reduced ferredoxin and NAD(P)H are particularly useful as they can serve as redox carriers for various Wood-Ljungdahl pathway and reductive TCA cycle enzymes.

Herein below the enzymes and the corresponding genes used for extracting redox from synags components are described. CODH is a reversible enzyme that interconverts CO and CO₂ at the expense or gain of electrons. The natural physiological role of the CODH in ACS/CODH complexes is to convert CO₂ to CO for incorporation into acetyl-CoA by acetyl-CoA synthase.

Nevertheless, such CODH enzymes are suitable for the extraction of reducing equivalents from CO due to the reversible nature of such enzymes. Expressing such CODH enzymes in the absence of ACS allows them to operate in the direction opposite to their natural physiological role (i.e., CO oxidation).

In *M. thermoacetica*, *C. hydrogenoformans*, *C. carboxidivorans* P7, and several other organisms, 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 *M. thermoacetica* gene (Genbank Accession Number: YP_430813) is expressed by itself in an operon and is believed to transfer electrons from CO to an external mediator like ferredoxin in a “Ping-pong” reaction. The reduced mediator then couples to other reduced nicotinamide adenine dinucleotide phosphate (NAD(P)H) carriers or ferredoxin-dependent cellular processes (Ragsdale, *Annals of the New York Academy of Sciences* 1125: 129–136 (2008)). The genes encoding the *C. hydrogenoformans* CODH-II and CooF, a neighboring protein, were cloned and sequenced (Gonzalez and Robb, *FEMS Microbiol Lett.* 191:243-247 (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.* 183:5134-5144 (2001)). The crystal structure of the CODH-II is also available (Dobbek et al., *Science* 293:1281-1285 (2001)). Similar ACS-free CODH enzymes can be found in a diverse array of organisms including *Geobacter metallireducens* GS-15, *Chlorobium phaeobacteroides* DSM 266, *Clostridium cellulolyticum* H10, *Desulfovibrio desulfuricans* subsp. *desulfuricans* str. ATCC 27774, *Pelobacter carbinolicus* DSM 2380, and *Campylobacter curvus* 525.92.

<u>Protein</u>	<u>GenBank ID</u>	<u>GI Number</u>	<u>Organism</u>
CODH (putative)	YP_430813	83590804	<i>Moorella thermoacetica</i>
CODH-II (CooS-II)	YP_358957	78044574	<i>Carboxydotherrmus hydrogenoformans</i>
CooF	YP_358958	78045112	<i>Carboxydotherrmus hydrogenoformans</i>
<u>CODH (putative)</u>	ZP_05390164.1	255523193	<i>Clostridium carboxidivorans</i> P7
CcarbDRAFT_0341	ZP_05390341.1	255523371	<i>Clostridium carboxidivorans</i> P7

CcarbDRAFT_1756	ZP_05391756.1	255524806	<i>Clostridium carboxidivorans</i> P7
CcarbDRAFT_2944	ZP_05392944.1	255526020	<i>Clostridium carboxidivorans</i> P7
CODH	YP_384856.1	78223109	<i>Geobacter metallireducens</i> GS-15
<i>Cpha266_0148</i> (cytochrome c)	YP_910642.1	119355998	<i>Chlorobium phaeobacteroides</i> DSM 266
<i>Cpha266_0149</i> (CODH)	YP_910643.1	119355999	<i>Chlorobium phaeobacteroides</i> DSM 266
<i>Ccel_0438</i>	YP_002504800.1	220927891	<i>Clostridium cellulolyticum</i> H10
<i>Ddes_0382</i> (CODH)	YP_002478973.1	220903661	<i>Desulfovibrio desulfuricans</i> subsp. <i>desulfuricans</i> str. ATCC 27774
<i>Ddes_0381</i> (CooC)	YP_002478972.1	220903660	<i>Desulfovibrio desulfuricans</i> subsp. <i>desulfuricans</i> str. ATCC 27774
<i>Pcar_0057</i> (CODH)	YP_355490.1	7791767	<i>Pelobacter carbinolicus</i> DSM 2380
<i>Pcar_0058</i> (CooC)	YP_355491.1	7791766	<i>Pelobacter carbinolicus</i> DSM 2380
<i>Pcar_0058</i> (HypA)	YP_355492.1	7791765	<i>Pelobacter carbinolicus</i> DSM 2380
<i>CooS</i> (CODH)	YP_001407343.1	154175407	<i>Campylobacter curvus</i> 525.92

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 and H₂O to CO₂ and H₂ (Fox et al., *J Bacteriol.* 178:6200-6208 (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: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.* 129:10328-10329 (2007)). The protein sequences of

exemplary CODH and hydrogenase genes can be identified by the following GenBank accession numbers.

<u>Protein</u>	<u>GenBank ID</u>	<u>GI Number</u>	<u>Organism</u>
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>Carboxydotherrmus hydrogenoformans</i>
CooX	YP_360649	78043124	<i>Carboxydotherrmus hydrogenoformans</i>
CooL	YP_360650	78043938	<i>Carboxydotherrmus hydrogenoformans</i>
CooK	YP_360651	78044700	<i>Carboxydotherrmus hydrogenoformans</i>
CooM	YP_360652	78043942	<i>Carboxydotherrmus hydrogenoformans</i>
CooC	YP_360654.1	78043296	<i>Carboxydotherrmus hydrogenoformans</i>
CooA-1	YP_360655.1	78044021	<i>Carboxydotherrmus hydrogenoformans</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>

Native to *E. coli* and other enteric bacteria are multiple genes encoding up to four hydrogenases (Sawers, G., *Antonie Van Leeuwenhoek* 66:57-88 (1994); Sawers et al., *J Bacteriol.* 164:1324-1331 (1985); Sawers and Boxer, *Eur.J Biochem.* 156:265-275 (1986); Sawers et al., *J Bacteriol.* 168:398-404 (1986)). Given the multiplicity of enzyme activities, *E. coli* or another host organism can provide sufficient hydrogenase activity to split incoming molecular hydrogen and reduce the corresponding acceptor. *E. coli* possesses two uptake hydrogenases, Hyd-1 and Hyd-2, encoded by the *hyaABCDEF* and *hybOABCDEFG* gene clusters, respectively (Lukey et al., How *E. coli* is equipped to oxidize hydrogen under different redox conditions, *J Biol Chem* published online Nov 16, 2009). Hyd-1 is oxygen-tolerant, irreversible, and is coupled to quinone reduction via the *hyaC* cytochrome. Hyd-2 is sensitive to O₂, reversible, and transfers electrons to the periplasmic ferredoxin *hybA* which, in turn, reduces a quinone via the *hybB* integral membrane protein. Reduced quinones can serve as the source of electrons for fumarate reductase in the reductive branch of the TCA cycle. Reduced ferredoxins can be used by enzymes such as NAD(P)H:ferredoxin oxidoreductases to generate NADPH or NADH. They can alternatively be used as the electron donor for reactions such as pyruvate ferredoxin oxidoreductase, AKG ferredoxin oxidoreductase, and 5,10-methylene-H4folate reductase.

<u>Protein</u>	<u>GenBank ID</u>	<u>GI Number</u>	<u>Organism</u>
HyaA	AAC74057.1	1787206	<i>Escherichia coli</i>
HyaB	AAC74058.1	1787207	<i>Escherichia coli</i>
HyaC	AAC74059.1	1787208	<i>Escherichia coli</i>
HyaD	AAC74060.1	1787209	<i>Escherichia coli</i>
HyaE	AAC74061.1	1787210	<i>Escherichia coli</i>
HyaF	AAC74062.1	1787211	<i>Escherichia coli</i>

<u>Protein</u>	<u>GenBank ID</u>	<u>GI Number</u>	<u>Organism</u>
HybO	AAC76033.1	1789371	<i>Escherichia coli</i>
HybA	AAC76032.1	1789370	<i>Escherichia coli</i>
HybB	AAC76031.1	2367183	<i>Escherichia coli</i>
HybC	AAC76030.1	1789368	<i>Escherichia coli</i>
HybD	AAC76029.1	1789367	<i>Escherichia coli</i>
HybE	AAC76028.1	1789366	<i>Escherichia coli</i>

HybF	AAC76027.1	1789365	<i>Escherichia coli</i>
HybG	AAC76026.1	1789364	<i>Escherichia coli</i>

The hydrogen-lyase systems of *E. coli* include 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 3 has been shown to be a reversible enzyme (Maeda et al., *Appl Microbiol Biotechnol* 76(5):1035-42 (2007)). Hydrogenase activity in *E. coli* is also dependent upon the expression of the *hyp* genes whose corresponding proteins are involved in the assembly of the hydrogenase complexes (Jacobi et al., *Arch. Microbiol* 158:444-451 (1992); Rangarajan et al., *J. Bacteriol.* 190:1447-1458 (2008)).

<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>

<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>

<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>

The *M. thermoacetica* hydrogenases are suitable for a host that lacks 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 Wood-Ljungdahl pathway (Drake, H. L., *J. Bacteriol.* 150:702-709 (1982); Drake and Daniel, *Res. Microbiol.* 155:869-883 (2004); Kellum and Drake, *J. Bacteriol.* 160:466-469 (1984)) (see Figure 21). *M. thermoacetica* has homologs to several *hyp*, *hyc*, and *hyf* genes from *E. coli*. The protein sequences encoded for by these genes are identified by the following GenBank accession numbers.

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>

Proteins in *M. thermoacetica* that are homologous to the *E. coli* Hydrogenase 3 and/or 4 proteins are listed in the following table.

<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>

In addition, several gene clusters encoding hydrogenase functionality are present in *M. thermoacetica* and their corresponding protein sequences are provided 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	83589663	<i>Moorella thermoacetica</i>
Moth_0812	YP_429673	83589664	<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>

Ralstonia eutropha H16 uses hydrogen as an energy source with oxygen as a terminal electron acceptor. Its membrane-bound uptake [NiFe]-hydrogenase is an “O₂-tolerant” hydrogenase (Cracknell, et al. Proc Nat Acad Sci, 106(49) 20681-20686 (2009)) that is periplasmically-oriented and connected to the respiratory chain via a b-type cytochrome (Schink and Schlegel, *Biochim. Biophys. Acta*, 567, 315-324 (1979); Bernhard et al., *Eur. J. Biochem.* 248, 179–186 (1997)). *R. eutropha* also contains an O₂-tolerant soluble hydrogenase encoded by the *Hox* operon which is cytoplasmic and directly reduces NAD⁺ at the expense of hydrogen (Schneider and Schlegel, *Biochim. Biophys. Acta* 452, 66–80 (1976); Burgdorf, *J. Bact.* 187(9) 3122-3132(2005)). Soluble hydrogenase enzymes are additionally present in several other organisms

including *Geobacter sulfurreducens* (Coppi, *Microbiology* 151, 1239–1254 (2005)), *Synechocystis* str. PCC 6803 (Germer, *J. Biol. Chem.*, 284(52), 36462–36472 (2009)), and *Thiocapsa roseopersicina* (Rakhely, *Appl. Environ. Microbiol.* 70(2) 722–728 (2004)). The *Synechocystis* enzyme is capable of generating NADPH from hydrogen. Overexpression of both the *Hox* operon from *Synechocystis* str. PCC 6803 and the accessory genes encoded by the *Hyp* operon from *Nostoc* sp. PCC 7120 led to increased hydrogenase activity compared to expression of the *Hox* genes alone (Germer, *J. Biol. Chem.* 284(52), 36462–36472 (2009)).

<u>Protein</u>	<u>GenBank ID</u>	<u>GI Number</u>	<u>Organism</u>
HoxF	NP_942727.1	38637753	<i>Ralstonia eutropha</i> H16
HoxU	NP_942728.1	38637754	<i>Ralstonia eutropha</i> H16
HoxY	NP_942729.1	38637755	<i>Ralstonia eutropha</i> H16
HoxH	NP_942730.1	38637756	<i>Ralstonia eutropha</i> H16
HoxW	NP_942731.1	38637757	<i>Ralstonia eutropha</i> H16
HoxI	NP_942732.1	38637758	<i>Ralstonia eutropha</i> H16

HoxE	NP_953767.1	39997816	<i>Geobacter sulfurreducens</i>
HoxF	NP_953766.1	39997815	<i>Geobacter sulfurreducens</i>
HoxU	NP_953765.1	39997814	<i>Geobacter sulfurreducens</i>
HoxY	NP_953764.1	39997813	<i>Geobacter sulfurreducens</i>
HoxH	NP_953763.1	39997812	<i>Geobacter sulfurreducens</i>
GSU2717	NP_953762.1	39997811	<i>Geobacter sulfurreducens</i>
HoxE	NP_441418.1	16330690	<i>Synechocystis</i> str. PCC 6803
HoxF	NP_441417.1	16330689	<i>Synechocystis</i> str. PCC 6803
Unknown function	NP_441416.1	16330688	<i>Synechocystis</i> str. PCC 6803
HoxU	NP_441415.1	16330687	<i>Synechocystis</i> str. PCC 6803
HoxY	NP_441414.1	16330686	<i>Synechocystis</i> str. PCC 6803
Unknown function	NP_441413.1	16330685	<i>Synechocystis</i> str. PCC 6803
Unknown function	NP_441412.1	16330684	<i>Synechocystis</i> str. PCC 6803
HoxH	NP_441411.1	16330683	<i>Synechocystis</i> str. PCC 6803
HypF	NP_484737.1	17228189	<i>Nostoc</i> sp. PCC 7120

HypC	NP_484738.1	17228190	<i>Nostoc</i> sp. PCC 7120
HypD	NP_484739.1	17228191	<i>Nostoc</i> sp. PCC 7120
Unknown function	NP_484740.1	17228192	<i>Nostoc</i> sp. PCC 7120
HypE	NP_484741.1	17228193	<i>Nostoc</i> sp. PCC 7120
HypA	NP_484742.1	17228194	<i>Nostoc</i> sp. PCC 7120
HypB	NP_484743.1	17228195	<i>Nostoc</i> sp. PCC 7120

Hox1E	AAP50519.1	37787351	<i>Thiocapsa roseopersicina</i>
Hox1F	AAP50520.1	37787352	<i>Thiocapsa roseopersicina</i>
Hox1U	AAP50521.1	37787353	<i>Thiocapsa roseopersicina</i>
Hox1Y	AAP50522.1	37787354	<i>Thiocapsa roseopersicina</i>
Hox1H	AAP50523.1	37787355	<i>Thiocapsa roseopersicina</i>

Several enzymes and the corresponding genes used for fixing carbon dioxide to either pyruvate or phosphoenolpyruvate to form the TCA cycle intermediates, oxaloacetate or malate are described below.

Carboxylation of phosphoenolpyruvate to oxaloacetate is catalyzed by phosphoenolpyruvate carboxylase. Exemplary PEP carboxylase enzymes are encoded by *ppc* in *E. coli* (Kai et al., *Arch. Biochem. Biophys.* 414:170-179 (2003)), *ppcA* in *Methylobacterium extorquens* AM1 (Arps et al., *J. Bacteriol.* 175:3776-3783 (1993)), and *ppc* in *Corynebacterium glutamicum* (Eikmanns et al., *Mol. Gen. Genet.* 218:330-339 (1989)).

<u>Protein</u>	<u>GenBank ID</u>	<u>GI Number</u>	<u>Organism</u>
<i>Ppc</i>	NP_418391	16131794	<i>Escherichia coli</i>
<i>ppcA</i>	AAB58883	28572162	<i>Methylobacterium extorquens</i>
<i>Ppc</i>	ABB53270	80973080	<i>Corynebacterium glutamicum</i>

An alternative enzyme for converting phosphoenolpyruvate to oxaloacetate is PEP carboxykinase, which simultaneously forms an ATP while carboxylating PEP. In most organisms PEP carboxykinase serves a gluconeogenic function and converts oxaloacetate to PEP

at the expense of one ATP. *S. cerevisiae* is one such organism whose native PEP carboxykinase, *PCK1*, serves a gluconeogenic role (Valdes-Hevia et al., *FEBS Lett.* 258:313-316 (1989)). *E. coli* is another such organism, as the role of PEP carboxykinase in producing oxaloacetate is believed to be minor when compared to PEP carboxylase, which does not form ATP, possibly due to the higher K_m for bicarbonate of PEP carboxykinase (Kim et al., *Appl. Environ. Microbiol.* 70:1238-1241 (2004)). Nevertheless, activity of the native *E. coli* PEP carboxykinase from PEP towards oxaloacetate has been recently demonstrated in *ppc* mutants of *E. coli* K-12 (Kwon et al., *J. Microbiol. Biotechnol.* 16:1448-1452 (2006)). These strains exhibited no growth defects and had increased succinate production at high NaHCO_3 concentrations. Mutant strains of *E. coli* can adopt Pck as the dominant CO_2 -fixing enzyme following adaptive evolution (Zhang et al. 2009). In some organisms, particularly rumen bacteria, PEP carboxykinase is quite efficient in producing oxaloacetate from PEP and generating ATP. Examples of PEP carboxykinase genes that have been cloned into *E. coli* include those from *Mannheimia succiniciproducens* (Lee et al., *Biotechnol. Bioprocess Eng.* 7:95-99 (2002)), *Anaerobiospirillum succiniciproducens* (Laivenieks et al., *Appl. Environ. Microbiol.* 63:2273-2280 (1997)), and *Actinobacillus succinogenes* (Kim et al. *supra*). The PEP carboxykinase enzyme encoded by *Haemophilus influenza* is effective at forming oxaloacetate from PEP.

<u>Protein</u>	<u>GenBank ID</u>	<u>GI Number</u>	<u>Organism</u>
<i>PCK1</i>	NP_013023	6322950	<i>Saccharomyces cerevisiae</i>
<i>pck</i>	NP_417862.1	16131280	<i>Escherichia coli</i>
<i>pckA</i>	YP_089485.1	52426348	<i>Mannheimia succiniciproducens</i>
<i>pckA</i>	O09460.1	3122621	<i>Anaerobiospirillum succiniciproducens</i>
<i>pckA</i>	Q6W6X5	75440571	<i>Actinobacillus succinogenes</i>
<i>pckA</i>	P43923.1	1172573	<i>Haemophilus influenza</i>

Pyruvate carboxylase (EC 6.4.1.1) directly converts pyruvate to oxaloacetate at the cost of one ATP. Pyruvate carboxylase enzymes are encoded by *PYC1* (Walker et al., *Biochem. Biophys. Res. Commun.* 176:1210-1217 (1991)) and *PYC2* (Walker et al., *supra*) in *Saccharomyces*

cerevisiae, and *pyc* in *Mycobacterium smegmatis* (Mukhopadhyay and Purwantini, *Biochim. Biophys. Acta* 1475:191-206 (2000)).

<u>Protein</u>	<u>GenBank ID</u>	<u>GI Number</u>	<u>Organism</u>
<i>PYC1</i>	NP_011453	6321376	<i>Saccharomyces cerevisiae</i>
<i>PYC2</i>	NP_009777	6319695	<i>Saccharomyces cerevisiae</i>
<i>Pyc</i>	YP_890857.1	118470447	<i>Mycobacterium smegmatis</i>

Malic enzyme can be applied to convert CO₂ and pyruvate to malate at the expense of one reducing equivalent. Malic enzymes for this purpose can include, without limitation, malic enzyme (NAD-dependent) and malic enzyme (NADP-dependent). For example, one of the *E. coli* malic enzymes (Takeo, *J. Biochem.* 66:379-387 (1969)) or a similar enzyme with higher activity can be expressed to enable the conversion of pyruvate and CO₂ to malate. By fixing carbon to pyruvate as opposed to PEP, malic enzyme allows the high-energy phosphate bond from PEP to be conserved by pyruvate kinase whereby ATP is generated in the formation of pyruvate or by the phosphotransferase system for glucose transport. Although malic enzyme is typically assumed to operate in the direction of pyruvate formation from malate, overexpression of the NAD-dependent enzyme, encoded by *maeA*, has been demonstrated to increase succinate production in *E. coli* while restoring the lethal Δ pfl- Δ ldhA phenotype under anaerobic conditions by operating in the carbon-fixing direction (Stols and Donnelly, *Appl. Environ. Microbiol.* 63(7) 2695-2701 (1997)). A similar observation was made upon overexpressing the malic enzyme from *Ascaris suum* in *E. coli* (Stols et al., *Appl. Biochem. Biotechnol.* 63-65(1), 153-158 (1997)). The second *E. coli* malic enzyme, encoded by *maeB*, is NADP-dependent and also decarboxylates oxaloacetate and other alpha-keto acids (Iwakura et al., *J. Biochem.* 85(5):1355-65 (1979)).

<u>Protein</u>	<u>GenBank ID</u>	<u>GI Number</u>	<u>Organism</u>
<i>maeA</i>	NP_415996	90111281	<i>Escherichia coli</i>
<i>maeB</i>	NP_416958	16130388	<i>Escherichia coli</i>
<i>NAD-ME</i>	P27443	126732	<i>Ascaris suum</i>

The enzymes used for converting oxaloacetate (formed from, for example, PEP carboxylase, PEP carboxykinase, or pyruvate carboxylase) or malate (formed from, for example, malic enzyme or malate dehydrogenase) to succinyl-CoA via the reductive branch of the TCA cycle are malate dehydrogenase, fumarate hydratase (fumarase), fumarate reductase, and succinyl-CoA transferase. The genes for each of the enzymes are described herein above.

Enzymes, genes and methods for engineering pathways from succinyl-CoA to various products into a microorganism are now known in the art. The additional reducing equivalents obtained from CO and/or H₂, as disclosed herein, improve the yields of a primary alcohol, a fatty acyl-CoA, a fatty ester, or a wax when utilizing carbohydrate-based feedstock. For example, a primary alcohol, a fatty acyl-CoA, a fatty ester, or a wax can be produced from succinyl-CoA via the reverse TCA cycle for production of acetyl-CoA, which is then converted to acyl-CoA by a ketoacyl-CoA acyltransferase or a ketoacyl-CoA thiolase, a 3-hydroxyacyl-CoA dehydrogenase, an enoyl-CoA hydratase and an enoyl-CoA reductase. Exemplary enzymes for the conversion acyl-CoA to a primary alcohol, a fatty acyl-CoA, a fatty ester, or a wax include an acyl-CoA reductase, an alcohol dehydrogenase, a wax ester synthase and an alcohol acetyltransferase.

Enzymes, genes and methods for engineering pathways from glycolysis intermediates to various products into a microorganism are known in the art. The additional reducing equivalents obtained from CO and H₂, as described herein, improve the yields of all these products on carbohydrates.

EXAMPLE V

Methods for Handling CO and Anaerobic Cultures

This example describes methods used in handling CO and anaerobic cultures.

A. Handling of CO in small quantities for assays and small cultures. CO is an odorless, colorless and tasteless gas that is a poison. Therefore, cultures and assays that utilized CO required special handling. Several assays, including CO oxidation, acetyl-CoA synthesis, CO concentration using myoglobin, and CO tolerance/utilization in small batch cultures, called for small quantities of the CO gas that were dispensed and handled within a fume hood. Biochemical assays called for saturating very small quantities (<2 mL) of the biochemical assay medium or buffer with CO and then performing the assay. All of the CO handling steps were

performed in a fume hood with the sash set at the proper height and blower turned on; CO was dispensed from a compressed gas cylinder and the regulator connected to a Schlenk line. The latter ensures that equal concentrations of CO were dispensed to each of several possible cuvettes or vials. The Schlenk line was set up containing an oxygen scrubber on the input side and an oil pressure release bubbler and vent on the other side. Assay cuvettes were both anaerobic and CO-containing. Therefore, the assay cuvettes were tightly sealed with a rubber stopper and reagents were added or removed using gas-tight needles and syringes. Secondly, small (~50 mL) cultures were grown with saturating CO in tightly stoppered serum bottles. As with the biochemical assays, the CO-saturated microbial cultures were equilibrated in the fume hood using the Schlenk line setup. Both the biochemical assays and microbial cultures were in portable, sealed containers and in small volumes making for safe handling outside of the fume hood. The compressed CO tank was adjacent to the fume hood.

Typically, a Schlenk line was used to dispense CO to cuvettes, each vented. Rubber stoppers on the cuvettes were pierced with 19 or 20 gage disposable syringe needles and were vented with the same. An oil bubbler was used with a CO tank and oxygen scrubber. The glass or quartz spectrophotometer cuvettes have a circular hole on top into which a Kontes stopper sleeve, Sz7 774250-0007 was fitted. The CO detector unit was positioned proximal to the fume hood.

B. Handling of CO in larger quantities fed to large-scale cultures. Fermentation cultures are fed either CO or a mixture of CO and H₂ to simulate syngas as a feedstock in fermentative production. Therefore, quantities of cells ranging from 1 liter to several liters can include the addition of CO gas to increase the dissolved concentration of CO in the medium. In these circumstances, fairly large and continuously administered quantities of CO gas are added to the cultures. At different points, the cultures are harvested or samples removed. Alternatively, cells are harvested with an integrated continuous flow centrifuge that is part of the fermenter.

The fermentative processes are carried out under anaerobic conditions. In some cases, it is uneconomical to pump oxygen or air into fermenters to ensure adequate oxygen saturation to provide a respiratory environment. In addition, the reducing power generated during anaerobic fermentation may be needed in product formation rather than respiration. Furthermore, many of the enzymes for various pathways are oxygen-sensitive to varying degrees. Classic acetogens

such as *M. thermoacetica* are obligate anaerobes and the enzymes in the Wood-Ljungdahl pathway are highly sensitive to irreversible inactivation by molecular oxygen. While there are oxygen-tolerant acetogens, the repertoire of enzymes in the Wood-Ljungdahl pathway might be incompatible in the presence of oxygen because most are metallo-enzymes, key components are ferredoxins, and regulation can divert metabolism away from the Wood-Ljungdahl pathway to maximize energy acquisition. At the same time, cells in culture act as oxygen scavengers that moderate the need for extreme measures in the presence of large cell growth.

C. Anaerobic chamber and conditions. Exemplary anaerobic chambers are available commercially (see, for example, Vacuum Atmospheres Company, Hawthorne CA; MBraun, Newburyport MA). Conditions included an O₂ concentration of 1 ppm or less and 1 atm pure N₂. In one example, 3 oxygen scrubbers/catalyst regenerators were used, and the chamber included an O₂ electrode (such as Teledyne; City of Industry CA). Nearly all items and reagents were cycled four times in the airlock of the chamber prior to opening the inner chamber door. Reagents with a volume >5mL were sparged with pure N₂ prior to introduction into the chamber. Gloves are changed twice/yr and the catalyst containers were regenerated periodically when the chamber displays increasingly sluggish response to changes in oxygen levels. The chamber's pressure was controlled through one-way valves activated by solenoids. This feature allowed setting the chamber pressure at a level higher than the surroundings to allow transfer of very small tubes through the purge valve.

The anaerobic chambers achieved levels of O₂ that were consistently very low and were needed for highly oxygen sensitive anaerobic conditions. However, growth and handling of cells does not usually require such precautions. In an alternative anaerobic chamber configuration, platinum or palladium can be used as a catalyst that requires some hydrogen gas in the mix. Instead of using solenoid valves, pressure release can be controlled by a bubbler. Instead of using instrument-based O₂ monitoring, test strips can be used instead.

D. Anaerobic microbiology. Small cultures were handled as described above for CO handling. In particular, serum or media bottles are fitted with thick rubber stoppers and aluminum crimps are employed to seal the bottle. Medium, such as Terrific Broth, is made in a conventional manner and dispensed to an appropriately sized serum bottle. The bottles are sparged with

nitrogen for ~30 min of moderate bubbling. This removes most of the oxygen from the medium and, after this step, each bottle is capped with a rubber stopper (such as Bellco 20 mm septum stoppers; Bellco, Vineland, NJ) and crimp-sealed (Bellco 20 mm). Then the bottles of medium are autoclaved using a slow (liquid) exhaust cycle. At least sometimes a needle can be poked through the stopper to provide exhaust during autoclaving; the needle needs to be removed immediately upon removal from the autoclave. The sterile medium has the remaining medium components, for example buffer or antibiotics, added via syringe and needle. Prior to addition of reducing agents, the bottles are equilibrated for 30 - 60 minutes with nitrogen (or CO depending upon use). A reducing agent such as a 100 x 150 mM sodium sulfide, 200 mM cysteine-HCl is added. This is made by weighing the sodium sulfide into a dry beaker and the cysteine into a serum bottle, bringing both into the anaerobic chamber, dissolving the sodium sulfide into anaerobic water, then adding this to the cysteine in the serum bottle. The bottle is stoppered immediately as the sodium sulfide solution generates hydrogen sulfide gas upon contact with the cysteine. When injecting into the culture, a syringe filter is used to sterilize the solution. Other components are added through syringe needles, such as B12 (10 μ M cyanocobalamin), nickel chloride (NiCl_2 , 20 microM final concentration from a 40 mM stock made in anaerobic water in the chamber and sterilized by autoclaving or by using a syringe filter upon injection into the culture), and ferrous ammonium sulfate (final concentration needed is 100 μ M—made as 100-1000x stock solution in anaerobic water in the chamber and sterilized by autoclaving or by using a syringe filter upon injection into the culture). To facilitate faster growth under anaerobic conditions, the 1 liter bottles were inoculated with 50 mL of a preculture grown anaerobically. Induction of the pA1-lacO1 promoter in the vectors was performed by addition of isopropyl β -D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.2 mM and was carried out for about 3 hrs.

Large cultures can be grown in larger bottles using continuous gas addition while bubbling. A rubber stopper with a metal bubbler is placed in the bottle after medium addition and sparged with nitrogen for 30 minutes or more prior to setting up the rest of the bottle. Each bottle is put together such that a sterile filter will sterilize the gas bubbled in and the hoses on the bottles are compressible with small C clamps. Medium and cells are stirred with magnetic stir bars. Once

all medium components and cells are added, the bottles are incubated in an incubator in room air but with continuous nitrogen sparging into the bottles.

EXAMPLE VI

CO oxidation (CODH) Assay

This example describes assay methods for measuring CO oxidation (CO dehydrogenase; CODH).

The 7 gene CODH/ACS operon of *Moorella thermoacetica* was cloned into *E. coli* expression vectors. The intact ~10 kbp DNA fragment was cloned, and it is likely that some of the genes in this region are expressed from their own endogenous promoters and all contain endogenous ribosomal binding sites. These clones were assayed for CO oxidation, using an assay that quantitatively measures CODH activity. Antisera to the *M. thermoacetica* gene products was used for Western blots to estimate specific activity. *M. thermoacetica* is Gram positive, and ribosome binding site elements are expected to work well in *E. coli*. This activity, described below in more detail, was estimated to be ~1/50th of the *M. thermoacetica* specific activity. It is possible that CODH activity of recombinant *E. coli* cells could be limited by the fact that *M. thermoacetica* enzymes have temperature optima around 55°C. Therefore, a mesophilic CODH/ACS pathway could be advantageous such as the close relative of *Moorella* that is mesophilic and does have an apparently intact CODH/ACS operon and a Wood-Ljungdahl pathway, *Desulfitobacterium hafniense*. Acetogens as potential host organisms include, but are not limited to, *Rhodospirillum rubrum*, *Moorella thermoacetica* and *Desulfitobacterium hafniense*.

CO oxidation is both the most sensitive and most robust of the CODH/ACS assays. It is likely that an *E. coli*-based syngas using system will ultimately need to be about as anaerobic as *Clostridial* (i.e., *Moorella*) systems, especially for maximal activity. Improvement in CODH should be possible but will ultimately be limited by the solubility of CO gas in water.

Initially, each of the genes was cloned individually into expression vectors. Combined expression units for multiple subunits/1 complex were generated. Expression in *E. coli* at the

protein level was determined. Both combined *M. thermoacetica* CODH/ACS operons and individual expression clones were made.

CO oxidation assay. This assay is one of the simpler, reliable, and more versatile assays of enzymatic activities within the Wood-Ljungdahl pathway and tests CODH (Seravalli et al., *Biochemistry* 43:3944-3955 (2004)). A typical activity of *M. thermoacetica* CODH specific activity is 500 U at 55°C or ~60U at 25°C. This assay employs reduction of methyl viologen in the presence of CO. This is measured at 578 nm in stoppered, anaerobic, glass cuvettes.

In more detail, glass rubber stoppered cuvettes were prepared after first washing the cuvette four times in deionized water and one time with acetone. A small amount of vacuum grease was smeared on the top of the rubber gasket. The cuvette was gassed with CO, dried 10 min with a 22 Ga. needle plus an exhaust needle. A volume of 0.98 mL of reaction buffer (50 mM Hepes, pH 8.5, 2mM dithiothreitol (DTT) was added using a 22 Ga. needle, with exhaust needed, and 100%CO. Methyl viologen (CH₃ viologen) stock was 1 M in water. Each assay used 20 microliters for 20 mM final concentration. When methyl viologen was added, an 18 Ga needle (partial) was used as a jacket to facilitate use of a Hamilton syringe to withdraw the CH₃ viologen. 4 -5 aliquots were drawn up and discarded to wash and gas equilibrate the syringe. A small amount of sodium dithionite (0.1 M stock) was added when making up the CH₃ viologen stock to slightly reduce the CH₃ viologen. The temperature was equilibrated to 55°C in a heated Olis spectrophotometer (Bogart GA). A blank reaction (CH₃ viologen + buffer) was run first to measure the base rate of CH₃ viologen reduction. Crude *E. coli* cell extracts of ACS90 and ACS91 (CODH-ACS operon of *M. thermoacetica* with and without, respectively, the first *cooC*). 10 microliters of extract were added at a time, mixed and assayed. Reduced CH₃ viologen turns purple. The results of an assay are shown in Table I.

Table I. 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	
Extract	Vol	OD/	U/ml	U/mg
ACS90	10 microliters	0.073	0.376	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	0.51	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			

Mta98/Mta99 are *E. coli* MG1655 strains that express methanol methyltransferase genes from *M. thermoacetica* and, therefore, are negative controls for the ACS90 ACS91 *E. coli* strains that contain *M. thermoacetica* CODH operons.

If ~ 1% of the cellular protein is CODH, then these figures would be approximately 100X less than the 500 U/mg activity of pure *M. thermoacetica* CODH. Actual estimates based on Western blots are 0.5% of the cellular protein, so the activity is about 50X less than for *M. thermoacetica* CODH. Nevertheless, this experiment demonstrates 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.

To estimate the final concentrations of CODH and Mtr 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* CODH-ACS and Mtr proteins and were visualized using an alkaline phosphatase-

linked goat-anti-rabbit secondary antibody. The Westerns were performed and results are shown in Figure 23. The amounts of CODH in ACS90 and ACS91 were estimated at 50 ng by comparison to the control lanes. Expression of CODH-ACS operon genes including 2 CODH subunits and the methyltransferase were confirmed via Western blot analysis. Therefore, the recombinant *E. coli* cells express multiple components of a 7 gene operon. In addition, both the methyltransferase and corrinoid iron sulfur protein were active in the same recombinant *E. coli* cells. These proteins are part of the same operon cloned into the same cells.

The CO oxidation assays were repeated using extracts of *Moorella thermoacetica* cells for the positive controls. Though CODH activity in *E. coli* ACS90 and ACS91 was measurable, it was at about 130 – 150 X lower than the *M. thermoacetica* control. The results of the assay are shown in Figure 24. Briefly, cells (*M. thermoacetica* or *E. coli* with the CODH/ACS operon; ACS90 or ACS91 or empty vector: pZA33S) were grown and extracts prepared as described herein. Assays were performed as described above at 55°C at various times on the day the extracts were prepared. Reduction of methylviologen was followed at 578 nm over a 120 sec time course.

These results describe the CO oxidation (CODH) assay and results. Recombinant *E. coli* cells expressed CO oxidation activity as measured by the methyl viologen reduction assay.

EXAMPLE VII

***E. coli* CO Tolerance Experiment and CO Concentration Assay (myoglobin assay)**

This example describes the tolerance of *E. coli* for high concentrations of CO. To test whether or not *E. coli* can grow anaerobically in the presence of saturating amounts of CO, cultures were set up in 120 ml serum bottles with 50 ml of Terrific Broth medium (plus reducing solution, NiCl₂, Fe(II)NH₄SO₄, cyanocobalamin, IPTG, and chloramphenicol) as described above for anaerobic microbiology in small volumes. One half of these bottles were equilibrated with nitrogen gas for 30 min. and one half was equilibrated with CO gas for 30 min. An empty vector (pZA33) was used as a control, and cultures containing the pZA33 empty vector as well as both ACS90 and ACS91 were tested with both N₂ and CO. All were inoculated and grown for 36 hrs with shaking (250 rpm) at 37°C. At the end of the 36 hour period,

examination of the flasks showed high amounts of growth in all. The bulk of the observed growth occurred overnight with a long lag.

Given that all cultures appeared to grow well in the presence of CO, the final CO concentrations were confirmed. This was performed using an assay of the spectral shift of myoglobin upon exposure to CO. Myoglobin reduced with sodium dithionite has an absorbance peak at 435 nm; this peak is shifted to 423 nm with CO. Due to the low wavelength and need to record a whole spectrum from 300 nm on upwards, quartz cuvettes must be used. CO concentration is measured against a standard curve and depends upon the Henry's Law constant for CO of maximum water solubility = 970 micromolar at 20°C and 1 atm.

For the myoglobin test of CO concentration, cuvettes were washed 10X with water, 1X with acetone, and then stoppered as with the CODH assay. N₂ was blown into the cuvettes for ~10 min. A volume of 1 ml of anaerobic buffer (HEPES, pH 8.0, 2mM DTT) was added to the blank (not equilibrated with CO) with a Hamilton syringe. A volume of 10 microliter myoglobin (~1 mM—can be varied, just need a fairly large amount) and 1 microliter dithionite (20 mM stock) were added. A CO standard curve was made using CO saturated buffer added at 1 microliter increments. Peak height and shift was recorded for each increment. The cultures tested were pZA33/CO, ACS90/CO, and ACS91/CO. Each of these was added in 1 microliter increments to the same cuvette. Midway through the experiment a second cuvette was set up and used. The results are shown in Table II.

Table II. Carbon Monoxide Concentrations, 36 hrs.

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

The results shown in Table II indicate that the cultures grew whether or not a strain was cultured in the presence of CO or not. These results indicate that *E. coli* can tolerate exposure to CO under anaerobic conditions and that *E. coli* cells expressing the CODH-ACS operon can metabolize some of the CO.

These results demonstrate that *E. coli* cells, whether expressing CODH/ACS or not, were able to grow in the presence of saturating amounts of CO. Furthermore, these grew equally well as the controls in nitrogen in place of CO. This experiment demonstrated that laboratory strains of *E. coli* are insensitive to CO at the levels achievable in a syngas project performed at normal atmospheric pressure. In addition, preliminary experiments indicated that the recombinant *E. coli* cells expressing CODH/ACS actually consumed some CO, probably by oxidation to carbon dioxide.

EXAMPLE VIII

The following example provides various gene candidates for enzymes disclosed herein for conversion of acyl-CoA to fatty alcohols and other related compounds.

Acetyl-CoA carboxylase

Gene	Organism	Accession number	GI Number
accA	Escherichia coli K-12	AAC73296.1	1786382
accB	Escherichia coli K-12	AAC76287.1	1789653
accC	Escherichia coli K-12	AAC76288.1	1789654
accD	Escherichia coli K-12	AAC75376.1	1788655
accA	Salmonella enterica	CAD08690.1	16501513
accB	Salmonella enterica	CAD07894.1	16504441
accC	Salmonella enterica	CAD07895.1	16504442
accD	Salmonella enterica	CAD07598.1	16503590
YMR207C	Saccharomyces cerevisiae	NP_013934.1	6323863
YNR016C	Saccharomyces	NP_014413.1	6324343

	cerevisiae		
YGR037C	Saccharomyces cerevisiae	NP_011551.1	6321474
YKL182W	Saccharomyces cerevisiae	NP_012739.1	6322666
YPL231W	Saccharomyces cerevisiae	NP_015093.1	6325025

Fatty acid synthase

Multiple genes are required for that step: 1. Acetyl-CoA: ACP transacylase (2.3.1.38)- for converting acetyl-CoA to acetyl-ACP, 2. malonyl-CoA:ACP transacylase that converts malonyl-CoA into malonyl-ACP (2.3.1.39), 3. acetyl[acp]:malonyl-[acp] C-acyl transferase (2.3.1.41) and others in fatty acid elongation

Gene	Accession number	GI number	Organism
<i>fabH</i>	AP_001717.1	89107937	<i>Escherichia coli</i> K12
<i>fabB</i>	NP_416826.1	16130258	<i>Escherichia coli</i> K12
<i>fabF</i>	NP_415613.1	16129058	<i>Escherichia coli</i> K12
<i>fabD</i>	NP_415610.1	16129055	<i>Escherichia coli</i> K12
<i>fabI</i>	NP_415804.1	16129249	<i>Escherichia coli</i> K12
<i>fabA</i>	NP_415474.1	16128921	<i>Escherichia coli</i> K12
<i>fabZ</i>	NP_414722.1	16128173	<i>Escherichia coli</i> K12
<i>fabG</i>	NP_415611.1	16129056	<i>Escherichia coli</i> K12
<i>FasII</i> (fatty acid synthase, alpha subunit)	AAA34601.1	171502	<i>Saccharomyces cerevisiae</i>
<i>FasI</i> (fatty acid synthase, beta subunit)	AAA34602.1	171506	<i>Saccharomyces cerevisiae</i>
<i>fas</i>	NP_217040.1	15609661	<i>Mycobacterium tuberculosis</i> H37Rv
<i>fas</i>	AAN25329.1	23326820	<i>Bifidobacterium longum</i> NCC2705
<i>fas</i>	YP_003971698.1	311064972	<i>Bifidobacterium bifidum</i>
<i>fas</i>	AEG82252.1	334697455	<i>Corynebacterium ulcerans</i>

Acyl acp thioesterase

Gene	Accession number	GI number	Organism
<i>fatA</i>	AEE76980	332643459	<i>Arabidopsis thaliana</i>
<i>fatA</i>	ACC41415	183176305	<i>Mycobacterium marinum M</i>
<i>fatA</i>	AAX54527	61741120	<i>Helianthus annuus</i>
<i>fatA</i>	CAC14164	10944734	<i>Brassica juncea</i>
<i>fatA</i>	ZP_04749108	240170449	<i>Mycobacterium kansasii</i> ATCC 12478
<i>fatA</i>	ZP_04384386.1	229490548	<i>Rhodococcus erythropolis SK121</i>
<i>fatA</i>	YP_885312.1	118472377	<i>Mycobacterium smegmatis</i> str. MC2 155
<i>fatB</i>	AAQ08202.1	33325193	<i>Helianthus annuus</i>
<i>fatB</i>	AEE28300.1	332190179	<i>Arabidopsis thaliana</i>
<i>fatB</i>	ABI18986.1	112455672	<i>Brassica juncea</i>
<i>tesA</i>	NP_415027.1	16128478	<i>Escherichia coli</i> K12

Acyl CoA synthetase and acyl CoA ligase

Gene	Accession number	GI number	Organism
<i>scs</i>	YP_135572.1	55377722	<i>Haloarcula marismortui</i>
<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>
<i>paaF</i>	AAC24333.2	22711873	<i>Pseudomonas putida</i>
<i>matB</i>	AAC83455.1	3982573	<i>Rhizobium leguminosarum</i>
<i>bioW</i>	NP_390902.2	50812281	<i>Bacillus subtilis</i>
<i>bioW</i>	CAA10043.1	3850837	<i>Pseudomonas mendocina</i>
<i>bioW</i>	P22822.1	115012	<i>Bacillus sphaericus</i>
<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>acs</i>	AAC77039.1	1790505	<i>Escherichia coli</i>
<i>acoE</i>	AAA21945.1	141890	<i>Ralstonia eutropha</i>

<i>acs1</i>	ABC87079.1	86169671	<i>Methanothermobacter thermautotrophicus</i>
<i>acs1</i>	AAL23099.1	16422835	<i>Salmonella enterica</i>
<i>ACS1</i>	Q01574.2	257050994	<i>Saccharomyces cerevisiae</i>
<i>AF1211</i>	NP_070039.1	11498810	<i>Archaeoglobus fulgidus</i>
<i>AF1983</i>	NP_070807.1	11499565	<i>Archaeoglobus fulgidus</i>

EXAMPLE IX Exemplary Carboxylic Acid Reductases

This example describes the use of carboxylic acid reductases to carry out the conversion of a carboxylic acid to an aldehyde. In particular, conversion of a fatty acid to a fatty aldehyde (see Figure 25 and Figure 26) can be carried out by these enzymes.

1.2.1.e Acid reductase.

The conversion of unactivated acids to aldehydes can be carried out by an acid reductase. Examples of such conversions include, but are not limited, the conversion of 4-hydroxybutyrate, succinate, alpha-ketoglutarate, and 4-aminobutyrate to 4-hydroxybutanal, succinate semialdehyde, 2,5-dioxopentanoate, and 4-aminobutanal, respectively. One notable carboxylic acid reductase can be found in *Nocardia iowensis* which 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 is encoded by the *car* gene and 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)).

Gene	Accession No.	GI No.	Organism
<i>car</i>	AAR91681.1	40796035	<i>Nocardia iowensis</i> (sp. NRRL

			5646)
<i>npt</i>	ABI83656.1	114848891	<i>Nocardia iowensis</i> (sp. NRRL 5646)

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

Gene	Accession No.	GI No.	Organism
<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

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-acetylamino-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.

Gene	Accession No.	GI No.	Organism
<i>griC</i>	182438036	YP_00182575 5.1	<i>Streptomyces griseus</i> subsp. <i>griseus</i> NBRC 13350
<i>griD</i>	182438037	YP_00182575 6.1	<i>Streptomyces griseus</i> subsp. <i>griseus</i> NBRC 13350
MSMEG_2956	YP_887275.1	YP_887275.1	<i>Mycobacterium smegmatis</i> MC2 155
MSMEG_5739	YP_889972.1	118469671	<i>Mycobacterium smegmatis</i> MC2 155
MSMEG_2648	YP_886985.1	118471293	<i>Mycobacterium smegmatis</i> MC2 155
MAP1040c	NP_959974.1	41407138	<i>Mycobacterium avium</i> subsp. <i>paratuberculosis</i> K-10
MAP2899c	NP_961833.1	41408997	<i>Mycobacterium avium</i> subsp. <i>paratuberculosis</i> K-10
MMAR_2117	YP_001850422.1	183982131	<i>Mycobacterium marinum</i> M
MMAR_2936	YP_001851230.1	183982939	<i>Mycobacterium marinum</i> M

MMAR_1916	YP_001850220.1	183981929	Mycobacterium marinum M
TpauDRAFT_33060	ZP_04027864.1	227980601	Tsukamurella paurometabola DSM 20162
TpauDRAFT_20920	ZP_04026660.1	227979396	Tsukamurella paurometabola DSM 20162
CPCC7001_1320	ZP_05045132.1	254431429	Cyanobium PCC7001
DDBDRAFT_0187729	XP_636931.1	66806417	Dictyostelium discoideum AX4

An enzyme with similar characteristics, alpha-amino adipate reductase (AAR, EC 1.2.1.31), participates in lysine biosynthesis pathways in some fungal species. This enzyme naturally reduces alpha-amino adipate to alpha-amino adipate 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 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.

Gene	Accession No.	GI No.	Organism
<i>LYS2</i>	AAA34747.1	171867	<i>Saccharomyces cerevisiae</i>
<i>LYS5</i>	P50113.1	1708896	<i>Saccharomyces cerevisiae</i>
<i>LYS2</i>	AAC02241.1	2853226	<i>Candida albicans</i>
<i>LYS5</i>	AAO26020.1	28136195	<i>Candida albicans</i>
<i>Lys1p</i>	P40976.3	13124791	<i>Schizosaccharomyces pombe</i>
<i>Lys7p</i>	Q10474.1	1723561	<i>Schizosaccharomyces pombe</i>
<i>Lys2</i>	CAA74300.1	3282044	<i>Penicillium chrysogenum</i>

Cloning and Expression of Carboxylic Acid Reductase. *Escherichia coli* is used as a target organism to engineer the pathway for primary alcohol, a fatty acyl-CoA, a fatty ester, or a wax or a primary alcohol, a fatty acyl-CoA, a fatty ester, or a wax. *E. coli* provides a good host for generating a non-naturally occurring microorganism capable of producing primary alcohol, a fatty acyl-CoA, a fatty ester, or a wax or a primary alcohol, a fatty acyl-CoA, a fatty ester, or a

wax. *E. coli* is amenable to genetic manipulation and is known to be capable of producing various intermediates and products effectively under various oxygenation conditions.

To generate a microbial organism strain such as an *E. coli* strain engineered to produce primary alcohol, a fatty acyl-CoA, a fatty ester, or a wax or a primary alcohol, a fatty acyl-CoA, a fatty ester, or a wax, nucleic acids encoding a carboxylic acid reductase and phosphopantetheine transferase are expressed in *E. coli* using well known molecular biology techniques (see, for example, Sambrook, supra, 2001; Ausubel supra, 1999). In particular, *car* genes from *Nocardia iowensis* (designated 720), *Mycobacterium smegmatis* mc(2)155 (designated 890), *Mycobacterium avium* subspecies *paratuberculosis* K-10 (designated 891) and *Mycobacterium marinum* M (designated 892) were cloned into pZS*13 vectors (Expressys, Ruelzheim, Germany) under control of PA1/lacO promoters. The *npt* (ABI83656.1) gene (i.e., 721) was cloned into the pKJL33S vector, a derivative of the original mini-F plasmid vector PML31 under control of promoters and ribosomal binding sites similar to those used in pZS*13.

The *car* gene (GNM_720) was cloned by PCR from *Nocardia* genomic DNA. Its nucleic acid and protein sequences are shown in Figures 27A and 27B, respectively. A codon-optimized version of the *npt* gene (GNM_721) was synthesized by GeneArt (Regensburg, Germany). Its nucleic acid and protein sequences are shown in Figures 28A and 28B, respectively. The nucleic acid and protein sequences for the *Mycobacterium smegmatis* mc(2)155 (designated 890), *Mycobacterium avium* subspecies *paratuberculosis* K-10 (designated 891) and *Mycobacterium marinum* M (designated 892) genes and enzymes can be found in Figures 29, 30, and 31, respectively. The plasmids are transformed into a host cell to express the proteins and enzymes required for primary alcohol, a fatty acyl-CoA, a fatty ester, or a wax or a primary alcohol, a fatty acyl-CoA, a fatty ester, or a wax production.

Additional CAR variants were generated. A codon optimized version of CAR 891 was generated and designated 891GA. The nucleic acid and amino acid sequences of CAR 891GA are shown in Figures 32A and 32B, respectively. Over 2000 CAR variants were generated. In particular, all 20 amino acid combinations were made at positions V295, M296, G297, G391, G421, D413, G414, Y415, G416, and S417, and additional variants were tested as well. Exemplary CAR variants include: E16K; Q95L; L100M; A1011T; K823E; T941S; H15Q;

D198E; G446C; S392N; F699L; V883I; F467S; T987S; R12H; V295G; V295A; V295S; V295T; V295C; V295V; V295L; V295I; V295M; V295P; V295F; V295Y; V295W; V295D; V295E; V295N; V295Q; V295H; V295K; V295R; M296G; M296A; M296S; M296T; M296C; M296V; M296L; M296I; M296M; M296P; M296F; M296Y; M296W; M296D; M296E; M296N; M296Q; M296H; M296K; M296R; G297G; G297A; G297S; G297T; G297C; G297V; G297L; G297I; G297M; G297P; G297F; G297Y; G297W; G297D; G297E; G297N; G297Q; G297H; G297K; G297R; G391G; G391A; G391S; G391T; G391C; G391V; G391L; G391I; G391M; G391P; G391F; G391Y; G391W; G391D; G391E; G391N; G391Q; G391H; G391K; G391R; G421G; G421A; G421S; G421T; G421C; G421V; G421L; G421I; G421M; G421P; G421F; G421Y; G421W; G421D; G421E; G421N; G421Q; G421H; G421K; G421R; D413G; D413A; D413S; D413T; D413C; D413V; D413L; D413I; D413M; D413P; D413F; D413Y; D413W; D413D; D413E; D413N; D413Q; D413H; D413K; D413R; G414G; G414A; G414S; G414T; G414C; G414V; G414L; G414I; G414M; G414P; G414F; G414Y; G414W; G414D; G414E; G414N; G414Q; G414H; G414K; G414R; Y415G; Y415A; Y415S; Y415T; Y415C; Y415V; Y415L; Y415I; Y415M; Y415P; Y415F; Y415Y; Y415W; Y415D; Y415E; Y415N; Y415Q; Y415H; Y415K; Y415R; G416G; G416A; G416S; G416T; G416C; G416V; G416L; G416I; G416M; G416P; G416F; G416Y; G416W; G416D; G416E; G416N; G416Q; G416H; G416K; G416R; S417G; S417A; S417S; S417T; S417C; S417V; S417L; S417I; S417M; S417P; S417F; S417Y; S417W; S417D; S417E; S417N; S417Q; S417H; S417K; and S417R.

The CAR variants were screened for activity, and numerous CAR variants were found to exhibit CAR activity.

This example describes the use of CAR for converting carboxylic acids to aldehydes.

Throughout this application various publications have been referenced. The disclosures of these publications in their entireties, including GenBank and GI number publications, 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 examples provided above, it should be understood that various modifications can be made without departing from the spirit of the invention.

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 primary alcohol pathway comprising at least one exogenous nucleic acid encoding a primary alcohol pathway enzyme expressed in a sufficient amount to produce a primary alcohol; said non-naturally occurring microbial organism further comprising:

(i) a reductive TCA pathway comprising at least one exogenous nucleic acid encoding a reductive TCA pathway enzyme, wherein said at least one exogenous nucleic acid is selected from an ATP-citrate lyase, a citrate lyase, a citryl-CoA synthetase, a citryl-CoA lyase, a fumarate reductase, and an alpha-ketoglutarate:ferredoxin oxidoreductase;

(ii) a reductive TCA pathway comprising at least one exogenous nucleic acid encoding a reductive TCA pathway enzyme, wherein said at least one exogenous nucleic acid is selected from a pyruvate:ferredoxin oxidoreductase, a phosphoenolpyruvate carboxylase, a phosphoenolpyruvate carboxykinase, a CO dehydrogenase, and an H₂ hydrogenase; or

(iii) at least one exogenous nucleic acid encodes an enzyme selected from a CO dehydrogenase, an H₂ hydrogenase, and combinations thereof;

wherein said primary alcohol pathway comprises:

(a) a malonyl-CoA-independent FAS pathway comprising ketoacyl-CoA acyltransferase or ketoacyl-CoA thiolase, 3-hydroxyacyl-CoA dehydrogenase, enoyl-CoA hydratase and enoyl-CoA reductase; and an acyl-reduction pathway comprising an acyl-CoA reductase and an alcohol dehydrogenase.

2. A non-naturally occurring microbial organism, comprising a microbial organism having a fatty acyl-CoA pathway comprising at least one exogenous nucleic acid encoding a fatty acyl-CoA pathway enzyme expressed in a sufficient amount to produce a fatty acyl-CoA; said non-naturally occurring microbial organism further comprising:

(i) a reductive TCA pathway comprising at least one exogenous nucleic acid encoding a reductive TCA pathway enzyme, wherein said at least one exogenous nucleic acid is selected

from an ATP-citrate lyase, a citrate lyase, a citryl-CoA synthetase, a citryl-CoA lyase, a fumarate reductase, and an alpha-ketoglutarate:ferredoxin oxidoreductase;

(ii) a reductive TCA pathway comprising at least one exogenous nucleic acid encoding a reductive TCA pathway enzyme, wherein said at least one exogenous nucleic acid is selected from a pyruvate:ferredoxin oxidoreductase, a phosphoenolpyruvate carboxylase, a phosphoenolpyruvate carboxykinase, a CO dehydrogenase, and an H₂ hydrogenase; or

(iii) at least one exogenous nucleic acid encodes an enzyme selected from a CO dehydrogenase, an H₂ hydrogenase, and combinations thereof;

wherein said fatty acyl-CoA pathway comprises:

(a) a malonyl-CoA-independent FAS pathway comprising ketoacyl-CoA acyltransferase or ketoacyl-CoA thiolase, 3-hydroxyacyl-CoA dehydrogenase, enoyl-CoA hydratase and enoyl-CoA reductase.

3. A non-naturally occurring microbial organism, comprising a microbial organism having a fatty ester pathway comprising at least one exogenous nucleic acid encoding a fatty ester pathway enzyme expressed in a sufficient amount to produce a fatty ester; said non-naturally occurring microbial organism further comprising:

(i) a reductive TCA pathway comprising at least one exogenous nucleic acid encoding a reductive TCA pathway enzyme, wherein said at least one exogenous nucleic acid is selected from an ATP-citrate lyase, a citrate lyase, a citryl-CoA synthetase, a citryl-CoA lyase, a fumarate reductase, and an alpha-ketoglutarate:ferredoxin oxidoreductase;

(ii) a reductive TCA pathway comprising at least one exogenous nucleic acid encoding a reductive TCA pathway enzyme, wherein said at least one exogenous nucleic acid is selected from a pyruvate:ferredoxin oxidoreductase, a phosphoenolpyruvate carboxylase, a phosphoenolpyruvate carboxykinase, a CO dehydrogenase, and an H₂ hydrogenase; or

(iii) at least one exogenous nucleic acid encodes an enzyme selected from a CO dehydrogenase, an H₂ hydrogenase, and combinations thereof;

wherein said fatty ester pathway comprises a pathway selected from:

(a) a malonyl-CoA-independent fatty acid synthesis (FAS) pathway and a wax ester synthase, wherein said malonyl-CoA-independent FAS pathway comprises ketoacyl-CoA acyltransferase or ketoacyl-CoA thiolase, 3-hydroxyacyl-CoA dehydrogenase, enoyl-CoA hydratase and enoyl-CoA reductase; and

(b) a malonyl-CoA-independent fatty acid synthesis (FAS) pathway and an alcohol acetyltransferase, wherein said malonyl-CoA-independent FAS pathway comprises ketoacyl-CoA acyltransferase or ketoacyl-CoA thiolase, 3-hydroxyacyl-CoA dehydrogenase, enoyl-CoA hydratase and enoyl-CoA reductase.

4. A non-naturally occurring microbial organism, comprising a microbial organism having a wax pathway comprising at least one exogenous nucleic acid encoding a wax pathway enzyme expressed in a sufficient amount to produce a wax; said non-naturally occurring microbial organism further comprising:

(i) a reductive TCA pathway comprising at least one exogenous nucleic acid encoding a reductive TCA pathway enzyme, wherein said at least one exogenous nucleic acid is selected from an ATP-citrate lyase, a citrate lyase, a citryl-CoA synthetase, a citryl-CoA lyase, a fumarate reductase, and an alpha-ketoglutarate:ferredoxin oxidoreductase;

(ii) a reductive TCA pathway comprising at least one exogenous nucleic acid encoding a reductive TCA pathway enzyme, wherein said at least one exogenous nucleic acid is selected from a pyruvate:ferredoxin oxidoreductase, a phosphoenolpyruvate carboxylase, a phosphoenolpyruvate carboxykinase, a CO dehydrogenase, and an H₂ hydrogenase; or

(iii) at least one exogenous nucleic acid encodes an enzyme selected from a CO dehydrogenase, an H₂ hydrogenase, and combinations thereof;

wherein said wax pathway comprises a pathway selected from:

(a) a malonyl-CoA-independent fatty acid synthesis (FAS) pathway and a wax ester synthase, wherein said malonyl-CoA-independent FAS pathway comprises ketoacyl-CoA

acyltransferase or ketoacyl-CoA thiolase, 3-hydroxyacyl-CoA dehydrogenase, enoyl-CoA hydratase and enoyl-CoA reductase; and

(b) a malonyl-CoA-independent fatty acid synthesis (FAS) pathway and an alcohol acetyltransferase, wherein said malonyl-CoA-independent FAS pathway comprises ketoacyl-CoA acyltransferase or ketoacyl-CoA thiolase, 3-hydroxyacyl-CoA dehydrogenase, enoyl-CoA hydratase and enoyl-CoA reductase.

5. The non-naturally occurring microbial organism of any one of claims 1-4, wherein said microbial organism comprising (i) further comprises an exogenous nucleic acid encoding an enzyme selected from a pyruvate:ferredoxin oxidoreductase, an aconitase, an isocitrate dehydrogenase, a succinyl-CoA synthetase, a succinyl-CoA transferase, a fumarase, a malate dehydrogenase, an acetate kinase, a phosphotransacetylase, an acetyl-CoA synthetase, an NAD(P)H:ferredoxin oxidoreductase, ferredoxin, and combinations thereof.

6. The non-naturally occurring microbial organism of any one of claims 1-4, wherein said microbial organism comprising (ii) further comprises an exogenous nucleic acid encoding an enzyme selected from an aconitase, an isocitrate dehydrogenase, a succinyl-CoA synthetase, a succinyl-CoA transferase, a fumarase, a malate dehydrogenase, and combinations thereof.

7. The non-naturally occurring microbial organism of any one of claims 1-4, wherein said microbial organism comprises two, three, four, five or six exogenous nucleic acids each encoding a primary alcohol, a fatty acyl-CoA, a fatty ester or a wax pathway enzyme.

8. The non-naturally occurring microbial organism of claim 1, wherein said microbial organism comprises:

(A) four exogenous nucleic acids encoding malonyl-CoA-independent FAS pathway enzymes comprising ketoacyl-CoA acyltransferase or ketoacyl-CoA thiolase, 3-hydroxyacyl-CoA dehydrogenase, enoyl-CoA hydratase and enoyl-CoA reductase; and two exogenous nucleic acids encoding acyl-reduction pathway enzymes comprising an acyl-CoA reductase and an alcohol dehydrogenase.

9. The non-naturally occurring microbial organism of claim 2, wherein said microbial organism comprises:

(A) four exogenous nucleic acids encoding malonyl-CoA-independent FAS pathway enzymes comprising ketoacyl-CoA acyltransferase or ketoacyl-CoA thiolase, 3-hydroxyacyl-CoA dehydrogenase, enoyl-CoA hydratase and enoyl-CoA reductase.

10. The non-naturally occurring microbial organism of claim 3, wherein said microbial organism comprises:

(A) four exogenous nucleic acids encoding malonyl-CoA-independent fatty acid synthesis (FAS) pathway enzymes and one exogenous nucleic acid encoding a wax ester synthase, wherein said malonyl-CoA-independent FAS pathway enzymes comprise a ketoacyl-CoA acyltransferase or a ketoacyl-CoA thiolase, a 3-hydroxyacyl-CoA dehydrogenase, an enoyl-CoA hydratase and an enoyl-CoA reductase; or

(B) four exogenous nucleic acids encoding malonyl-CoA-independent fatty acid synthesis (FAS) pathway enzymes and one exogenous nucleic acid encoding an alcohol acetyltransferase, wherein said malonyl-CoA-independent FAS pathway enzymes comprise a ketoacyl-CoA acyltransferase or a ketoacyl-CoA thiolase, a 3-hydroxyacyl-CoA dehydrogenase, an enoyl-CoA hydratase and an enoyl-CoA reductase.

11. The non-naturally occurring microbial organism of claim 4, wherein said microbial organism comprises:

(A) four exogenous nucleic acids encoding malonyl-CoA-independent fatty acid synthesis (FAS) pathway enzymes and one nucleic acid encoding a wax ester synthase, wherein said malonyl-CoA-independent FAS pathway enzymes comprise a ketoacyl-CoA acyltransferase or a ketoacyl-CoA thiolase, a 3-hydroxyacyl-CoA dehydrogenase, an enoyl-CoA hydratase and an enoyl-CoA reductase; or

(B) four exogenous nucleic acids encoding malonyl-CoA-independent fatty acid synthesis (FAS) pathway enzymes and one nucleic acid encoding an alcohol acetyltransferase, wherein said malonyl-CoA-independent FAS pathway enzymes comprise a ketoacyl-CoA

acyltransferase or a ketoacyl-CoA thiolase, a 3-hydroxyacyl-CoA dehydrogenase, an enoyl-CoA hydratase and an enoyl-CoA reductase.

12. The non-naturally occurring microbial organism of any one of claims 1-4, wherein said microbial organism comprising (i) comprises two, three or four exogenous nucleic acids each encoding a reductive TCA pathway enzyme.

13. The non-naturally occurring microbial organism of any one of claims 1-4, wherein said microbial organism comprising (ii) comprises two, three or four exogenous nucleic acids each encoding a reductive TCA pathway enzyme.

14. The non-naturally occurring microbial organism of claims 1-4, wherein said microbial organism comprising (i) comprises three exogenous nucleic acids encoding an ATP-citrate lyase or a citrate lyase; a fumarate reductase; and an alpha-ketoglutarate:ferredoxin oxidoreductase;

wherein said microbial organism comprising (ii) comprises four exogenous nucleic acids encoding a pyruvate:ferredoxin oxidoreductase; a phosphoenolpyruvate carboxylase or a phosphoenolpyruvate carboxykinase; a CO dehydrogenase; and an H₂ hydrogenase; or

wherein said microbial organism comprising (iii) comprises two exogenous nucleic acids encoding a CO dehydrogenase and an H₂ hydrogenase.

15. The non-naturally occurring microbial organism of any one of claims 1-4, wherein said at least one exogenous nucleic acid is a heterologous nucleic acid.

16. The non-naturally occurring microbial organism of any one of claims 1-4, wherein said non-naturally occurring microbial organism is in a substantially anaerobic culture medium.

17. A method for producing a primary alcohol, comprising culturing the non-naturally occurring microbial organism of any one of claims 1, 5-8 and 12-15 under conditions and for a sufficient period of time to produce a primary alcohol.

18. A method for producing a fatty acyl-CoA, comprising culturing the non-naturally occurring microbial organism of any one of claims 2, 5-7, 9 and 12-15 under conditions and for a sufficient period of time to produce a fatty acyl-CoA.

19. A method for producing a fatty ester, comprising culturing the non-naturally occurring microbial organism of any one of claims 3, 5-7, 10 and 12-15 under conditions and for a sufficient period of time to produce a fatty ester.

20. A method for producing a wax, comprising culturing the non-naturally occurring microbial organism of any one of claims 4-7 and 11-15 under conditions and for a sufficient period of time to produce a wax.

21. The method of any one of claims 17-20, wherein said non-naturally occurring microbial organism is in a substantially anaerobic culture medium.

22. The non-naturally occurring microbial organism of claim 1, wherein said exogenous nucleic acid encoding an acyl-reduction pathway enzyme comprises an enzyme having acyl-CoA reductase and alcohol dehydrogenase activity.

23. The non-naturally occurring microbial organism of claim 22, wherein said enzyme having acyl-CoA reductase and alcohol dehydrogenase activity comprises fatty alcohol forming acyl-CoA reductase (FAR).

24. The non-naturally occurring microbial organism of claim 1, wherein said primary alcohol is produced in amounts at least 10% greater levels compared to a microbial organism lacking said exogenous nucleic acid encoding a malonyl-CoA-independent FAS pathway enzyme.

25. The non-naturally occurring microbial organism of claim 1, wherein said primary alcohol comprises an alcohol having between 4-24 carbon atoms.

26. The non-naturally occurring microbial organism of claim 25, wherein said alcohol having between 4-24 carbon atoms is selected from butanol, hexanol, heptanol, octanol, nonanol, decanol, dodecanol, tetradecanol or hexadecanol.

27. A non-naturally occurring microbial organism, comprising a microbial organism having an acyl-ACP pathway comprising at least one exogenous nucleic acid encoding a acyl-ACP pathway enzyme expressed in a sufficient amount to produce acyl-ACP; said non-naturally occurring microbial organism further comprising:

(i) a reductive TCA pathway comprising at least one exogenous nucleic acid encoding a reductive TCA pathway enzyme, wherein said at least one exogenous nucleic acid is selected from an ATP-citrate lyase, a citrate lyase, a citryl-CoA synthetase, a citryl-CoA lyase, a fumarate reductase, and an alpha-ketoglutarate:ferredoxin oxidoreductase;

(ii) a reductive TCA pathway comprising at least one exogenous nucleic acid encoding a reductive TCA pathway enzyme, wherein said at least one exogenous nucleic acid is selected from a pyruvate:ferredoxin oxidoreductase, a phosphoenolpyruvate carboxylase, a phosphoenolpyruvate carboxykinase, a CO dehydrogenase, and an H₂ hydrogenase; or

(iii) at least one exogenous nucleic acid encodes an enzyme selected from a CO dehydrogenase, an H₂ hydrogenase, and combinations thereof;

wherein said acyl-ACP pathway comprises a pathway selected from:

(A) acetyl-CoA carboxylase and a fatty acid synthase; and

(B) a fatty acid synthase.

28. The non-naturally occurring microbial organism of claim 27, further comprising a fatty acid pathway comprising an exogenous nucleic acid encoding a thioesterase expressed in a sufficient amount to produce fatty acid.

29. The non-naturally occurring microbial organism of claim 28, further comprising a acyl-CoA pathway comprising an exogenous nucleic acid encoding an acyl-CoA synthetase or a acyl-CoA ligase expressed in a sufficient amount to produce acyl-CoA.

30. The non-naturally occurring microbial organism of claim 29, further comprising a fatty aldehyde pathway comprising an exogenous nucleic acid encoding an acyl-CoA reductase

(aldehyde forming) or carboxylic acid reductase expressed in a sufficient amount to produce a fatty aldehyde.

31. The non-naturally occurring microbial organism of claim 30, further comprising a fatty alcohol pathway comprising an exogenous nucleic acid encoding an aldehyde reductase expressed in a sufficient amount to produce a fatty alcohol.

32. The non-naturally occurring microbial organism of claim 29, further comprising a fatty alcohol pathway comprising an exogenous nucleic acid encoding an acyl-CoA reductase (alcohol forming) expressed in a sufficient amount to produce a fatty alcohol.

33. The non-naturally occurring microbial organism of any one of claims 27-32, wherein said microbial organism comprising (i) further comprises an exogenous nucleic acid encoding an enzyme selected from a pyruvate:ferredoxin oxidoreductase, an aconitase, an isocitrate dehydrogenase, a succinyl-CoA synthetase, a succinyl-CoA transferase, a fumarase, a malate dehydrogenase, an acetate kinase, a phosphotransacetylase, an acetyl-CoA synthetase, an NAD(P)H:ferredoxin oxidoreductase, ferredoxin, and combinations thereof.

34. The non-naturally occurring microbial organism of any one of claims 27-32, wherein said microbial organism comprising (ii) further comprises an exogenous nucleic acid encoding an enzyme selected from an aconitase, an isocitrate dehydrogenase, a succinyl-CoA synthetase, a succinyl-CoA transferase, a fumarase, a malate dehydrogenase, and combinations thereof.

35. The non-naturally occurring microbial organism of any one of claims 27-32, wherein said microbial organism comprises two, three, four, five, or six exogenous nucleic acids each encoding a acyl-ACP, fatty acid, acyl-CoA, fatty aldehyde or fatty alcohol pathway enzyme.

36. The non-naturally occurring microbial organism of any one of claims 27-32, wherein said microbial organism comprising (i) comprises two, three or four exogenous nucleic acids each encoding a reductive TCA pathway enzyme.

37. The non-naturally occurring microbial organism of any one of claims 27-32, wherein said microbial organism comprising (ii) comprises two, three or four exogenous nucleic acids each encoding a reductive TCA pathway enzyme.

38. The non-naturally occurring microbial organism of any one of claims 27-32, wherein said at least one exogenous nucleic acid is a heterologous nucleic acid.

39. The non-naturally occurring microbial organism of claims 27-32, wherein said microbial organism comprising (i) comprises three exogenous nucleic acids encoding an ATP-citrate lyase or a citrate lyase; a fumarate reductase; and an alpha-ketoglutarate:ferredoxin oxidoreductase;

wherein said microbial organism comprising (ii) comprises four exogenous nucleic acids encoding a pyruvate:ferredoxin oxidoreductase; a phosphoenolpyruvate carboxylase or a phosphoenolpyruvate carboxykinase; a CO dehydrogenase; and an H₂ hydrogenase; or

wherein said microbial organism comprising (iii) comprises two exogenous nucleic acids encoding a CO dehydrogenase and an H₂ hydrogenase.

40. The non-naturally occurring microbial organism of any one of claims 27-32, wherein said non-naturally occurring microbial organism is in a substantially anaerobic culture medium.

41. A method for producing acyl-ACP, comprising culturing the non-naturally occurring microbial organism of claim 27 under conditions and for a sufficient period of time to produce acyl-ACP.

42. A method for producing fatty acid, comprising culturing the non-naturally occurring microbial organism of claim 28 under conditions and for a sufficient period of time to produce fatty acid.

43. A method for producing acyl-CoA, comprising culturing the non-naturally occurring microbial organism of claim 29 under conditions and for a sufficient period of time to produce acyl-CoA.

44. A method for producing fatty aldehyde, comprising culturing the non-naturally occurring microbial organism of claim 31 under conditions and for a sufficient period of time to produce fatty aldehyde.

45. A method for producing fatty alcohol, comprising culturing the non-naturally occurring microbial organism of claims 31 or 32 under conditions and for a sufficient period of time to produce fatty alcohol.

46. The method of any one of claims 41-45, wherein said non-naturally occurring microbial organism is in a substantially anaerobic culture medium.

47. The non-naturally occurring microbial organism of claim 1, where said microbial organism further comprises an acyl-reduction pathway comprising an acyl-CoA hydrolase, an acyl-CoA transferase or an acyl-CoA ligase; a carboxylic acid reductase and an alcohol dehydrogenase.

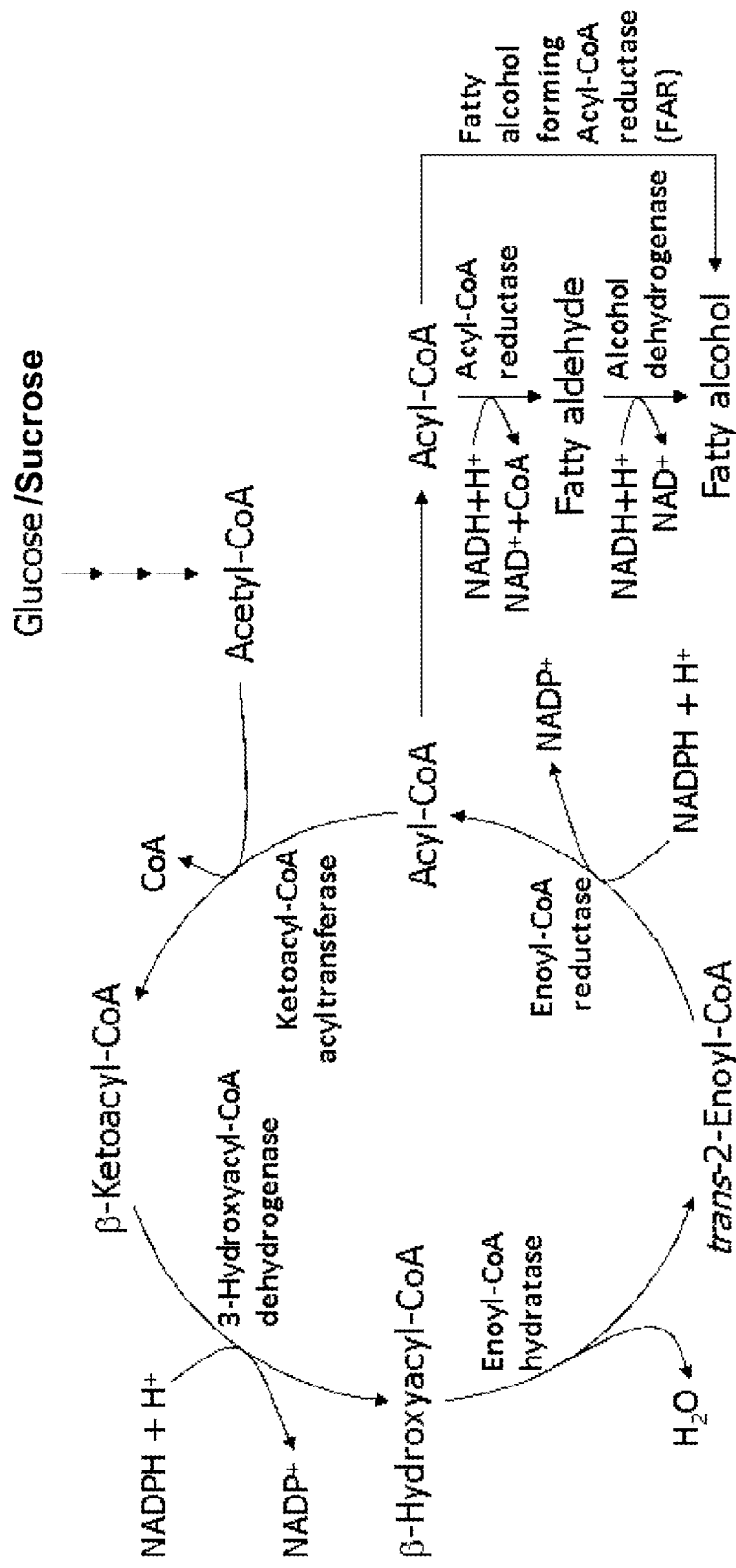


Figure 1

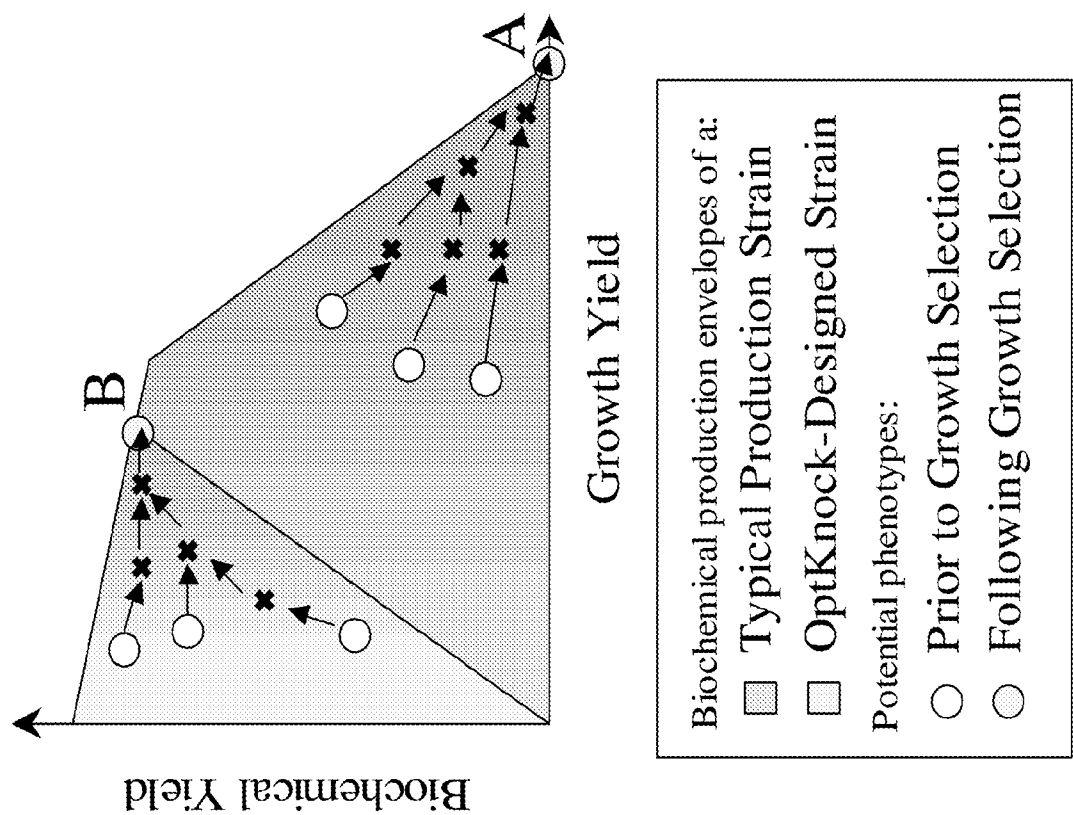


Figure 2

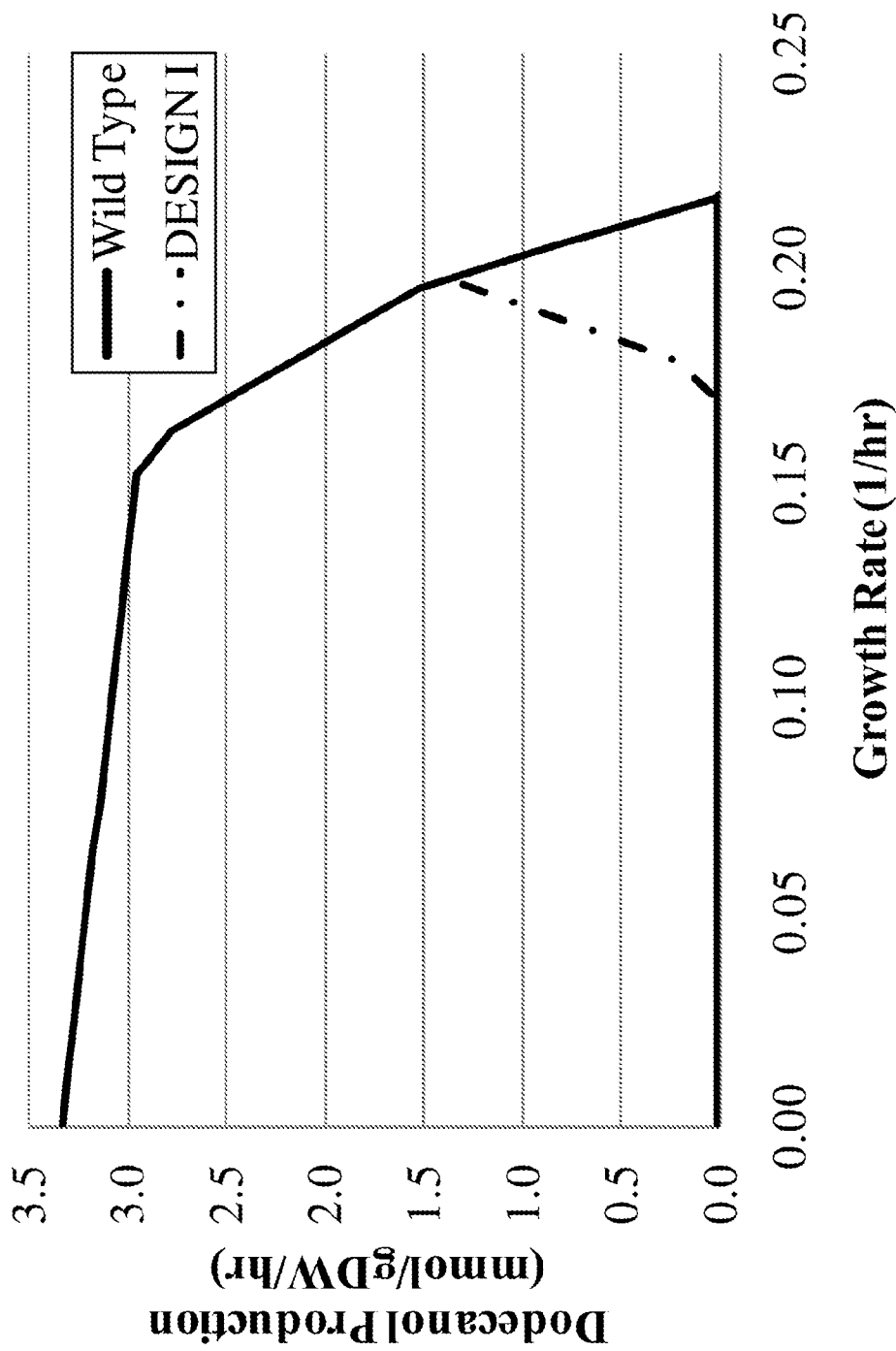


Figure 3

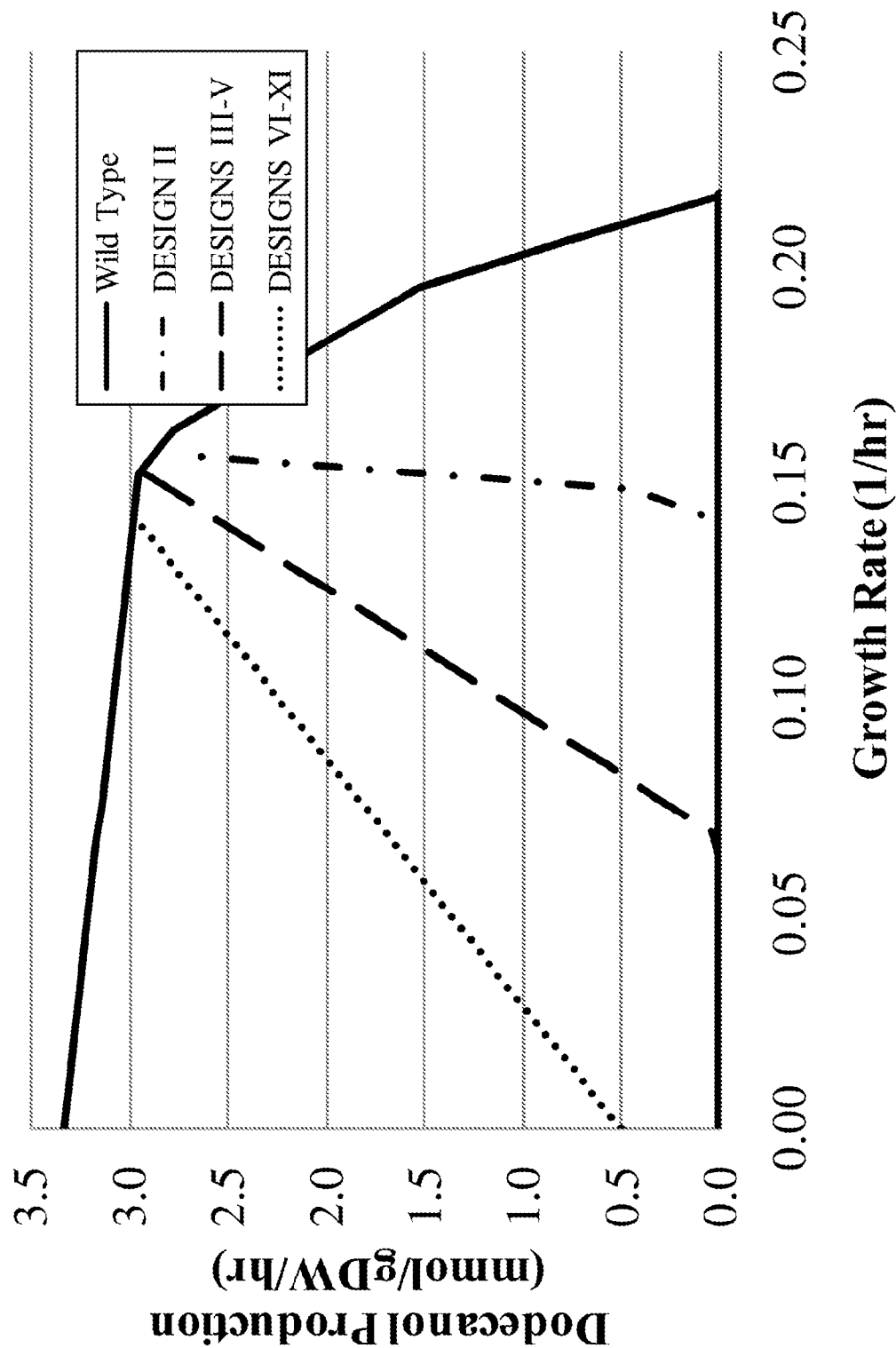


Figure 4

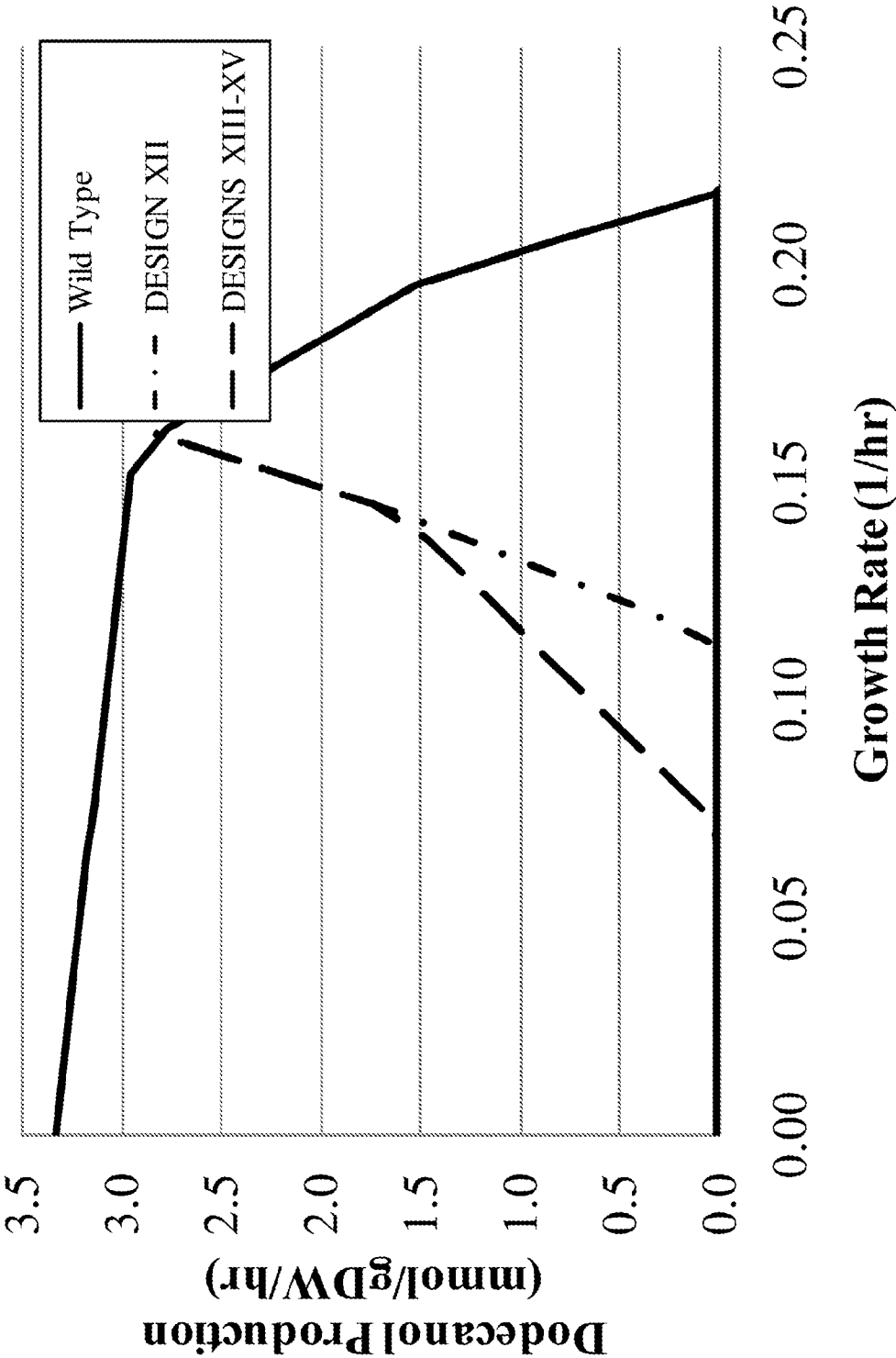


Figure 5

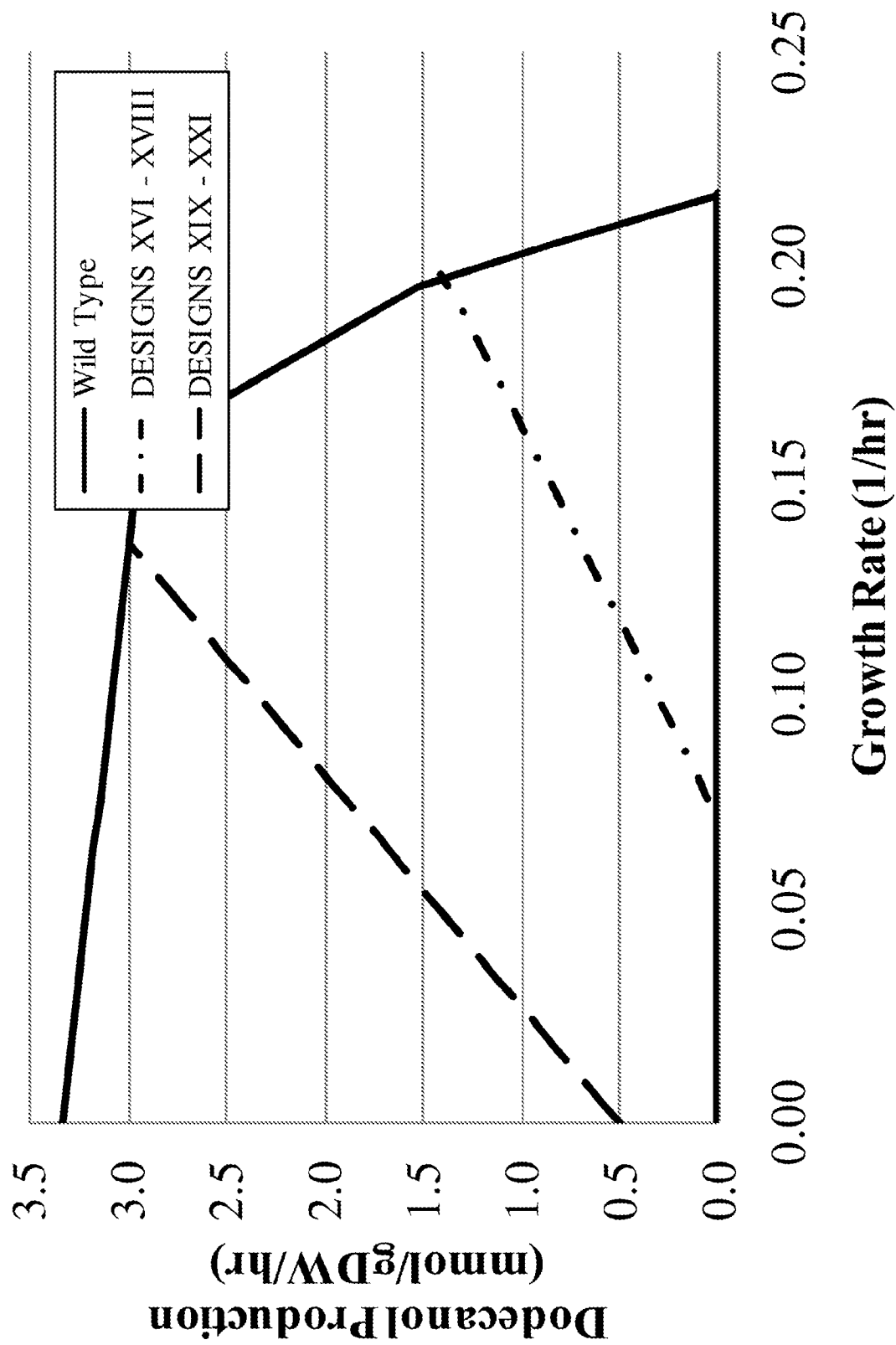


Figure 6

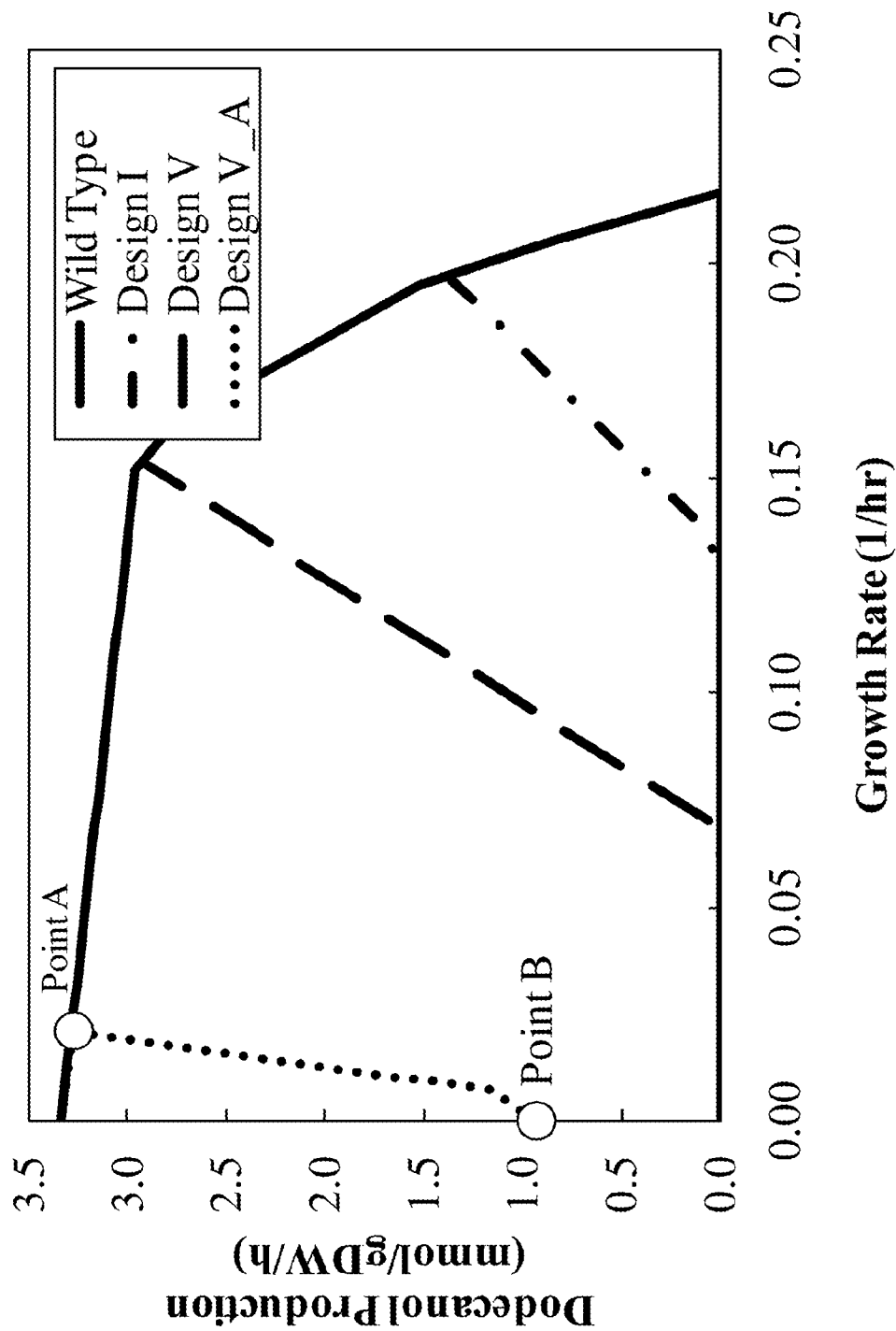


Figure 7

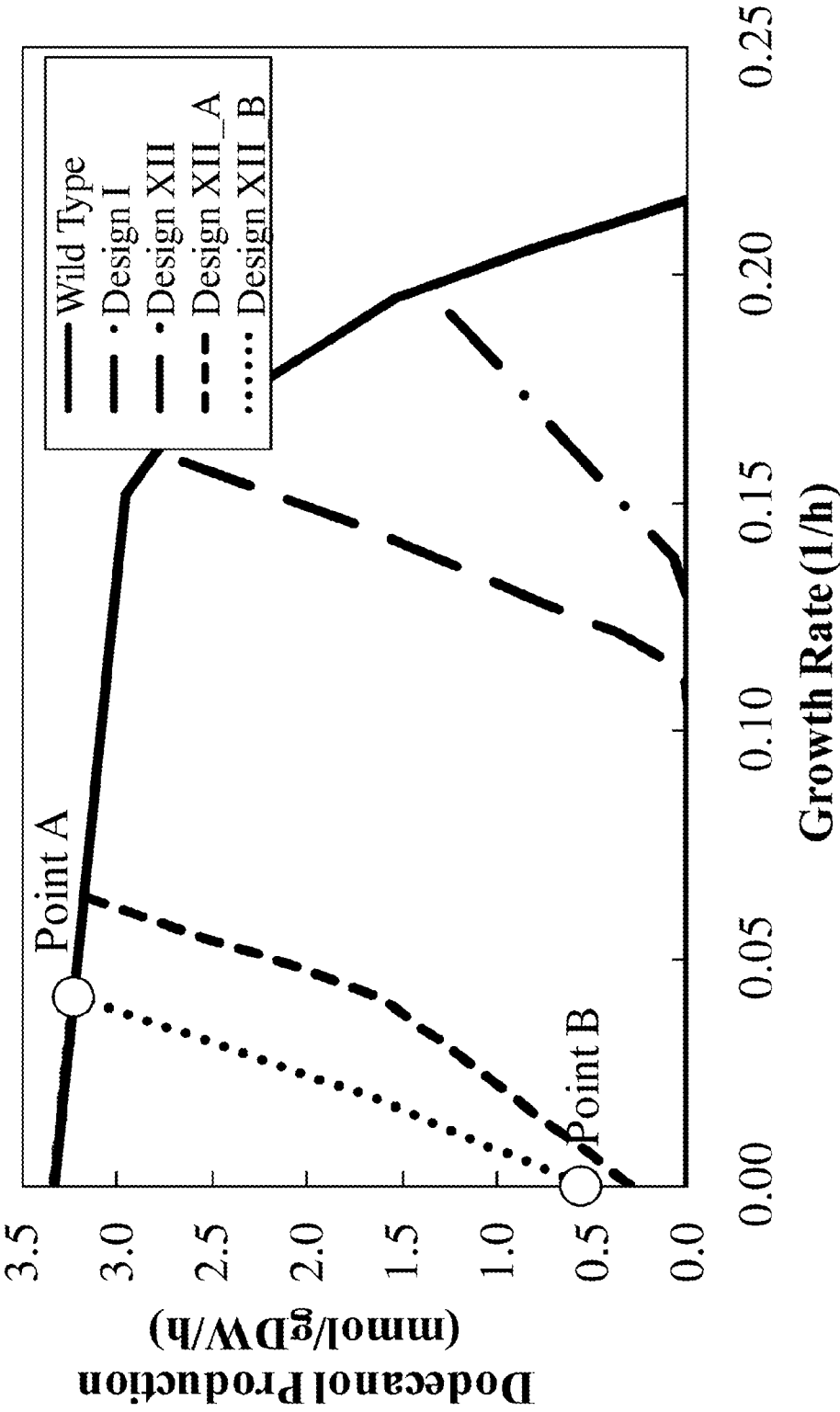


Figure 8

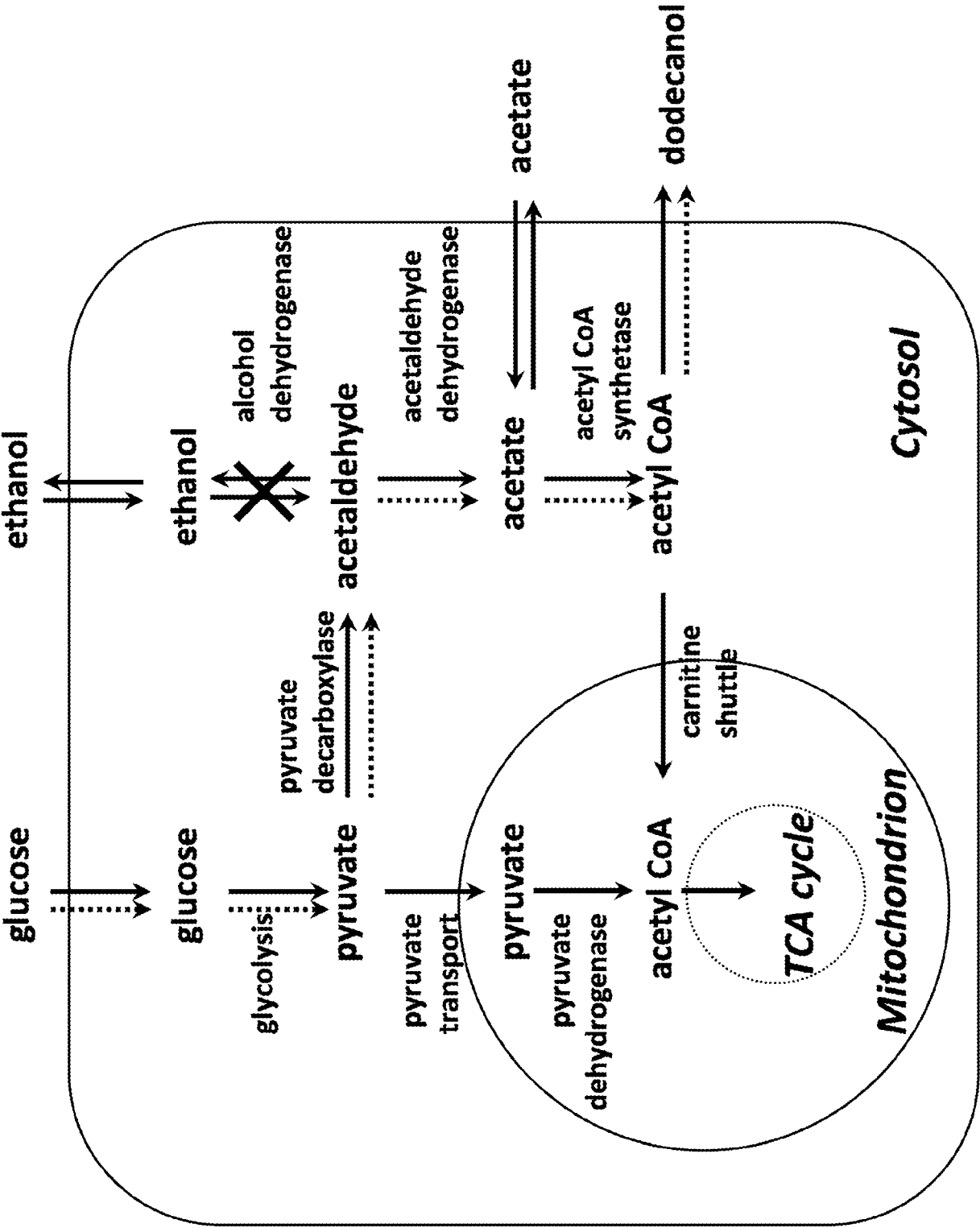


Figure 9a

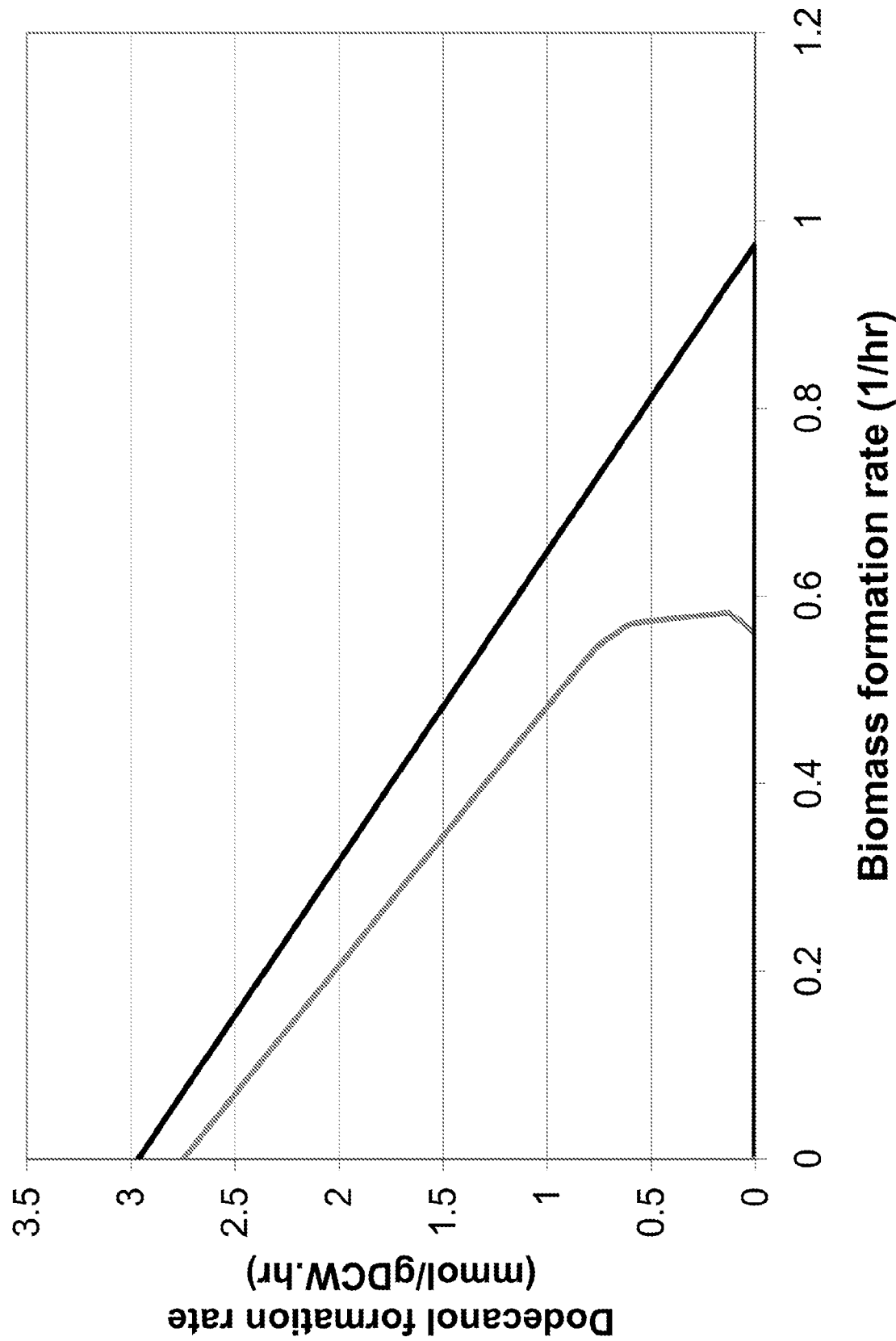


Figure 9b

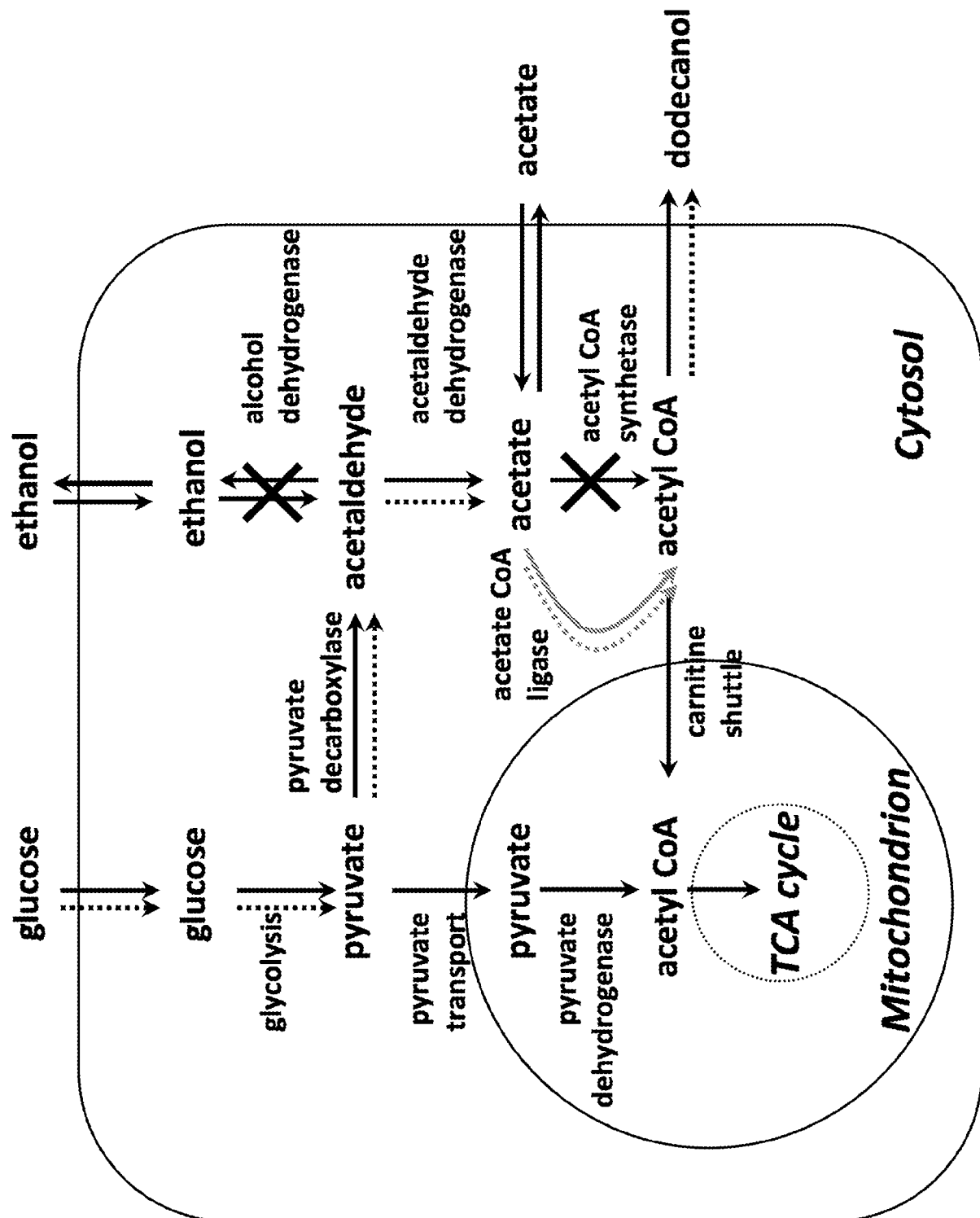


Figure 10a

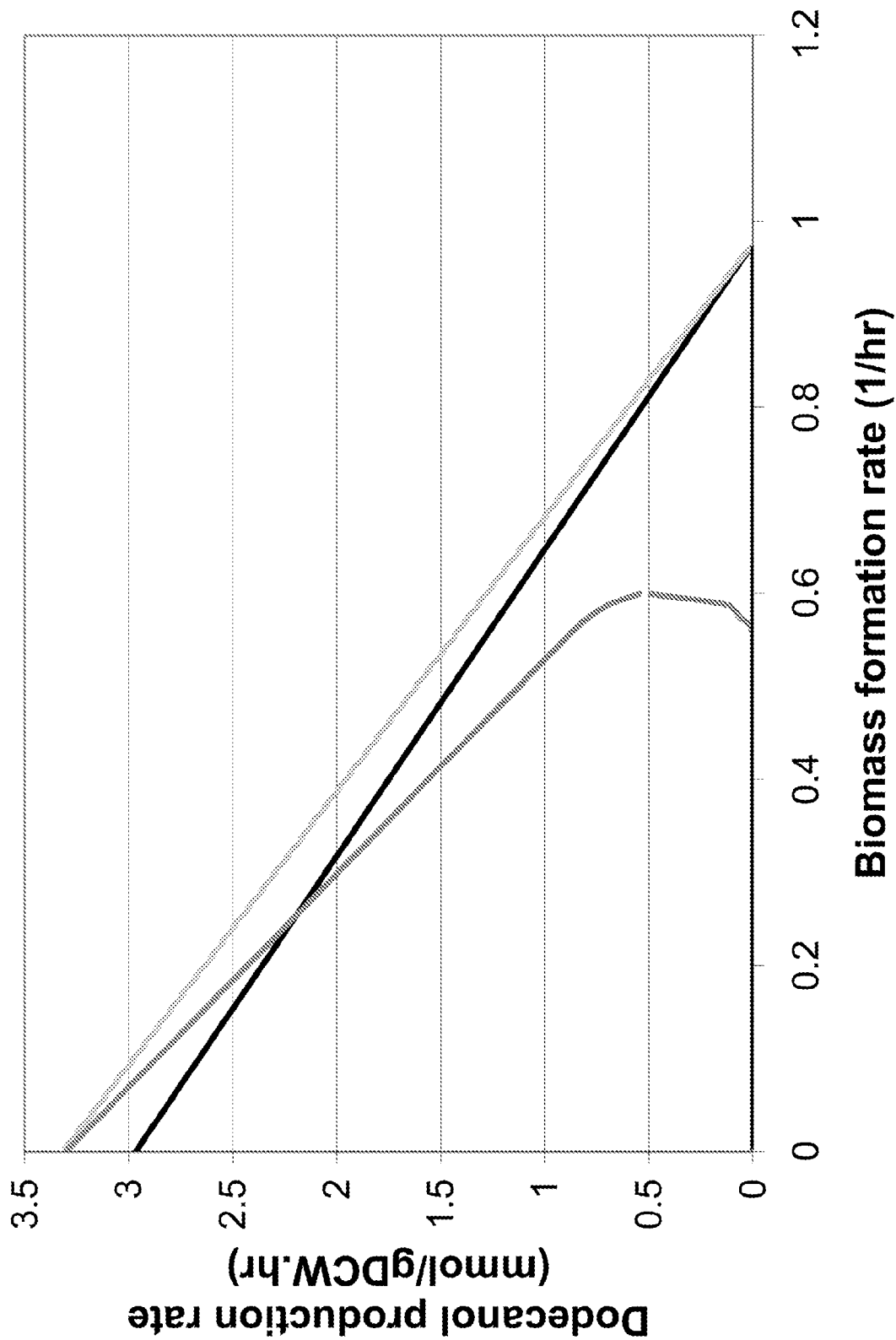


Figure 10b

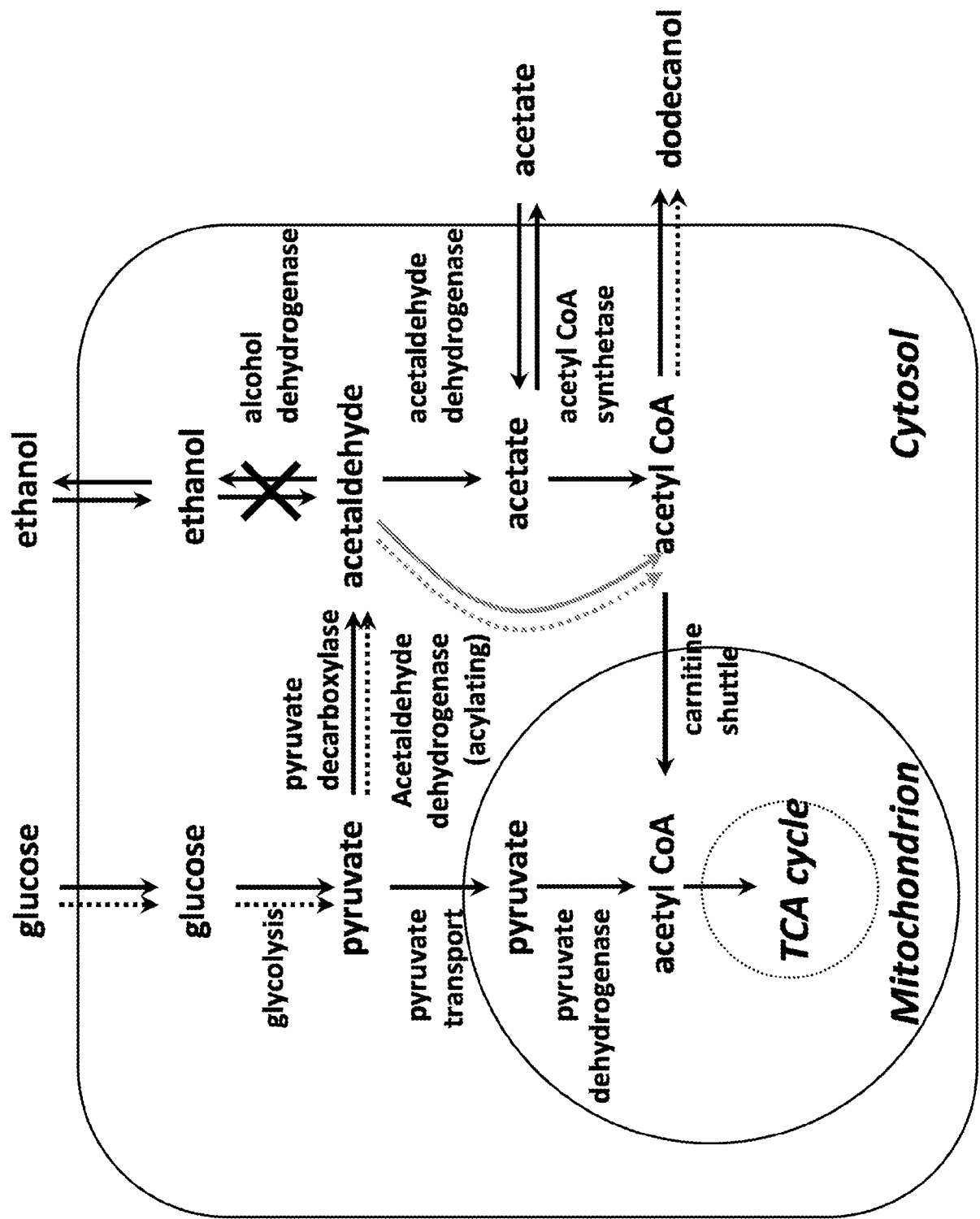


Figure 11a

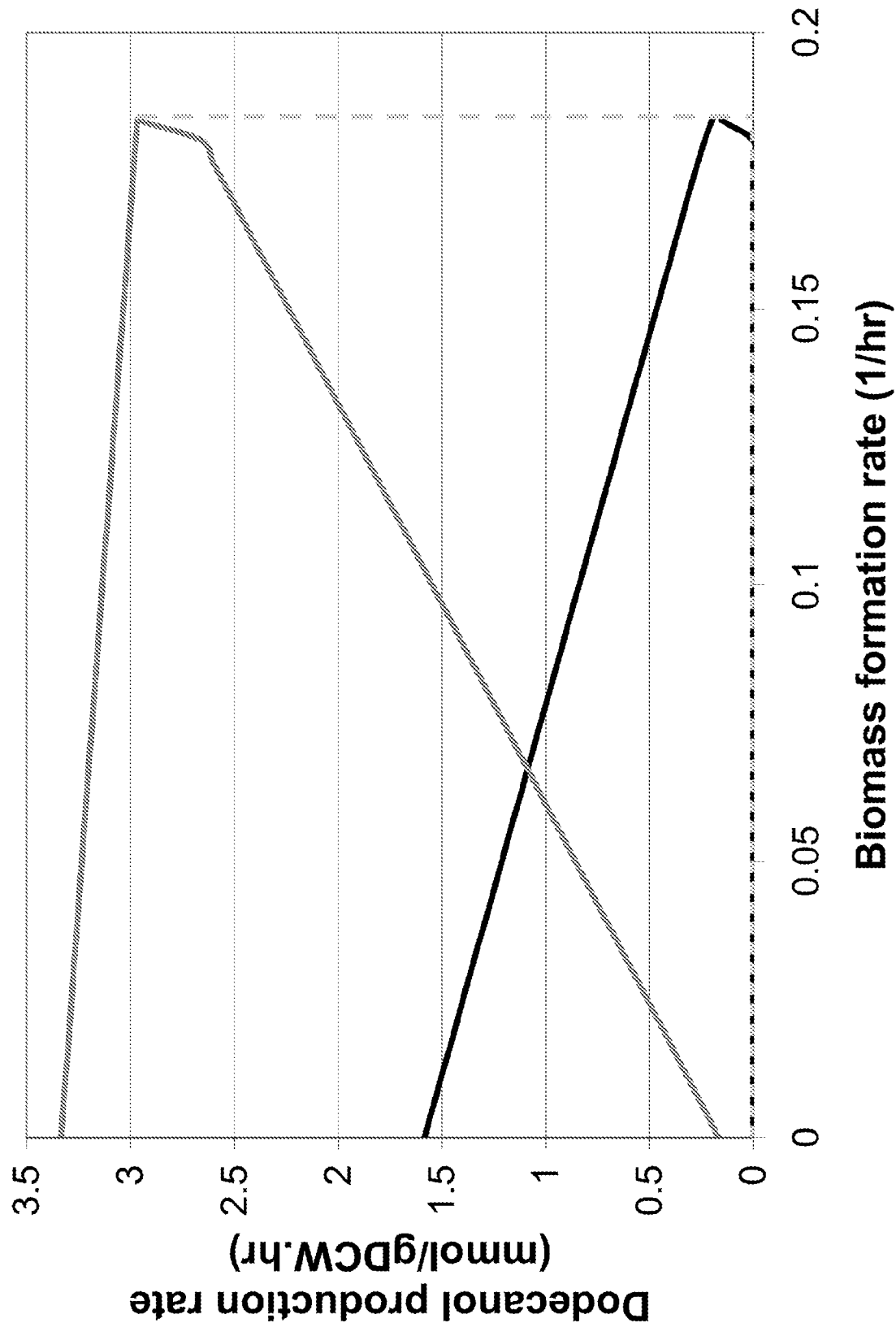


Figure 11b

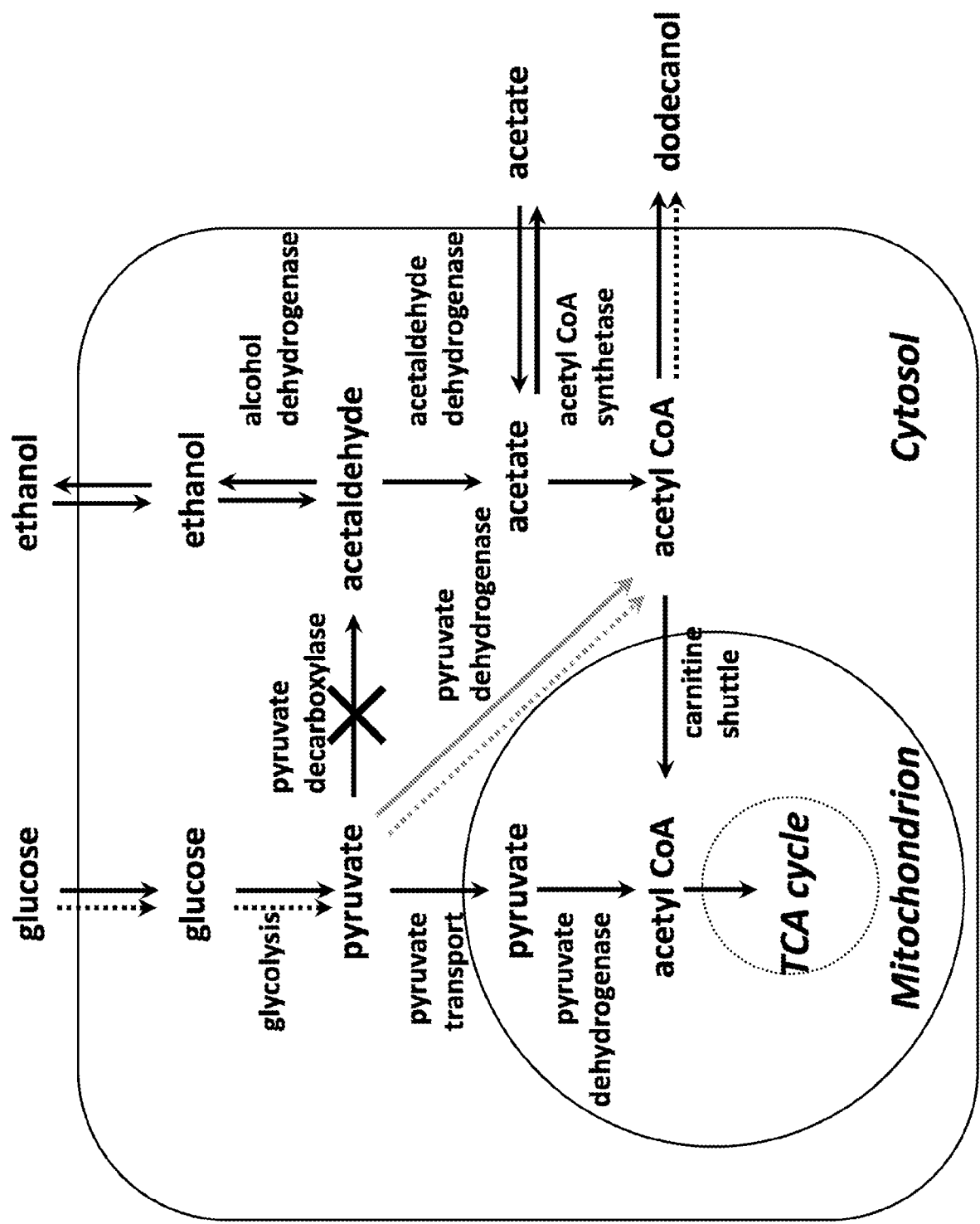


Figure 12

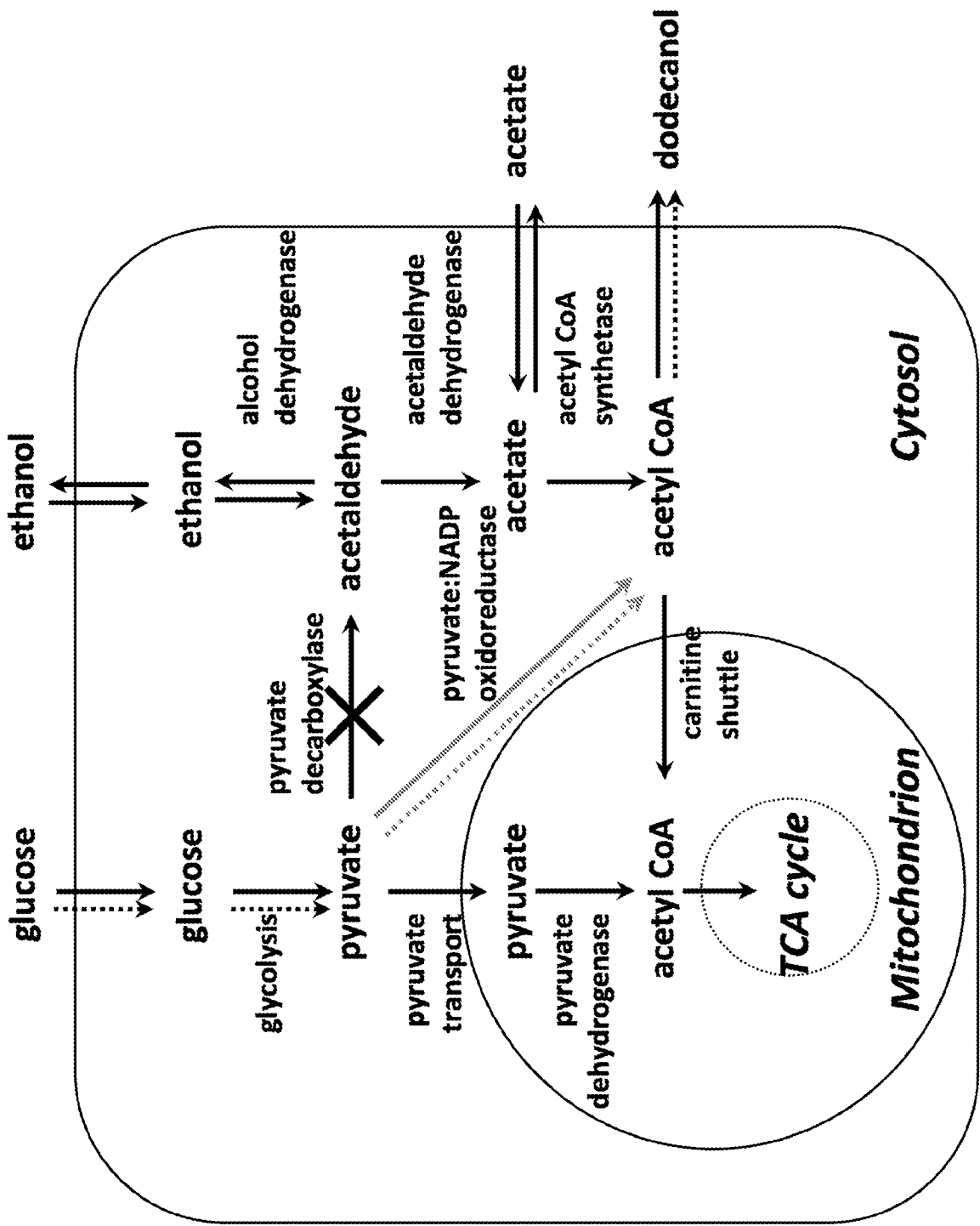


Figure 13

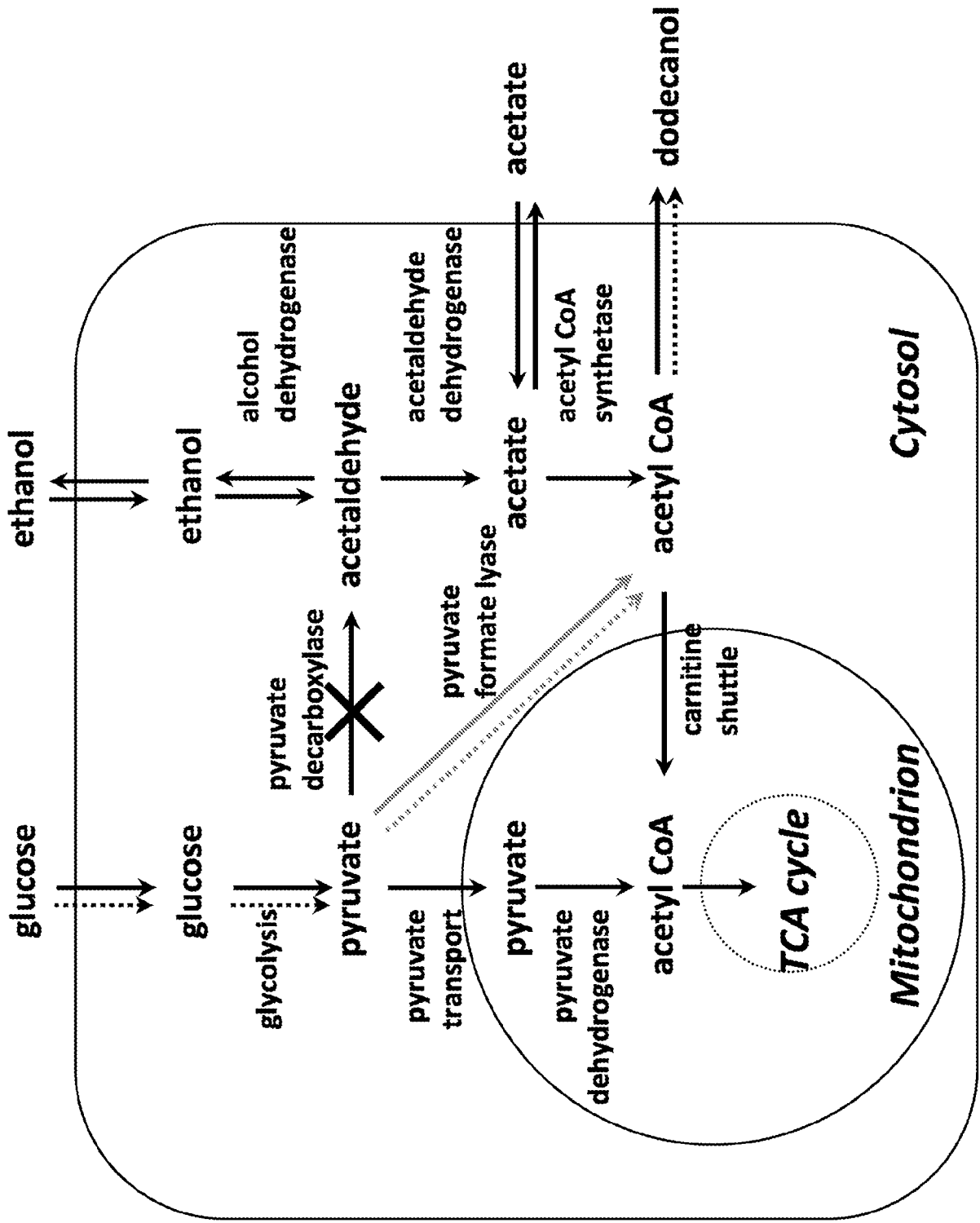


Figure 14

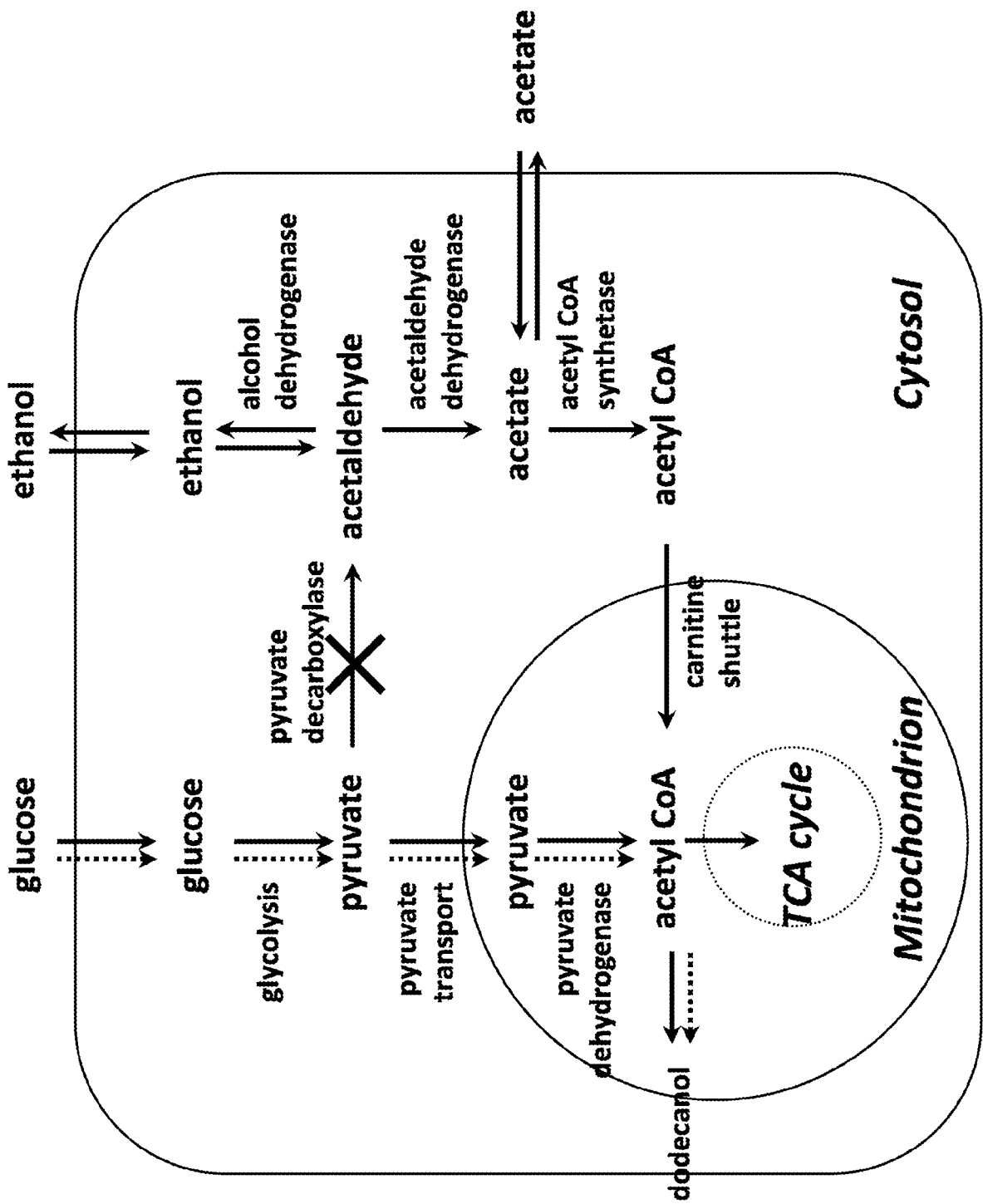


Figure 15a

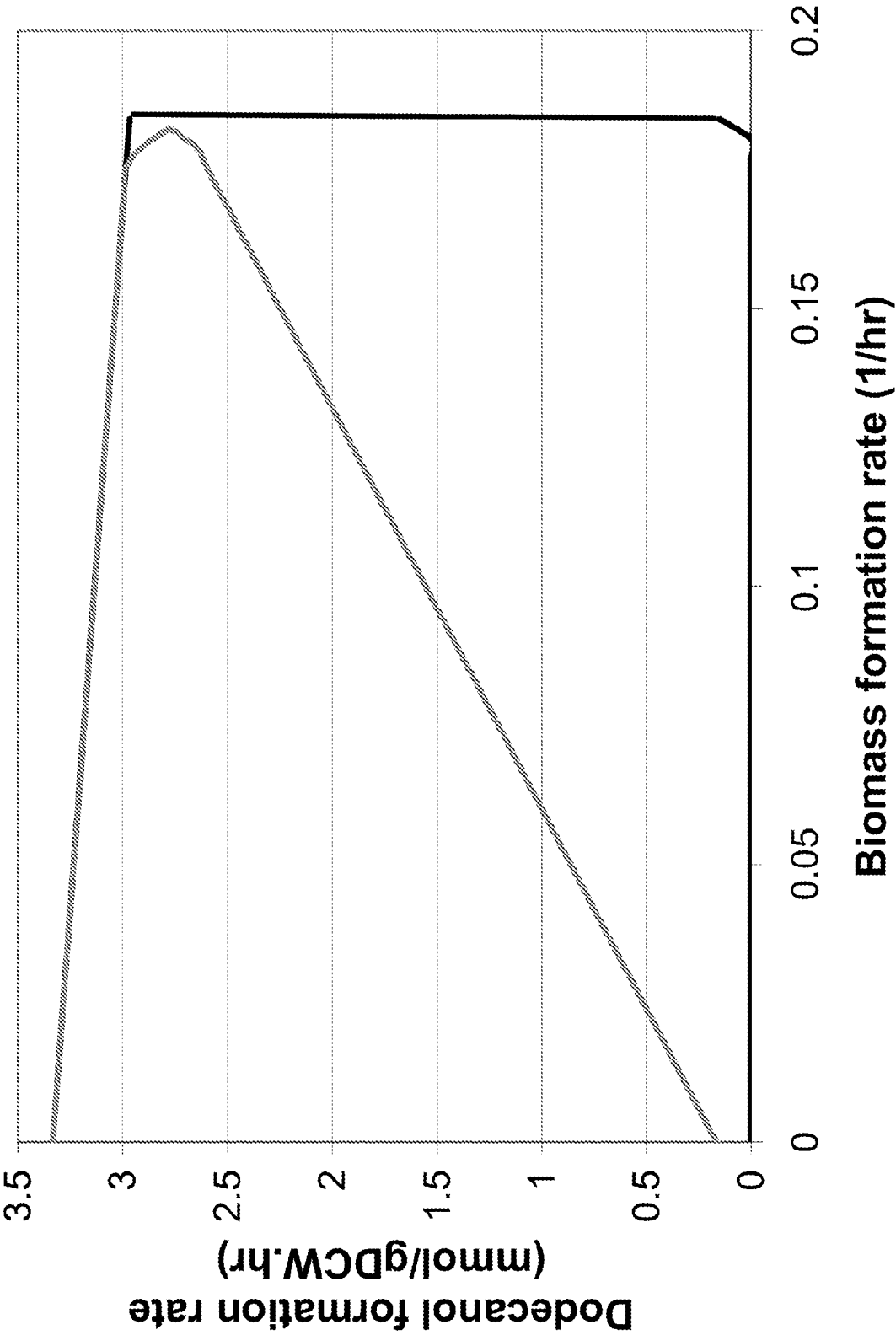


Figure15b

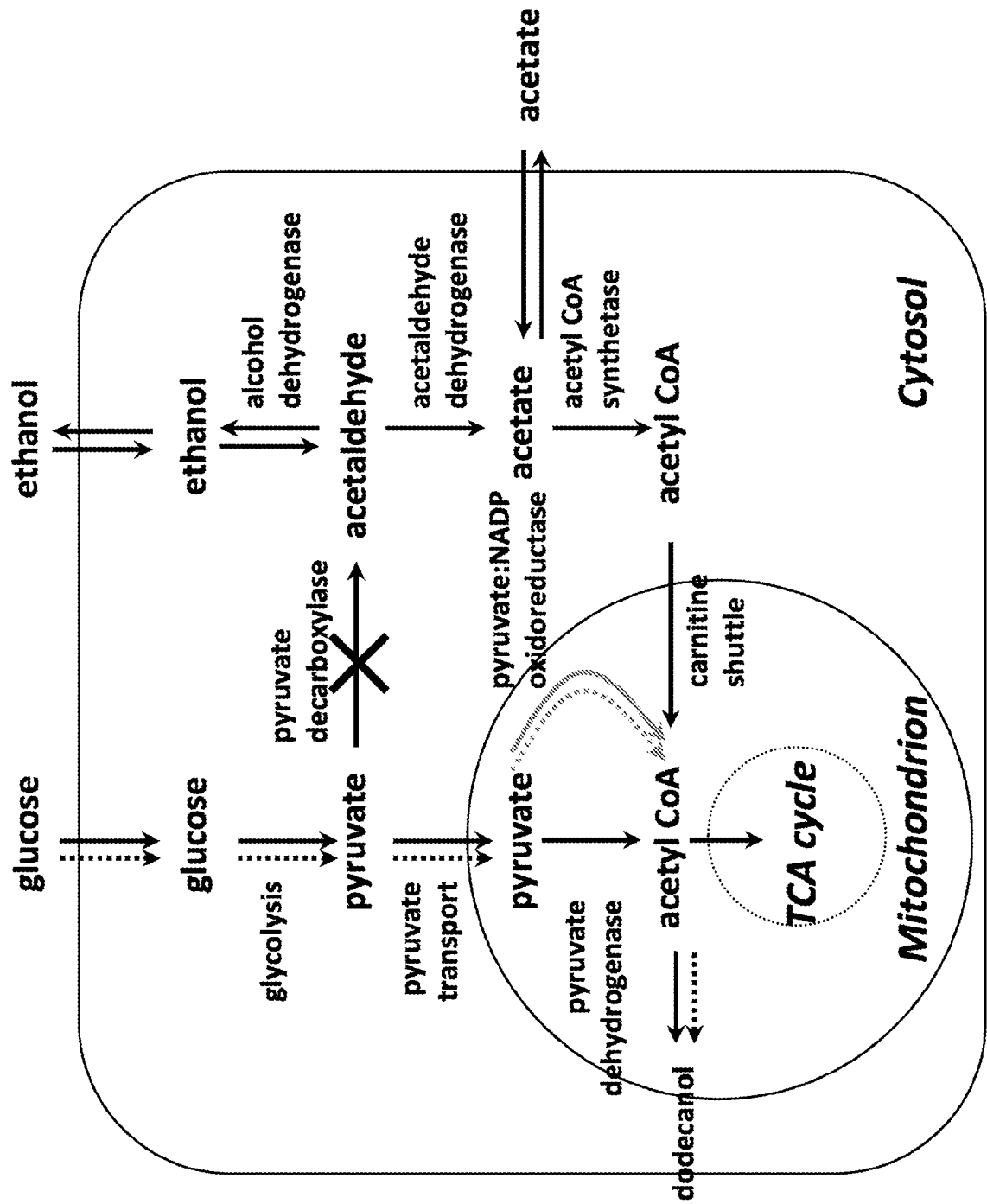


Figure 16

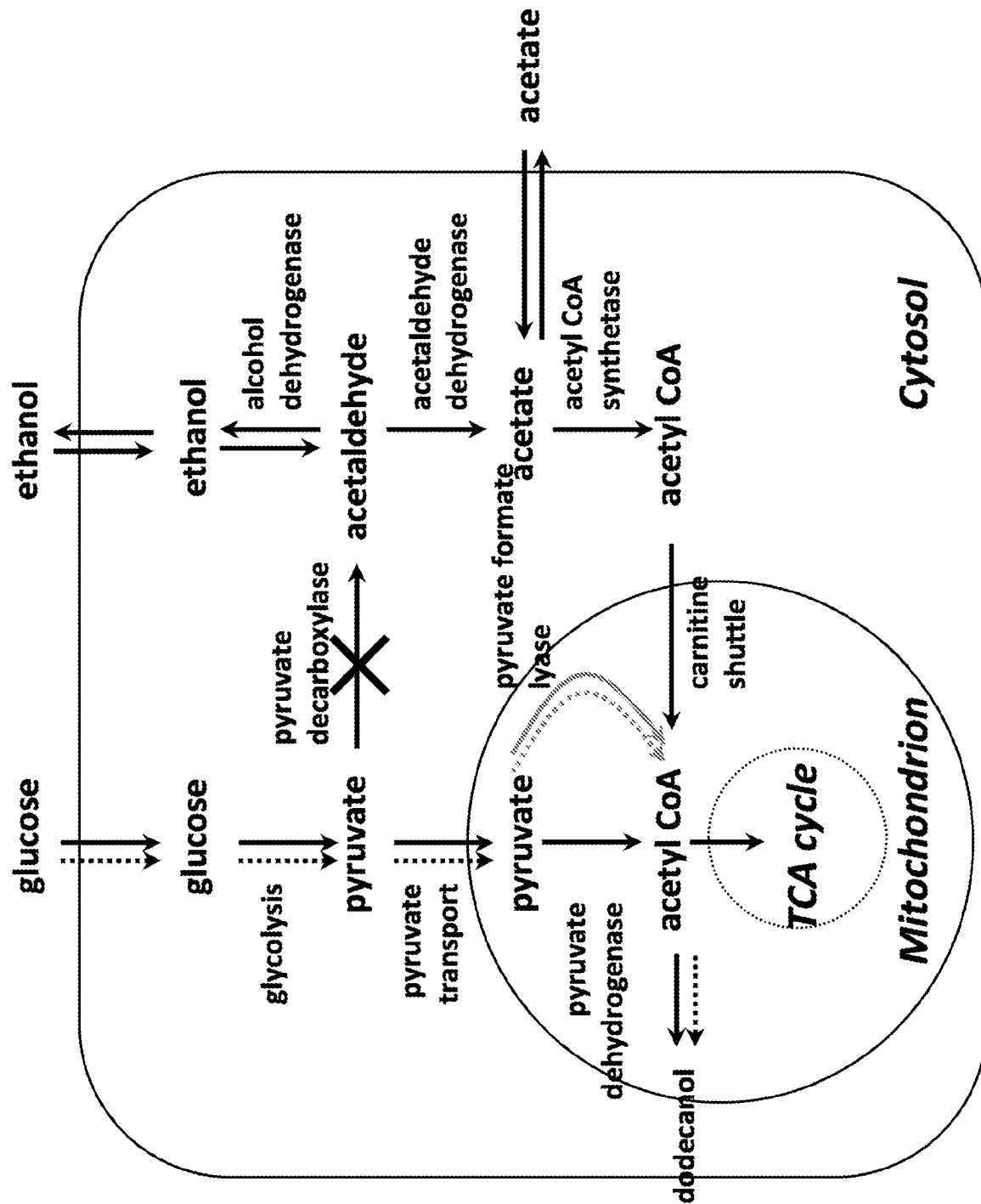


Figure 17

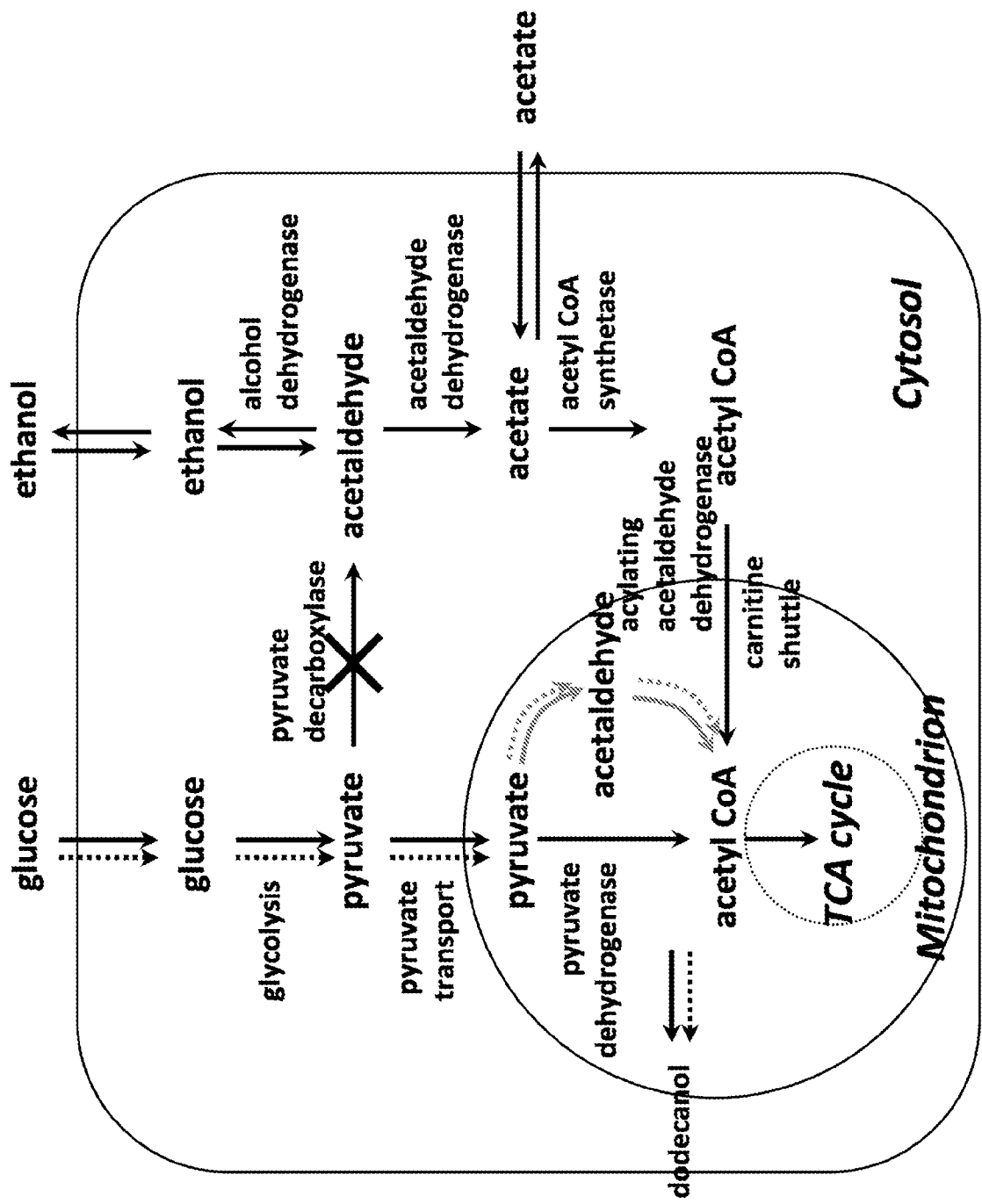


Figure 18

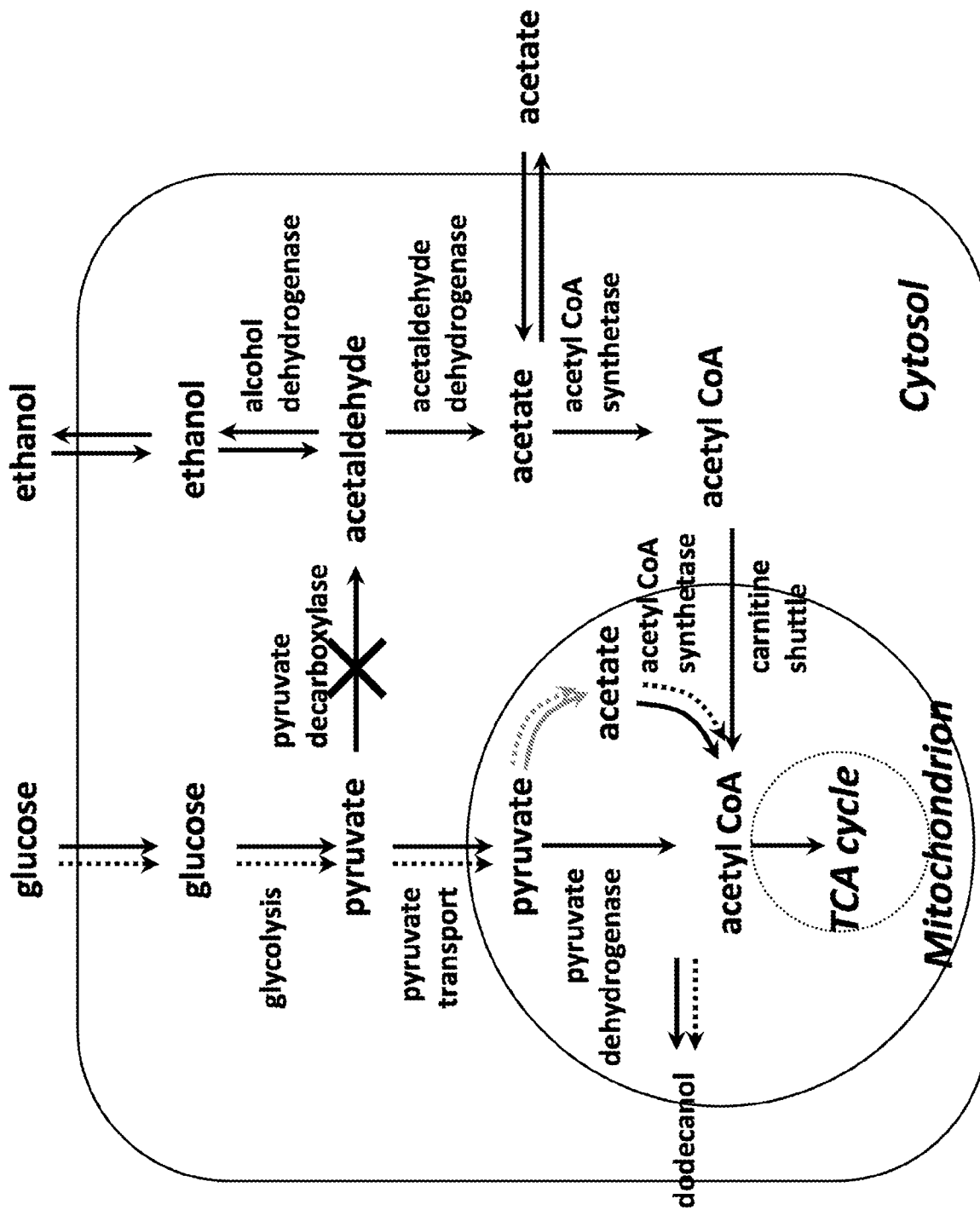


Figure 19a

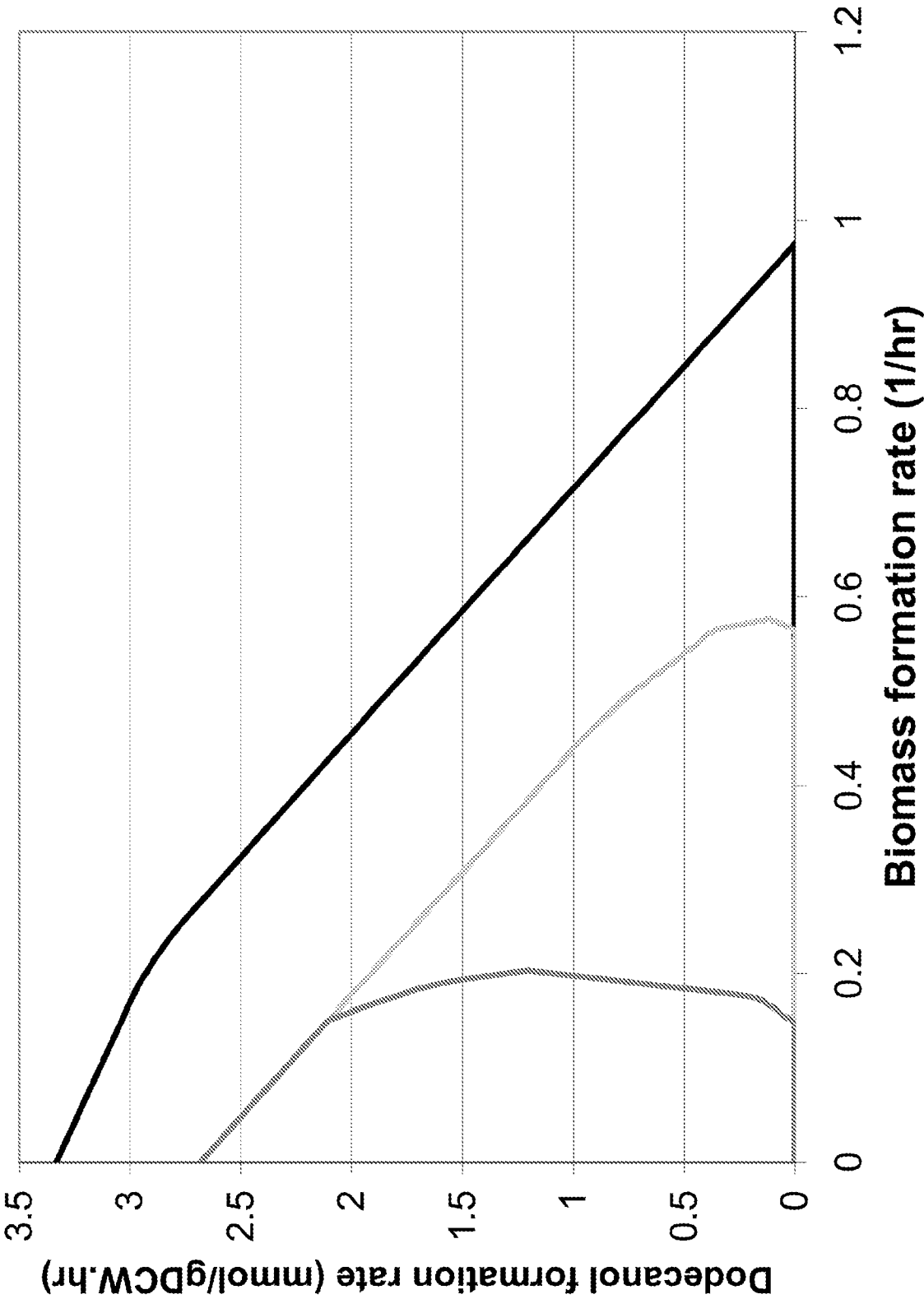


Figure 19b

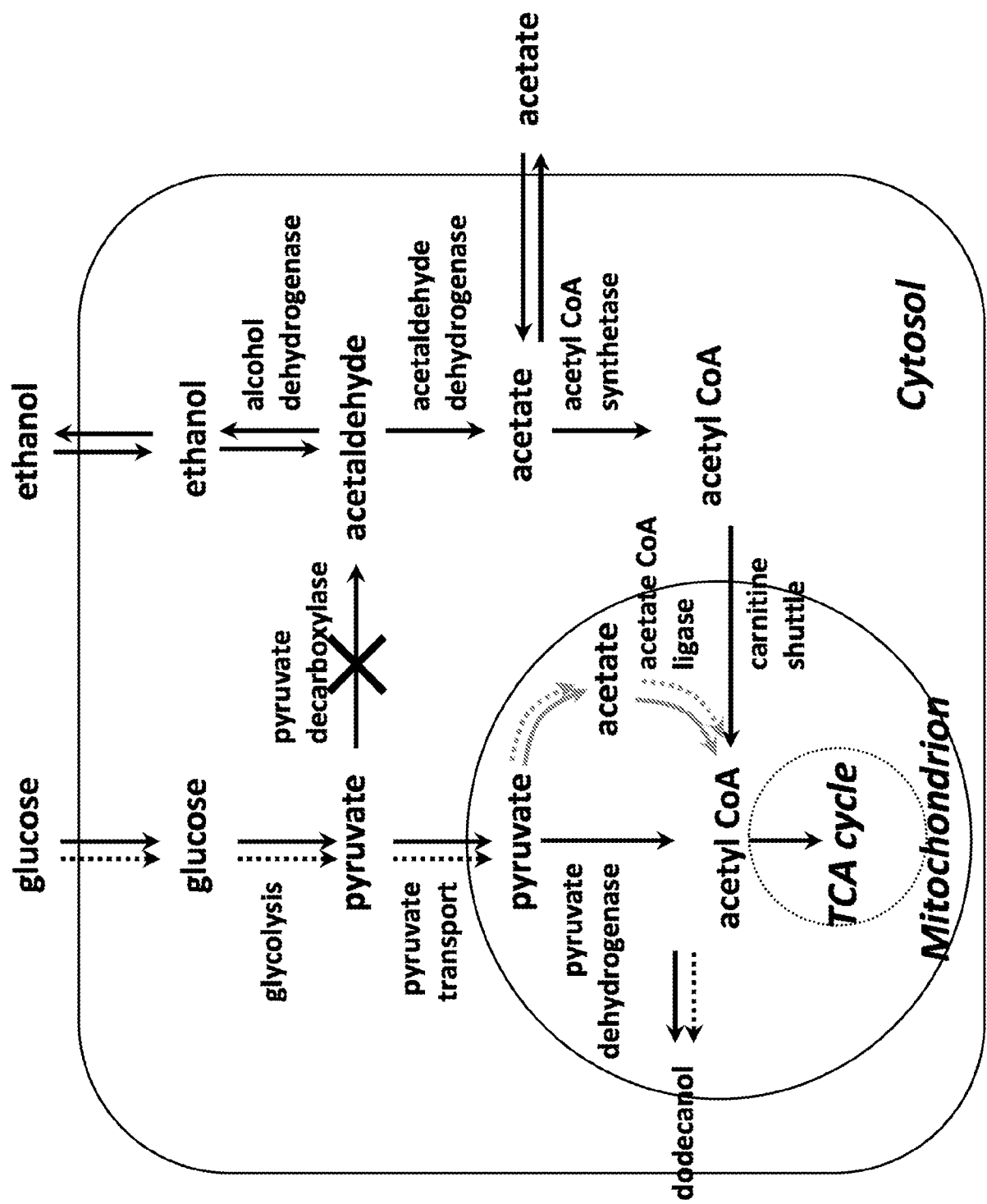


Figure 20

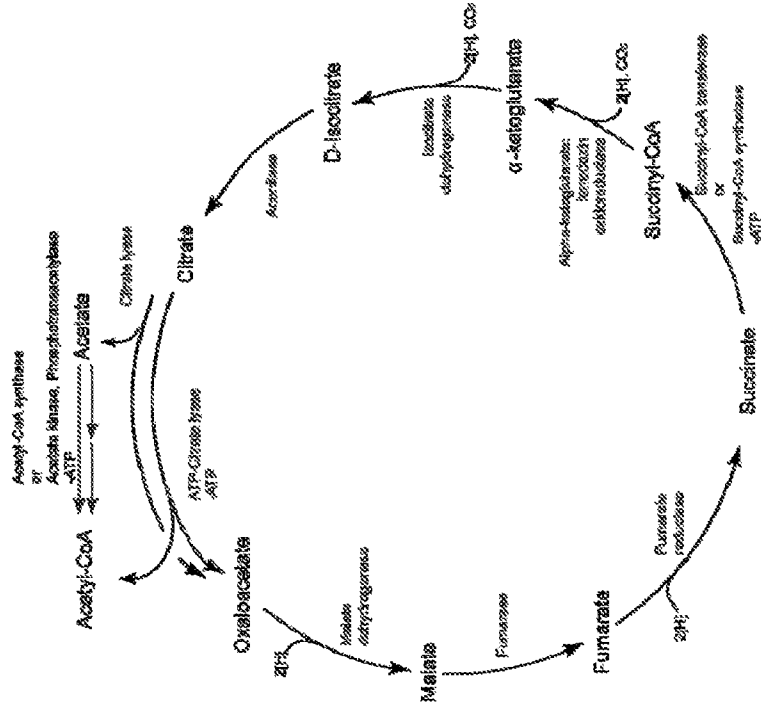


Figure 21

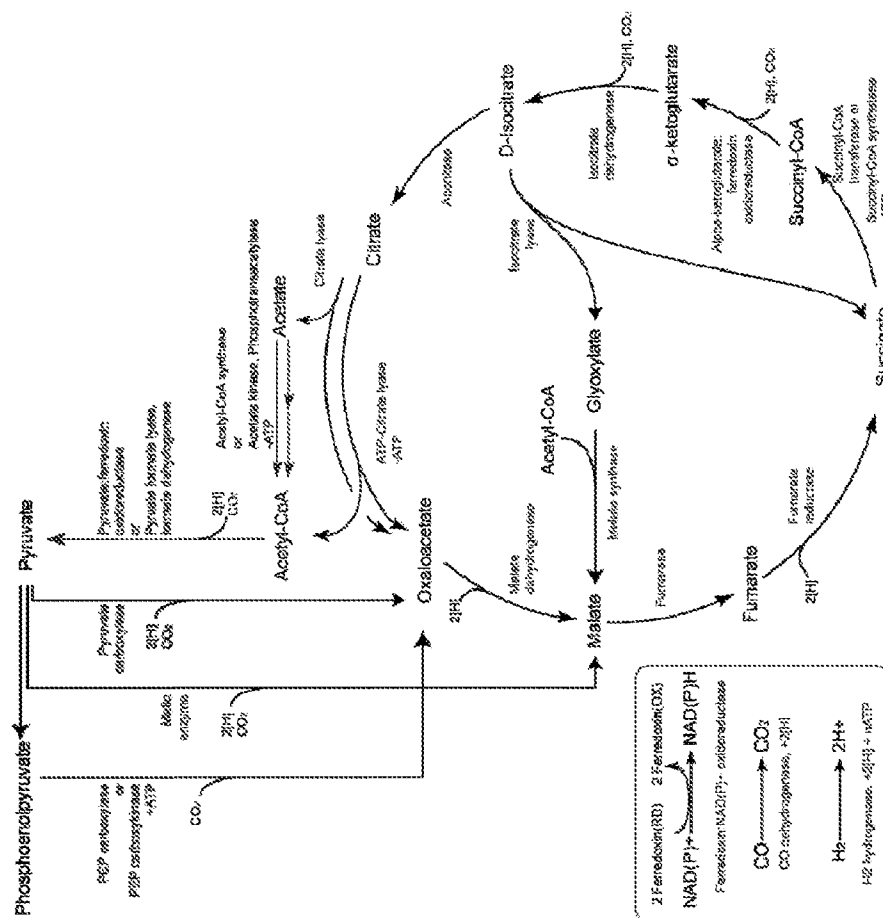


Figure 22

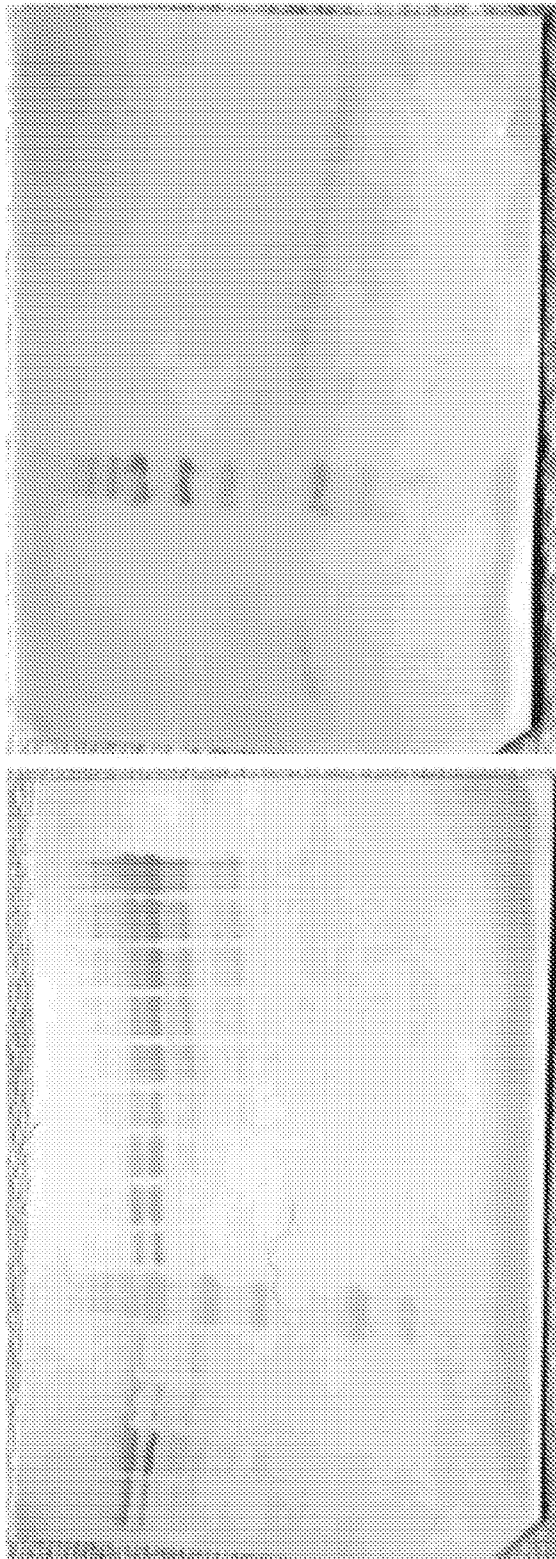


Figure 23

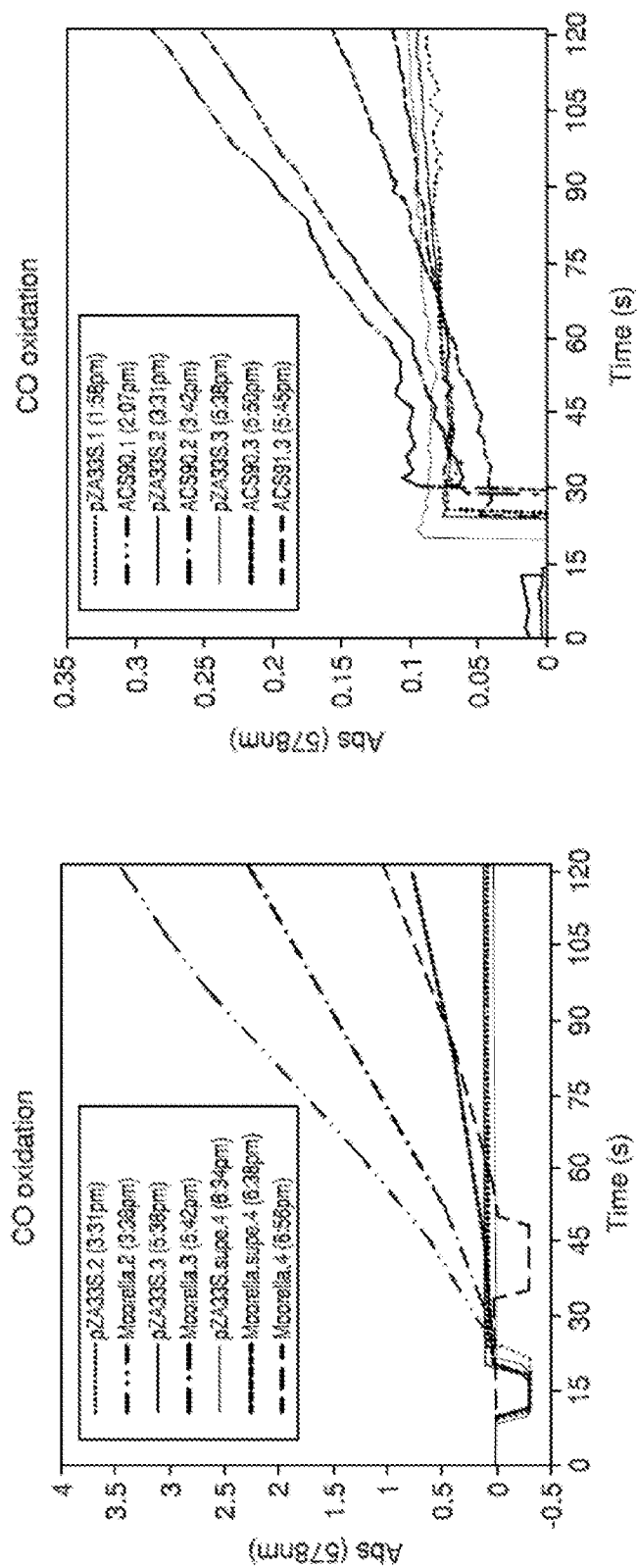


Figure 24

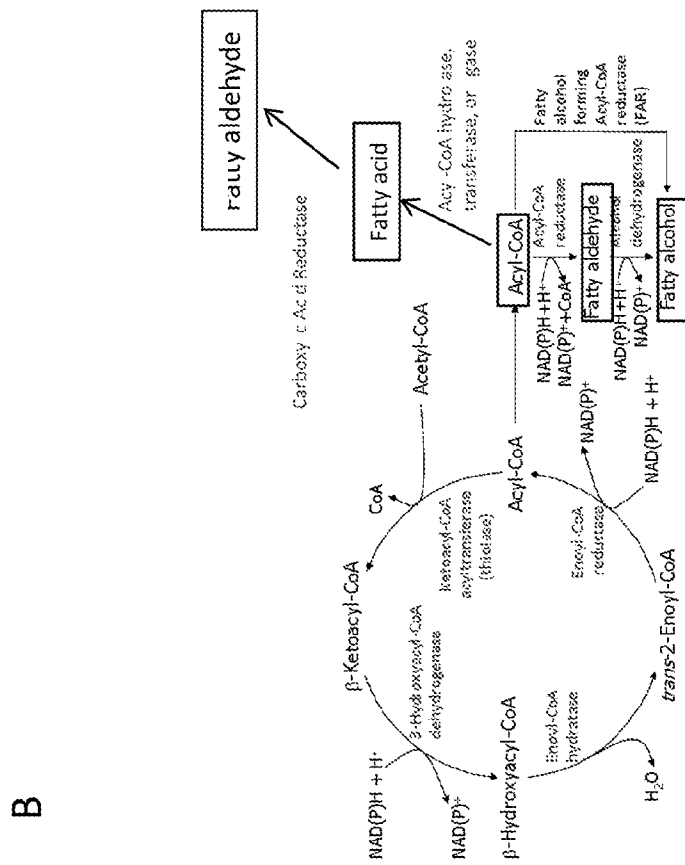
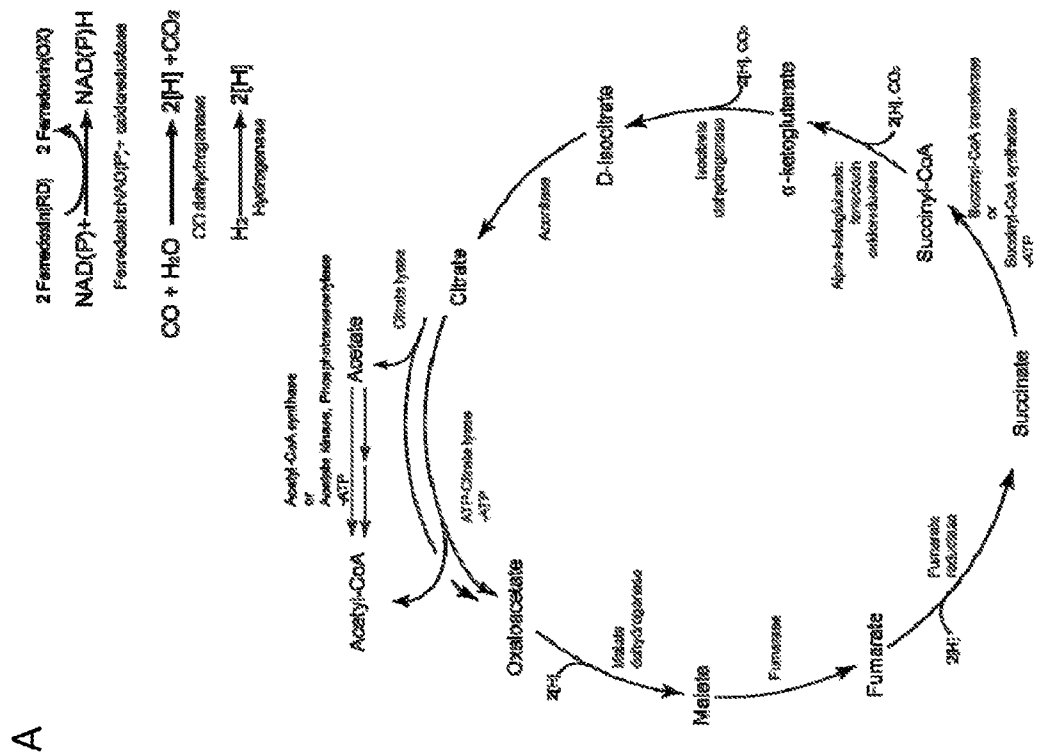
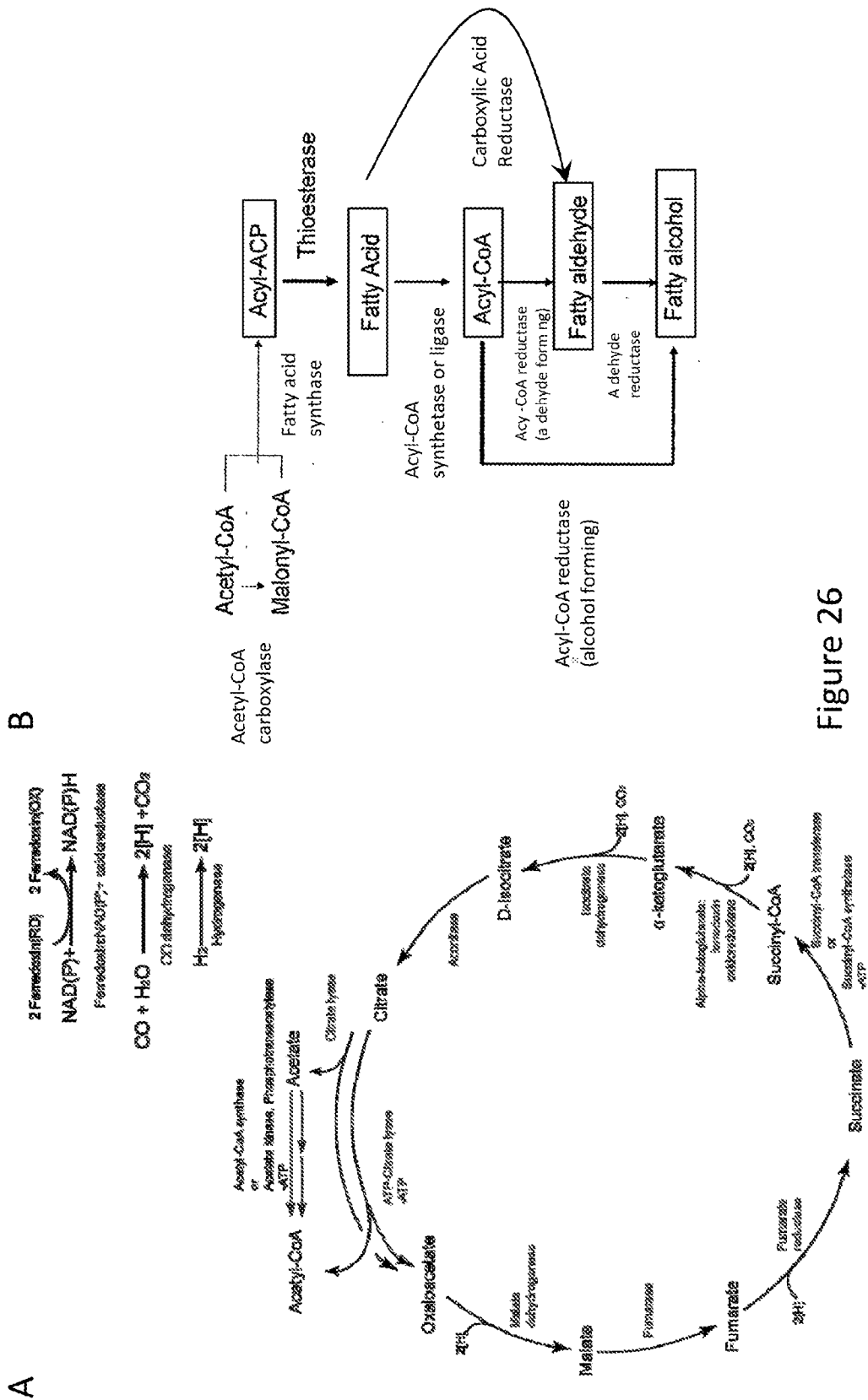


Figure 25



ATGGCAGTGGATTACCCGGATGAGCGGCTACAGCGCCGCATTGCACAGTTGTTTGCAGAAGATG
AGCAGGTCAAGGCCGCACGTCCGCTCGAAGCGGTGAGCGCGCGGTGAGCGCGCCCGGTATGCG
GCTGGCGCAGATCGCCGCCACTGTTATGGCGGGTTACGCCGACCGCCCGGCCGCGGGCAGCGT
GCGTTTCGAACCTGAACACCGACGACGCGACGGGCCGCACCTCGCTGCGGTTACTTCCCCGATTTCG
AGACCATCACCTATCGCGAACTGTGGCAGCGAGTCGGCGAGGTTGCCGCGGCCTGGCATCATGA
TCCCCGAGAACCCCTTGCGCGCAGGTGATTTCGTGCGCCCTGCTCGGCTTCACCAGCATCGACTAC
GCCACCCTCGACCTGGCCGATATCCACCTCGGCGCGGTTACCGTGCCGTTGCAGGCCAGCGCGG
CGGTGTCCAGCTGATCGCTATCCTCACCGAGACTTCGCCGCGGCTGCTCGCCTCGACCCCGGA
GCACCTCGATGCGGCGGTTCGAGTGCCTACTCGCGGGCACCACACCGGAACGACTGGTGGTCTTC
GACTACCACCCCGAGGACGACGACCAGCGTGCGGCCCTTCGAATCCGCCCGCCGCGCCTTGCCG
ACGCGGGCAGCTTGTTGATCGTCGAAACGCTCGATGCCGTGCGTGCCCGGGGCCGCGACTTACC
GGCCGCGCCACTGTTGTTCCCGACACCGACGACGACCCGCTGGCCCTGCTGATCTACACCTCC
GGCAGCACCGGAACGCCGAAGGGCGCGATGTACACCAATCGGTTGGCCGCCACGATGTGGCAGG
GGAACCTCGATGCTGCAGGGGAACTCGCAACGGGTGCGGATCAATCTCAACTACATGCCGATGAG
CCACATCGCCGGTCGCATATCGCTGTTGCGCGTGCTCGCTCGCGGTGGCACCGCATACTTCGCG
GCCAAGAGCGACATGTGACACTGTTTCGAAGACATCGGCTTGTTACGTCCCACCGAGATCTTCT
TCGTCCCGCGCGTGTGCGACATGGTCTTCAGCGCTATCAGAGCGAGCTGGACCGGCGCTCGGT
GGCGGGCGCCGACCTGGACACGCTCGATCGGGAAGTGAAAGCCGACCTCCGGCAGAACTACCTC
GGTGGGCGCTTCCTGGTGGCGGTGCTCGGCAGCGCGCCGCTGGCCGCGGAGATGAAGACGTTCA
TGGAGTCCGTCCTCGATCTGCCACTGCACGACGGGTACGGGTTCGACCGAGGCGGGCGCAAGCGT
GCTGCTCGACAACCAGATCCAGCGGCCCGCGGTGCTCGATTACAAGCTCGTCGACGTGCCCGAA
CTGGGTTACTTCCGCACCGACCGGCCGCATCCGCGCGGTGAGCTGTTGTTGAAGGCGGAGACCA
CGATTCCGGGCTACTACAAGCGGCCCGAGGTACCGCGGAGATCTTCGACGAGGACGGCTTCTA
CAAGACCGGCGATATCGTGGCCGAGCTCGAGCACGATCGGCTGGTCTATGTGACCGTTCGCAAC
AATGTGCTCAAACCTGTCGAGGGCGAGTTCGTGACCGTTCGCCCATCTCGAGGCCGTGTTTCGCCA
GCAGCCCGCTGATCCGGCAGATCTTCATCTACGGCAGCAGCGAACGTTCTATCTGCTCGCGGT
GATCGTCCCCACCGACGACGCGCTGCGCGGCCGCGACACCGCCACCTTGAAATCGGCACTGGCC
GAATCGATTTCAGCGCATCGCCAAGGACGCGAACCTGCAGCCCTACGAGATTCCGCGCGATTTC
TGATCGAGACCGAGCCGTTACCATCGCCAACGGACTGCTCTCCGGCATCGCGAAGCTGCTGCG
CCCCAATCTGAAGGAACGCTACGGCGCTCAGCTGGAGCAGATGTACACCGATCTCGCGACAGGC
CAGGCCGATGAGCTGCTCGCCCTGCGCCGCGAAGCCGCGACCTGCCGGTGCTCGAAACCGTCA
GCCGGGCAGCGAAAGCGATGCTCGGCGTCGCCTCCGCCGATATGCGTCCCGACGCGCACTTCAC
CGACCTGGGCGGCGATTCCCTTTCCGCGCTGTGCTTCTCGAACCTGCTGCACGAGATCTTCGGG
GTCGAGGTGCCGGTGGGTGTCGTCGTCAGCCCGGCGAACGAGCTGCGCGATCTGGCGAATTACA
TTGAGGCGGAACGCAACTCGGGCGCGAAGCGTCCACCTTCACCTCGGTGCACGGCGGCGGTTTC
CGAGATCCGCGCCGCCGATCTGACCCTCGACAAGTTCATCGATGCCCGCACCTGGCCGCCGCC
GACAGCATTCGCGACGCGCCGGTGCCAGCGCAGACGGTGCTGCTGACCGGCGCGAACGGCTACC
TCGGCCGGTTCTGTGCCTGGAATGGCTGGAGCGGCTGGACAAGACGGGTGGCACGCTGATCTG
CGTCGTGCGCGGTAGTGACGCGGCCGCGGCCCGTAAACGGCTGGACTCGGCGTTCGACAGCGGC
GATCCCGGCCTGCTCGAGCACTACCAGCAACTGGCCGCACGGACCTGGAAGTCTTCGCCGGTG
ATATCGGCGACCCGAATCTCGGTCTGGACGACGCGACTTGCGAGCGGTTGGCCGAAACCGTCGA
CCTGATCGTCCATCCCGCCGCGTTGGTCAACCACGTCCTTCCCTACACCCAGCTGTTTCGGCCCC

Figure 27A

AATGTCGTCGGCACCGCCGAAATCGTCCGGTTGGCGATCACGGCGCGGCGCAAGCCGGTCACCT
ACCTGTCGACCGTCGGAGTGGCCGACCAGGTGACCCGGCGGAGTATCAGGAGGACAGCGACGT
CCGCGAGATGAGCGCGGTGCGCGTCGTGCGCGAGAGTTACGCCAACGGCTACGGCAACAGCAAG
TGGGCGGGGGAGGTCTCTGCTGCGCGAAGCACACGATCTGTGTGGCTTGCCGGTTCGCGGTGTTCC
GTTCCGGACATGATCCTGGCGCACAGCCGGTACGCGGGTCAGCTCAACGTCCAGGACGTGTTTAC
CCGGCTGATCCTCAGCCTGGTCGCCACCGGCATCGCGCCGTA CTCTTCTACCGAACCGACGCG
GACGGCAACCGGCAGCGGGCCCCACTATGACGGCTTGCCGGCGGACTTCACGGCGGCGGCGATCA
CCGCGCTCGGCATCCAAGCCACCGAAGGCTTCCGGACCTACGACGTGCTCAATCCGTACGACGA
TGGCATCTCCCTCGATGAATTCGTGCGACTGGCTCGTGAATCCGGCCACCCGATCCAGCGCATC
ACCGACTACAGCGACTGGTTCCACCGTTTTCGAGACGGCGATCCGCGCGCTGCCGGAAGCAAC
GCCAGGCCTCGGTGCTGCCGTTGCTGGACGCCTACCGCAACCCCTGCCCGGCGGTCCGCGGCGC
GATACTCCCGGCCAAGGAGTTCCAAGCGGCGGTGCAAACAGCCAAAATCGGTCCGGAACAGGAC
ATCCCGCATTTCTCCGCGCCACTGATCGATAAGTACGTACGCGATCTGGAAGTCTTCTCAGCTGC
TCTAA

Figure 27A (cont'd)

mavdspderlqrriaqlfaedeqvkaarpleavsaavsapgmrlaqiaatvmagyadrpaagqr
afelntddatgrtslrlprfetityrelwqrvgevaaawhhdpenplragdfvallgftsdy
atldladihlgavtvpqlqasaavsqliailtetsprllastpehladaavecllagttperlvvf
dyhpdeddgraafesarrlradagslvivetldavrargrdlpaaplfvpdtdddplalliys
gstgtpkgamytnrlaatmwqgnsmlqgnsqrvginlnympmshiagrislfvgvlarggtayfa
aksdmstlfediglvrpteiffvprvcdmvfqryqseldrrsvagadldtldrevkadlrqnyl
ggrflvavvgsaplaaemktfmesvldlplhdgygsteagasvlldnqigrppvldyklvdvpe
lgyfrtdrphprgelllkaettipgyykrpevtaeifdedgfyktgdivaelehdrlvyvdrn
nvlklsqgefvtvahleavfassplirqifiygssersyllavivptddalrgrdtatlksala
esiqriakdanlqpyeiprdflietepftiangllsgiakllrpnlkerygaqleqmytdlatg
qadellalrreaadlpvletvsraakamlgvasadmrpdahftdlggdslsalsfsnllheifg
vevpvgvvvspanelrdlanyiearnsgakrptftsvhgggseiraadltldkfidartlaaa
dsiphapvpaqtvlltgangylgrflclewlerldktggtlicvvrgsdaaaarkrldsafds
dpgllehyqqlaartlevlagdigdpnlgllddatwqrlaetvdlivhpaalvnhvlpvtqlfgp
nvvgtaeivrlaitarrkpvtylstvgvadqvdpaeyqedsdvremsavrvvresyangygnsk
wagevllreahdlcglpvavfrsdmilahsryagqlnvqdvftrllslslvatgiapysfyrt
da dgnrqrhydglpadftaaaitalgigategfrtydvlnpyddgisldfvdwlvsgphiqri
tdysdwfhrfetairalpekqrqasvlppldayrnpavrgailpakefqaavqtakigpeqd
iphlsaplidyvsdlellqll*

Figure 27B

ATGATTGAAACCATTTCTGCCTGCAGGCGTTGAAAGCGCAGAACTGCTGGAATATCCGGAAGATC
TGAAAGCACATCCGGCAGAAGAACATCTGATTGCCAAAAGCGTTGAAAAACGTCGTCGTGATTT
TATTGGTGCACGTCATTGTGCACGTCTGGCACTGGCAGAACTGGGTGAACCTCCGGTTGCAATT
GGTAAAGGTGAACGTGGTGCACCGATTGGCCTCGTGGTGTGTTGGTAGCCTGACCCATTGTG
ATGGTTATCGTGCAGCAGCAGTTGCACATAAAATGCGCTTTCGCAGCATTGGTATTGATGCAGA
ACCGCATGCAACCCTGCCGGAAGGTGTTCTGGATAGCGTTAGCCTGCCGCCGGAACGTGAATGG
CTGAAAACCACCGATAGCGCACTGCATCTGGATCGTCTGCTGTTTTGTGCAAAAAGAAGCCACCT
ATAAAGCCTGGTGGCCGCTGACAGCACGTTGGCTGGGTTTTGAAGAAGCCCATATTACCTTTGA
AATTGAAGATGGTAGCGCAGATAGCGGTAATGGCACCTTTCATAGCGAACTGCTGGTTCCGGGT
CAGACCAATGATGGTGGTACACCGCTGCTGAGCTTTGATGGTCGTTGGCTGATTGCAGATGGTT
TTATTCTGACCGCAATTGCCTATGCCTAA

Figure 28A

mietilpagvesaelleyedlkahpaeehliaksvekrerrdfigarhcarlalaedelgeppvai
gkgergapiwprgvvgsllthcdgyraaavahkmrfrsigidaephatlpegvldsvslpperew
lkttdsalhldrllfcakeatykawwpltarwlgfeeahitfeiedgsadsgngtfhsellvpg
qtn dggtpllsfdgrwliadgfiltaiaya*

Figure 28B

atgaccagcgatgttcacgacgcccacagacggcggtaccgaaaccgcactcgacgacgagcagtcgacccgcccgc
cgccgagctgtacgccaccgatcccaggttcgcgcgcgcgcgcacccgttgcccgccgtgggtcgacgcggcgacacaaac
cggggctgcggctggcagagatcctgcagaccctgttcacccggctacggtgaccgcccggcgctgggataccgcgc
cgtgaactggccaccgacgagggcgggcgccacccgtgacgcgtctgtgcgcgcgggttcgacacccctcacctacgcca
gggtgtgggtgcgcgcgtgcaagcgggtgcgcgcggccctgcgcacaaacttcgcgcagccgatctacccggcgacgcgc
tcgcgcagatcgggttcgcgcagtcgccgattacctgacgcgtggatctcgtatgcgcctacctgggcctcgtgagtggt
ccgctgcagcacaacgcacccgggtcagccgggtgcgcgcgatcctggccgaggtcgaaaccgggatcctcacccgtgag
cgccgaataacctcgacctcgagtcgaatcgtgcgggacgtcaactcgggtgtgcgcagctcgtgggtgttcgaccatc
accccgaggtcgacgaccacgcgcgacgcactggcccgcgcggtgaacaaactcgccgggaagggcatcgccgtcac
accctggacgcgatcgccgacgagggcgccgggtgcgcgcgaaccgatctacaccgcccagccatgatcagcgcct
cgcgatgatcctgtacacctcgggttcacccggcgaccccaagggtgcgatgtacaccgaggcgatgggtggcgcggc
tgtggaccatgtcgttcatcacgggtgacccacgcgcggtcatcaacgtcaacttcatgcgcgtcaaccacctgggc
gggcgcgatccccatttcacccgcgcgtgcagaacgggtggaaccagttacttcgtaccggaatccgacatgtccacgct
gttcgaggatctcgcgctgggtgcgcgcgacccaactcggcctgggttcgcgcgcgtgcgcgacatgctctaccagcac
acctgcgcacccgtcgacgcgcctgggtcacgcagggcgccgacgaactgaccgcccgaagcagggcgggtgccaactg
cgtgagcaggtgtcgcgcgacgcgtgatcacccgattcgtcagcaccgcacccgtggcgcgaggatgagggcggtt
cctcgacatcacccctgggcgcacacatcgtcgacgggtacgggtcacccgagacccggcgcggtgacacgcgacgggtg
tgatcgtgcggccaccgggtgatcgactacaagctgatcgacgttcccgaactcgggtacttcagcaccgacaagccc
taccgcgcgtggcgaactgctgggtcaggtcgcaacgcgtgactcccgggtactacaagcgcgccgaggtcacgcgcgag
cgtcttcgacccgggacgggtactaccacaccggcgacgtcatggccgagaccgcacccgaccacctgggtgtacgtgg
accgtcgcaacaacgtcctcaaactcgcgagggcgagttcgtggcggtcgccaaacctggaggcggtgttctcgggc
gcggcgctgggtgcgcacgatcttcgtgtacggcaacagcgagcgcagtttcttctggcgctgggtgggtcccgaacgc
ggaggcgctcgagcagtagatccggccgcgcgtcaaggccgcgcgtggccgactcgtgcagcgcacccgcacgcgacg
ccgaactgcaatcctacgaggtgcccggccgatttcatcgtcgagacccgagccgttcagcgcgcgccaacgggtgctg
tcgggtgtcggaactgctgcggcccaacctcaaagaccgctacgggcagcgccctggagcagatgtacgcgcgat
cgccggccacgcagggccaaccagttgcgcgaactgcgcgcgcgcgcgcgcacacacacccgggtgatcgacacccctcaccc
aggccgctgccaogatcctcggcacccgggagcgaggtggcatccgacgcgccacttcacccgacctgggcggggatcc
ctgtcggcgctgacactttcgaacctgctgagcgatttcttcgggtttcgaagttcccgctcggcaccatcgtgaaccc
ggccaccaacctcgcccaactcgccagcacatcgaggcgacgcgcacccgggggtgaccgcagggccgagtttacc
cgtgcaogggcgggacgcacccagagatccgggcgagtgagctgacctggacaagttcatcgacgcgcaaacgctc
cgggcgcgcacccgggtctgcccgaaggtcaccacccgagccacggacgggtgttgctctcgggcgcgcaacgggtgggtggg
ccgggttctcagcttgagtggtggaacgcctggcacctgtcggcggcacccctcatcacgatcgtgcggggccgcg
acgacgcgcgcggcccgcgacgggtgacccaggctacgacacccgatcccgagttgtccgcgcgttcgcccagctg
gcccagccgcacccctgcgggtggtgcgcgggtgacatcggcgacccgaatctgggcctcacaccgagatctggcaccg
gctcgccgcgcgaggtcgacctgggtggtgcatccggcagcgctgggtcaaccaegtgtccctaccggcagctgttcg
gccccaaactcgtgggcacggccgaggtgatcaagctggccctcacccgaacggatcaagcccgctcacgtacctgtcc
accgtgtcgggtggccatggggatccccgacttcgaggaggacggcgacatccggaccgtgagcccggtgcgcgcgcgt
cgacggcggtatcgccaaacgggtacggcaacagcaagtgggccggcgaggtgctgctgaggggagggccacgatctgt
gcggggtgcccgtggcgacgttcgcgtcgacatgatcctggcgcatccgcgctacccgggtcaggtcaacgtgcca
gacatgttcacgcgactcctggtgagcctcttgatcacccggcgtcgcgccgcgggtcgttctacatcgagacgggtga
gcgcgcgcggggcgactaccccgccctgacggtcgatttcgtggccgagggcggtcacgacgctcggcgcgcgagcgc
gcgagggatacgtgtcctacgacgtgatgaaccgcacgacgacgggatctccctggatgtgttcgtggactggctg
atccgggcggggccatccgatcgaccgggtcgacgactacgacgactgggtgcgtcggttcgagaccgcgttgaccgc
gcttcccagagaagcgccgcgcacagaccgtactgcgcgtgctgcacgcgttcgcgcgtccgcagggcacccgttcgcgc
gcgcacccgaacccacggaggtgttcacgcgcgcgggtgcgcacccgcgaaggtgggcgcgggagacatccgcgacctc
gacgagggcgctgatcgacaagtagatcacgcgatctgcgtgagttcgggtctgatctaa

Figure 29A

Figure 29B

MTSDVHDATDGVTTETALDDEQSTRRIAELYATDPEFAAAAPLPAVVDAAHKPGLRLAEILQTLFTGYGDRPALGYRA
RELATDEGGRTVTRLLPRFDTLTYAQWMSRVQAVAAALRHNFAQPIYPGDAVATIGFASPDYLTLDLVCAYLGLVSV
PLQHNAPVSR LAPILAEEVEPRILTVSAEYLDLAVESVRDVNSVSQLVVFDDHPEVDDHRDALARAREQLAGKGI AVT
TLDAIADEGAGLPAEPIYTADHDQRLAMILYTSGSTGAPKGAMYTEAMVARLWTMSFITGDPPTVINVNFMPLNHLG
GRIPISTAVQNGGTSYFVPESDMSTLFEDLALVRPTLGLVPRVADMLYQHHLATVDRILVTQGADELTAEKQAGAEEL
REQVLGGRVITGFVSTAPLAAEMRAFLDITLGAHIVDCYGLTETGAVTRDGVIVRPPVIDYKLLIDVPELGYFSTDKP
YPRGELLVRSQTLTPGYKKRPEVTASVFD RDGYHTGDVMAETA PDHLVYVDRRNNVLKLAQGEFVAVANLEAVFSG
AALVRQIFVYGNSESRSPILLAVVPTPEALEQYDPAALKAALADSLQRTARDAELQSYEVPADFIVETEPFSAANGLL
SGVGKLLRPNLKDRYGQRLEQMYADIAATQANQLRELRRAAATQPVIDTLTQAAATILGTGSEVASDAHFTDLGGDS
LSALTLNLLSDDFFGFEPVPGTIVNPA TNLAQLAQHIEAQRTAGDRRPSFTTVHGADATEIRASELTLDKFIDAETL
RAAPGLPKVTTTEPRTVLLSGANGWLGRFLTQWLERLAPVGGTLITIVGRDDAAARARLTQAYDTPDPELSRRFAEL
ADRLRVVWAGDIGDPNLGLTPEIWHRLAAEVDLVVHPAALVNHVLPYRQLFGPNVVGTAEVIKIALTERIKPVTYLS
TVSVAMGIPDFEEDGDIRTVSPVRPLDGGYANGYNSKWAGEVLLREAHDL CGLPVATFRSDMILAHPRYRGQVNV P
DMFTRLLLSLLITGVAPRSFYIGDGERPRAHY PGLTVDFVAEAVTTLGAQQREGYVSVDVMNPHDDGISLDVFDWL
IRAGHPIDRVDDYDDWVRRFFETALTALPEKRRAQTVLPLLHAFRAPQAPLRGAPEPTEVFHAAVRTAKVGP GDI PHL
DEALIDKYIRD LREFGLI

atgtcgactgccacccatgaacgaactcgaccgtcgcgctccacgaactcatcgccaccgacccgcaatttcgcccgc
cgcccaaccccgacccggcgatcacccgcgcctcgaacagcccggtcgcgctgccgcagatcatccgcaccgtgc
tcgacggctacgcccagccggcgcgctgggacagcgcggtggagggttcgtcacggacgccaagaccggggcgcaog
tcggcgagctgctcccccgcttcgagaccatcacgtacagcgaagtagcgagcggtgttcggcgctggggcgcg
cctgtccgacgacgcggtgcaccccgcgacccgggtgtgctgctgggcttcaacagcgctcgactacgccaccatcg
acatggcgctggggcgccatcgggcgcgcttcgggtgccgctgcagaccagcgcggaatcagctcgctgcagccgatc
gtggccgagacccgagccaccctgatcgcgctccagcggtgaaccagctgtccgacgcggtgcagctgatcacccggcg
cgagcaggcgcccacccggctgggtgggtgttcgactaccacccgcaggtcgacgaccagcgcgagggcgctccaggacg
ccggcgcgcggtgtccagcaccggcggtggcgctccagacgctggccgagctgctggagcgcggaaggacctgccc
gocgtcgcgagcgcccgcgacgaggactcgctggccctgctgatctacacctccgggtccacggcgcccccaa
ggggcgatgtacccacagagcaacgtcggaagatgtggcgcgcgcgagcaagaactgggttcggcgagagcgccg
cgtcgatcacctgaacttcatgcgatgagccaagtgatggggcggaagcatcctctacggcacgctgggcaacggc
ggcaccgcctacttcgcccgcgcagcgacctgtccacctgcttgaggacctcgagctggtgcggccccacggagct
caacttcgtcccgcgatctgggagacgctgtacggcgaattccagcgctcaggtcgagcgcgcgctctccgagggcg
gggacgcccgcgaacgtcgcgcgctcgagggcgaggtgctggcgagcagcgccagtagctgctggggcgggcggttc
accttcgcatgacgggtcgcgccccatctcgccggagctgcgcaactgggtcgagtcgctgctcgaaatgcacct
gatggacggctacgggtccacccgagggcggaatgggtgtgttcgacggggagattcagcgcccccgcggtgatcgact
acaagctggtcgacgtgccggacctgggtacttcagcaccgacccggcgcatccgcgcggcgagctgctgctgcgc
accgagaacatgttcccggtactacaagcggggcgaaaccaccccgggcgctcttcgacgaggacggctactaccg
caccggcgagctgttcgcccagatcgccccggacccggctgggtctacgtcgaccgcgcgaacaacgtgctcaagctgg
cgagggcggaattcgtcacgctggccaaagctggagggcggttcggcgaacagcccgctgatccgcccagatctacgtc
tacggcaacagcgccccagccctacctgctggcggtcggtgggtgccaccgaggaggcgctggcctcggggtgacccga
gacgctcaagcccaagatcgccgactcgctgcagcaggtcgccaaggaggcgcgctgcagtcctacgaggtgcgc
gcgacttcatcatcgagaccaccccgcttcagcctggaaaacgggtctgctgaccgggatccggaagctggcggtggcg
aaactgaagcagcactacggggaacggctggagcagatgtacggcgacctggccgcgggacaggccaacgagctggc
cgagctgcgcgcgaacgggtgcccaggcgcggtgttgacagaccgtgagccgcgcggcgggcgccatgctgggttcgg
ccgctccgacctgtcccccgacgcccacttcaccgatctgggcggagactcgtgtcgcgcttgacattcggaac
ctgctgcgcgagatcttcgacgtcgacgtgcgggtaggcggtgatcgtcagcccgcccaacgacctggcgggccatcgc
gagctacatcgagggcgagcgggcagggcagcaagcgcccgacgttcgctcggtgcaacggccgggacgcgacccgtgg
tgcgcgcccgacacctgacgtgggacaagttcctcgacgcccagacgctggccgcggcgccgaacctgcccgaagccg
gccaccgaggtgcgcaacgtgctgctgaccggcgccaccgggttcctgggcccgtacctggccctggaatggctgga
gcggtggacatgggtggacggcaaggtcatcgccctgggtccgggcccgtccgacgaggaggaacgcgcggcggtgg
acaagaccttcgacagcgggcgaccgaaactgctcgcgactaccagcagctggccgcccgatcacctggaggtcatc
gccggcgacaagggcgaggccaatctgggcctgggccaagacgttgggcaacgactggccgacacgggtcgacgtgat
cgtcgaccccgcccgctgggtcaaccacgtgttgccgtacagcgagctgttcggggcccaacgcctgggcaacggcg
agctgatccggctggcgctgacgtccaagcagaagccgtacacctacgtgtccaccatcgcggtggggcgaccagatc
gagccgggcaagttcgtcgagaacgcgcgacatccggcagatgagcgccaccggggcgatcaacgacagctacgcca
cggctatggcaacagcaagtgggcgggcgaggtgctgctgcgcgaggcgacgacctgtgcgggtgcccgtcgcg
tgttcgctgcgacatgatcctggcgacaccacgtatgcggggcagctcaacctgcgggacatgttcaaccggctg
atgctgagcctgggtggccacccgggatcgcgccggctcgttctacgagctcgacgcccagcggaacccggcgagggc
gcactacgacggcctgcgggtcgagttcatcgccggcgcgatctcgacgctgggttcgcagatcacggacagcgaca
ccggcttcagacctaccagtgatgaaccctacgatgacggcgctcggtctggacgagtacgtcgattggctgggtg
gacgcccgtattcgatcgagcggtatcgccgactcgaatgggtcgggcggttcgagacctcgctcggggccc
gcccggacggcagcgccagtagctcgctgctgcgcgctgacacactaccgacgcccggagagccgatcaacgggt
cgatagctccacccgacgtgttcggggcagcggtgcaggaggcgaaaatcgggcccgacaaagacattccgcaagtg
tcgcccgggtcatcgtcaagtagatcacggacctgcagctgctcgggctgctctaa

Figure 30A

Figure 30B

MSTATHERLDRRVHELIA TDPQFAAAQPDPAITAALEQPGRLRPQIIRTVLDGYADRPALGQRVVEFVTDKTGRT
SAQLLPRFETITYSEVAQRVSALGRALSDDAVHPGDRVCVLGFNSVDYATIDMALGAIGAVSVPLQTSAAISSLQPI
VAETEPTLIASSVNQLSDAVQLITGAEQAPTRLVFDYHPQVDDQREAVQDAAARLSSGTGAVQTLAELLERKDLPL
AVAEPPADEDSLALLIYTSSTGAPKAMYPQSNVGMWRGSKNWFGEASAASITLNFMPMSHVMGSRSLYGTILGNG
GTAYFAARSDLSLLELLELVPRTELNFVPRIWETLYGEFQQRVERRLSEAGDAGERRAVEAEVLAEQRQYLLGGRF
TFAMTGSAPI SPELRNWVESLLEMLMDGYGSTAGMVLFDGEIQRPPIVYKLVDPDLYGYFSTDRPHPRGELLRL
TENMFFGYKRAETTAGVFDEDDGYRTGDFVFAEIA PDRLVYVDRNNVLKLAQGEFVTLAKLEAVFGNSPLIRQIYV
YGNSAQPYLLAVVPTTEALASGDPETLKPKIADSLQQVAKAQLQSYEVPRDFIETTPFSLENGLLTGIRKLAWP
KLKQHYGERLEQMYADLAAGQANELAE LRNGAQAPVLQTVSRAAGAMLGSAASDLSPDAHFTDLGGDSLSALTFFGN
LLREIFDQVDPVGVIVSPANDLAAIASYIEAERQGSKRPTFASVHGRDATVVRADLTLDKFLDAETLAAAPNLPKP
ATEVRTVLLTGATGFLGRYLALEWLERMDMVDGKVI ALVRARSDEEARARLDKTFDSGDPKLLAHYQQLAADHLEVI
AGDKGEANLGLGQDVWQRLADTVDIVDPAALVNHVLPYSELFGPNALGTAELIRLALTSKQKPYTYVSTIGVDQI
EPGKFVENADIRQMSATRAINDSYANGYGNK WAGEVLLREAHDLGCLPVAVFRCDMILADTYAGQLNLPDMFTRL
MLSLVATGIAPGSFYELDADGNRQRAHYDGLPVEFIAAAISTLGSQITDSDTGTQTYHVMNPYDDGVGLDEYVDWL
DAGYSIERIADYSEWLRRFETSLRALPDRQRQYSLPLLHNRYRTPEKPINGSIAPTDFRAAVQEAKIGPDKDIPHV
SPPVIVKYITDLQLLGLL

atgtcgccaatcacgcgtgaagagcggctcgagcgcgcgcacccaggacctctacgccaaacgacccgcagttcgccgc
cgccaaacccgccacggcgatcacgcgagcaatcgagcggcgggtctaccgctacccagatcatcgagaccgtca
tgaccggatagccgatcgccgggtctctcgctcagcgcctcggtcgaattcgtgaccgacgcgggcacccggccacacc
acgctcgcgactgctccccacttcgaaaccatcagctacggcgagctttgggaccgcatcagcgcactggccgacgt
gctcagcaccgaacagacgggtgaaacccgggcgacccgggtctgcttgttgggcttcaacagcgtcgactacgccacga
tcgacatgactttggcgcggtggggcgcggtggccgtaccactgcagaccagcgcggcgataacccagctgcagccg
atcgtcgccgagacccagccaccatgatcgccggccagcgtcgacgcactcgtgacgccacccaattgggtctgtc
cggtcagaccgctacccaggtcctgggtgttcgaccaccacccggcaggttgacgcacacccgcgcagcggtcgaatccg
cccgggagcgcctggccggctcgccgggtcgtcgaaacccctggcggaggccatcgccgcggcgacgtgcccccggggt
gctcgcccggtcgccggcccgccagatgtgtccgacgactcgtcgcgctactgatctacacctcgggcagccac
gggtcgccccaaggcgcgatgtacccccgacgcaacgttgcgaccttctggcgcaagcgcacctgggtcgaaggcg
gctacgagccgtcgatcacgctgaacttcatgccaatgagccacgtcatgggcgcgcaaatcctgtacggcacgctg
tgcaatggcggcacccgctacttctggcgcaaaagcgatctctccacctgttcgaagacctggcgctgggtcgccgc
caccgagctgaccttctgcccgcgcgtgtgggacatgggtgttcgacgagtttcagagtgggtcgaccgcgcctgg
tcgacggcgccgacccgggtcgccgtcgaagccaggtcaaggccgagatacgcacacgacgtgctcggtggacgggtat
accagcgcactgacccggctcgcgcctatctccgacgagatgaaggcgtgggtcgaggagctgctcgacatgcactc
gggtcgagggtacgggtccacccgagggccgggatgatcctgatcgacggagccattcgccgcggcggggtactcgact
acaagctgggtcgatgttcccgacctgggttacttctgaccgacccggccacatccgcggggcgagttgctgggtcaag
accgatagtgttgggtcccggtactaccagcgagccgaagtaccgcgcgacgtgttcgatgctgacggcttctaccg
gacccggcgacatcatggccgaggtcgcccccgaacagttcgtgtacctcgaccgcgcgaacaacgtgttgaaagctgt
cgcaggcgaggttcgtcacgctctccaaactcgaagcgggtgtttggcgacagcccaactgggtacggcagatctacatc
tacggcaacagcgcgcctgacctacctgttggcggtgatcgtccccaccaggaggcgctggacggcgctgctgtcga
ggagctcaaggcgcggtggggcgactcgtgcaagaggtcgcaaggccgcggcgctgcagtcctacgagatccgc
gagacttcatcatcgaaacaacaccatggaagcgtggagaacggcctgctcacgggcacccgcaagtggccaggccg
cagctgaaaaagcattacggcgagcttctcgagcagatctacacggacctggcacacggccaggccgacgaactgcg
ctcgtgcgcgaaggcggtgcccgatgcgcgggtgctgggtgacgggtgtgccgtgcggcgccgcgcgtgttggggcgga
gcgcctctgacgtccagcccgatgcgcacttaccgatttgggcggcgactcgtgtcgccgctgtcgttcaccaac
ctgctgcacgagatcttcgacatcgaagtgcgggtggcgctcatcgtcagcccccgaacgacttgcaggccctggc
cgactacgtcgaggcggtcgcaaaacccggctcgtcacggccgaccttcgcctcggtccacggcgccctcgaatgggc
aggtcaccgaggtgcatgcccgtgacctgtcctggacaaattcatcgatgccgcaaacctggccgaagctccccgg
ctgcccgcgcgaacaacccaaagtgcgcacctgctgctgacggcgccaccggcttctcgggcgctacctggccct
ggaatggctggagcggatggacctgggtcgacggcaaacgtgatcgtcgtgggtccgggcaagtccgacaccgaagcac
ggcgcggttggaacaagacgttcgacagcggcgaccccgaaactgctggccactaccgcgcactggccggcgaccac
ctcgagggtgctcgccgggtgacaaggcggaagccgacctcgactggaccggcagacctgggcaacgcctggccgacac
ggtcgacctgatcgtcgaccccgccggccctgggtcaaccacgtactgccatacagccagctgttcggggcccaacgcgc
tgggcaccgcgcgagctgctgcgggtggcgctcacctccaagatcaagccctacagctacacctcgacaatcggtgtc
gccgaccagatccgcgcgtcgccgttcaccgaggacgcccagatccgggtcatcagcgcacccgcgcgggtcgacga
cagctacgccaatggctactcgaacagcaagtgggcggcgaggtgctgttcgcgcaggcgcatgacctgtgtggcc
tgccggttgccgtgttcgctgcgacatgatcctggccgacaccacatggcggggacagctcaatgtgcgggacatg
ttaccccggatgatcctgagcctggcgccaccgggtatcgcccggggtcgttctatgagcttgcggcgacggcgcc
ccggcaacgcgcgccactatgacggtctgcccgtcgagttcatcgccgagggcatttcgactttgggtgcgcgagacc
aggatgggtttccacacgtatcacgtgatgaacccctacgacgacggcatcgactcgacgagttcgtcgactggctc
aacgagtcgggttgcccatccagcgcacgcgtgactatggcgactggctgcagcgttcgaaaccgcactgcgcgc
actgccgatcggcagcggcagcagctcactgctgcccgtgttgcaaaactatcggcagcggagcggcccgctccgcg
ggtcgatcgccctaccgatcgttccgggcagcgggtgcaagaggccaaagatcgcccccgaacaagacatccgcac
gtcgccgcgcggatcatcgtgaagtacgtcagcgacctgcgcctactcgccctgctctaa

Figure 31A

Figure 31B

MSPITREERLERRIQDLYANDPQFAAAKPAATAITAAIERPGLPLPQIIETVMTGYADRPALAAQRSVEFVTDAGTGH
 TLRLLPHFETISYGELWDRISALADVLSTEQTVKPGDRVCLLGFNSVDYATIDMTLARLGAVAVPLQTSAATQLQP
 IVAETQPTMIAASVDALADATELALSGQTATRVLVFDHHRQVDAHRAAAVESARERLAGSAVETLAEAIARGDVPRG
 ASAGSAPGTDVSDDSLALLIYTSGSTGAPKGAMYPRRNVAATFWRKRTWFEGGYEPSITLNFMPMSHVMGRQILYGTLL
 CNGGTAYFVAKSDSLSTLFEDLALVRPTELTFFVPRVWDMVFDEFQSEVDRRLVDGADRVALEAQVKAIEIRNDVLGGRY
 TSALTSAPISDEMKAWVEELLDMHLVEGYGSTEAGMILIDGAI RRPAPVLDYKLVDPDLGYFLTDRPHPRGELLVK
 TDSLFPGYYYQRAEVTADVFADGFFYRTGDIMAEVGPEQFVYLDRRNNVLKLSQGEFVTVSKLEAVFGDSPLVRQIYI
 YGNSARAYLLAVI VPTQEALDAVPVEELKARLGDSLQEVAKAAGLQSYEIPRDFI IETTPWTLENGLLTGIRKRLARP
 QLKHHYGELLEQI IYTDLAHQADELRSLRSGADAPVLVTVCRAAAALLGGSASDVQPDHFTDLGGDSLALSFTN
 LLHEIFDIEVPVGVIVSPANDLQALADYVEAARKPGSSRPTFASVHCASNGQVTEVHAGDLSLDKFI DAATLAEAPR
 LPAANTQVRTVLLTGATGFLGRYLALWLERMDLVDGKLI CLVRAKSDTEARARLDKTFDSGDPELLAHYRALAGDH
 LEVLAGDKGEADLGLDRQTWQRLADTVDLIVDPAALVNHVLPYSQLFGNALGTAELLRLALTSKIKPYSYSTIGV
 ADQIPPSAFTEDADIRV ISATRAVDDSYANGYSNSKWAGEVLLREAHDL CGLPVAVFRCDMI LADTTWAGQLNV PDM
 FTRMILSLAATGIA PGSFYELAADGARQRAHYDGLPVEFIAEAI STLGAQSODGFHTYHVMNPYDDGIGLDEFVDWL
 NESGCPIQRIADYGDWLQRFETALRALPDRQRHSSLLPLLHNRYRQPERPVRGSIAPTDRFRAAVQEA KIGPKDIPH
 VGAPIIVKYVSDLRLLGLL

atgagcaccgcaacccatgatgaacgtctggatcgctcggttcatgaactgattgcaaccgatc
cgcagtttgccagcagcacagccggatcctgcaattaccgcagcactggaacagccctgggtctgcg
tctgccgcagattattcgtaccgttctggatgggttatgcagatcgctccggcactgggtcagcgt
gttgttgaatttggtaccgatgcaaaaaccggctcgtagccagcgcacagctgctgcctcggtttg
aaaccattacctatagcgaagttgcacagcgtggttagcgcactgggtcgctgcactgagtgatga
tgcagttcatccgggtgatcggtttgtgttctgggttttaatagcgttgattatgccaccatt
gatatggcactgggtgcaattgggtgcagtttagcgttccgctgcagaccagcgcagcaattagca
gcctgcagccgattgttgagaaaaccgaaccgaccctgattgcaagcagcgttaatcagctgtc
agatgcagttcagctgattaccggtgcagaacaggcaccgaccgctctgggtgtttttgattat
catccgcaggttgatgatcagcgtgaagcagttcaggatgcagcagcagcgtctgagcagcaccg
gtgttgccagttcagaccctggcagaactgctggaacgtggtaaagatctgcctgcagttgcaga
accgctgcagatgaagatagcctggcactgctgatttataaccagcggtagcacaggtgcaccg
aaaggtgcaatgtatccgcagagcaatgttggtaaaatgtggcgctcgtagcaaaaattgggt
ttgggtgaaagcgcagcaagcattaccctgaatttcatgccgatgagccatgttatgggtcgtag
cattctgtatggcaccctgggtaatgggtggcaccgcataattttgccagcagctagcgatctgagc
accctgctggaagatctggaactgggttcgtccgaccgaactgaattttgttccgcgtattttggg
aaaccctgtatgggtgaatttccagcgtcaggttgaaacgtcgctctgagcgaagctggcgatgccgg
tgaacgtcgctgcagttgaagcagaagttctggcagaacagcgtcagtatctgctgggtgggtcgt
tttacctttgcaatgaccggttagcgcaccgatttagtccggaactgcgtaatgggttgaaagcc
tgctggaaatgcacatctgatggatggctatggtagcaccgaagcaggtatgggttctgtttgatgg
cgaaattcagcgtccgcctgtgattgattataaactgggtgatgttccggatctgggttatttt
agcaccgatcgctccgcacatccgcgtgggtgaactgctgctgcgtaccgaaaatatgtttccgggtt
attataaacgtgcagaaaccaccgcagggcgtttttgatgaagatgggttattatcgtagccggtga
tgtgtttgcagaaattgcaccggatcgctctgggttatgttgatcgctcgtaataatgttctgaaa
ctggcacaggggtgaattttgtgaccctggccaaactggaagcagtttttggtaatagtcgcctga
ttcgtcagatttatgtgtatggtaatagcgcacagccgtatctgctggcagttgttgttccgac
cgaagaggcactggcaagcgggtgatccggaaaccctgaaaccgaaaattgcagatagcctgcag
caggttgcaaaaagaagcaggtctgcagagctatgaagttccgcgtgattttattattgaaacca
ccccgtttagcctggaaaatgggtctgctgaccggtattcgtaaaactggcatggccgaaactgaa
acagcattatgggtgaacgcctggaacaaatgtatgcagatctggcagcaggtcaggcaaatgaa
ctggccgaactgcgtcgtaatgggtgcacaggcaccggttctgcagaccggttagccgtgcagccg
gtgcaatgctgggttagcgcagccagcagatctgagtcggatgcacattttaccgatctgggtgg
tgatagcctgagcgcactgacctttggtaatctgctgcgtgaaatttttgatgttgatgtgccc
gttgggtgttattgttagtccggctaatgatctggcagccattgcaagctatatattgaagcagaac
gtcagggtagcaaacgtccgacctttgcaagcgttcatgggtcgatgcaaccggttgttctgtgc
agcagatctgacctggataaaatcttggtatgcagaaaccctggcagcagcaccgaatctgccg
aaaccggcaaccgaagttcgtaccgtgctgctgacaggtgcaaccggttttctgggtcggttacc
tggcactggaatggctggaacgtatggatattgggtgatggtaaagttattgcactgggtcgctgc
ccgtagtgatgaagaagcacgcgcacgtctggataaaacctttgatagtgggtgatccgaaactg
ctggcacattatcagcagctggctgcagatcatctggaagttattgccggtgataaaggtgaag
caaatctgggtctgggtcaggtatgtttggcagcgtctggcagataccggttgatgttattgtgga
tccggcagcactgggttaatcatgttctgccgtatagcgaactgtttgggtccgaatgcactgggc

FIGURE 32A

accgcagaactgattcgtctggcactgaccagcaaacagaaaccgtatacctatgttagcacca
ttgggtgttggcgatcagattgaaccgggtaaatttgttgaaaatgccgatattcgtcagatgag
cgcaaccggtgcaattaatgatagctatgcaaatggctacggcaatagcaaatgggcaggcgaa
gttctgctgcgcggaagcacatgatctgtgtggtctgccggttgacgtttttcgttgtgatatga
ttctggccgataccacctatgcaggctcagctgaatctgccggatatgtttaccgctctgatgct
gagcctgggttgcaaccgggtattgcaccgggtagctttttatgaactggatgcagatggtaatcgt
cagcgtgcacattatgatggcctgccggttgaaatttattgcagcagccatttagcaccctgggtt
cacagattaccgatagcgataaccggttttcagacctatcatgttatgaaccgctatgatgatgg
tggttggtctggatgaatatgttgattggctgggtgatgccggttatagcattgaacgtattgca
gattatagcgaatggctgcgtgcgtttgaaacctcactgcgtgcactgccggatcgctcagcgcc
agtatagcctgctgccgctgctgcacaattatcgtaaccggaaaaaccgattaatggtagcat
tgcaccgaccgatgtttttcgtgcagccgttcaagaagccaaaattggtcgggataaagatatt
ccgcatgttagccctccggtgattgttaaataatattaccgatctgcagctgctgggtctgctgt
aa

FIGURE 32A cont.

MSTATHDERLDRRVHELIAIDPQFAAAQPDPAITAALQPGRLRPQIIRTVDGYADRPALGQR
VVEFVTDKTRTSAGLLPRFETITYSEVAQRVSALGRALSDDAVHPGDRVCVLGFNSVDYATI
DMALGAIGAVSVPLQTSAAISSLPQIVAEETEPTLIASSVNQLSDAVQLITGAEQAPTRLVFDY
HPQVDDQREAVQDAAARLSSTGVAVQTLAELLERKDLPAVAEPPADEDSLALLIYTSGSTGAP
KGAMY PQSNVGMWRRGSKNWFGEASAASITLNFMPMSHVMGRSILYGTGNGGTAYFAARSDLS
TLLEDLELVRPTELNFVPRIWETLYGEFQRQVERRLSEAGDAGERRAVEAEVLAEQRQYLLGGR
FTFAMTGSAPISPELRNWNVESLLEMHLMDGYGSTEAGMVLFDGEIQRPPVIDYKLVDVPDLGYF
STDRPHPRGELLRLTENMFPGYKRAETTAGVFDEDDGYRTGDVFAEIAPDRLVYVDRRNNVLK
LAQGEFVTLAKLEAVFGNSPLIRQIYVYGNSAQPYLLAVVVPTEEALASGDPETLKPPIADSLQ
QVAKEAGLQSYEVPRDFI IETTPFSLNGLLTGIRKLAWPKLKQHYGERLEQMYADLAAGQANE
LAELRRNGAQAPVLQTVSRAAGAMLGSAASDLSPDAHFTDLGGDSLALTFGNLLREIFDVDVP
VGVIVSPANDLAAIASYIEAERQGSKRPTFASVHGRDATVVRAADLTLDKFLDAETLAAAPNLP
KPATEVRTVLLTGATGFLGRYLALEWLERMDMVDGKVIALVRARSDEEARARLDKTFDSGDPKL
LAHYQQLAADHLEVIAGDKGEANLGLGQDVWQRLADTVDVIVDPAALVNHVLPYSELEFGPNALG
TAEILRLALTSKQKPYTYVSTIGVGDQIEPGKFVENADIRQMSATRAINDSYANGYGNSKWAGE
VLLREAHDLCLGPVAVFRCDMILADTTYAGQLNLPDMFTRLMLSLVATGIAPGSFYELDADGNR
QRAHYDGLPVEFIAAAISTLGSQITDSDTGFTYHVMNPYDDGVGLDEYVDWLVDAGYSIERIA
DYSEWLRRFETSLRALPDRQRQYSLPLLNHYRTPEKPINGSIAPTDVFRAAVQEAKIGPKDI
PHVSPPVIVKYITDLQLLGLL

FIGURE 32B

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 12/43293

A. CLASSIFICATION OF SUBJECT MATTER

IPC(8) - C12P 5/02, C12N 15/74, C12N 1/00 (2012.01)

USPC - 435/167, 435/471, 435/252.3

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC(8) - C12P 5/02, C12N 15/74, C12N 1/00 (2012.01)

USPC - 435/167, 435/471, 435/252.3

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
435/243, 435/166, 435/440, 435/252.1

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

PubWEST, PatBase, Google Scholar

search terms: Exogenous, recombin\$, transfect\$, enginee\$, microbe, microorganism, malonyl, CoA, independent, tricarboxylic acid cycle, TCA cycle, krebs cycle, citric acid cycle, citryl-CoA synthetase, citryl-CoA lyase, citrate lyase, fumarate reductase, alpha-ketoglutarate:

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X — Y	US 2009/0275097 A1 (SUN et al.) 05 November 2009 (05.11.2009) para [0007]; [0008]; [0039]; [0040]; [0044]; [0046]; [0073]-[0075]; [0139]; [0143]; [0212]; claim 25.	1, 5, 7, 8, 14-16, 22-26 ----- 6, 12, 13
Y	US 2009/0203070 A1 (DEVROE et al.) 13 August 2009 (13.08.2009) para [0009]; [0055]; [0088]; [0146]-[0148]; [0182]-[0185]; Table 1.	6, 12, 13

☐ Further documents are listed in the continuation of Box C.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

09 November 2012 (09.11.2012)

Date of mailing of the international search report

20 NOV 2012

Name and mailing address of the ISA/US

Mail Stop PCT, Attn: ISA/US, Commissioner for Patents
P.O. Box 1450, Alexandria, Virginia 22313-1450

Facsimile No. 571-273-3201

Authorized officer:

Lee W. Young

PCT Helpdesk: 571-272-4300
PCT OSP: 571-272-7774

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 12/43293

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☒ Claims Nos.: 17-21
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:
This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1. In order for all inventions to be examined, the appropriate additional examination fees must be paid.

Group I: claims 1, 5-8, 12-16, 22-26, drawn to a non-naturally occurring microbial organism having a primary alcohol pathway

Group II: claims 2, 5-7, 9, 12-16, drawn to a non-naturally occurring microbial organism having a fatty acyl-CoA pathway

Group III: claims 3, 5-7, 10, 12-16, drawn to a non-naturally occurring microbial organism having a fatty ester pathway

Group IV: claims 4-7, 11-16, drawn to a non-naturally occurring microbial organism having a wax pathway

Group V: claims 27-47, drawn to a non-naturally occurring microbial organism having an acyl-ACP pathway

---please see continuation on extra sheet---

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1, 5-8, 12-16, 22-26

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.

Continuation of Box No III Observations where unity of invention is lacking

The inventions listed as Groups I-V do not relate to a single general inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

The shared technical feature of the inventions listed as Groups I-V is a non-naturally occurring microbial organism comprising a pathway and an exogenous nucleic acid encoding a pathway enzyme. This shared technical feature fails to provide a contribution over the prior art, as evidenced by US 2009/0275097 A1 to Sun et al. (published on November 5, 2009; hereinafter 'Sun'). Sun teaches a non-naturally occurring microbial organism, comprising a microbial organism having a primary alcohol pathway comprising at least one exogenous nucleic acid encoding a primary alcohol pathway enzyme expressed in a sufficient amount to produce a primary alcohol (para [0007]-[0008]);

said non-naturally occurring microbial organism further comprising: (iii) at least one exogenous nucleic acid encodes an enzyme selected from a CO dehydrogenase (para [0139] and [0143] - "carbon monoxide dehydrogenase");

wherein said primary alcohol pathway comprises: (a) a malonyl-CoA-independent FAS pathway comprising ketoacyl-CoA acyltransferase or ketoacyl-CoA thiolase, 3-hydroxyacyl-CoA dehydrogenase, enoyl-CoA hydratase, and enoyl-CoA reductase; and an acyl-reduction pathway comprising an acyl-CoA reductase and an alcohol dehydrogenase (para [0007]). In the absence of a contribution over the prior art, the shared technical feature is not a shared special technical feature.

Further, the special technical feature of the inventions listed as Group I is a primary alcohol pathway. This special technical feature is not shared by the inventions of Groups II-V. The special technical feature of the inventions listed as Group II is a fatty acyl-CoA pathway. This special technical feature is not shared by the inventions of Groups I and III-V. The special technical feature of the inventions listed as Group III is a fatty ester pathway. This special technical feature is not shared by the inventions of Groups I-II and IV-V. The special technical feature of the inventions listed as Group IV is a wax pathway. This special technical feature is not shared by the inventions of Groups I. The special technical feature of the inventions listed as Group V is an acyl-ACP pathway. This special technical feature is not shared by the inventions of Groups I-IV.

The inventions of Groups I-V lack the same or corresponding special technical features and therefore lack unity with one another.