



(51) International Patent Classification:

C12N 1/21 (2006.01) *C12P 7/04* (2006.01)
C12N 15/52 (2006.01) *C12R 1/145* (2006.01)
C12P 7/06 (2006.01)

(21) International Application Number:

PCT/US2012/031628

(22) International Filing Date:

30 March 2012 (30.03.2012)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

61/471,029 1 April 2011 (01.04.2011) US

(71) Applicant (for all designated States except US): **THE REGENTS OF THE UNIVERSITY OF CALIFORNIA** [US/US]; 1111 Franklin Street, Oakland, CA 94607-5200 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **LIAO, James, C.** [US/US]; 10573 Wellworth Avenue, Los Angeles, CA 90024 (US). **DEKISHIMA, Yasumasa** [JP/JP]; 5-4-29-502 Nakamachidai, Tsuzuki-ku, Yokohama 224-0041 (JP).

(74) Agent: **BAKER, Joseph, R., Jr.**; Gavrilovich, Dodd & Lindsey, LLP, 4660 La Jolla Village Drive, Suite 750, San Diego, CA 92122 (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, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

- without international search report and to be republished upon receipt of that report (Rule 48.2(g))
- with sequence listing part of description (Rule 5.2(a))

(54) Title: ALCOHOL PRODUCTION FROM RECOMBINANT MICROORGANISMS



Figure 1A

(57) Abstract: Provided are microorganisms that catalyze the synthesis of biofuels from a suitable carbon dioxide. Also provided are methods of generating such organisms and methods of synthesizing biofuels using such organisms. Provided are microorganisms comprising non- naturally occurring metabolic pathway for the production of higher alcohols.



ALCOHOL PRODUCTION FROM RECOMBINANT MICROORGANISMS

CROSS REFERENCE TO RELATED APPLICATIONS

[0001] This application claims priority to U.S. Provisional Application Serial No. 61/471,029, filed April 1, 2011, the disclosure of which is incorporated herein by reference in its entirety.

TECHNICAL FIELD

[0002] Metabolically-modified microorganisms and methods of producing such organisms are provided. Also provided are methods of producing various alcohols including n-hexanol and n-octanol by contacting a suitable substrate with a metabolically-modified microorganism and enzymatic preparations of the disclosure.

BACKGROUND

[0003] With the prospect of unstable and rising price for petroleum, there has been an increasing interest on the development of sustainable manufacturing processes to supply chemicals and fuels. Production of biofuels, in particular, has been the focus of many groups with successful outcomes (for review see, Jang et al. *Biotechnol Adv.*, 2011; Mainguet and Liao, *Biotechnol J.* 5, 1297-308 2010; Yan and Liao, *J Ind Microbiol Biotechnol.* 36, 471-9, 2009). Several metabolic pathways have been engineered to produce higher alcohols (Atsumi et al., *Metab Eng.* 10, 305-11, 2008; Atsumi et al., *Nature.* 451, 86-9, 2008), alkane (Schirmer et al., *Science.* 329, 559-62, 2010), and biodiesel (Steen et al., *Nature.* 463, 559-62, 2010) in *Escherichia coli* as well as in others hosts such as *Corynebacterium glutamicum* (Smith et al., *Appl Microbiol Biotechnol.* 87, 1045-55, 2010), *Clostridium cellulolyticum* (Higashide et al., *Appl Environ Microbiol.* 77, 2727-33, 2011), *Bacillus subtilis* (Li et al., *Appl Microbiol Biotechnol.* 91, 577-89, 2011), and cyanobacteria (Atsumi et al., *Nat Biotechnol.* 27, 1177-80, 2009; Lan and Liao, *Metab Eng.* 13, 353-63, 2011). In addition, the production of 1-butanol using native *Clostridium* producers has continued to make significant strides (Lee et al., *Biotechnol J.* 4, 1432-40, 2009; Wang and Blaschek, *Bioresour Technol.* 102, 9985-90, 2011).

SUMMARY

[0004] Production of green chemicals and fuels using metabolically engineered organisms is a promising alternative to petroleum-based production. Higher chain alcohols (C4-C8) are of interest because they can be used as chemical feedstocks as well as fuels.

[0005] The methods and compositions described herein extend the CoA-dependent 1-butanol synthetic reaction sequence to produce n-hexanol. The disclosure demonstrates n-hexanol production from glucose by the engineered *E. coli* strain. Increasing the BktB activity towards butyryl-CoA while eliminating the AdhE2 activity towards the same substrate is one direction to improve n-hexanol production. Furthermore, *Clostridium sp.* BS-1 isolated by Joen et al. can produce 1.73 g/L of hexanoic acid, but no n-hexanol production was reported. The biosynthesis of hexanoic acid in this organism is also presumed to start from acetyl-CoA using the similar scheme as in Scheme 1. It appears that this organism may have more efficient enzymes to extend the acyl-CoA chain length to hexanoyl-CoA. Transferring the corresponding genes from this organism to *E. coli* appears to be an interesting direction for improvement. The kinetics of the enzymes involved remains to be characterized. Nevertheless, the strategy described in this study may be extended further for the production of other even number longer chain alcohols.

[0006] The disclosure provides a recombinant microorganism that produces a C6-8 alcohol comprising a recombinant pathway including a heterologous beta-ketothiolase and one or more enzymes selected from the group consisting of an acetyl-CoA acetyltransferase, a 3-hydroxybutyryl-CoA dehydrogenase, a crotonase, a trans-2-enoyl-CoA reductase and a bifunctional aldehyde/alcohol dehydrogenase. In one embodiment, the recombinant microorganism comprises a heterologous polypeptide having beta ketothiolase activity and a heterologous polypeptide having acetyl-coA acetyltransferase activity. In a further embodiment the recombinant microorganism comprises a heterologous expression or elevated expression of a polypeptide having trans-2-enoyl-CoA reductase activity, as compared to a parental microorganism, wherein the recombinant microorganism

produces a metabolite comprising butyryl-CoA from a substrate that includes crotonyl-CoA. In any of the foregoing embodiments, the trans-2-enoyl-CoA reductase (*ter*) is derived from a *Treponema denticola* or *F. succinogenes*. In another embodiment of any of the foregoing embodiments, the polypeptide having trans-2-enoyl-CoA (*Ter*) polypeptide activity is at least 50% identical to a sequence as set forth in SEQ ID NO:69, 70, 71, or 72. In another embodiment, the *Ter* comprises an M11K substitution. In yet another embodiment of any of the foregoing, the recombinant microorganism comprises a heterologous expression or elevated expression of a polypeptide having acetyl-CoA acetyltransferase activity, as compared to a parental microorganism, wherein the recombinant microorganism produces a metabolite comprising acetoacetyl-CoA from a substrate comprising acetyl-CoA. In a further embodiment, the polypeptide having acetyl-CoA acetyltransferase activity is encoded by a polynucleotide having at least about 50% identity to a sequence as set forth in SEQ ID NO:29. In another embodiment of any of the foregoing embodiments, the polypeptide having acetyl-CoA acetyltransferase activity is encoded by an *atoB* gene or homolog thereof, or a *fadA* gene or homolog thereof. In one embodiment, the *atoB* gene or *fadA* gene is derived from the genus *Escherichia* (e.g., *E. coli*). In yet another embodiment of any of the foregoing embodiments, the recombinant microorganism comprises a heterologous expression or elevated expression of a polypeptide having hydroxybutyryl-CoA dehydrogenase activity, as compared to a parental microorganism, wherein the recombinant microorganism produces a metabolite comprising 3-hydroxybutyryl-CoA from a substrate comprising acetoacetyl-CoA. In one embodiment, the polypeptide having hydroxybutyryl-CoA activity is encoded by a polynucleotide having at least about 50% identity to a sequence as set forth in SEQ ID NO:35, 73, 74, 75, 76, 77 or 78. In another embodiment, the hydroxybutyryl-CoA dehydrogenase is encoded by an *hbd* gene or homolog thereof or a *paahl* gene or homolog thereof. In yet a further embodiment, the *hbd* gene is derived from a microorganism selected from the group consisting of *Clostridium acetobutylicum*, *Clostridium difficile*, *Dastricha ruminatum*, *Butyrivibrio fibrisolvens*, *Treponema phagedemes*, *Acidaminococcus fermentans*,

Clostridium kluyveri, *Syntrophospora bryanti*, and *Thermoanaerobacterium thermosaccharolyticum*. In a specific embodiment, the microorganism from which the hbd is derived is *Clostridium acetobutylicum*. In another embodiment of any of the foregoing, the recombinant microorganism comprises a heterologous expression or elevated expression of a polypeptide having crotonase activity, as compared to a parental microorganism, wherein the recombinant microorganism produces a metabolite comprising crotonyl-CoA from a substrate comprising 3-hydroxybutyryl-CoA. In one embodiment, the polypeptide having crotonase activity is encoded by a polynucleotide having at least about 50% identity to a sequence as set forth in SEQ ID NO:51. In another embodiment, the crotonase is encoded by a *crt* gene or homolog thereof. In yet another embodiment, the *crt* gene is derived from a microorganism selected from the group consisting of *Clostridium acetobutylicum*, *Butyrivibrio fibrisolvens*, *Thermoanaerobacterium thermosaccharolyticum*, and *Clostridium difficile*. In yet a further embodiment, the microorganism from which the *crt* is derived is *Clostridium acetobutylicum*. In another embodiment of any of the foregoing embodiments, the microorganism comprises a heterologous expression or elevated expression of a polypeptide having aldehyde/alcohol dehydrogenase activity, as compared to a parental microorganism, wherein the recombinant microorganism produces a metabolite comprising butyraldehyde from a substrate comprising butyryl-CoA. In a further embodiment, the polypeptide having aldehyde/alcohol dehydrogenase activity is encoded by a polynucleotide having at least about 50% identity to a sequence as set forth in SEQ ID NO:67. In another embodiment, the polypeptide having aldehyde/alcohol dehydrogenase is encoded by an *aad* gene or homolog thereof, or an *adhE2* gene or homolog thereof. In a further embodiment, the *aad* gene or *adhE2* gene is derived from *Clostridium acetobutylicum*. In another embodiment, based on any of the foregoing embodiment, the recombinant microorganism comprises elevated expression of a polypeptide having beta-ketothiolase activity, as compared to a parental microorganism, wherein the recombinant microorganism produces a metabolite comprising 3-ketohexanoyl-CoA from substrates comprising butyryl-CoA and acetyl-

CoA. In a further embodiment, the polypeptide having beta-ketothiolase activity is encoded by a *bktB* gene or homolog thereof. In yet a further embodiment, the *bktB* is derived from *R. eutropha*. In a yet further embodiment, the *bktB* gene or homolog thereof comprises a sequence having at least 50% identity to SEQ ID NO:37. In addition to the recombinant expression of any of the foregoing genes/polypeptides, the microorganism may further comprise a knockout in a gene selected from the group consisting of *ldhA*, *adhE*, *frdBC* and any combination thereof. In a specific embodiment, the microorganism expresses polypeptides having beta-ketothiolase activity, acetyl-CoA acetyltransferase activity, 3-hydroxybutyryl-CoA dehydrogenase activity, crotonase activity, trans-2-enoyl-CoA reductase activity and aldehyde/alcohol dehydrogenase activity. For example, the microorganism comprises an *atoB*, *BktB*, *hbd* or *paaH1*, *crt* and a *ter* gene. In an embodiment, of the disclosure the the C6-C8 alcohol are selected from n-hexanol and n-octanol.

[0007] The disclosure also provides a method of making a recombinant microorganism as set forth above, comprising transforming a parental cell with one or more vectors for expressing a polypeptide having beta-ketothiolase activity, acetyl-CoA acetyltransferase activity, 3-hydroxybutyryl-CoA dehydrogenase activity, crotonase activity, trans-2-enoyl-CoA reductase activity and/or bifunctional aldehyde/alcohol dehydrogenase activity.

[0008] The disclosure also provides a method of making n-hexanol and/or octanol comprising culturing a recombinant microorganism as set forth above with a substrate and under conditions to produce n-hexanol and/or n-octanol.

[0009] The disclosure also provides a selection platform that allows selection or enrichment of enzymes that showed increased synthesis of C6 and C8 linear alcohols from mutated protein libraries or enzyme variants in nature. In addition, the disclosure identifies 3-hydroxy-acyl-CoA dehydrogenase (*Hbd*) as a limiting step in the synthesis of n-hexanol and octanol using a recombinant pathway for production of these alcohols. Replacement of the original enzyme *Clostridium acetobutylicum* *Hbd* with *Ralstonia eutropha* homologue *PaaH1* increased production of n-hexanol by at least 10-fold. Further directed evolution by random mutagenesis of

PaaH1 and the selection scheme identified several PaaH1 variants that showed improvement in n-hexanol and n-octanol production. This selection platform are useful for production of long-chain alcohols using the synthetic reverse β -oxidation pathway.

[0010] The disclosure provides a recombinant microorganism that produced n-hexanol from a suitable carbon source comprising beta-ketothiolase (e.g., BktB), acetyl-CoA acetyltransferase (e.g., AtoB), 3-hydroxybutyryl-CoA dehydrogenase (e.g., Hbd or PaaH1), crotonase (e.g., Crt), and trans-enoyl-CoA reductase (e.g., Ter) from various organisms. The disclosure also provides a recombinant microorganism that can produced at least 27 mg/L of n-hexanol secreted to the fermentation medium under anaerobic conditions. Furthermore, by co-expressing formate dehydrogenase n-hexanol titer can be increased to 47 mg/L or more.

[0011] The disclosure provides a recombinant microorganism that produced n-octanol from a suitable carbon source comprising beta-ketothiolase (e.g., BktB), acetyl-CoA acetyltransferase (e.g., AtoB), 3-hydroxybutyryl-CoA dehydrogenase (e.g., Hbd or PaaH1), crotonase (e.g., Crt), and trans-enoyl-CoA reductase (e.g., Ter) from various organisms.

[0012] The details of one or more embodiments of the disclosure are set forth in the accompanying drawings and the description below. Other features, objects, and advantages will be apparent from the description and drawings, and from the claims.

BRIEF DESCRIPTION OF THE DRAWINGS

[0013] The accompanying drawings, which are incorporated into and constitute a part of this specification, illustrate one or more embodiments of the disclosure and, together with the detailed description, serve to explain the principles and implementations of the invention.

[0014] **Figure 1A-G** shows HPLC analysis of each enzymatic reaction of hexenoyl-CoA with (A) no enzymes and cofactors, (B) Crt, (C) Crt, Hbd, and NAD^+ , (D) Crt, Hbd, BktB, NAD^+ , and CoA. (Right) HPLC analysis of EgTer and TdTer reaction with hexenoyl-CoA. Hexenoyl-CoA was incubated with (E) no enzyme as control; (F) EgTer and NADH; or (G) TdTer and NADH.

[0015] **Figure 2A-B** shows characterization of AtoB activity to C6 substrate by HPLC. Hexenoyl-CoA was incubated with (A) Crt, Hbd, and NAD⁺, (B) Crt, Hbd, AtoB, NAD⁺, and CoA.

[0016] **Figure 3A-C** shows an HPLC analysis of enzymatic reaction of hexenoyl-CoA with (A) NADH, (B) TdTer and NADH, and (C) EgTer and NADH.

[0017] **Figure 4A-D** shows gas chromatogram analysis of enzymatic reaction of hexanoyl-CoA with (A) cell lysate from JCL166 and NADH, (B) cell lysate from JCL166/pCS38_ *adhE2* and NADH. Abbreviations: Internal standard (I.S.). (C) Shows a diagram of a recombinant pathway of the disclosure. (D) Shows a diagram of the anaerobic growth rescue system and higher alcohol production in *E. coli*. In the presence of AdhE, both n-butanol and n-hexanol are produced in *E. coli* under anaerobic conditions (connected lines). Elimination of AdhE induces cell growth arrest due to the accumulation of NAD⁺ and acyl-CoA intermediates. To rescue cellular growth, a long-chain acyl-CoA thioesterase (mBACH, dotted line) was introduced, promoting the consumption of NADH and longer-chain acyl-CoA intermediates to produce fatty acids (hexanoic acid). Abbreviations: Fdh, formate dehydrogenase; AtoB, acetyl-CoA acetyltransferase; BktB, β -ketothiolase; Hbd, 3-hydroxy-acyl-CoA dehydrogenase; Crt, crotonase; Ter, trans-enoyl-CoA reductase; AdhE, aldehyde/alcohol dehydrogenase; mBACH, mouse brain acyl-CoA hydrolase.

[0018] **Figure 5A-C** shows a gas chromatogram analysis of n-hexanol production in engineered *E. coli*: (A) JCL299/pEL11/pEL102, (B) JCL166/pEL11/pEL102, (C) JCL166/pEL11/pIM8 (I.S. - internal standard (2-methyl-1-pentanol)).

[0019] **Figure 6** shows a GC-MS analysis of n-hexanol in the culture broth of JCL299/pEL11/pEL102/pCS138. Total ion chromatogram (left) and EI mass spectra (right) were shown. Comparison of the n-hexanol standard and fermentation sample regarding retention time and EI mass spectra confirms the presence of n-hexanol in the culture medium.

[0020] **Figure 7** is a time course of n-hexanol production and cell density in JCL299/pEL11/pEL102/ pCS138 cultivation.

[0021] **Figure 8** depicts SEQ ID NO:29, a nucleic acid sequence derived from an *atoB* gene encoding a polypeptide having keto thiolase activity (SEQ ID NO:30 is the encoded polypeptide).

[0022] **Figure 9** depicts SEQ ID NO:31, a nucleic acid sequence derived from a *thlA* gene encoding a polypeptide having acetyl-CoA acetyltransferase activity. (SEQ ID NO:32 is the encoded polypeptide).

[0023] **Figure 10** depicts SEQ ID NO:33, a nucleic acid sequence derived from a *crt* gene encoding a polypeptide having crotonase activity. (SEQ ID NO:34 is the encoded polypeptide).

[0024] **Figure 11** depicts SEQ ID NO:35 (SEQ ID NO:36 is the encoded polypeptide) and 37 (SEQ ID NO:38 is the encoded polypeptide), a nucleic acid sequence derived from a *hbd* gene encoding a polypeptide having hydroxybutyryl CoA dehydrogenase activity and a nucleic acid sequence encoding a beta-ketothiolase (BktB), respectively.

[0025] **Figure 12** depicts SEQ ID NO:39, a nucleic acid sequence derived from a *bcd* gene encoding a polypeptide having butyryl-CoA dehydrogenase activity. (SEQ ID NO:40 is the encoded polypeptide).

[0026] **Figure 13** depicts SEQ ID NO:41, a nucleic acid sequence derived from an *etfA* gene encoding an ETF polypeptide. (SEQ ID NO:42 is the encoded polypeptide).

[0027] **Figure 14** depicts SEQ ID NO:43, a nucleic acid sequence derived from an *etfB* gene encoding an ETF polypeptide. (SEQ ID NO:44 is the encoded polypeptide).

[0028] **Figure 15** depicts SEQ ID NO:45, a nucleic acid sequence derived from a *bcd* gene encoding a polypeptide having butyryl-CoA dehydrogenase activity. (SEQ ID NO:46 is the encoded polypeptide).

[0029] **Figure 16** depicts SEQ ID NO:47, a nucleic acid sequence derived from an *etfA* gene encoding an ETF polypeptide. (SEQ ID NO:48 is the encoded polypeptide).

[0030] **Figure 17** depicts SEQ ID NO:49, a nucleic acid sequence derived from an *etfB* gene encoding an ETF polypeptide. (SEQ ID NO:50 is the encoded polypeptide), and SEQ ID NO:51, a *crt* from *C. acetobutylicum* (SEQ ID NO:52 is the encoded polypeptide).

[0031] **Figure 18** depicts SEQ ID NO:52, a nucleic acid sequence derived from a *ccr* gene encoding a polypeptide having crotonyl CoA reductase activity. (SEQ ID NO:54 is the encoded polypeptide).

[0032] **Figure 19** depicts SEQ ID NO:55, a nucleic acid sequence derived from a *ccr* gene encoding a polypeptide having crotonyl CoA reductase activity. (SEQ ID NO:56 is the encoded polypeptide).

[0033] **Figure 20** depicts SEQ ID NO:57, a nucleic acid sequence derived from a *ccr* gene encoding a polypeptide having crotonyl CoA reductase activity. (SEQ ID NO:58 is the encoded polypeptide).

[0034] **Figure 21** depicts SEQ ID NO:59, a nucleic acid sequence derived from a *ccr* gene encoding a polypeptide having crotonyl CoA reductase activity. (SEQ ID NO:60 is the encoded polypeptide).

[0035] **Figure 22** depicts SEQ ID NO:61, a nucleic acid sequence derived from a *ccr* gene encoding a polypeptide having crotonyl CoA reductase activity. (SEQ ID NO:62 is the encoded polypeptide).

[0036] **Figure 23** depicts SEQ ID NO:63, a nucleic acid sequence derived from a *ccr* gene encoding a polypeptide having crotonyl CoA reductase activity. (SEQ ID NO:64 is the encoded polypeptide).

[0037] **Figure 24** depicts SEQ ID NO:65, a nucleic acid sequence derived from a *ccr* gene encoding a polypeptide having crotonyl CoA reductase activity. (SEQ ID NO:66 is the encoded polypeptide).

[0038] **Figure 25** depicts SEQ ID NO:67, a nucleic acid sequence derived from a *adhE2* gene encoding a polypeptide having alcohol dehydrogenase activity. (SEQ ID NO:68 is the encoded polypeptide).

[0039] **Figure 26A-B** shows (A) Multiple sequence alignments of Ter homologues from various organisms using ClustalW. The M11K amino acid substitution found in the *F. succinogenes* Ter mutants is shaded. Fully conserved residues are noted with asterisks. TD: *T. denticola* (SEQ ID NO:69), TV: *T. vincentii* (SEQ ID NO:70), FS: *F. succinogenes* (SEQ ID NO:71), FJ: *F. johnsonia* (SEQ ID NO:72). (B) Multiple sequence alignments of PaaH1 mutants (SEQ ID NOs: 73-78). Amino acid substitutions are highlighted. WT, wild type (SEQ ID NO:73).

[0040] **Figure 27** shows cell density of different strains after anaerobic incubation. Cells were grown for 24 h at 30 °C in LB medium supplemented with 1% glucose, 100 µM IPTG and the appropriate antibiotics, as described with more details in Material and Methods.

Except for control (Ctrl, JCL166 without plasmids), all other strains carry the plasmid pHL101 (Tde.Ter BktB Egl.Ter), and a second plasmid: pHM10H (AtoB mBACH Crt Hbd), pHM10C (AtoB mBACH Crt1 Hbd), or pHM10P (AtoB mBACH Crt PaaH1). Asterisk indicates that these strains carry either the wild type gene (PaaH1) or its variants (P01, P042, P06, P09, and P13).

[0041] **Figure 28** shows growth rescue of PaaH1 mutant pools after two rounds of enrichment under anaerobic conditions. Columns represent the cellular growth density of *E. coli* JCL166 bearing plasmids pHL101 and pHM10P* mutant library grown in LB medium supplemented with 1% glucose, 30 µg/mL kanamycin, 150 µg/mL ampicillin, 100 µM IPTG, after incubation for 48 hours at 30 °C and 240 rpm, under anaerobic conditions. Control culture is JCL166 carrying pHL101 and the parental plasmid pHM10P. Five different clones were selected for further testing and are indicated above their respective library origin.

[0042] **Figure 29** shows fatty acid production by PaaH1 variants in JCL299. Butanoic and hexanoic acids were determined by GC analysis as described in Material and Methods. Except for control, all other strains carry the plasmid pHL101 which expresses Tde.Ter, BktB, and Egl.Ter; and a second plasmid expressing AtoB, mBACH, Crt, and PaaH1 and its variants as indicated. Strains: Control, JCL299 strain without plasmids; Hbd, *C. acetobutylicum* Hdb; PaaH1, *R. eutropha* PaaH1; P01, P042, P06, P09, P13, PaaH1 variants. The values represent the mean ± SD of two independent cultures. ** P < 0.01 and *** P < 0.001 when compared to the PaaH1 wild-type by One-Way ANOVA followed by Tukey's Test.

[0043] **Figure 30A-D** shows production of higher-chain alcohols by *E. coli* JCL299 strains. Determination of alcohol production and culture conditions are described in Material and Methods. Except for control, all other strains carry the plasmid pHL101 which expresses Tde.Ter, BktB, and Egl.Ter; and a second plasmid expressing AtoB, mBACH, Crt, and PaaH1 and its variants as indicated. Strains: Control, JCL299 strain without plasmids; PaaH1, *R. eutropha* PaaH1 wild type; P01, P042, P06, P09, and P13, PaaH1 variants. Results are mean ± SD of at least two independent cultures.

DETAILED DESCRIPTION

[0044] As used herein and in the appended claims, the singular forms "a," "and," and "the" include plural referents unless the context clearly dictates otherwise. Thus, for example, reference to "a polynucleotide" includes a plurality of such polynucleotides and reference to "the microorganism" includes reference to one or more microorganisms, and so forth.

[0045] Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood to one of ordinary skill in the art to which this disclosure belongs. Although methods and materials similar or equivalent to those described herein can be used in the practice of the disclosed methods and compositions, the exemplary methods, devices and materials are described herein.

[0046] Also, the use of "or" means "and/or" unless stated otherwise. Similarly, "comprise," "comprises," "comprising" "include," "includes," and "including" are interchangeable and not intended to be limiting.

[0047] It is to be further understood that where descriptions of various embodiments use the term "comprising," those skilled in the art would understand that in some specific instances, an embodiment can be alternatively described using language "consisting essentially of" or "consisting of."

[0048] Any publications discussed above and throughout the text are provided solely for their disclosure prior to the filing date of the present application. Nothing herein is to be construed as an admission that the inventors are not entitled to antedate such disclosure by virtue of prior disclosure.

[0049] The disclosure provides recombinant organisms comprising metabolically engineered biosynthetic pathways that utilize an organism's CoA pathway for the production of higher alcohols. Higher alcohol production utilizing the organism's CoA pathway offers several advantages. Not only does it avoid the difficulty of expressing a large set of foreign genes but it also minimizes the possible accumulation of toxic intermediates. Contrary to the butanol production pathway found in many species of *Clostridium*, the engineered amino acid biosynthetic routes for higher alcohol

production circumvent the need to involve oxygen-sensitive enzymes and intermediates.

[0050] In one embodiment, the disclosure provides a recombinant microorganism comprising elevated expression of at least one target enzyme as compared to a parental microorganism or encodes an enzyme not found in the parental organism. In another or further embodiment, the microorganism comprises a reduction, disruption or knockout of at least one gene encoding an enzyme that competes with a metabolite necessary for the production of a desired higher alcohol product or which produces an unwanted product. The recombinant microorganism produces at least one metabolite involved in a biosynthetic pathway for the production of n-hexanol, butanol and/or n-octanol. In general, the recombinant microorganisms comprises at least one recombinant metabolic pathway that comprises a target enzyme and may further include a reduction in activity or expression of an enzyme in a competitive biosynthetic pathway. The pathway acts to modify a substrate or metabolic intermediate in the production of, for example, n-hexanol, butanoyl-CoA and/or n-octanol. The target enzyme is encoded by, and expressed from, a polynucleotide derived from a suitable biological source. In some embodiments, the polynucleotide comprises a gene derived from a bacterial or yeast source and recombinantly engineered into the microorganism of the disclosure.

[0051] As used herein, an "activity" of an enzyme is a measure of its ability to catalyze a reaction resulting in a metabolite, *i.e.*, to "function", and may be expressed as the rate at which the metabolite of the reaction is produced. For example, enzyme activity can be represented as the amount of metabolite produced per unit of time or per unit of enzyme (*e.g.*, concentration or weight), or in terms of affinity or dissociation constants.

[0052] The term "biosynthetic pathway", also referred to as "metabolic pathway", refers to a set of anabolic or catabolic biochemical reactions for converting (transmuting) one chemical species into another. Gene products belong to the same "metabolic pathway" if they, in parallel or in series, act on the same substrate, produce the same product, or act on or produce a metabolic intermediate (*i.e.*, metabolite) between the same substrate

and metabolite end product. The disclosure provides recombinant microorganism having a metabolically engineered pathway for the production of a desired product or intermediate.

[0053] Accordingly, metabolically "engineered" or "modified" microorganisms are produced via the introduction of genetic material into a host or parental microorganism of choice thereby modifying or altering the cellular physiology and biochemistry of the microorganism. Through the introduction of genetic material the parental microorganism acquires new properties, e.g. the ability to produce a new, or greater quantities of, an intracellular metabolite. In an illustrative embodiment, the introduction of genetic material into a parental microorganism results in a new or modified ability to produce n-hexanol and/or n-octanol. The genetic material introduced into the parental microorganism contains gene(s), or parts of gene(s), coding for one or more of the enzymes involved in a biosynthetic pathway for the production of n-hexanol and/or n-octanol, and may also include additional elements for the expression and/or regulation of expression of these genes, e.g. promoter sequences.

[0054] An engineered or modified microorganism can also include in the alternative or in addition to the introduction of a genetic material into a host or parental microorganism, the disruption, deletion or knocking out of a gene or polynucleotide to alter the cellular physiology and biochemistry of the microorganism. Through the reduction, disruption or knocking out of a gene or polynucleotide the microorganism acquires new or improved properties (e.g., the ability to produce a new or greater quantities of an intracellular metabolite, improve the flux of a metabolite down a desired pathway, and/or reduce the production of undesirable by-products).

[0055] An "enzyme" means any substance, preferably composed wholly or largely of amino acids making up a protein or polypeptide that catalyzes or promotes, more or less specifically, one or more chemical or biochemical reactions.

[0056] As used herein, the term "metabolically engineered" or "metabolic engineering" involves rational pathway design and assembly of biosynthetic genes, genes associated with operons, and

control elements of such polynucleotides, for the production of a desired metabolite, such as an acetoacetyl-CoA or higher alcohol, in a microorganism. "Metabolically engineered" can further include optimization of metabolic flux by regulation and optimization of transcription, translation, protein stability and protein functionality using genetic engineering and appropriate culture condition including the reduction of, disruption, or knocking out of, a competing metabolic pathway that competes with an intermediate leading to a desired pathway. A biosynthetic gene can be heterologous to the host microorganism, either by virtue of being foreign to the host, or being modified by mutagenesis, recombination, and/or association with a heterologous expression control sequence in an endogenous host cell. In one aspect, where the polynucleotide is xenogenetic to the host organism, the polynucleotide can be codon optimized.

[0057] A "metabolite" refers to any substance produced by metabolism or a substance necessary for or taking part in a particular metabolic process that gives rise to a C6-C8 alcohol. A metabolite can be an organic compound that is a starting material (e.g., glucose or pyruvate), an intermediate in (e.g., acetyl-coA), or an end product (e.g., n-hexanol) of metabolism. Metabolites can be used to construct more complex molecules, or they can be broken down into simpler ones. Intermediate metabolites may be synthesized from other metabolites, perhaps used to make more complex substances, or broken down into simpler compounds, often with the release of chemical energy.

[0058] A "native" or "wild-type" protein, enzyme, polynucleotide, gene, or cell, means a protein, enzyme, polynucleotide, gene, or cell that occurs in nature.

[0059] A "parental microorganism" refers to a cell used to generate a recombinant microorganism. The term "parental microorganism" describes a cell that occurs in nature, i.e. a "wild-type" cell that has not been genetically modified. The term "parental microorganism" also describes a cell that has been genetically modified but which does not express or over-express a target enzyme e.g., an enzyme involved in the biosynthetic pathway

for the production of a desired metabolite such as n-butanol, n-hexanol or n-octanol.

[0060] For example, a wild-type microorganism can be genetically modified to express or over express a first target enzyme such as acetyl-coA acetyl transferase. This microorganism can act as a parental microorganism in the generation of a microorganism modified to express or over-express a second target enzyme *e.g.*, 3-hydroxybutyryl-coA dehydrogenase. In turn, the microorganism modified to express or over express *e.g.*, crotonase and trans-2-enoyl-CoA reductase can be modified to express or over express a third target enzyme, *e.g.*, beta-ketothiolase.

[0061] Accordingly, a parental microorganism functions as a reference cell for successive genetic modification events. Each modification event can be accomplished by introducing a nucleic acid molecule in to the reference cell. The introduction facilitates the expression or over-expression of a target enzyme. It is understood that the term "facilitates" encompasses the activation of endogenous polynucleotides encoding a target enzyme through genetic modification of *e.g.*, a promoter sequence in a parental microorganism. It is further understood that the term "facilitates" encompasses the introduction of exogenous polynucleotides encoding a target enzyme in to a parental microorganism.

[0062] A "protein" or "polypeptide", which terms are used interchangeably herein, comprises one or more chains of chemical building blocks called amino acids that are linked together by chemical bonds called peptide bonds. A protein or polypeptide can function as an enzyme.

[0063] Polynucleotides that encode enzymes useful for generating metabolites (*e.g.*, keto thiolase, acetyl-CoA acetyltransferase, hydroxybutyryl-CoA dehydrogenase, crotonase, crotonyl-CoA reductase, butyryl-CoA dehydrogenase, trans-enoyl-CoA reductase, alcohol dehydrogenase) including homologs, variants, fragments, related fusion proteins, or functional equivalents thereof, are used in recombinant nucleic acid molecules that direct the expression of such polypeptides in appropriate host cells, such as bacterial or yeast cells. Figures 8 through 26B provide exemplary polynucleotide sequences encoding polypeptides useful in the methods described

herein. It is understood that the addition of sequences which do not alter the encoded activity of a nucleic acid molecule, such as the addition of a non-functional or non-coding sequence, is a conservative variation of the basic nucleic acid.

[0064] It is understood that a polynucleotide described above include "genes" and that the nucleic acid molecules described above include "vectors" or "plasmids." For example, a polynucleotide encoding a keto thiolase can comprise an *atoB* gene or homolog thereof, or an *fadA* gene or homolog thereof. Accordingly, the term "gene", also called a "structural gene" refers to a polynucleotide that codes for a particular polypeptide comprising a sequence of amino acids, which comprise all or part of one or more proteins or enzymes, and may include regulatory (non-transcribed) DNA sequences, such as promoter region or expression control elements, which determine, for example, the conditions under which the gene is expressed. The transcribed region of the gene may include untranslated regions, including introns, 5'-untranslated region (UTR), and 3'-UTR, as well as the coding sequence. The term "polynucleotide," "nucleic acid" or "recombinant nucleic acid" refers to polynucleotides such as deoxyribonucleic acid (DNA), and, where appropriate, ribonucleic acid (RNA). The term "expression" with respect to a gene or polynucleotide refers to transcription of the gene or polynucleotide and, as appropriate, translation of the resulting mRNA transcript to a protein or polypeptide. Thus, as will be clear from the context, expression of a protein or polypeptide results from transcription and translation of the open reading frame.

[0065] Those of skill in the art will recognize that, due to the degenerate nature of the genetic code, a variety of codons differing in their nucleotide sequences can be used to encode a given amino acid. A particular polynucleotide or gene sequence encoding a biosynthetic enzyme or polypeptide described above are referenced herein merely to illustrate an embodiment of the disclosure, and the disclosure includes polynucleotides of any sequence that encode a polypeptide comprising the same amino acid sequence of the polypeptides and proteins of the enzymes utilized in the methods of the disclosure. In similar fashion, a polypeptide can typically

tolerate one or more amino acid substitutions, deletions, and insertions in its amino acid sequence without loss or significant loss of a desired activity. The disclosure includes such polypeptides with alternate amino acid sequences, and the amino acid sequences encoded by the DNA sequences shown herein merely illustrate preferred embodiments of the disclosure.

[0066] The disclosure provides polynucleotides in the form of recombinant DNA expression vectors or plasmids, as described in more detail elsewhere herein, that encode one or more target enzymes. Generally, such vectors can either replicate in the cytoplasm of the host microorganism or integrate into the chromosomal DNA of the host microorganism. In either case, the vector can be a stable vector (*i.e.*, the vector remains present over many cell divisions, even if only with selective pressure) or a transient vector (*i.e.*, the vector is gradually lost by host microorganisms with increasing numbers of cell divisions). The disclosure provides DNA molecules in isolated (*i.e.*, not pure, but existing in a preparation in an abundance and/or concentration not found in nature) and purified (*i.e.*, substantially free of contaminating materials or substantially free of materials with which the corresponding DNA would be found in nature) form.

[0067] A polynucleotide of the disclosure can be amplified using cDNA, mRNA or alternatively, genomic DNA, as a template and appropriate oligonucleotide primers according to standard PCR amplification techniques and those procedures described in the Examples section below. The nucleic acid so amplified can be cloned into an appropriate vector and characterized by DNA sequence analysis. Furthermore, oligonucleotides corresponding to nucleotide sequences can be prepared by standard synthetic techniques, *e.g.*, using an automated DNA synthesizer.

[0068] It is also understood that an isolated polynucleotide molecule encoding a polypeptide homologous to the enzymes described herein can be created by introducing one or more nucleotide substitutions, additions or deletions into the nucleotide sequence encoding the particular polypeptide, such that one or more amino acid substitutions, additions or deletions are introduced into the encoded protein. Mutations can be introduced into the

polynucleotide by standard techniques, such as site-directed mutagenesis and PCR-mediated mutagenesis. In contrast to those positions where it may be desirable to make a non-conservative amino acid substitutions (see above), in some positions it is preferable to make conservative amino acid substitutions.

[0069] As will be understood by those of skill in the art, it can be advantageous to modify a coding sequence to enhance its expression in a particular host. The genetic code is redundant with 64 possible codons, but most organisms typically use a subset of these codons. The codons that are utilized most often in a species are called optimal codons, and those not utilized very often are classified as rare or low-usage codons. Codons can be substituted to reflect the preferred codon usage of the host, a process sometimes called "codon optimization" or "controlling for species codon bias."

[0070] Optimized coding sequences containing codons preferred by a particular prokaryotic or eukaryotic host (see also, Murray *et al.* (1989) Nucl. Acids Res. 17:477-508) can be prepared, for example, to increase the rate of translation or to produce recombinant RNA transcripts having desirable properties, such as a longer half-life, as compared with transcripts produced from a non-optimized sequence. Translation stop codons can also be modified to reflect host preference. For example, typical stop codons for *S. cerevisiae* and mammals are UAA and UGA, respectively. The typical stop codon for monocotyledonous plants is UGA, whereas insects and *E. coli* commonly use UAA as the stop codon (Dalphin *et al.* (1996) Nucl. Acids Res. 24: 216-218). Methodology for optimizing a nucleotide sequence for expression in a plant is provided, for example, in U.S. Pat. No. 6,015,891, and the references cited therein.

[0071] The term "recombinant microorganism" and "recombinant host cell" are used interchangeably herein and refer to microorganisms that have been genetically modified to express or over-express endogenous polynucleotides, or to express non-endogenous sequences, such as those included in a vector. The polynucleotide generally encodes a target enzyme involved in a metabolic pathway for producing a desired metabolite as described above, but may also include protein factors necessary for regulation

or activity or transcription. Accordingly, recombinant microorganisms described herein have been genetically engineered to express or over-express target enzymes not previously expressed or over-expressed by a parental microorganism. It is understood that the terms "recombinant microorganism" and "recombinant host cell" refer not only to the particular recombinant microorganism but to the progeny or potential progeny of such a microorganism.

[0072] The term "substrate" or "suitable substrate" refers to any substance or compound that is converted or meant to be converted into another compound by the action of an enzyme. The term includes not only a single compound, but also combinations of compounds, such as solutions, mixtures and other materials which contain at least one substrate, or derivatives thereof. Further, the term "substrate" encompasses not only compounds that provide a carbon source suitable for use as a starting material, but also intermediate and end product metabolites used in a pathway associated with a metabolically engineered microorganism as described herein.

[0073] "Transformation" refers to the process by which a vector is introduced into a host cell. Transformation (or transduction, or transfection), can be achieved by any one of a number of means including electroporation, microinjection, biolistics (or particle bombardment-mediated delivery), or agrobacterium mediated transformation.

[0074] A "vector" generally refers to a polynucleotide that can be propagated and/or transferred between organisms, cells, or cellular components. Vectors include viruses, bacteriophage, proviruses, plasmids, phagemids, transposons, and artificial chromosomes such as YACs (yeast artificial chromosomes), BACs (bacterial artificial chromosomes), and PLACs (plant artificial chromosomes), and the like, that are "episomes," that is, that replicate autonomously or can integrate into a chromosome of a host cell. A vector can also be a naked RNA polynucleotide, a naked DNA polynucleotide, a polynucleotide composed of both DNA and RNA within the same strand, a poly-lysine -conjugated DNA or RNA, a peptide-conjugated DNA or RNA, a liposome-conjugated DNA, or the like, that are not episomal in nature, or it can be an organism which comprises

one or more of the above polynucleotide constructs such as an agrobacterium or a bacterium.

[0075] The various components of an expression vector can vary widely, depending on the intended use of the vector and the host cell(s) in which the vector is intended to replicate or drive expression. Expression vector components suitable for the expression of genes and maintenance of vectors in *E. coli*, yeast, *Streptomyces*, and other commonly used cells are widely known and commercially available. For example, suitable promoters for inclusion in the expression vectors of the disclosure include those that function in eukaryotic or prokaryotic host microorganisms. Promoters can comprise regulatory sequences that allow for regulation of expression relative to the growth of the host microorganism or that cause the expression of a gene to be turned on or off in response to a chemical or physical stimulus. For *E. coli* and certain other bacterial host cells, promoters derived from genes for biosynthetic enzymes, antibiotic-resistance conferring enzymes, and phage proteins can be used and include, for example, the galactose, lactose (*lac*), maltose, tryptophan (*trp*), beta-lactamase (*bla*), bacteriophage lambda PL, and T5 promoters. In addition, synthetic promoters, such as the *tac* promoter (U.S. Pat. No. 4,551,433, which is incorporated herein by reference in its entirety), can also be used. For *E. coli* expression vectors, it is useful to include an *E. coli* origin of replication, such as from pUC, p1P, p1, and pBR.

[0076] Thus, recombinant expression vectors contain at least one expression system, which, in turn, is composed of at least a portion of a gene coding sequences operably linked to a promoter and optionally termination sequences that operate to effect expression of the coding sequence in compatible host cells. The host cells are modified by transformation with the recombinant DNA expression vectors of the disclosure to contain the expression system sequences either as extrachromosomal elements or integrated into the chromosome.

[0077] The disclosure provides methods for the heterologous expression of one or more of the biosynthetic genes or polynucleotides involved in n-hexanol and n-octanol biosynthesis and recombinant DNA expression vectors useful in the method. Thus,

included within the scope of the disclosure are recombinant expression vectors that include such nucleic acids.

[0078] Recombinant microorganisms provided herein can express a plurality of target enzymes involved in pathways for the production of n-hexanol and/or n-octanol from a suitable carbon substrate such as, for example, glucose.

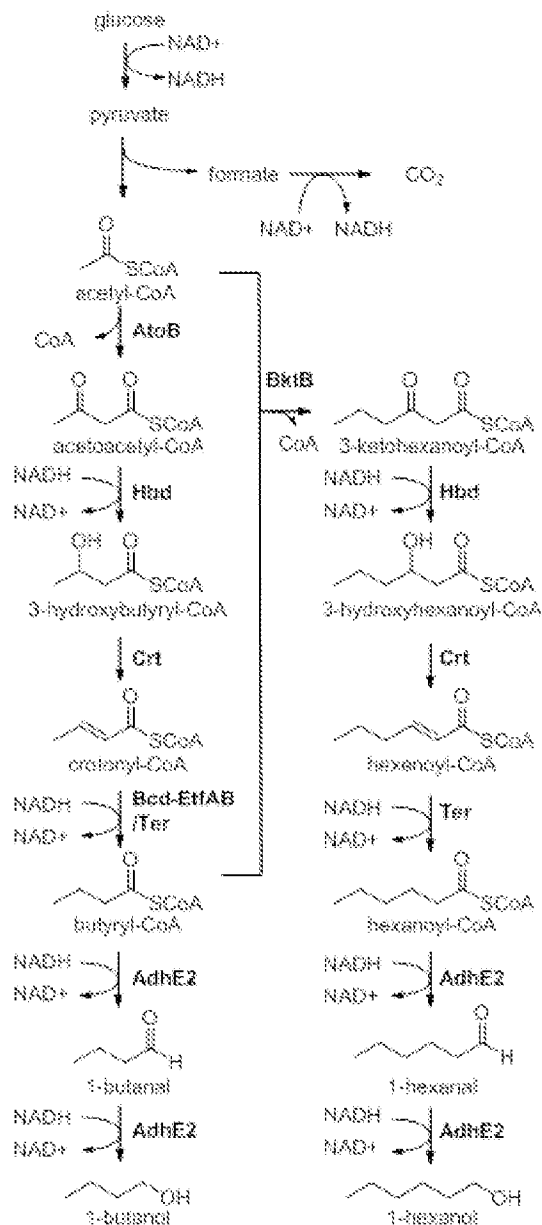
[0079] The disclosure demonstrates that the expression or overexpression of one or more heterologous polynucleotide or overexpression of one or more native polynucleotides encoding (i) a polypeptide that catalyzes the production of acetoacetyl-coA from two molecules of acetyl-coA; (ii) a polypeptide that catalyzes the conversion of acetoacetyl-coA to 3-hydroxybutyryl-CoA; (iii) a polypeptide that catalyzes the conversion of 3-hydroxybutyryl-CoA to crotonyl-CoA; (iv) a polypeptide (or polypeptide combination) that catalyzes the reduction of crotonyl-CoA to butyryl-CoA; (v) a polypeptide that catalyzes the conversion of butyryl-CoA to 3-ketohexanoyl-CoA; (vi) a polypeptide that converts 3-ketohexanoyl-CoA to 3-hydroxy-hexanoyl-CoA; (vii) a polypeptide that converts 3-hydroxy-hexanoyl-CoA to trans-2-hexenoyl-CoA; (viii) a polypeptide that converts trans-2-hexenoyl-CoA to hexanoyl-CoA; and (ix) a polypeptide that catalyzes the conversion of hexanoyl-CoA to 1-hexanal and n-hexanol. For example, the disclosure demonstrates that with expression of the heterologous *atoB* or *thl*, *hbd* or *Paah1*, *crt*, *Ter*, *BktB*, and *adhE2* genes in *Escherichia* (e.g., *E. coli*) the production of n-hexanol and n-octanol can be obtained.

[0080] Microorganisms provided herein are modified to produce metabolites in quantities not available in the parental microorganism.

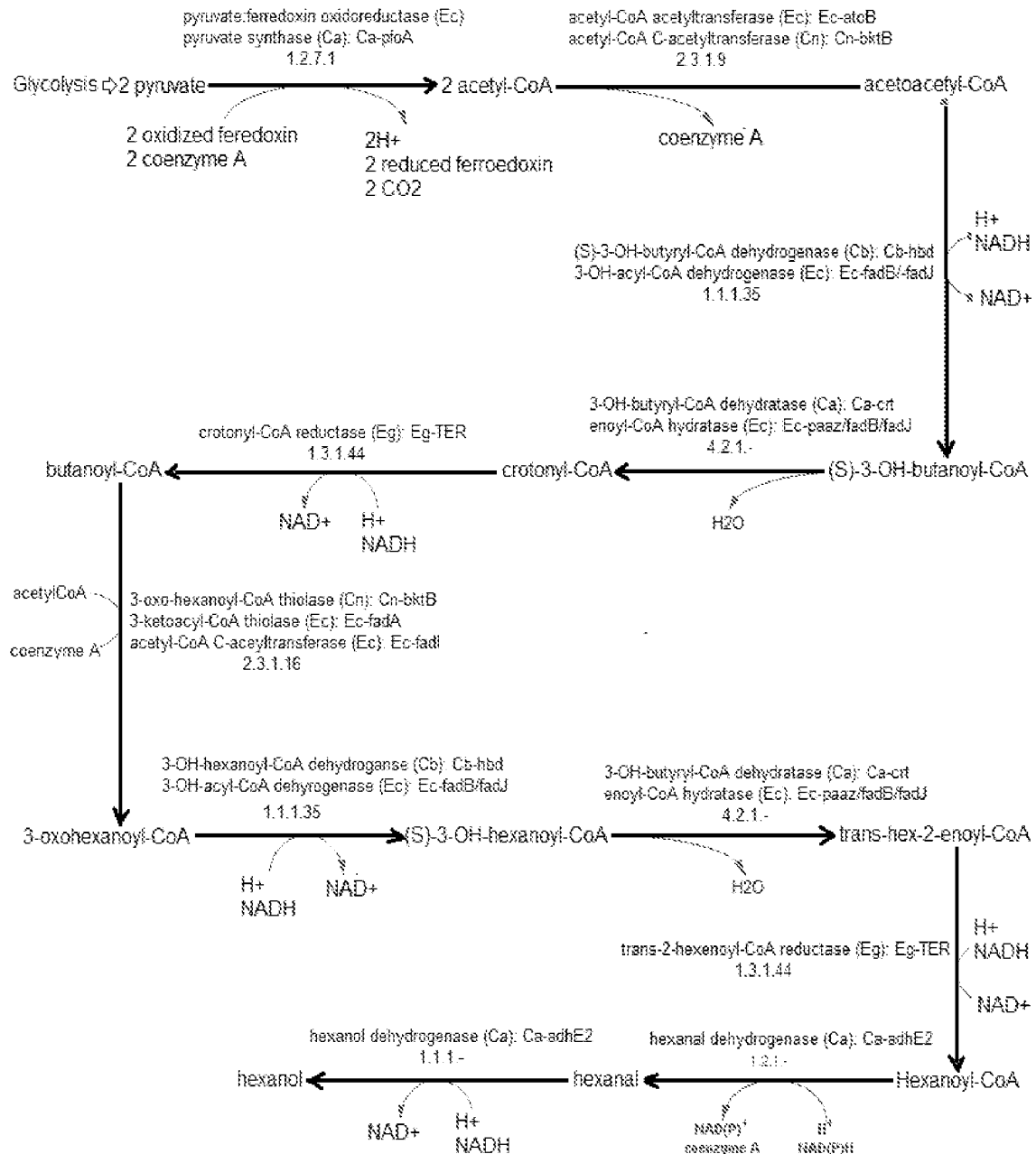
[0081] Accordingly, the disclosure provides a recombinant microorganism that produces n-hexanol and/or n-octanol and includes the expression or elevated expression of target enzymes such as an acetyl-coA acetyl transferase (e.g., *atoB*), a 3-hydroxybutyryl-coA dehydrogenase (e.g., *hbd* or *Paah1*), a crotonase (e.g., *crt*), trans-2-enoyl-CoA reductase (*Ter*), beta-ketothiolase (*BktB*) and an aldehyde/alcohol dehydrogenase (e.g., *adhE2*), or any combination thereof, as compared to a parental microorganism. In addition, the microorganism may include a disruption, deletion or knockout of

expression of an alcohol/acetoaldehyde dehydrogenase that preferentially uses acetyl-coA as a substrate (e.g. *adhE* gene), as compared to a parental microorganism. In some embodiments, further knockouts may include knockouts in a lactate dehydrogenase (e.g., *ldh*) and *frdBC*. It will be recognized that organism that inherently have one or more (but not all) of the foregoing enzymes can be utilized as a parental organism.

[0082] Accordingly, a recombinant microorganism provided herein includes the elevated expression of at least one target enzyme, such as AtoB or BktB. In other embodiments, a recombinant microorganism can express a plurality of target enzymes involved in a pathway to produce n-hexanol or n-octanol as depicted in Scheme 1 and 2 from a sugar intermediate. The plurality of enzymes can include keto thiolase, acetyl-CoA acetyltransferase, hydroxybutyryl CoA dehydrogenase, crotonase, trans-2-enoyl-CoA reductase, beta-ketothiolase and alcohol dehydrogenase (ADHE2), or any combination thereof.



Scheme 1. Artificial pathway for *n*-hexanol biosynthesis in *Escherichia coli*. Abbreviations: AtoB, acetyl-CoA acetyltransferase; Hbd, 3-hydroxybutyryl-CoA dehydrogenase; Crt, crotonase; Bcd-EtfAB, butyryl-CoA dehydrogenase-electron transfer protein; Ter, trans-enoyl-CoA reductase, AdhE2; alcohol/aldehyde dehydrogenase, BktB; β -ketothiolase



Scheme 2

[0083] As previously noted, the target enzymes described throughout this disclosure generally produce metabolites. In addition, the target enzymes described throughout this disclosure are encoded by polynucleotides. For example, an acetyl-CoA acetyltransferase can be encoded by an *atoB* gene, polynucleotide or homolog thereof. The *atoB* gene can be derived from any biologic source that provides a suitable nucleic acid sequence encoding a suitable enzyme having acetyl-CoA acetyltransferase activity.

[0084] Accordingly, in one embodiment, a recombinant microorganism provided herein includes elevated expression of an acetyl-CoA acetyltransferase as compared to a parental microorganism. This expression may be combined with the expression or over-expression with other enzymes in the metabolic pathway for the production of n-hexanol and/or n-octanol as described herein below. The recombinant microorganism produces a metabolite that includes an acetoacetyl-CoA from a substrate that includes 2 acetyl-CoA molecules. The acetyl-CoA acetyltransferase can be encoded by an *atoB* gene, polynucleotide or homolog thereof. For example, an *atoB* gene can be derived from *E. coli* or *C. acetobutylicum*.

[0085] In another embodiment, a recombinant microorganism provided herein includes elevated expression of a hydroxybutyryl CoA dehydrogenase as compared to a parental microorganism. This expression may be combined with the expression or over-expression with other enzymes in the metabolic pathway for the production of n-hexanol and/or n-octanol as described herein above and below. The recombinant microorganism produces a metabolite that includes a 3-hydroxybutyryl-CoA from a substrate that includes acetoacetyl-CoA. The microorganism can also produce 3-hydroxy-hexanoyl-CoA from 3-keto-hexanoyl-CoA using hydroxybutyryl CoA dehydrogenase. The hydroxybutyryl CoA dehydrogenase can be encoded by an *hbd* gene, polynucleotide or homolog thereof. The *hbd* gene can be derived from various microorganisms including *Clostridium acetobutylicum*, *Clostridium difficile*, *Dastricha ruminatum*, *Butyrivibrio fibrisolvens*, *Treponema phagedemes*, *Acidaminococcus fermentans*, *Clostridium kluyveri*, *Syntrophospora bryanti*, and *Thermoanaerobacterium thermosaccharolyticum*. In other embodiments, a 3-hydroxy-acyl-coA reductase (PaaH1) from *R. eutropha*, or homolog or variant thereof) can be used. Such variants can include those having improved activity compared to a wild-type PaaH1 such as those set forth in Figure 26 (see below).

[0086] In another embodiment, a recombinant microorganism provided herein includes elevated expression of crotonase as compared to a parental microorganism. This expression may be combined with the expression or over-expression with other enzymes in the metabolic pathway for the production of n-hexanol and/or n-

octanol as described herein above and below. The recombinant microorganism produces a metabolite that includes crotonyl-CoA from a substrate that includes 3-hydroxybutyryl-CoA. The microorganism can also produce trans-2-hexenoyl-CoA from 3-hydroxyhexanoyl-CoA using the crotonase. The crotonase can be encoded by a *crt* gene, polynucleotide or homolog thereof. The *crt* gene or polynucleotide can be derived from various microorganisms including *Clostridium acetobutylicum*, *Butyrivibrio fibrisolvens*, *Thermoanaerobacterium thermosaccharolyticum*, and *Clostridium difficile*.

[0087] In yet another embodiment, a recombinant microorganism provided herein includes elevated expression of a crotonyl-CoA reductase as compared to a parental microorganism. This expression may be combined with the expression or over-expression with other enzymes in the metabolic pathway for the production of n-hexanol and/or n-octanol as described herein above and below. The microorganism produces a metabolite that includes butyryl-CoA from a substrate that includes crotonyl-CoA. The crotonyl-CoA reductase can be encoded by a *ccr* gene, polynucleotide or homolog thereof. The *ccr* gene or polynucleotide can be derived from the genus *Streptomyces*. Alternatively, or in addition to, the microorganism provided herein includes elevated expression of a trans-2-hexenoyl-CoA reductase as compared to a parental microorganism. The microorganism produces a metabolite that includes butyryl-CoA from a substrate that includes crotonyl-CoA. The trans-2-hexenoyl-CoA reductase can also convert trans-2-hexenoyl-CoA to hexanoyl-CoA. The trans-2-hexenoyl-CoA reductase can be encoded by a *ter* gene, polynucleotide or homolog thereof. The *ter* gene or polynucleotide can be derived from the genus *Euglena*. The *ter* gene or polynucleotide can be derived from *Treponema denticola*. The enzyme from *Euglena gracilis* acts on crotonoyl-CoA and, more slowly, on trans-hex-2-enoyl-CoA and trans-oct-2-enoyl-CoA.

[0088] In yet another embodiment, a recombinant microorganism provided herein includes elevated expression of a butyryl-CoA dehydrogenase as compared to a parental microorganism. This expression may be combined with the expression or over-expression with other enzymes in the metabolic pathway for the production of n-hexanol and/or n-octanol as described herein above and below. The

recombinant microorganism produces a metabolite that includes butyryl-CoA from a substrate that includes crotonyl-CoA. The butyryl-CoA dehydrogenase can be encoded by a *bcd* gene, polynucleotide or homolog thereof. The *bcd* gene, polynucleotide can be derived from *Clostridium acetobutylicum*, *Mycobacterium tuberculosis*, or *Megasphaera elsdenii*.

[0089] In yet another embodiment, a recombinant microorganism provided herein includes elevated expression of an alcohol dehydrogenase as compared to a parental microorganism. This expression may be combined with the expression or over-expression with other enzymes in the metabolic pathway for the production of n-hexanol and/or n-octanol as described herein above and below. The recombinant microorganism produces a metabolite that includes butanol from a substrate that includes butyryl-CoA or can produce n-hexanol from hexanoyl-CoA. The alcohol dehydrogenase can be encoded by an *aad* gene, polynucleotide or homolog thereof, or an *adhE2* gene, polynucleotide or homolog thereof. The *aad* gene or *adhE2* gene or polynucleotide can be derived from *Clostridium acetobutylicum*. In some embodiments, the alcohol dehydrogenase is an NADH-dependent alcohol dehydrogenase.

[0090] In yet another embodiment, a recombinant microorganism provided herein includes expression of a beta-ketothiolase (*bktB*) as compared to a parental microorganism. This expression may be combined with the expression or over-expression with other enzymes in the metabolic pathway for the production of n-hexanol and/or n-octanol as described herein above and below. The recombinant microorganism produces a metabolite that includes 3-ketohexanoyl-CoA from a substrate that includes butyryl-CoA and acetyl-CoA. The beta-ketothiolase can be encoded by a *bktB* gene, polynucleotide or homolog thereof. The *bktB* gene or polynucleotide can be derived from *R. eutropha*.

[0091] In addition, homologs of enzymes useful for generating metabolites are encompassed by the microorganisms and methods provided herein. The term "homologs" used with respect to an original enzyme or gene of a first family or species refers to distinct enzymes or genes of a second family or species which are determined by functional, structural or genomic analyses to be an

enzyme or gene of the second family or species which corresponds to the original enzyme or gene of the first family or species. Most often, homologs will have functional, structural or genomic similarities. Techniques are known by which homologs of an enzyme or gene can readily be cloned using genetic probes and PCR. Identity of cloned sequences as homolog can be confirmed using functional assays and/or by genomic mapping of the genes.

[0092] A protein has "homology" or is "homologous" to a second protein if the nucleic acid sequence that encodes the protein has a similar sequence to the nucleic acid sequence that encodes the second protein. Alternatively, a protein has homology to a second protein if the two proteins have "similar" amino acid sequences. (Thus, the term "homologous proteins" is defined to mean that the two proteins have similar amino acid sequences).

[0093] As used herein, two proteins (or a region of the proteins) are substantially homologous when the amino acid sequences have at least about 30%, 40%, 50%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identity. To determine the percent identity of two amino acid sequences, or of two nucleic acid sequences, the sequences are aligned for optimal comparison purposes (e.g., gaps can be introduced in one or both of a first and a second amino acid or nucleic acid sequence for optimal alignment and non-homologous sequences can be disregarded for comparison purposes). In one embodiment, the length of a reference sequence aligned for comparison purposes is at least 30%, typically at least 40%, more typically at least 50%, even more typically at least 60%, and even more typically at least 70%, 80%, 90%, 100% of the length of the reference sequence. The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are identical at that position (as used herein amino acid or nucleic acid "identity" is equivalent to amino acid or nucleic acid "homology"). The percent identity between the two sequences is a function of the number of identical positions shared by the sequences, taking into account the number of gaps, and the length of

each gap, which need to be introduced for optimal alignment of the two sequences.

[0094] When "homologous" is used in reference to proteins or peptides, it is recognized that residue positions that are not identical often differ by conservative amino acid substitutions. A "conservative amino acid substitution" is one in which an amino acid residue is substituted by another amino acid residue having a side chain (R group) with similar chemical properties (e.g., charge or hydrophobicity). In general, a conservative amino acid substitution will not substantially change the functional properties of a protein. In cases where two or more amino acid sequences differ from each other by conservative substitutions, the percent sequence identity or degree of homology may be adjusted upwards to correct for the conservative nature of the substitution. Means for making this adjustment are well known to those of skill in the art (see, e.g., Pearson *et al.*, 1994, hereby incorporated herein by reference).

[0095] A "conservative amino acid substitution" is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art. These families include amino acids with basic side chains (e.g., lysine, arginine, histidine), acidic side chains (e.g., aspartic acid, glutamic acid), uncharged polar side chains (e.g., glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (e.g., alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (e.g., threonine, valine, isoleucine) and aromatic side chains (e.g., tyrosine, phenylalanine, tryptophan, histidine). The following six groups each contain amino acids that are conservative substitutions for one another: 1) Serine (S), Threonine (T); 2) Aspartic Acid (D), Glutamic Acid (E); 3) Asparagine (N), Glutamine (Q); 4) Arginine (R), Lysine (K); 5) Isoleucine (I), Leucine (L), Methionine (M), Alanine (A), Valine (V), and 6) Phenylalanine (F), Tyrosine (Y), Tryptophan (W).

[0096] Sequence homology for polypeptides, which can also be referred to as percent sequence identity, is typically measured

using sequence analysis software. See, e.g., the Sequence Analysis Software Package of the Genetics Computer Group (GCG), University of Wisconsin Biotechnology Center, 910 University Avenue, Madison, Wis. 53705. Protein analysis software matches similar sequences using measure of homology assigned to various substitutions, deletions and other modifications, including conservative amino acid substitutions. For instance, GCG contains programs such as "Gap" and "Bestfit" which can be used with default parameters to determine sequence homology or sequence identity between closely related polypeptides, such as homologous polypeptides from different species of organisms or between a wild type protein and a mutein thereof. See, e.g., GCG Version 6.1.

[0097] A typical algorithm used comparing a molecule sequence to a database containing a large number of sequences from different organisms is the computer program BLAST (Altschul, 1990; Gish, 1993; Madden, 1996; Altschul, 1997; Zhang, 1997), especially blastp or tblastn (Altschul, 1997). Typical parameters for BLASTp are: Expectation value: 10 (default); Filter: seg (default); Cost to open a gap: 11 (default); Cost to extend a gap: 1 (default); Max. alignments: 100 (default); Word size: 11 (default); No. of descriptions: 100 (default); Penalty Matrix: BLOWSUM62.

[0098] When searching a database containing sequences from a large number of different organisms, it is typical to compare amino acid sequences. Database searching using amino acid sequences can be measured by algorithms other than blastp known in the art. For instance, polypeptide sequences can be compared using FASTA, a program in GCG Version 6.1. FASTA provides alignments and percent sequence identity of the regions of the best overlap between the query and search sequences (Pearson, 1990, hereby incorporated herein by reference). For example, percent sequence identity between amino acid sequences can be determined using FASTA with its default parameters (a word size of 2 and the PAM250 scoring matrix), as provided in GCG Version 6.1, hereby incorporated herein by reference.

[0099] The following table and the disclosure provides non-limiting examples of genes and homologs. In addition, the disclosure provides various sequences and further provides accession

numbers for variants and homologs of genes useful in the methods and compositions of the disclosure. One of skill in the art can readily identify the sequences from the accession numbers, clone and PCR such sequences and additional homologs using techniques known in the art.

Table 1

Enzyme	Exemplary Gene (s)	Exemplary Organism
acetyl-coA acetyltransferase	atoB	<i>C. acetobutylicum</i>
acetoacetyl-coA thiolase	thl, thlA, thlB	<i>E. coli</i> , <i>C. acetobutylicum</i>
3-hydroxybutyryl-CoA dehydrogenase	Hbd	<i>C. acetobutylicum</i>
crotonase	Crt	<i>C. acetobutylicum</i>
butyryl-CoA dehydrogenase	Bcd	<i>C. acetobutylicum</i> , <i>M. elsdenii</i>
electron transfer flavoprotein	etfAB	<i>C. acetobutylicum</i> , <i>M. elsdenii</i>
aldehyde/alcohol dehydrogenase	adhE2	<i>C. acetobutylicum</i>
crotonyl-coA reductase	Ccr	<i>S. coelicolor</i>
trans-2-enoyl-CoA reductase	Ter	<i>T. denticola</i> , <i>F. succinogenes</i>
Bet-ketothiolase	BktB	<i>R. eutropha</i> , <i>C. taiwanensis</i>

[00100] The disclosure provides accession numbers for various genes, homologs and variants useful in the generation of recombinant microorganism described herein. It is to be understood that homologs and variants described herein are exemplary and non-limiting. Additional homologs, variants and sequences are available to those of skill in the art using various databases including, for example, the National Center for Biotechnology Information (NCBI) access to which is available on the World-Wide-Web.

[00101] Trans-2-enoyl-CoA reductase or TER is a protein that is capable of catalyzing the conversion of crotonyl-CoA to butyryl-CoA, and trans-2-hexenoyl-CoA to hexanoyl-CoA. In certain embodiments, the recombinant microorganism expresses a TER which catalyzes the same reaction as Bcd/EtfA/EtfB from Clostridia and other bacterial species. Mitochondrial TER from *E. gracilis* has been described, and many TER proteins and proteins with TER activity derived from a number of species have been identified forming a TER protein family (see, e.g., U.S. Pat. Appl. 2007/0022497 to Cirpus et al.; and Hoffmeister et al., J. Biol. Chem., 280:4329-4338, 2005, both of

which are incorporated herein by reference in their entirety). A truncated cDNA of the *E. gracilis* gene has been functionally expressed in *E. coli*. This cDNA or the genes of homologues from other microorganisms can be expressed together with *atoB*, *hbd* (or *paaH1*), *bktB*, *crt*, and *adhE2* to produce n-hexanol and/or octanol in *E. coli*, *S. cerevisiae* or other hosts.

[00102] TER proteins can also be identified by generally well known bioinformatics methods, such as BLAST. Examples of TER proteins include, but are not limited to, TERs from species such as: *Euglena* spp. including, but not limited to, *E. gracilis*, *Aeromonas* spp. including, but not limited, to *A. hydrophila*, *Psychromonas* spp. including, but not limited to, *P. ingrahamii*, *Photobacterium* spp. including, but not limited, to *P. profundum*, *Vibrio* spp. including, but not limited, to *V. angustum*, *V. cholerae*, *V. alginolyticus*, *V. parahaemolyticus*, *V. vulnificus*, *V. fischeri*, *V. splendidus*, *Shewanella* spp. including, but not limited to, *S. amazonensis*, *S. woodyi*, *S. frigidimarina*, *S. paeleana*, *S. baltica*, *S. denitrificans*, *Oceanospirillum* spp., *Xanthomonas* spp. including, but not limited to, *X. oryzae*, *X. campestris*, *Chromohalobacter* spp. including, but not limited, to *C. salexigens*, *Idiomarina* spp. including, but not limited, to *I. baltica*, *Pseudoalteromonas* spp. including, but not limited to, *P. atlantica*, *Alteromonas* spp., *Saccharophagus* spp. including, but not limited to, *S. degradans*, *S. marine gamma proteobacterium*, *S. alpha proteobacterium*, *Pseudomonas* spp. including, but not limited to, *P. aeruginosa*, *P. putida*, *P. fluorescens*, *Burkholderia* spp. including, but not limited to, *B. phytofirmans*, *B. cenocepacia*, *B. cepacia*, *B. ambifaria*, *B. vietnamensis*, *B. multivorans*, *B. dolosa*, *Methylobacillus* spp. including, but not limited to, *M. flageliatus*, *Stenotrophomonas* spp. including, but not limited to, *S. maltophilia*, *Congregibacter* spp. including, but not limited to, *C. litoralis*, *Serratia* spp. including, but not limited to, *S. proteamaculans*, *Marinomonas* spp., *Xytella* spp. including, but not limited to, *X. fastidiosa*, *Reinekea* spp., *Colweffia* spp. including, but not limited to, *C. psychrerythraea*, *Yersinia* spp. including, but not limited to, *Y. pestis*, *Y. pseudotuberculosis*, *Methylobacillus* spp. including, but not limited to, *M. flagellatus*, *Cytophaga* spp. including, but not

limited to, *C. hutchinsonii*, *Flavobacterium* spp. including, but not limited to, *F. johnsoniae*, *Microscilla* spp. including, but not limited to, *M. marina*, *Polaribacter* spp. including, but not limited to, *P. irgensii*, *Clostridium* spp. including, but not limited to, *C. acetobutylicum*, *C. beijerenckii*, *C. cellulolyticum*, *Coxiella* spp. including, but not limited to, *C. burnetii*.

[00103] In addition to the foregoing, the terms "trans-2-enoyl-CoA reductase" or "TER" refer to proteins that are capable of catalyzing the conversion of crotonyl-CoA to butyryl-CoA, or trans-2-hexenoyl-CoA to hexanoyl-CoA and which share at least about 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or greater sequence identity, or at least about 50%, 60%, 70%, 80%, 90%, 95%, 96%, 97%, 98%, 99% or greater sequence similarity, as calculated by NCBI BLAST, using default parameters, to either or both of the truncated *E. gracilis* TER or the full length *A. hydrophila* TER.

[00104] In a specific embodiment, the trans-2-enoyl-CoA reductase is encoded in *T. denticola* *F.succinogens*, *T. vincentii* or *F. johnsoniae ter* gene. In *T. denticola* TER has the accession number Q73Q47 (see also Figure 26A). In one embodiment the *F. succinogens* Ter comprises the sequence set forth in Figure 26A and has a Met11Lys mutation. Other Ter polypeptides are set forth in Figure 26A.

[00105] Beta-ketothiolase enzymes (bktB) catalyzing the formation of beta-ketovalerate from acetyl-CoA and propionyl-CoA can also catalyze the formation of 3-oxopimeloyl-CoA and 3-ketohexanoyl-CoA. For example, the *Ralstonia eutropha* BktB and PhbB genes catalyze the condensation of butyryl-CoA and acetyl-CoA to form beta-ketohexanoyl-CoA and the reduction of beta-keto-hexanoyl-CoA to 3-hydroxy-hexanoyl-CoA (Fukui et al., *Biomacromolecules* 3:618-624 (2002)). The protein sequences for exemplary gene products can be found using the following GenBank accession numbers: bktB YP_002005382.1 *Cupriavidus taiwanensis*; bktB YP_725948 and YP_295567 *R. eutropha*.

[00106] In addition to the foregoing, the terms "beta ketothiolase" or "BktB" refer to proteins that are capable of catalyzing the conversion of beta-ketovalerate from acetyl-CoA and

propeionyl-CoA, and can also catalyze the formation of 3-oxopimeloyl-CoA and 3-ketohexanoyl-CoA, and which share at least about 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or greater sequence identity, or at least about 50%, 60%, 70%, 80%, 90%, 95%, 96%, 97%, 98%, 99% or greater sequence similarity, as calculated by NCBI BLAST, using default parameters, to SEQ ID NO:38.

[00107] Phosphate acetyltransferase is encoded in *E.coli* by *pta*. PTA is involved in conversion of acetate to acetyl-CoA. Specifically, PTA catalyzes the conversion of acetyl-coA to acetyl-phosphate. PTA homologs and variants are known. There are approximately 1075 bacterial phosphate acetyltransferases available on NCBI. For example, such homologs and variants include phosphate acetyltransferase Pta (*Rickettsia felis* URRWXCal2) gi|67004021|gb|AAV60947.1|(67004021); phosphate acetyltransferase (*Buchnera aphidicola* str. Cc (Cinara cedri)) gi|116256910|gb|ABJ90592.1|(116256910); pta (*Buchnera aphidicola* str. Cc (Cinara cedri)) gi|116515056|ref|YP_802685.1|(116515056); pta (*Wigglesworthia glossinidia endosymbiont* of *Glossina brevipalpis*) gi|25166135|dbj|BAC24326.1|(25166135); Pta (*Pasteurella multocida* subsp. *multocida* str. Pm70) gi|12720993|gb|AAK02789.1|(12720993); Pta (*Rhodospirillum rubrum*) gi|25989720|gb|AAN75024.1|(25989720); pta (*Listeria welshimeri* serovar 6b str. SLCC5334) gi|116742418|emb|CAK21542.1|(116742418); Pta (*Mycobacterium avium* subsp. *paratuberculosis* K-10) gi|41398816|gb|AAS06435.1|(41398816); phosphate acetyltransferase (pta) (*Borrelia burgdorferi* B31) gi|15594934|ref|NP_212723.1|(15594934); phosphate acetyltransferase (pta) (*Borrelia burgdorferi* B31) gi|2688508|gb|AAB91518.1|(2688508); phosphate acetyltransferase (pta) (*Haemophilus influenzae* Rd KW20) gi|1574131|gb|AAC22857.1|(1574131); Phosphate acetyltransferase Pta (*Rickettsia bellii* RML369-C) gi|91206026|ref|YP_538381.1|(91206026); Phosphate acetyltransferase Pta (*Rickettsia bellii* RML369-C) gi|91206025|ref|YP_538380.1|(91206025); phosphate acetyltransferase pta (*Mycobacterium tuberculosis* F11) gi|148720131|gb|ABR04756.1|(148720131); phosphate acetyltransferase pta (*Mycobacterium tuberculosis* str. Haarlem)

gi|134148886|gb|EBA40931.1|(134148886); phosphate acetyltransferase pta (*Mycobacterium tuberculosis* C)
 gi|124599819|gb|EAY58829.1|(124599819); Phosphate acetyltransferase Pta (*Rickettsia bellii* RML369-C)
 gi|91069570|gb|ABE05292.1|(91069570); Phosphate acetyltransferase Pta (*Rickettsia bellii* RML369-C)
 gi|91069569|gb|ABE05291.1|(91069569); phosphate acetyltransferase (pta) (*Treponema pallidum* subsp. *pallidum* str. Nichols)
 gi|15639088|ref|NP_218534.1|(15639088); and phosphate acetyltransferase (pta) (*Treponema pallidum* subsp. *pallidum* str. Nichols) gi|3322356|gb|AAC65090.1|(3322356), each sequence associated with the accession number is incorporated herein by reference in its entirety.

[00108] Pyruvate-formate lyase (Formate acetyltransferase) is an enzyme that catalyzes the conversion of pyruvate to acetyl-coA and formate. It is induced by pfl-activating enzyme under anaerobic conditions by generation of an organic free radical and decreases significantly during phosphate limitation. Formate acetyltransferase is encoded in *E.coli* by *pflB*. PFLB homologs and variants are known. For examples, such homologs and variants include, for example, Formate acetyltransferase 1 (Pyruvate formate-lyase 1) gi|129879|sp|P09373.2|PFLB_ECOLI(129879); formate acetyltransferase 1 (*Yersinia pestis* CO92)
 gi|16121663|ref|NP_404976.1|(16121663); formate acetyltransferase 1 (*Yersinia pseudotuberculosis* IP 32953)
 gi|51595748|ref|YP_069939.1|(51595748); formate acetyltransferase 1 (*Yersinia pestis* biovar Microtus str. 91001)
 gi|45441037|ref|NP_992576.1|(45441037); formate acetyltransferase 1 (*Yersinia pestis* CO92) gi|115347142|emb|CAL20035.1|(115347142); formate acetyltransferase 1 (*Yersinia pestis* biovar Microtus str. 91001) gi|45435896|gb|AAS61453.1|(45435896); formate acetyltransferase 1 (*Yersinia pseudotuberculosis* IP 32953)
 gi|51589030|emb|CAH20648.1|(51589030); formate acetyltransferase 1 (*Salmonella enterica* subsp. *enterica* serovar Typhi str. CT18)
 gi|16759843|ref|NP_455460.1|(16759843); formate acetyltransferase 1 (*Salmonella enterica* subsp. *enterica* serovar Paratyphi A str. ATCC 9150) gi|56413977|ref|YP_151052.1|(56413977); formate

acetyltransferase 1 (*Salmonella enterica* subsp. *enterica* serovar Typhi) gi|16502136|emb|CAD05373.1|(16502136); formate
acetyltransferase 1 (*Salmonella enterica* subsp. *enterica* serovar Paratyphi A str. ATCC 9150) gi|56128234|gb|AAV77740.1|(56128234);
formate acetyltransferase 1 (*Shigella dysenteriae* Sd197) gi|82777577|ref|YP_403926.1|(82777577); formate acetyltransferase 1 (*Shigella flexneri* 2a str. 2457T) gi|30062438|ref|NP_836609.1|(30062438); formate acetyltransferase 1 (*Shigella flexneri* 2a str. 2457T) gi|30040684|gb|AAP16415.1|(30040684); formate acetyltransferase 1 (*Shigella flexneri* 5 str. 8401) gi|110614459|gb|ABF03126.1|(110614459); formate acetyltransferase 1 (*Shigella dysenteriae* Sd197) gi|81241725|gb|ABB62435.1|(81241725); formate acetyltransferase 1 (*Escherichia coli* O157:H7 EDL933) gi|12514066|gb|AAG55388.1|AE005279_8(12514066); formate acetyltransferase 1 (*Yersinia pestis* KIM) gi|22126668|ref|NP_670091.1|(22126668); formate acetyltransferase 1 (*Streptococcus agalactiae* A909) gi|76787667|ref|YP_330335.1|(76787667); formate acetyltransferase 1 (*Yersinia pestis* KIM) gi|21959683|gb|AAM86342.1|AE013882_3(21959683); formate acetyltransferase 1 (*Streptococcus agalactiae* A909) gi|76562724|gb|ABA45308.1|(76562724); formate acetyltransferase 1 (*Yersinia enterocolitica* subsp. *enterocolitica* 8081) gi|123441844|ref|YP_001005827.1|(123441844); formate acetyltransferase 1 (*Shigella flexneri* 5 str. 8401) gi|110804911|ref|YP_688431.1|(110804911); formate acetyltransferase 1 (*Escherichia coli* UTI89) gi|91210004|ref|YP_539990.1|(91210004); formate acetyltransferase 1 (*Shigella boydii* Sb227) gi|82544641|ref|YP_408588.1|(82544641); formate acetyltransferase 1 (*Shigella sonnei* Ss046) gi|74311459|ref|YP_309878.1|(74311459); formate acetyltransferase 1 (*Klebsiella pneumoniae* subsp. *pneumoniae* MGH 78578) gi|152969488|ref|YP_001334597.1|(152969488); formate acetyltransferase 1 (*Salmonella enterica* subsp. *enterica* serovar Typhi Ty2) gi|29142384|ref|NP_805726.1|(29142384) formate acetyltransferase 1 (*Shigella flexneri* 2a str. 301) gi|24112311|ref|NP_706821.1|(24112311); formate acetyltransferase 1

(*Escherichia coli* O157:H7 EDL933)
gi|15800764|ref|NP_286778.1|(15800764); formate acetyltransferase 1
(*Klebsiella pneumoniae* subsp. *pneumoniae* MGH 78578)
gi|150954337|gb|ABR76367.1|(150954337); formate acetyltransferase 1
(*Yersinia pestis* CA88-4125)
gi|149366640|ref|ZP_01888674.1|(149366640); formate
acetyltransferase 1 (*Yersinia pestis* CA88-4125)
gi|149291014|gb|EDM41089.1|(149291014); formate acetyltransferase 1
(*Yersinia enterocolitica* subsp. *enterocolitica* 8081)
gi|122088805|emb|CAL11611.1|(122088805); formate acetyltransferase 1
(*Shigella sonnei* Ss046) gi|73854936|gb|AAZ87643.1|(73854936);
formate acetyltransferase 1 (*Escherichia coli* UTI89)
gi|91071578|gb|ABE06459.1|(91071578); formate acetyltransferase 1
(*Salmonella enterica* subsp. *enterica* serovar Typhi Ty2)
gi|29138014|gb|AAO69575.1|(29138014); formate acetyltransferase 1
(*Shigella boydii* Sb227) gi|81246052|gb|ABB66760.1|(81246052);
formate acetyltransferase 1 (*Shigella flexneri* 2a str. 301)
gi|24051169|gb|AAN42528.1|(24051169); formate acetyltransferase 1
(*Escherichia coli* O157:H7 str. Sakai)
gi|13360445|dbj|BAB34409.1|(13360445); formate acetyltransferase 1
(*Escherichia coli* O157:H7 str. Sakai)
gi|15830240|ref|NP_309013.1|(15830240); formate acetyltransferase I
(pyruvate formate-lyase 1) (*Photorhabdus luminescens* subsp.
laumondii TTo1) gi|36784986|emb|CAE13906.1|(36784986); formate
acetyltransferase I (pyruvate formate-lyase 1) (*Photorhabdus*
luminescens subsp. *laumondii* TTo1)
gi|37525558|ref|NP_928902.1|(37525558); formate acetyltransferase
(*Staphylococcus aureus* subsp. *aureus* Mu50)
gi|14245993|dbj|BAB56388.1|(14245993); formate acetyltransferase
(*Staphylococcus aureus* subsp. *aureus* Mu50)
gi|15923216|ref|NP_370750.1|(15923216); Formate acetyltransferase
(Pyruvate formate-lyase)
gi|81706366|sp|Q7A7X6.1|PFLB_STAAN(81706366); Formate
acetyltransferase (Pyruvate formate-lyase)
gi|81782287|sp|Q99WZ7.1|PFLB_STAAM(81782287); Formate
acetyltransferase (Pyruvate formate-lyase)
gi|81704726|sp|Q7A1W9.1|PFLB_STAAW(81704726); formate

acetyltransferase (*Staphylococcus aureus* subsp. *aureus* Mu3)
gi|156720691|dbj|BAF77108.1|(156720691); formate acetyltransferase
(*Erwinia carotovora* subsp. *atroseptica* SCRI1043)
gi|50121521|ref|YP_050688.1|(50121521); formate acetyltransferase
(*Erwinia carotovora* subsp. *atroseptica* SCRI1043)
gi|49612047|emb|CAG75496.1|(49612047); formate acetyltransferase
(*Staphylococcus aureus* subsp. *aureus* str. Newman)
gi|150373174|dbj|BAF66434.1|(150373174); formate acetyltransferase
(*Shewanella oneidensis* MR-1) gi|24374439|ref|NP_718482.1|(24374439);
formate acetyltransferase (*Shewanella oneidensis* MR-1)
gi|24349015|gb|AAN55926.1|AE015730_3(24349015); formate
acetyltransferase (*Actinobacillus pleuropneumoniae* serovar 3 str.
JL03) gi|165976461|ref|YP_001652054.1|(165976461); formate
acetyltransferase (*Actinobacillus pleuropneumoniae* serovar 3 str.
JL03) gi|165876562|gb|ABY69610.1|(165876562); formate
acetyltransferase (*Staphylococcus aureus* subsp. *aureus* MW2)
gi|21203365|dbj|BAB94066.1|(21203365); formate acetyltransferase
(*Staphylococcus aureus* subsp. *aureus* N315)
gi|13700141|dbj|BAB41440.1|(13700141); formate acetyltransferase
(*Staphylococcus aureus* subsp. *aureus* str. Newman)
gi|151220374|ref|YP_001331197.1|(151220374); formate
acetyltransferase (*Staphylococcus aureus* subsp. *aureus* Mu3)
gi|156978556|ref|YP_001440815.1|(156978556); formate
acetyltransferase (*Synechococcus* sp. JA-2-3B'a(2-13))
gi|86607744|ref|YP_476506.1|(86607744); formate acetyltransferase
(*Synechococcus* sp. JA-3-3Ab) gi|86605195|ref|YP_473958.1|(86605195);
formate acetyltransferase (*Streptococcus pneumoniae* D39)
gi|116517188|ref|YP_815928.1|(116517188); formate acetyltransferase
(*Synechococcus* sp. JA-2-3B'a(2-13))
gi|86556286|gb|ABD01243.1|(86556286); formate acetyltransferase
(*Synechococcus* sp. JA-3-3Ab) gi|86553737|gb|ABC98695.1|(86553737);
formate acetyltransferase (*Clostridium novyi* NT)
gi|118134908|gb|ABK61952.1|(118134908); formate acetyltransferase
(*Staphylococcus aureus* subsp. *aureus* MRSA252)
gi|49482458|ref|YP_039682.1|(49482458); and formate
acetyltransferase (*Staphylococcus aureus* subsp. *aureus* MRSA252)
gi|49240587|emb|CAG39244.1|(49240587), each sequence associated with

the accession number is incorporated herein by reference in its entirety.

[00109] FNR transcriptional dual regulators are transcription regulators responsive to oxygen content. FNR is an anaerobic regulator that represses the expression of PDHc. Accordingly, reducing FNR will result in an increase in PDHc expression. FNR homologs and variants are known. For examples, such homologs and variants include, for example, DNA-binding transcriptional dual regulator, global regulator of anaerobic growth (*Escherichia coli* W3110) gi|1742191|dbj|BAA14927.1|(1742191); DNA-binding transcriptional dual regulator, global regulator of anaerobic growth (*Escherichia coli* K12) gi|16129295|ref|NP_415850.1|(16129295); DNA-binding transcriptional dual regulator, global regulator of anaerobic growth (*Escherichia coli* K12) gi|1787595|gb|AAC74416.1|(1787595); DNA-binding transcriptional dual regulator, global regulator of anaerobic growth (*Escherichia coli* W3110) gi|89108182|ref|AP_001962.1|(89108182); fumarate/nitrate reduction transcriptional regulator (*Escherichia coli* UTI89) gi|162138444|ref|YP_540614.2|(162138444); fumarate/nitrate reduction transcriptional regulator (*Escherichia coli* CFT073) gi|161486234|ref|NP_753709.2|(161486234); fumarate/nitrate reduction transcriptional regulator (*Escherichia coli* O157:H7 EDL933) gi|15801834|ref|NP_287852.1|(15801834); fumarate/nitrate reduction transcriptional regulator (*Escherichia coli* APEC O1) gi|117623587|ref|YP_852500.1|(117623587); fumarate and nitrate reduction regulatory protein gi|71159334|sp|P0A9E5.1|FNR_ECOLI(71159334); transcriptional regulation of aerobic, anaerobic respiration, osmotic balance (*Escherichia coli* O157:H7 EDL933) gi|12515424|gb|AAG56466.1|AE005372_11(12515424); Fumarate and nitrate reduction regulatory protein gi|71159333|sp|P0A9E6.1|FNR_ECOL6(71159333); Fumarate and nitrate reduction Regulatory protein (*Escherichia coli* CFT073) gi|26108071|gb|AAN80271.1|AE016760_130(26108071); fumarate and nitrate reduction regulatory protein (*Escherichia coli* UTI89) gi|91072202|gb|ABE07083.1|(91072202); fumarate and nitrate reduction regulatory protein (*Escherichia coli* HS)

gi|157160845|ref|YP_001458163.1|(157160845); fumarate and nitrate reduction regulatory protein (Escherichia coli E24377A)
 gi|157157974|ref|YP_001462642.1|(157157974); fumarate and nitrate reduction regulatory protein (Escherichia coli E24377A)
 gi|157080004|gb|ABV19712.1|(157080004); fumarate and nitrate reduction regulatory protein (Escherichia coli HS)
 gi|157066525|gb|ABV05780.1|(157066525); fumarate and nitrate reduction regulatory protein (Escherichia coli APEC O1)
 gi|115512711|gb|ABJ00786.1|(115512711); transcription regulator Fnr (Escherichia coli O157:H7 str. Sakai)
 gi|13361380|dbj|BAB35338.1|(13361380) DNA-binding transcriptional dual regulator (Escherichia coli K12)
 gi|16131236|ref|NP_417816.1|(16131236), to name a few, each sequence associated with the accession number is incorporated herein by reference in its entirety.

[00110] An acetoacetyl-coA thiolase (also sometimes referred to as an acetyl-coA acetyltransferase) catalyzes the production of acetoacetyl-coA from two molecules of acetyl-coA. Depending upon the organism used a heterologous acetoacetyl-coA thiolase (acetyl-coA acetyltransferase) can be engineered for expression in the organism. Alternatively a native acetoacetyl-coA thiolase (acetyl-coA acetyltransferase) can be overexpressed. Acetoacetyl-coA thiolase is encoded in *E.coli* by *thl*. Acetyl-coA acetyltransferase is encoded in *C. acetobutylicum* by *atoB*. THL and AtoB homologs and variants are known. For examples, such homologs and variants include, for example, acetyl-coa acetyltransferase (thiolase) (*Streptomyces coelicolor* A3(2))
 gi|21224359|ref|NP_630138.1|(21224359); acetyl-coa acetyltransferase (thiolase) (*Streptomyces coelicolor* A3(2))
 gi|3169041|emb|CAA19239.1|(3169041); Acetyl CoA acetyltransferase (thiolase) (*Alcanivorax borkumensis* SK2)
 gi|110834428|ref|YP_693287.1|(110834428); Acetyl CoA acetyltransferase (thiolase) (*Alcanivorax borkumensis* SK2)
 gi|110647539|emb|CAL17015.1|(110647539); acetyl CoA acetyltransferase (thiolase) (*Saccharopolyspora erythraea* NRRL 2338)
 gi|133915420|emb|CAM05533.1|(133915420); acetyl-coa acetyltransferase (thiolase) (*Saccharopolyspora erythraea* NRRL 2338)

gi|134098403|ref|YP_001104064.1|(134098403); acetyl-coA
acetyltransferase (thiolase) (Saccharopolyspora erythraea NRRL 2338)
gi|133911026|emb|CAM01139.1|(133911026); acetyl-CoA
acetyltransferase (thiolase) (Clostridium botulinum A str. ATCC
3502) gi|148290632|emb|CAL84761.1|(148290632); acetyl-CoA
acetyltransferase (thiolase) (Pseudomonas aeruginosa UCBPP-PA14)
gi|115586808|gb|ABJ12823.1|(115586808); acetyl-CoA acetyltransferase
(thiolase) (Ralstonia metallidurans CH34)
gi|93358270|gb|ABF12358.1|(93358270); acetyl-CoA acetyltransferase
(thiolase) (Ralstonia metallidurans CH34)
gi|93357190|gb|ABF11278.1|(93357190); acetyl-CoA acetyltransferase
(thiolase) (Ralstonia metallidurans CH34)
gi|93356587|gb|ABF10675.1|(93356587); acetyl-CoA acetyltransferase
(thiolase) (Ralstonia eutropha JMP134)
gi|72121949|gb|AAZ64135.1|(72121949); acetyl-CoA acetyltransferase
(thiolase) (Ralstonia eutropha
JMP134)gi|72121729|gb|AAZ63915.1|(72121729); acetyl-CoA
acetyltransferase (thiolase) (Ralstonia eutropha JMP134)
gi|72121320|gb|AAZ63506.1|(72121320); acetyl-CoA acetyltransferase
(thiolase) (Ralstonia eutropha JMP134)
gi|72121001|gb|AAZ63187.1|(72121001); acetyl-CoA acetyltransferase
(thiolase) (Escherichia coli) gi|2764832|emb|CAA66099.1|(2764832),
each sequence associated with the accession number is incorporated
herein by reference in its entirety.

[00111] In addition to the foregoing, the terms " acetoacetyl-coA
thiolase" or "atoB" refer to proteins that are capable of catalyzing
the production of acetoacetyl-coA from two molecules of acetyl-coA,
and which share at least about 40%, 45%, 50%, 55%, 60%, 65%, 70%,
75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or greater sequence
identity, or at least about 50%, 60%, 70%, 80%, 90%, 95%, 96%, 97%,
98%, 99% or greater sequence similarity, as calculated by NCBI
BLAST, using default parameters, to SEQ ID NO:30.

[00112] 3 hydroxy-butyryl-coA-dehydrogenase catalyzes the
conversion of acetoacetyl-coA to 3-hydroxybutyryl-CoA. Depending
upon the organism used a heterologous 3-hydroxy-butyryl-coA-
dehydrogenase can be engineered for expression in the organism.
Alternatllively a native 3-hydroxy-butyryl-coA-dehydrogenase can be

overexpressed. 3-hydroxy-butyryl-coA-dehydrogenase is encoded in *C.acetobuylicum* by *hbd*. HBD homologs and variants are known. For examples, such homologs and variants include, for example, 3-hydroxybutyryl-CoA dehydrogenase (*Clostridium acetobutylicum* ATCC 824) gi|15895965|ref|NP_349314.1|(15895965); 3-hydroxybutyryl-CoA dehydrogenase (*Bordetella pertussis* Tohama I) gi|33571103|emb|CAE40597.1|(33571103); 3-hydroxybutyryl-CoA dehydrogenase (*Streptomyces coelicolor* A3(2)) gi|21223745|ref|NP_629524.1|(21223745); 3-hydroxybutyryl-CoA dehydrogenase gi|1055222|gb|AAA95971.1|(1055222); 3-hydroxybutyryl-CoA dehydrogenase (*Clostridium perfringens* str. 13) gi|18311280|ref|NP_563214.1|(18311280); 3-hydroxybutyryl-CoA dehydrogenase (*Clostridium perfringens* str. 13) gi|18145963|dbj|BAB82004.1|(18145963) each sequence associated with the accession number is incorporated herein by reference in its entirety.

[00113] In addition to the foregoing, the terms "3 hydroxy-butyryl-coA-dehydrogenase" or "hbd" refer to proteins that are capable of catalyzing the conversion of acetoacetyl-coA to 3-hydroxybutyryl-CoA, and which share at least about 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or greater sequence identity, or at least about 50%, 60%, 70%, 80%, 90%, 95%, 96%, 97%, 98%, 99% or greater sequence similarity, as calculated by NCBI BLAST, using default parameters, to SEQ ID NO:36.

[00114] Crotonase catalyzes the conversion of 3-hydroxybutyryl-CoA to crotonyl-CoA. Depending upon the organism used a heterologous Crotonase can be engineered for expression in the organism. Alternatnlively a native Crotonase can be overexpressed. Crotonase is encoded in *C.acetobuylicum* by *crt*. CRT homologs and variants are known. For examples, such homologs and variants include, for example, crotonase (butyrate-producing bacterium L2-50) gi|119370267|gb|ABL68062.1|(119370267); crotonase gi|1055218|gb|AAA95967.1|(1055218); crotonase (*Clostridium perfringens* NCTC 8239) gi|168218170|ref|ZP_02643795.1|(168218170); crotonase (*Clostridium perfringens* CPE str. F4969) gi|168215036|ref|ZP_02640661.1|(168215036); crotonase (*Clostridium perfringens* E str. JGS1987)

gi|168207716|ref|ZP_02633721.1|(168207716); crotonase (Azoarcus sp. EbN1) gi|56476648|ref|YP_158237.1|(56476648); crotonase (Roseovarius sp. TM1035) gi|149203066|ref|ZP_01880037.1|(149203066); crotonase (Roseovarius sp. TM1035) gi|149143612|gb|EDM31648.1|(149143612); crotonase; 3-hydroxybutyryl-CoA dehydratase (Mesorhizobium loti MAFF303099) gi|14027492|dbj|BAB53761.1|(14027492); crotonase (Roseobacter sp. SK209-2-6) gi|126738922|ref|ZP_01754618.1|(126738922); crotonase (Roseobacter sp. SK209-2-6) gi|126720103|gb|EBA16810.1|(126720103); crotonase (Marinobacter sp. ELB17) gi|126665001|ref|ZP_01735984.1|(126665001); crotonase (Marinobacter sp. ELB17) gi|126630371|gb|EBA00986.1|(126630371); crotonase (Azoarcus sp. EbN1) gi|56312691|emb|CAI07336.1|(56312691); crotonase (Marinomonas sp. MED121) gi|86166463|gb|EAQ67729.1|(86166463); crotonase (Marinomonas sp. MED121) gi|87118829|ref|ZP_01074728.1|(87118829); crotonase (Roseovarius sp. 217) gi|85705898|ref|ZP_01036994.1|(85705898); crotonase (Roseovarius sp. 217) gi|85669486|gb|EAQ24351.1|(85669486); crotonase gi|1055218|gb|AAA95967.1|(1055218); 3-hydroxybutyryl-CoA dehydratase (Crotonase) gi|1706153|sp|P52046.1|CRT_CLOAB(1706153); Crotonase (3-hydroxybutyryl-CoA dehydratase) (Clostridium acetobutylicum ATCC 824) gi|15025745|gb|AAK80658.1|AE007768_12(15025745) each sequence associated with the accession number is incorporated herein by reference in its entirety.

[00115] In addition to the foregoing, the terms "crotonase" or "crt" refer to proteins that are capable of catalyzing the conversion of 3-hydroxybutyryl-CoA to crotonyl-CoA, and which share at least about 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or greater sequence identity, or at least about 50%, 60%, 70%, 80%, 90%, 95%, 96%, 97%, 98%, 99% or greater sequence similarity, as calculated by NCBI BLAST, using default parameters, to SEQ ID NO:34.

[00116] Butyryl-coA dehydrogenase is an enzyme in the protein pathway that catalyzes the reduction of crotonyl-CoA to butyryl-CoA. A butyryl-CoA dehydrogenase complex (Bcd/EtfAB) couples the reduction of crotonyl-CoA to butyryl-CoA with the reduction of ferredoxin. Depending upon the organism used a heterologous

butyryl-CoA dehydrogenase can be engineered for expression in the organism. Alternatively, a native butyryl-CoA dehydrogenase can be overexpressed. Butyryl-coA dehydrogenase is encoded in *C.acetobutylicum* and *M.elsdenii* by *bcd*. BCD homologs and variants are known. For examples, such homologs and variants include, for example, butyryl-CoA dehydrogenase (*Clostridium acetobutylicum* ATCC 824) gi|15895968|ref|NP_349317.1|(15895968); Butyryl-CoA dehydrogenase (*Clostridium acetobutylicum* ATCC 824) gi|15025744|gb|AAK80657.1|AE007768_11(15025744); butyryl-CoA dehydrogenase (*Clostridium botulinum* A str. ATCC 3502) gi|148381147|ref|YP_001255688.1|(148381147); butyryl-CoA dehydrogenase (*Clostridium botulinum* A str. ATCC 3502) gi|148290631|emb|CAL84760.1|(148290631), each sequence associated with the accession number is incorporated herein by reference in its entirety. BCD can be expressed in combination with a flavoprotein electron transfer protein. Useful flavoprotein electron transfer protein subunits are expressed in *C.acetobutylicum* and *M.elsdenii* by a gene *etfA* and *etfB* (or the operon *etfAB*). ETFA, B, and AB homologs and variants are known. For examples, such homologs and variants include, for example, putative a-subunit of electron-transfer flavoprotein gi|1055221|gb|AAA95970.1|(1055221); putative b-subunit of electron-transfer flavoprotein gi|1055220|gb|AAA95969.1|(1055220), each sequence associated with the accession number is incorporated herein by reference in its entirety.

[00117] In addition to the foregoing, the terms "Butyryl-coA dehydrogenase" or "bcd" refer to proteins that catalyzes the reduction of crotonyl-CoA to butyryl-CoA, and which share at least about 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or greater sequence identity, or at least about 50%, 60%, 70%, 80%, 90%, 95%, 96%, 97%, 98%, 99% or greater sequence similarity, as calculated by NCBI BLAST, using default parameters, to SEQ ID NO:40 or 46.

[00118] Aldehyde/alcohol dehydrogenase catalyzes the conversion of butyryl-CoA to butyraldehyde and butyraldehyde to 1-butanol or . In one aspect, the aldehyde/alcohol dehydrogenase preferentially catalyzes the conversion of butyryl-CoA to butyraldehyde and

butyraldehyde to 1-butanol. Depending upon the organism used a heterologous aldehyde/alcohol dehydrogenase can be engineered for expression in the organism. Alternatively, a native aldehyde/alcohol dehydrogenase can be overexpressed. aldehyde/alcohol dehydrogenase is encoded in *C.acetobutylicum* by *adhE* (e.g., an *adhE2*). ADHE (e.g., ADHE2) homologs and variants are known. For examples, such homologs and variants include, for example, aldehyde-alcohol dehydrogenase (*Clostridium acetobutylicum*) gi|3790107|gb|AAD04638.1|(3790107); aldehyde-alcohol dehydrogenase (*Clostridium botulinum* A str. ATCC 3502) gi|148378348|ref|YP_001252889.1|(148378348); Aldehyde-alcohol dehydrogenase (Includes: Alcohol dehydrogenase (ADH) Acetaldehyde dehydrogenase (acetylating) (ACDH) gi|19858620|sp|P33744.3|ADHE_CLOAB(19858620); Aldehyde dehydrogenase (NAD+) (*Clostridium acetobutylicum* ATCC 824) gi|15004865|ref|NP_149325.1|(15004865); alcohol dehydrogenase E (*Clostridium acetobutylicum*) gi|298083|emb|CAA51344.1|(298083); Aldehyde dehydrogenase (NAD+) (*Clostridium acetobutylicum* ATCC 824) gi|14994477|gb|AAK76907.1|AE001438_160(14994477); aldehyde/alcohol dehydrogenase (*Clostridium acetobutylicum*) gi|12958626|gb|AAK09379.1|AF321779_1(12958626); Aldehyde-alcohol dehydrogenase, ADHE1 (*Clostridium acetobutylicum* ATCC 824) gi|15004739|ref|NP_149199.1|(15004739); Aldehyde-alcohol dehydrogenase, ADHE1 (*Clostridium acetobutylicum* ATCC 824) gi|14994351|gb|AAK76781.1|AE001438_34(14994351); aldehyde-alcohol dehydrogenase E (*Clostridium perfringens* str. 13) gi|18311513|ref|NP_563447.1|(18311513); aldehyde-alcohol dehydrogenase E (*Clostridium perfringens* str. 13) gi|18146197|dbj|BAB82237.1|(18146197), each sequence associated with the accession number is incorporated herein by reference in its entirety.

[00119] In addition to the foregoing, the terms "aldehyde/alcohol dehydrogenase" or "adhE2" refer to proteins that catalyzes the conversion of butyryl-CoA to 1 butanol; or hexanoyl-CoA to hexanal and hexanal to hexanol, and which share at least about 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or greater sequence identity, or at least about 50%, 60%, 70%, 80%,

90%, 95%, 96%, 97%, 98%, 99% or greater sequence similarity, as calculated by NCBI BLAST, using default parameters, to SEQ ID NO:68.

[00120] Crotonyl-coA reductase catalyzes the reduction of crotonyl-CoA to butyryl-CoA. Depending upon the organism used a heterologous Crotonyl-coA reductase can be engineered for expression in the organism. Alternatively, a native Crotonyl-coA reductase can be overexpressed. Crotonyl-coA reductase is encoded in *S.coelicolor* by *ccr*. CCR homologs and variants are known. For examples, such homologs and variants include, for example, crotonyl CoA reductase (*Streptomyces coelicolor* A3(2)) gi|21224777|ref|NP_630556.1| (21224777); crotonyl CoA reductase (*Streptomyces coelicolor* A3(2)) gi|4154068|emb|CAA22721.1| (4154068); crotonyl-CoA reductase (*Methylobacterium* sp. 4-46) gi|168192678|gb|ACA14625.1| (168192678); crotonyl-CoA reductase (*Dinoroseobacter shibae* DFL 12) gi|159045393|ref|YP_001534187.1| (159045393); crotonyl-CoA reductase (*Salinispora arenicola* CNS-205) gi|159039522|ref|YP_001538775.1| (159039522); crotonyl-CoA reductase (*Methylobacterium extorquens* PA1) gi|163849740|ref|YP_001637783.1| (163849740); crotonyl-CoA reductase (*Methylobacterium extorquens* PA1) gi|163661345|gb|ABY28712.1| (163661345); crotonyl-CoA reductase (*Burkholderia ambifaria* AMMD) gi|115360962|ref|YP_778099.1| (115360962); crotonyl-CoA reductase (*Parvibaculum lavamentivorans* DS-1) gi|154252073|ref|YP_001412897.1| (154252073); Crotonyl-CoA reductase (*Silicibacter* sp. TM1040) gi|99078082|ref|YP_611340.1| (99078082); crotonyl-CoA reductase (*Xanthobacter autotrophicus* Py2) gi|154245143|ref|YP_001416101.1| (154245143); crotonyl-CoA reductase (*Nocardioides* sp. JS614) gi|119716029|ref|YP_922994.1| (119716029); crotonyl-CoA reductase (*Nocardioides* sp. JS614) gi|119536690|gb|ABL81307.1| (119536690); crotonyl-CoA reductase (*Salinispora arenicola* CNS-205) gi|157918357|gb|ABV99784.1| (157918357); crotonyl-CoA reductase (*Dinoroseobacter shibae* DFL 12) gi|157913153|gb|ABV94586.1| (157913153); crotonyl-CoA reductase (*Burkholderia ambifaria* AMMD) gi|115286290|gb|ABI91765.1| (115286290); crotonyl-CoA reductase (*Xanthobacter autotrophicus* Py2) gi|154159228|gb|ABS66444.1| (154159228); crotonyl-CoA reductase (*Parvibaculum lavamentivorans* DS-1) gi|154156023|gb|ABS63240.1| (154156023); crotonyl-CoA reductase (*Methylobacterium radiotolerans* JCM 2831)

gi|170654059|gb|ACB23114.1|(170654059); crotonyl-CoA reductase (Burkholderia graminis C4D1M) gi|170140183|gb|EDT08361.1|(170140183); crotonyl-CoA reductase (Methylobacterium sp. 4-46) gi|168198006|gb|ACA19953.1|(168198006); crotonyl-CoA reductase (Frankia sp. EAN1pec) gi|158315836|ref|YP_001508344.1|(158315836), each sequence associated with the accession number is incorporated herein by reference in its entirety.

[00121] In addition to the foregoing, the terms "crotonyl-CoA reductase" or "ccr" refer to proteins that catalyzes the reduction of crotonyl-CoA to butyryl-CoA, and which share at least about 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or greater sequence identity, or at least about 50%, 60%, 70%, 80%, 90%, 95%, 96%, 97%, 98%, 99% or greater sequence similarity, as calculated by NCBI BLAST, using default parameters, to SEQ ID NO:54, 56, 58, 60, 62, 64, or 66.

[00122] Culture conditions suitable for the growth and maintenance of a recombinant microorganism provided herein are described in the Examples below. The skilled artisan will recognize that such conditions can be modified to accommodate the requirements of each microorganism. Appropriate culture conditions useful in producing a 1-butanol, n-hexanol, and/or octanol products comprise conditions of culture medium pH, ionic strength, nutritive content, etc.; temperature; oxygen/CO₂/nitrogen content; humidity; light and other culture conditions that permit production of the compound by the host microorganism, *i.e.*, by the metabolic action of the microorganism. Appropriate culture conditions are well known for microorganisms that can serve as host cells.

[00123] It is understood that a range of microorganisms can be modified to include a recombinant metabolic pathway suitable for the production of n-butanol, n-hexanol and octanol. It is also understood that various microorganisms can act as "sources" for genetic material encoding target enzymes suitable for use in a recombinant microorganism provided herein. The term "microorganism" includes prokaryotic and eukaryotic microbial species from the Domains Archaea, Bacteria and Eucarya, the latter including yeast and filamentous fungi, protozoa, algae, or higher Protista. The

terms "microbial cells" and "microbes" are used interchangeably with the term microorganism. Of particularly use are cyanobacterium.

[00124] The term "prokaryotes" is art recognized and refers to cells which contain no nucleus or other cell organelles. The prokaryotes are generally classified in one of two domains, the Bacteria and the Archaea. The definitive difference between organisms of the Archaea and Bacteria domains is based on fundamental differences in the nucleotide base sequence in the 16S ribosomal RNA.

[00125] The term "Archaea" refers to a categorization of organisms of the division Mendosicutes, typically found in unusual environments and distinguished from the rest of the procaryotes by several criteria, including the number of ribosomal proteins and the lack of muramic acid in cell walls. On the basis of ssrRNA analysis, the Archaea consist of two phylogenetically-distinct groups: Crenarchaeota and Euryarchaeota. On the basis of their physiology, the Archaea can be organized into three types: methanogens (prokaryotes that produce methane); extreme halophiles (prokaryotes that live at very high concentrations of salt ([NaCl])); and extreme (hyper) thermophilus (prokaryotes that live at very high temperatures). Besides the unifying archaeal features that distinguish them from Bacteria (*i.e.*, no murein in cell wall, ester-linked membrane lipids, etc.), these prokaryotes exhibit unique structural or biochemical attributes which adapt them to their particular habitats. The Crenarchaeota consists mainly of hyperthermophilic sulfur-dependent prokaryotes and the Euryarchaeota contains the methanogens and extreme halophiles.

[00126] "Bacteria", or "eubacteria", refers to a domain of prokaryotic organisms. Bacteria include at least 11 distinct groups as follows: (1) Gram-positive (gram+) bacteria, of which there are two major subdivisions: (1) high G+C group (Actinomycetes, Mycobacteria, Micrococcus, others) (2) low G+C group (Bacillus, Clostridia, Lactobacillus, Staphylococci, Streptococci, Mycoplasmas); (2) Proteobacteria, *e.g.*, Purple photosynthetic +non-photosynthetic Gram-negative bacteria (includes most "common" Gram-negative bacteria); (3) Cyanobacteria, *e.g.*, oxygenic phototrophs; (4) Spirochetes and related species; (5) Planctomyces; (6)

Bacteroides, Flavobacteria; (7) Chlamydia; (8) Green sulfur bacteria; (9) Green non-sulfur bacteria (also anaerobic phototrophs); (10) *Radioresistant micrococci* and relatives; and (11) *Thermotoga* and *Thermosipho thermophiles*.

[00127] "Gram-negative bacteria" include cocci, nonenteric rods, and enteric rods. The genera of Gram-negative bacteria include, for example, *Neisseria*, *Spirillum*, *Pasteurella*, *Brucella*, *Yersinia*, *Francisella*, *Haemophilus*, *Bordetella*, *Escherichia*, *Salmonella*, *Shigella*, *Klebsiella*, *Proteus*, *Vibrio*, *Pseudomonas*, *Bacteroides*, *Acetobacter*, *Aerobacter*, *Agrobacterium*, *Azotobacter*, *Spirilla*, *Serratia*, *Vibrio*, *Rhizobium*, *Chlamydia*, *Rickettsia*, *Treponema*, and *Fusobacterium*.

[00128] "Gram positive bacteria" include cocci, nonsporulating rods, and sporulating rods. The genera of gram positive bacteria include, for example, *Actinomyces*, *Bacillus*, *Clostridium*, *Corynebacterium*, *Erysipelothrix*, *Lactobacillus*, *Listeria*, *Mycobacterium*, *Myxococcus*, *Nocardia*, *Staphylococcus*, *Streptococcus*, and *Streptomyces*.

[00129] In another embodiment, a method of producing a recombinant microorganism that converts a suitable carbon substrate to n-hexanol and/or octanol is provided. The method includes transforming a microorganism with one or more recombinant polynucleotides encoding polypeptides that include keto thiolase or acetyl-CoA acetyltransferase activity, hydroxybutyryl CoA dehydrogenase activity, crotonase activity, crotonyl-CoA reductase or butyryl-CoA dehydrogenase activity, trans-enoyl-CoA reductase and alcohol dehydrogenase activity.

[00130] In another embodiment, a method for producing n-hexanol and/or octanol is provided. The method includes culturing a recombinant microorganism as provided herein in the presence of a suitable carbon substrate and under conditions suitable for the conversion of the substrate to n-hexanol and/or octanol.

[00131] The n-hexanol and/or octanol produced by a microorganism provided herein can be detected by any method known to the skilled artisan. Such methods include mass spectrometry.

[00132] As previously discussed, general texts which describe molecular biological techniques useful herein, including the use of

vectors, promoters and many other relevant topics, include Berger and Kimmel, Guide to Molecular Cloning Techniques, Methods in Enzymology Volume 152, (Academic Press, Inc., San Diego, Calif.) ("Berger"); Sambrook et al., Molecular Cloning--A Laboratory Manual, 2d ed., Vol. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1989 ("Sambrook") and Current Protocols in Molecular Biology, F. M. Ausubel et al., eds., Current Protocols, a joint venture between Greene Publishing Associates, Inc. and John Wiley & Sons, Inc., (supplemented through 1999) ("Ausubel"), each of which is incorporated herein by reference in its entirety.

[00133] Examples of protocols sufficient to direct persons of skill through *in vitro* amplification methods, including the polymerase chain reaction (PCR), the ligase chain reaction (LCR), Q β -replicase amplification and other RNA polymerase mediated techniques (e.g., NASBA), e.g., for the production of the homologous nucleic acids of the disclosure are found in Berger, Sambrook, and Ausubel, as well as in Mullis et al. (1987) U.S. Pat. No. 4,683,202; Innis et al., eds. (1990) PCR Protocols: A Guide to Methods and Applications (Academic Press Inc. San Diego, Calif.) ("Innis"); Arnheim & Levinson (Oct. 1, 1990) C&EN 36-47; The Journal Of NIH Research (1991) 3: 81-94; Kwok et al. (1989) Proc. Natl. Acad. Sci. USA 86: 1173; Guatelli et al. (1990) Proc. Nat'l. Acad. Sci. USA 87: 1874; Lomell et al. (1989) J. Clin. Chem 35: 1826; Landegren et al. (1988) Science 241: 1077-1080; Van Brunt (1990) Biotechnology 8: 291-294; Wu and Wallace (1989) Gene 4:560; Barringer et al. (1990) Gene 89:117; and Sooknanan and Malek (1995) Biotechnology 13:563-564.

[00134] Improved methods for cloning *in vitro* amplified nucleic acids are described in Wallace et al., U.S. Pat. No. 5,426,039.

[00135] Improved methods for amplifying large nucleic acids by PCR are summarized in Cheng et al. (1994) Nature 369: 684-685 and the references cited therein, in which PCR amplicons of up to 40 kb are generated. One of skill will appreciate that essentially any RNA can be converted into a double stranded DNA suitable for restriction digestion, PCR expansion and sequencing using reverse transcriptase and a polymerase. See, e.g., Ausubel, Sambrook and Berger, all *supra*.

[00136] The invention is illustrated in the following examples, which are provided by way of illustration and are not intended to be limiting.

EXAMPLES

[00137] All chemicals were purchase from Thermo Scientific (Pittsburgh, PA) or Sigma-Aldrich (Saint Louis, MO). Restriction enzymes, T4 DNA polymerase, Phusion High-Fidelity DNA polymerase were acquired from New England Biolabs (Ipswich, MA). Deoxynucleotides were purchased from Fermentas/Thermo Scientific (Pittsburgh, PA). Oligonucleotides were obtained from IDT (San Diego, CA). Bugbuster® and Lysonase® were purchased from EMD/Novagen (San Diego, CA).

[00138] **DNA Manipulations.** All strains and plasmids used in this study are listed in Table 2. *E. coli* XL-1 Blue (Stratagene) was used to propagate plasmids. *E. coli* BL21(DE3) (Invitrogen), JCL166 and JCL299 were used as the host strain to allow the expression of genes and n-hexanol production. Constructions of strain JCL166 and JCL299 were described in previously (Atsumi et al., Metabol. Engineering, 10:305-311, 2008; and Shen et al., Appl. Environ. Microbiol., 77(9): 2906-2915, 2011, incorporated herein by reference).

[00139] In the modified LIC procedure both vector and inserts (gene of interest) were generated by PCR using a high fidelity DNA polymerase (Phusion DNA polymerase). Oligonucleotides for inserts were designed with 18 to 21 nucleotides at the 5'-end overlapping with the flanking sequences of the vector DNA. This method utilizes the exonuclease property of T4 DNA polymerase in the absence of nucleotides to generate single stranded DNA stretches in the vector and insert fragments ends which can anneal holding the fragments in place and then be repaired in vivo after transformation in an *E. coli* strain (Aslanidis and de Jong, 1990; Li and Evans, 1997; Li and Elledge, 2007). PCR products were purified using a PCR Purification Kit (Invitrogen, Carlsbad, CA, or Zymo Research, Irvine, CA) and the DNA concentrations were determined spectrophotometrically at 260 nm. Next, the vector and PCR product insert were mixed (1:1 to 1:2 ratio, respectively) in a total DNA amount of 400-500 ng in 10 µL 1x NEB Buffer 2 (New England Biolabs) containing 0.25 µL T4 DNA polymerase (0.75 U). The reaction was incubated for 5-10 min at

room temperature, followed by 5-10 min incubation on ice, and a 3 μ L aliquot was used to transform 50 μ L *E. coli* XL1-Blue competent cells according to a standard transformation protocol. The transformed cells were plated on LB plates containing the required antibiotics and incubated at 37 °C for 24 h. The presence of correctly cloned inserts was determined by colony PCR and DNA sequencing.

[00140] Plasmid pHM06 was created by amplifying the *Mus musculus* mBACH gene from cDNA prepared from mouse C57/Bl6 brain tissue and assembled onto PCR amplified pET28a. To create plasmid pHM10H, pEL11 was used as template and amplified by PCR deleting the adhE which was replaced by a PCR fragment carrying mBACH from pHM06. Plasmids pHM10C and pHM10P were constructed in a similar replacing the *C. acetobutylicum* crt and hbd genes in pHM10H by the *C. kluyveri* crt1 and *R. eutropha* paaH1, respectively. *C. kluyveri* crt1 was synthesized by Genewiz (South Plainfield, NJ) with codon optimization for *E. coli*. *R. eutropha* paaH1 gene was amplified from strain H16 genomic DNA. pDK019 was created by replacing the *C. acetobutylicum* hbd gene in pEL11 by the *R. eutropha* paaH1. pHL101 was created by replacing the ColA origin of replication in pEL102 by a PCR fragment containing the CloDF13 replicon from pCDFDuet-1 (EMD/Novagen). pHM10-P mutant series were created using error-prone PCR and the EZClone protocol as described below. pHM25-P series were constructed by replacing the mBACH gene in pHM10-P mutant plasmids by the *C. acetobutylicum* adhE PCR fragment.

[00141] All plasmids were verified by DNA sequencing performed by Laragen (Los Angeles, CA).

Table 2. *Strains and plasmids used in this study.**Plasmids*

pCDFDuet-1	T7lac :: CloDF13-derived CDF ori; Sm ^R	Novagen
pCDFDuet-1_ <i>bktB</i>	From pCDFDuet-1, T7lac :: <i>bktB</i> (RE)	this study
pCDFDuet-1_ <i>hbd</i>	From pCDFDuet-1, T7lac :: <i>hbd</i> (CA)	this study
pCDFDuet-1_ <i>crt</i>	From pCDFDuet-1, T7lac :: <i>crt</i> (CA)	this study
pCDFDuet-1_ <i>tdTer</i>	From pCDFDuet-1, T7lac :: <i>tdTer</i> (TD)	this study
pCDFDuet-1_ <i>egTer</i>	From pCDFDuet-1, T7lac :: <i>egTer</i> (EG)	this study
pEL11	PLlacO1:: <i>atoB</i> (EC)- <i>adhE2</i> (CA)- <i>crt</i> (CA)- <i>hbd</i> (CA); ColE1 ori; Amp ^R	Shen et al.
pIM8	PLlacO1:: <i>ter</i> (TD); <i>Cola</i> ori; Kan ^R	Shen et al.
pEL100	From pIM8, PLlacO1:: <i>bktB</i> (RE)- <i>egTer</i> (EG)- <i>hbd</i> (CA)- <i>crt</i> (CA); <i>Cola</i> ori; Kan	this study
pEL102	From pIM8, PLlacO1:: <i>tdTer</i> (TD)- <i>bktB</i> (RE)- <i>egTer</i> (EG)	this study
pCS138	PLlacO1:: <i>fdh</i> (CB); pSC101 ori; Cm ^R	Shen et al.
pCS38_ <i>adhE2</i>	PLlacO1:: <i>adhE2</i> (CA); ColE1 ori; Sm ^R	this study
pHL101	λ lacOp1:: <i>Tde .ter</i> <i>Reu .bktB</i> <i>Egl .ter</i> ; CloDF13 ori; Kan ^R	This work
pHM06	pET28a; T7lacp :: 6His:: <i>Mmu.mBACH</i> ; pBR322 ori, Kan ^R	This work
pHM10C	λ lacOp1:: <i>Eco .atoB</i> <i>Mmu.mBACH</i> <i>Ckl .crt1</i> <i>Cac .hbd</i> ; ColE1 ori; Amp ^R	This work
pHM10H	λ lacOp1:: <i>Eco .atoB</i> <i>Mmu.mBACH</i> <i>Cac .crt</i> <i>Cac .hbd</i> ; ColE1 ori; Amp ^R	This work
pHM10P	λ lacOp1:: <i>Eco .atoB</i> <i>Mmu.mBACH</i> <i>Cac .crt</i> <i>Reu .paaH1</i> ; ColE1 ori; Amp ^R	This work
pHM10-P series	λ lacOp1:: <i>Eco .atoB</i> <i>Mmu.mBACH</i> <i>Cac .crt</i> <i>Reu .paaH1*</i> ; ColE1 ori; Amp ^R	This work
pHM25-P series	λ lacOp1:: <i>Eco .atoB</i> <i>Cac .adhE</i> <i>Cac .crt</i> <i>Reu .paaH1*</i> ; ColE1 ori; Amp ^R	This work

[00142] pCDFDuet-1 (Novagen) was used to provide individual expression of gene under a T7lac promoter and ribosome-binding site. Constructions of plasmid pEL11 and pIM8 were described in Shen et al, *supra*.

[00143] **pCDFDuet-1_ *hbd* (*crt*/*bktB*/*hbd*/*tdTer*/*egTer*)**. DNA fragment containing *C. acetobutylicum* *hbd* gene (*C. acetobutylicum* *crt*/*R. eutropha* *bktB*/*E. coli* *atoB*/*T. denticola* *tdTer*/ *E. gracilis* *egTer*) was amplified from each templates using primers and digested with restriction enzymes listed Table 3. And then these fragments were ligated into plasmid pCDFDuet-1 previously digested with corresponding restriction enzymes.

[00144] **pEL100**. DNA fragments of *R. eutropha* *bktB*, *E. gracilis* *egTer*, *C. acetobutylicum* *crt*/*hbd*, and vector fragment containing ColE1 replicon and ampicillin resistant gene were amplified from pCDFDuet-1_ *bktB*, pCDFDuet-1_ *egTer*, pCDFDuet-1_ *egTer*, pCDFDuet-1_ *egTer*, and pEL11 using primers listed Table 3. These five DNA fragments were assembled using the method reported by Gibson et al.

[00145] **pEL102**. *bktB*-*egTer* gene fragment and vector fragment containing *T. denticola* *tdTer* gene, *Cola* replicon, and kanamycin

resistant gene were amplified from pIM8 as template using primers listed Table 3. These two fragments were assembled as mentioned above.

[00146] pCS38_adhE2. pCS38 plasmid was previously made in the laboratory which contains multiple cloning sites (MCS) preceded by a P_{LlacO1} promoter. The vector carries ColE1 replicon and spectinomycin resistance gene. adhE2 gene was amplified using primers and digested with restriction enzymes listed Table S4. And then this fragment was ligated into pCS38 vector previously digested with Acc65I and MluI.

Table 3. Primers used in this study.

Name	Sequences (5'-3') (SEQ ID NO: in parenthesis)	Target gene	Template
For pCDFDuet-1			
bktB fwd (BamHI)	GGGAAAGGATCCGACGCGTGAAGTGGTAGTGGTAAGCG (1)	bktB (RE)	R. eutropha genomic DNA
bktB rev (SacI)	GGGAAAGAGCTCTCAGATACGCTCGAAGATGGCGGCA(2)		
hbd fwd (BamHI)	GGGAAAGGATCCGAAAAAGGTATGTGTTATAGGTGCAGGTACTATG	hbd (CA)	pEL11
hbd rev (PstI)	GGGAAACTGCAGTTATTTGAATAATCGTAGAAACCTTTTCCTG (4)		
crt fwd (BamHI)	ACCATCATCACCAAGCCAGGATCCGGAACTAAACAATGTCATCC (5)	crt (CA)	pEL11
crt rev (HindIII)	AAGCATTATGCGGCCGCAAGCTTATTATCTATTTTGAAGCCTTC (6)		
atoB fwd (BamHI)	GGGAAAGGATCCGAAAAATTGTGTCATCGTCAGTGC GG (7)	atoB (EC)	pEL11
atoB rev (PstI)	GGGAAACTGCAGTTAATTCAACCGTTCAATCACCATCGC (8)		
tdTer fwd (BamHI)	CGGGATCCGATTGTGTAACCAATGGTTAGGAACAA (9)	tdTer (TD)	pIM8
tdTer rev (SalI)	ACGCGTCGACCTAAATCCTGTGCAACCTTTCTACCT (10)		
egTer fwd (BamHI)	CGGGATCCGGCAATGTTCACTACTACCGCTAAAGTC (11)	egTer (EG)	Codon optmied synthetic gene (Genewiz)
egTer rev (SalI)	ACGCGTCGACATTATTGTTGAGCGGCAGAAGGCAGATCC (12)		
For pCS38_adhE2			
adhE2 fwd (BsaI)	GCCACCGGTCTCCGTACCATGAAAGTTACAAATCAAAAAGAACTAA	adhE2 (CA)	pEL11
adhE2 rev (MullI)	ATCAACGCGTTTAAAATGATTTT (14)		
For pEL100			
pEL100 bktB fwd	CCGAATTCATTAAAGAGGAGAAAGGTACCATGACGCGTGAAGTGG (16)	bktB (RE)	pCDFDuet-1_bktB
pEL100 bktB rev	CATTGCCATGGTATATCTCTGGATCCATTAGATACGCTCGAAGA (16)		
pEL100 tdTer fwd	AATGGATCCAGGAGATATACCATGGCAATGTTCACTACTACCGC (17)	egTer (EG)	pCDFDuet-1_egTer
pEL100 tdTer rev	TGGTATATCTCTGCGGCCGCATTATTGTTGAGCGGCAGAAGGC (18)		
pEL100 hbd fwd	AACAATAATGCGGCCGAGGAGATATACCATGAAAAAGGTATGTG (19)	hbd (CA)	pCDFDuet-1_hbd
pEL100 hbd rev	TTAGTTCATGGTATATCTCTCCCGGGATTATTTGAATAATCG (20)		
pEL100 crt fwd	AATCCCGGGAGGAGATATACCATGGAATAACAATGTCATCC (21)	crt (CA)	pCDFDuet-1_crt
pEL100 crt rev	GCCTCTAGAATTATCTATTTTGAAGCCTTC (22)		
pEL100 fwd	GAAGGCTTCAAAAATAGATAATTCTAGAGGCATCAAATAAAACG (23)	pEL100 vector	pIM8
pEL100 rev	TCATGGTACCTTTCTCTCTTTAATGAATTCGG (24)		
For pEL102			
bktB-tdTer fwd	TCGATAAGCTTAGGAGATATACCATGACGCGTGAAGTGGTAGTGG (25)	bktB (RE)- egTer (EG)	pEL100
bktB-tdTer rev	ATCAAGCTTATTATTGTTGAGCGGCAGAAGGCAG (26)		
pEL102 fwd	TAATAAGCTTGATATCGAATTCCTGCAGCCCGGGGATCCCATGG (27)	pEL102 vector	pIM8
pEL102 rev	TATCTCCTAAGCTTATCGATACCGTCGACTAAATCCTGTCGAACC (28)		

[00147] paaH1 directed evolution. The *R. eutropha* paaH1 gene in pHM10P was randomly mutagenized by error-prone PCR using the GeneMorph II® EZClone Domain Mutagenesis Kit (Agilent/Stratagene, La Jolla, CA) with a medium mutation rate according to the manufacturer's instructions. Next, aliquots of 2 µL of the EZClone reaction were used to transform 50 µL electrocompetent *E. coli* strain JCL166 carrying plasmid pHL101. The cells were recovered in 1 mL LB medium for 1 hour at 37 °C and 250 rpm and then added 4 mL

LB medium, 1% glucose, 100 μ M IPTG, 100 μ g/mL ampicillin, 50 μ g/mL kanamycin, transferred to a 10-mL Vacutainer tube and subjected to anaerobiosis. The tubes were incubated at 30 °C and 250 rpm for 96 h. Next, a second round of anaerobic growth rescue was performed by diluting the culture from the first round 200 fold in 5 mL of fresh LB 1% glucose, 100 μ M IPTG and antibiotics, and again subjecting them to anaerobiosis and incubation at 30 °C and 250 rpm for 48 h. The cultures that showed increased growth compared to control were selected for plasmid isolation by minipreps (Qiagen, Valencia, CA), transformed in *E. coli* XL1-Blue, and selected in LB-agar plates containing 100 μ g/mL ampicillin. Two colonies from each plate were selected for individual clones isolation. These clones were again transformed into JCL166/pHL101 to confirm their ability to rescue growth under anaerobic conditions. Those which confirmed the growth rescue phenotype were sequenced to identify the mutations.

[00148] Synthesis of hexenoyl-CoA. Preparation of hexenoyl-CoA was accomplished with little modifying a procedure described by Kopp. Coenzyme A trilithium salt (0.025 mmol), trans-2-hexanoic acid (0.05 mmol), PyBOP (0.04 mmol), and K₂CO₃ (0.10 mmol) were dissolved in 2 mL of THF/water (1:1) and incubated for 2 h at room temperature. After evaporation to remove THF by vacuum centrifugation, the rest of solution was purified by HPLC (Agilent, 1100 series) on Nova-Pak C18 column (Waters, 3.6 x 150 mm). Product elution was monitored at 254 nm. Pooled fractions were concentrated again to approximately 1 mM based on absorbance at 254 nm of crotonyl-CoA (Sigma) as standard. The solutions was kept in this form for sub-sequent enzyme reactions and stored at -20 °C. The identity was confirmed by liquid chromatography/electrospray ionization-mass spectrometry (LC/ESI-MS).

[00149] Preparation of enzyme solutions. In order to characterize BktB, Hbd, Crt, TdTer and EgTer, these coding genes were cloned into pCDFDuet-1 as described above and expressed as N-terminal fusion protein with His6 in BL21(DE3). Overnight culture in LB medium at 37°C was inoculated 1% into 3 ml of fresh LB medium. LB medium was supplemented with 50 μ g/mL Spectinomycin (Sm). The cultures were grown at 37 °C to an OD₆₀₀ of 0.7-0.8 then induced with 0.5 mM IPTG and grown for another 3-4 hours aerobically. Thereafter the cells

were collected by centrifugation and the resulting pelleted cells were resuspended with 0.2 ml of the lysis buffer (50mM Tris-HCl at pH 7.5, Bugbuster® and Lysonase™ (Novagen)). Each lysate was allowed to go for 10-20 minutes until the cell resuspension turned clear and fusion protein was purified by His-Affinity column (Zymo Research, His-Spin Protein Miniprep™) and used for enzyme reaction.

[00150] For preparation of AdhE2 containing cell lysate, this coding gene was cloned into pCS38 and expressed in JCL166. Overnight culture in LB medium at 37°C was inoculated 1% into 3-4 ml of fresh LB medium. LB medium was supplemented with 50µg/mL Sm. The cultures were grown at 37°C to an OD600 of 0.7-0.8 then induced with 0.1 mM IPTG and grown for another 3-4 hours anaerobically in anaerobic bag (GasPak™ EZ (BD)). Cell lysate was prepared as described above except adding 1 mM DTT in lysis buffer. Resulting lysate was used for enzyme assay and reaction without neither centrifugation nor purification step.

[00151] Enzymatic Reaction. Purified enzyme solutions were added appropriately. Cofactors (NAD⁺, NADH or CoA) were added when required by the respective enzyme.

[00152] For BktB, Hbd and Crt characterization, the reactions were carried out independently on a stepwise manner with each enzyme being added progressively based on the reverse reaction from hexenoyl-CoA. 100 µM Hexenoyl-CoA in 100 mM Tris-HCl pH7.5 were incubated with Crt (Fig. 1B), Crt, Hbd, and 2 mM NAD⁺ (Fig. 1C) and Crt, Hbd, 2 mM NAD⁺, BktB, and 1 mM CoA (Fig. 1D). Each reaction mixture (100 µL) was incubated overnight at 37 °C and filtrated using 10,000 MWCO ultrafiltration membrane (Amicon) for HPLC analysis. In AtoB characterization, same condition was used as mentioned above.

[00153] For Ter characterization, 100 µM hexenoyl-CoA and 2 mM NADH in 100 mM Tris-HCl pH7.5 was incubated with TdTer (Fig. 3B) and EgTer (Fig. 3C) overnight at 37 °C and then filtrated for HPLC analysis.

[00154] For AdhE2 characterization, 1.5 mM hexanoyl-CoA (Sigma), 4 mM NADH in 50 mM Tris-HCl pH7.5 was incubated with cell lysate from JCL166 (Fig. 4A) and JCL166/pCS38_adhE2 (Fig. 4B) overnight at 37 °C and then filtrated as mentioned above for GC analysis.

[00155] Enzyme Assay of AdhE2. Aldehyde/alcohol dehydrogenase (AdhE2) activity was measured at 340 nm. The reaction mixture contained 100 mM Tris-HCl buffer pH 7.5, 0.3 mM NADH, 0.3 mM Acyl-CoA substrates, and crude cell extract. The reaction was initiated by the addition of Acyl-CoA substrates.

[00156] Detection of Enzymatic reaction products. In enzymatic reaction analysis of BktB, Hbd, Crt and Ter, products were analyzed by HPLC (Agilent Technologies, 1100 series) on a C18 column (TOSOH, TSKgel ODS-100Z 4.6 x 150 mm, 5 μ m) at a flow rate of 1 mL/min with the gradient condition (Table 4 and 5). Product elution was monitored at 254 nm.

[00157] AdhE2 reaction products were analyzed by GC equipped with flame ionization detector (Agilent, 6890N). Filtrated samples were injected in split injection mode (1:15 split ratio) using 2-methyl-1-pentanol as the internal standard.

Table 4. Gradient elution system for analyzing synthesized hexenoyl-CoA by LC/ESI-MS.

Buffer A	Buffer B	Gradient	
		Time (min)	%B
25 mM Ammonium formate pH 5.0	Acetonitrile	0	5
		5	5
		10	30
		15	30

Table 5. Gradient elution system for the separation of CoA compounds from enzymatic reactions.

Buffer A	Buffer B	Gradient	
		Time (min)	%B
50 mM Sodium phosphate buffer pH 5.0	Acetonitrile	0	10
		2	10
		8	30
		11	30
		11.5	10
		13	10

[00158] Quantification of alcohols. Supernatant of culture broth was analyzed by GC as described above.

[00159] GC-MS analysis. 15 μ L of 10 M NaOH was added to 500 μ L of culture broth supernatant, increasing solution alkalinity (pH 10).

200 µL of hexane was then added to the solution and mixed for extraction of n-hexanol. Organic layer was analyzed by GC-MS system (Agilent Technologies, 6890N GC/5973N MSD) equipped with a HP-5MS capillary column (Agilent Technologies, 30 m; 0.25 inside diameter; 0.25 µm film thickness). Helium (constant flow 1 mL/min) was used as a carrier gas. The temperature of the injector was 250 °C. The oven program was as follows: 50 °C for 3 min, ramp to 100 °C at 5 °C/min,

[00160] n-hexanol production media. Productions of n-hexanol were carried out in Terrific Broth (TB) (12 g tryptone, 24 g yeast extract, 2.31 g KH₂PO₄, 12.54 g K₂HPO₄, 4 ml glycerol per liter of water) supplemented with 2% glucose.

[00161] Culture condition for n-hexanol production. Culture condition used in the high-titer 1-butanol production were used. Overnight culture in LB broth at 37 °C was inoculated 1% into 5 ml (test tube) of fresh TB + 2% glucose medium. Antibiotics were appropriately added to the following final concentrations: ampicillin (Amp), 100 µg/mL; chloramphenicol (Cm), 50 µg/mL; kanamycin (km), 50 µg/mL. The cultures were grown at 37 °C to an OD₆₀₀ of 0.4-0.6 then induced with 0.1 mM IPTG for another 1-2 hours aerobically.

[00162] For anaerobic condition, the 5 ml induced cultures were transferred from test tubes to the 10 ml BD Vacutainer sealed tubes. Oxygen in the headspace and media was then evacuated through the needle by repeated vacuuming and refilling of nitrogen and hydrogen in the anaerobic transfer chamber. The needles were taken off from the caps inside the anaerobic chamber. The sealed tubes were then taken outside and wrapped with parafilm and tape to prevent bursting of the caps (when cells grow, pressure built up due to CO₂ released). Cultures were then incubated at 37 °C in a rotary shaker (250 rpm). Samples were taken everyday inside the anaerobic chamber to maintain anaerobicity. If a time course was taken, the media was refreshed after every 24 hours in order to maintain pH at around 7.0 by extracting 500 µL of broth and feeding same amount of newly prepared TB media containing 1.5% glucose and NaOH.

[00163] Culture medium and conditions. *E. coli* strains were normally grown in LB medium containing the appropriate antibiotics at 37 °C and rotatory agitation at 250 rpm. Antibiotics when

required were used at the following concentrations: ampicillin, 100 µg/mL; kanamycin, 50 µg/mL; tetracycline, 10 µg/mL.

[00164] Anaerobic growth rescue. The anaerobic growth rescue was performed as described previously (Shen et al., 2011). Briefly, 10 µL of overnight cultures of *E. coli* strain JCL166 and its derivatives are used to inoculate 5 mL of LB medium containing 1% glucose, 100 µM isopropyl-β-D-thiogalactopyranoside (IPTG), and antibiotic as needed, in a 10-mL BD Vacutainer tube (San Jose, CA). The tubes are sealed and a needle attached to a Millipore PES 0.22 µm filter is introduced. Next, the headspace is vacuumed and replaced by an anaerobic mix of 5% H₂ and 95% N₂. The needle is removed inside an anaerobic chamber, and the sealed tubes are incubated for 24 h or longer at 37 °C and 250 rpm.

[00165] Culture condition for hexanoic acid production. Overnight cultures of *E. coli* strain JCL299 carrying the desired plasmids were prepared from single colonies and 50 µL were used to inoculate 5 mL LB-1% glucose, and incubated at 37 °C and 250 rpm until A₆₀₀ nm of 0.6 to 0.8. The cultures were then induced with 100 µM IPTG and incubated at 30 °C and 250 rpm for 2 h. Next, the cultures were transferred to a 10-mL BD Vacutainer tube and vacuumed as described above for anaerobic growth rescue. For production, the cultures were incubated at 30 °C and 250 rpm for 48 h.

[00166] Culture condition for alcohol production. For alcohol production, the procedure is the same as described above for hexanoic acid with the exception that the cells were grown in Terrific Broth (12 g tryptone, 24 g yeast extract, 2.31 g KH₂PO₄, 12.54 g K₂HPO₄, and 4 mL glycerol per liter of water) supplemented with 2% glucose.

[00167] Reverse enzymatic reactions from hexenoyl-CoA to butyryl-CoA. A first step in constructing the n-hexanol pathway is to express the potential genes and detect the enzymatic activities. beta-ketothiolase (BktB) from *Ralstonia eutropha*, and *Clostridium acetobutylicum* Hbd and Crt were expressed to catalyze the first three steps after butyryl-CoA. Each enzyme was expressed as N-terminal fusion protein with His6 and purified by Ni-affinity column. Owing to the unavailability of 3-ketohexanoyl-CoA and 3-hydroxyhexanoyl-CoA, the enzymatic activities of these three

reactions were tested in the reverse direction using hexenoyl-CoA as the substrate. Hexenoyl-CoA was chemically synthesized and partially purified by preparative HPLC. The substrate was then incubated with Crt, Hbd, and BktB with appropriate cofactors in a stepwise fashion, and the resulting products were analyzed by HPLC (Fig. 1). Hexenoyl-CoA was detected at 10.3 min. In the presence of Crt, hexenoyl-CoA was consumed and a new peak at 8.7 min appeared (Fig. 1B), which was presumably 3-hydroxyhexanoyl-CoA. In the next step, the addition of Hbd and NAD⁺ showed a very similar pattern on the HPLC chart as the previous reaction. After the addition of BktB and CoA, acetyl-CoA and butyryl-CoA were detected as two peaks resolving at 4.3 and 8.3 min, respectively.

[00168] The appearance of the expected products (acetyl-CoA and butyryl-CoA, Fig. 1D) suggests that all enzymes were active towards C6 substrates. Considering that the three-enzyme combination was able to dissociate hexenoyl-CoA into acetyl-CoA and butyryl-CoA and these enzyme reactions are reversible, these enzymes are expected to be able to produce hexenoyl-CoA using acetyl-CoA and butyryl-CoA under normal condition. Slater *et al.* reported that BktB from *R. eutropha* catalyzes the cleavage of 3-ketohexanoyl-CoA as well as acetoacetyl-CoA. Waterson *et al.* reported that *C. acetobutylicum* Crt was active towards hexenoyl-CoA, although the activity was lower than that of crotonyl-CoA by more than 100-fold. The results herein were consistent with these previous findings.

[00169] On the other hand, *E. coli* AtoB is in the same enzyme family as BktB. Therefore, AtoB activity was also characterized to C6 substrate by HPLC analysis as mentioned above. Results showed no significant difference between reactions of with AtoB and without AtoB (Fig. 2), implying that AtoB could not catalyze the condensation reaction of acetyl-CoA and butyryl-CoA efficiently.

[00170] **Reduction of hexenoyl-CoA to hexanoyl-CoA with trans-enoyl-CoA reductases.** Production of 1-butanol at high concentration by recruiting trans-enoyl-CoA reductase (Ter) in lieu of *C. acetobutylicum* (Bcd) on the crotonyl-CoA reduction step has been obtained. Ter from *Euglena gracilis* (EgTer) is previously reported to be able to convert hexenoyl-CoA to hexanoyl-CoA. In contrast, *Treponema denticola* (TdTer), which was used in the high-titer 1-

butanol production, was reported not to possess reductive activity against C6 substrates. To confirm these results, hexenoyl-CoA was incubated with purified EgTer or TdTer and analyzed the products. Interestingly, HPLC data shown in Fig. 3 show that both EgTer and TdTer were able to reduce hexenoyl-CoA to hexanoyl-CoA using NADH as the reducing cofactor, suggesting the broad substrate specificity of these enzymes.

[00171] Reduction of hexanoyl-CoA to n-hexanol with aldehyde/alcohol dehydrogenase. *C. acetobutylicum* aldehyde/alcohol dehydrogenase (AdhE2) is known to reduce butyryl-CoA to butylaldehyde and butanol. This activity was tested by a spectrophotometric assay. As shown in Table 6, AdhE2 has higher activity to butyryl-CoA compared to acetyl-CoA, (consistent with previous data) and has also significant activity for hexanoyl-CoA. Interestingly activity to octanoyl-CoA is higher than hexanoyl-CoA (Table 1). To confirm that AdhE2 convert hexanoyl-CoA to n-hexanol (or hexanal) more directly, cell lysate containing AdhE2 was incubated with hexanoyl-CoA and NADH, and then reaction solution was analyzed by Gas Chromatography (GC). As shown Fig. 3, n-hexanol was detected only in the reaction mixture containing AdhE2.

Table 6. Specific activities of AdhE2 for acetyl-CoA, butyryl-CoA, hexanoyl-CoA, and octanoyl-CoA. Cells were cultured under anaerobic condition. JCL166 was tested as a control.

	Acetyl-CoA	Butyryl-CoA	Hexanoyl-CoA	Octanoyl-CoA
JCL166	1.1	5.1	n.d.	3.7
JCL166/pCS38 <i>adhE2</i>	30 ± 7.2	41 ± 6.9	15 ± 9.5	36 ± 6.9

Enzymatic activities are given as nmol/min/mg

n.d., not detectable

[00172] Demonstration of n-hexanol production by engineered *E. coli*. Having characterized BktB, Hbd, Crt, Ter (TdTer and EgTer) and AdhE2 activities toward C6 substrates, the production of n-hexanol by fermentation from glucose was then tested. The host strain used was JCL166 or JCL299, which have *adhE*, *ldhA*, and *frdAB* genes deleted to increase the NADH availability to drive the reaction. In addition, JCL299 has an additional *pta* gene deletion, which increases the acetyl-CoA availability. The genes for n-hexanol synthesis were expressed from various plasmids: pEL11 (expressing *atoB*, *adhE2*, *crt*, and *hbd*), pIM8 (expressing *tdTer*), pEL102 (expressing *tdTer*, *bktB*, and *egTer*). The strains JCL166/pEL11/pIM8,

JCL166/pEL11/pEL102 and JCL299/pEL11/pEL102 were cultivated in TB+2% glucose media (5 mL) under anaerobic condition for 68 hours. After centrifugation of broths to exclude the cells, supernatants were analyzed by GC.

[00173] In the culture of JCL166/pEL11/pIM8, there was no detectable amount of n-hexanol, but in cultures of JCL166/pEL11/pEL102 and JCL299/pEL11/pEL102, 23±11 mg/L and 27±15 mg/L of n-hexanol, respectively, were detected in 68 hours (Fig. 5). In order to confirm the chemical identity of n-hexanol, these samples were further analyzed by GC-MS. The retention time and MS spectra of the samples were identical to those of the n-hexanol standard (Fig. 30)

[00174] Fdh was overexpressed in addition to the n-hexanol synthesis genes, and the resulting strain JCL299/pEL11/pEL102/pCS138 was examined. Furthermore, during the fermentation, the media was refreshed after every 24 hours in order to maintain pH at around 7.0 by extracting 10% of broth and feeding same amount of newly prepared TB media containing adequate glucose and NaOH. Under aerobic condition cells grew up to 2.4 of OD₆₀₀ (0-7 hours), but no n-hexanol production was detected. After switching to anaerobic condition, production of n-hexanol started and reached at 47 mg/L (55 hours) (Fig. 7). This strain also produced 1-butanol 5.1 g/L and 6.5 g/L at 55 hours and 75 hours.

[00175] The methods and compositions described herein extend the CoA-dependent 1-butanol synthetic reaction sequence to produce n-hexanol. The disclosure demonstrates n-hexanol production from glucose by the engineered *E. coli* strain. Increasing the BktB activity towards butyryl-CoA while eliminating the AdhE2 activity towards the same substrate is one direction to improve n-hexanol production. Furthermore, *Clostridium sp.* BS-1 isolated by Joen *et al.* can produce 1.73 g/L of hexanoic acid, but no n-hexanol production was reported. The biosynthesis of hexanoic acid in this organism is also presumed to start from acetyl-CoA using the similar scheme as in Scheme 1. It appears that this organism may have more efficient enzymes to extend the acyl-CoA chain length to hexanoyl-CoA. Transferring the corresponding genes from this organism to *E. coli* appears to be an interesting direction for improvement. The

kinetics of the enzymes involved remains to be characterized. Nevertheless, the strategy described in this study may be extended further for the production of other even number longer chain alcohols.

[00176] Anaerobic growth rescue scheme using long chain acyl-CoA thioesterase. Previously, a selection scheme was constructed (Shen et al., 2011) to couple anaerobic growth rescue with n-butanol production based on NADH recycling. This scheme uses an *E. coli* strain, JCL166 (Δ ldhA Δ adhE Δ frdBC), which does not have any fermentative pathway to recycle NADH produced in glycolysis. This strain cannot grow anaerobically in medium containing glucose, unless an NADH-consuming pathway is functional or an electron acceptor is provided. The n-hexanol pathway is NADH-consuming, just like the n-butanol pathway, and in principle could rescue JCL166 in anaerobic growth in glucose.

[00177] However, to create a selection method that favors the production of C6 or higher compounds, the presence of the multi-functional alcohol dehydrogenase AdhE as the last enzyme in the pathway constitutes an obstacle since this enzyme produces more n-butanol (C4) rather than n-hexanol (C6) in the pathway. Because production of n-butanol already rescues anaerobic growth of JCL166, this scheme does not provide a preferential selection for chain elongation.

[00178] Replacing AdhE with an enzyme that preferentially metabolizes C6 or higher acyl-CoA compounds would be necessary for this system to select for longer than C4 compounds. Thus, instead of AdhE, acyl-CoA thioesterase that preferentially consumes longer chain acyl-CoA as a substrate was used to produce linear fatty acids. This scheme allows the selection for chain elongation. Although the final product is an acid, this scheme can be used to select or enrich the upstream enzymes that favor chain elongation. Then the selected enzyme can be used in the alcohol producing pathway that uses AdhE or other alcohol dehydrogenases.

[00179] With this purpose in mind, the activity of the long chain acyl-CoA thioesterase was investigated. In both human and mouse, the long-chain acyl-CoA thioesterase catalyze the hydrolysis of fatty acyl-CoA into their corresponding free fatty acid and CoA.

These enzymes show broad specificity being able to hydrolyze acyl-CoA with carbon-chain length of C6 to C20. Very low or no activity for C4 or lower has been observed by these enzymes (Kuramochi et al., 2002; Yamada et al., 1999). Therefore, it was hypothesized that a long-chain specific acyl-CoA thioesterase would eliminate the production of C4 compounds and could relieve the metabolic stress created in the anaerobic NADH stress system and partially rescue cellular growth, and providing a platform to select enzymes that favor synthesis of higher chain-length-acyl-CoA compounds rather than C4 compounds (Figure 4D).

[00180] To test this hypothesis, mouse brain acyl-CoA hydrolase (mBACH) was cloned and expressed from mouse cDNA and tested its activity as His-tag purified protein, confirming the enzyme preference to hydrolyze C6 to C12 acyl-CoA substrates.

[00181] Next, the mBACH gene was used to replace adhE in the plasmid pEL11 creating the expression vector pHM10H. The *E. coli* host JCL166 without plasmids was unable to grow in anaerobic conditions under the test conditions (Fig. 27, first bar). However when plasmids pHM10H (expressing mBACH) and pHL101 were introduced in the cell some growth is observed (Fig. 27, 2nd bar). This result indicates that mBACH was able to rescue anaerobic growth of JCL166. Since mBACH favors longer chain acyl-CoA as a substrate, this strain construct can be used as a platform for chain elongation selection.

[00182] Screening for enzymes for the chain elongation pathway. Any of the enzymes in the second round of chain elongation could limit the pathway flux, and thus the ability to rescue growth in JCL166. Therefore, some enzymes were replaced in the pathway and test were performed to determine whether they would be able to rescue growth under anaerobic conditions in JCL166 strain. A new plasmid carrying *C. kluyveri* crt1 was constructed to replace *C. acetobutylicum* crt gene. *C. kluyveri* is a strict anaerobe able to ferment ethanol and acetate to produce butyric acid and hexanoic acid (Seedorf et al., 2008). Its natural ability to produce hexanoic acid suggests its crotonase enzyme might be more active towards C6 compounds than *C. acetobutylicum* crt. A plasmid carrying *R. eutropha* paaH1 was constructed to replace *C. acetobutylicum* hbd. The 3-OH-acyl-CoA reductase (PaaH1) from *R. eutropha* paaH1 has been shown to

be active against C4 to C10 substrates (Haywood et al., 1988), and therefore, it seemed a promising alternative to *C. acetobutylicum* hbd. When tested for anaerobic growth rescue, *C. kluyveri* crt1 decreased the cell density at 24 h (or 48 h, not shown). On the other hand *R. eutropha* paaH1 behaved similarly to the original homologue *C. acetobutylicum* hbd (2nd and 4th bars in Fig. 27).

[00183] However, when tested for hexanoic acid production in *E. coli* strain JCL299, the presence of *R. eutropha* paaH1 allowed production of 91 mg/L compared to undetectable amounts by *C. acetobutylicum* hbd. This result suggested that PaaH1 may be more efficient in vivo than in vitro and that the reduction of 3-ketohexanoyl-CoA in the pathway may be a limiting step to improve chain elongation. Based on this hypothesis, the anaerobic growth rescue platform was used to isolate evolved variants of PaaH1 that would improve growth rescue and hence production of C6 or higher chain-length products.

[00184] Selection of evolved PaaH1 by anaerobic growth rescue. Error-prone PCR was performed on the paaH1 gene in plasmid pHM10P using a commercially available low fidelity DNA polymerase (Mutazyme II, Agilent) and electroporated directly into a JCL166 strain carrying pHL101, generating a library of 26 pools of at least 2,000 clones each. Immediately after electroporation and 1 h cell recovery in LB medium, the library cultures were supplemented with 1% glucose, 100 μ M IPTG and antibiotics, and subjected to the first round of anaerobic growth rescue procedure. After 96 h incubation at 30 °C and rotatory agitation at 250 rpm, the cultures were diluted 160 fold and subjected to a second round of anaerobic growth rescue. Fig. 28 shows the results after two rounds of growth rescue. Plasmids were isolated from thirteen library cultures with optical densities equal or higher than that observed for library plib25. After transformation in *E. coli* XL1-Blue strain, two independent plasmids from each selected library were isolated and again screened for anaerobic growth rescue in JCL166. After DNA sequencing, nine libraries were analyzed, the two isolated clones were identical, while in the other four libraries the two individual clones showed different mutations. These results showed that the selection platform is able to greatly enrich a particular mutant

from a mixed mutant population based on its ability improve the cell survival and growth under the tested conditions. Five different clones from five libraries were selected for further testing and renamed these variants as P01, P042, P06, P09, and P13. Their ability to rescue growth of JCL166 under anaerobic conditions were tested (Fig. 27). The protein sequence alignment for the selected clones is shown in Fig. 26B. Interestingly, PaaH1-P01 variant contains only a silent mutation on a leucine codon, CTG>CTA. This codon change from a higher codon usage to a rare one in *E. coli* seemed to be enough to improve the strain fitness under the selection strategy, perhaps by avoiding depletion of the cognate tRNA and allowing more efficient translation (Kane, 1995; Kurland and Gallant, 1996).

[00185] PaaH1 variants improved hexanoic acid production. The addition of a thioesterase (mBACH) at the last step in the selection platform pathway should result in the production of hexanoic acid. Therefore, the amount of hexanoic acid produced by the selected paaH1 mutants was analyzed compared to the wild type gene (Fig. 29). Production was carried out in *E. coli* JCL299 which has been previously shown to produce higher titers of n-butanol than JCL166 (Shen et al., 2011). As expected, increased production of hexanoic acid was observed for most PaaH1 variants compared to the wild type PaaH1 after 60 h fermentation.

[00186] Improved n-hexanol and n-octanol production by *E. coli*. The evolved PaaH1 mutants were tested to determine if they improve production of n-hexanol. For this purpose, the *C. acetobutylicum* adhE gene was used to replace mBACH in the pHM10 plasmid series to create the pHM25 plasmid series. These plasmids were transformed into *E. coli* JCL299 in conjunction with PHL101, which carries the other n-hexanol pathway genes, and tested for alcohol production after 48 h fermentation on glucose (Fig. 30). Replacing the original *C. acetobutylicum* hbd by the *R. eutropha* paaH1 wild-type gene already improves the titer on n-hexanol production by 10 fold, from ~30 mg/L to 280 mg/L after 48 h fermentation. This result was not surprising since increased hexanoic acid production was previously observed on the strain carrying paaH1 compared to hbd (Fig. 29). In a similar way, an

increase of 67% to 469 mg/L n-hexanol for the strains expressing the paaH1 mutants compared to that expressing the wild type gene was observed.

[00187] The methods of the disclosure were used to develop an anaerobic growth rescue selection platform for chain elongation in the reverse β -oxidation pathway for linear alcohol synthesis. This method is based on redox imbalance-induced growth inhibition in *E. coli* (Shen et al., 2011). The host strain, JCL166 (Δ ldhA Δ adhE Δ frdBC) in which the fermentative pathways were deleted, cannot grow anaerobically in glucose medium unless an NADH consuming pathway is introduced or an external electron acceptor is present. For instance, introduction of the n-hexanol pathway consumes six moles of NADH to produce one mole of n-hexanol thus recycling NAD⁺ and CoA (Dekishima et al., 2011). However, this n-hexanol pathway relies on the activity of the *C. acetobutylicum* bifunctional acetaldehyde-CoA/alcohol dehydrogenase (AdhE, also named AdhE2) at the last step to convert hexanoyl-CoA to hexaldehyde and finally hexanol. Owing to AdhE's low specificity, production of n-butanol outcompetes for the production of n-hexanol, and could not select for chain elongation. Elimination of AdhE results in accumulation of intermediates containing CoA, which are generally toxic. To create a platform that would allow for increased chain length of linear alcohols, a mammalian long-chain specific thioesterase, mBACH, was introduced to produce longer-chain fatty acid, e.g., hexanoic acid, and consume the accumulated acyl-CoA and allow recycling of CoA cofactor and NADH, and at the same time, avoiding or reducing the production of C4 products.

[00188] Using this approach to enrich for evolved mutants after a single round of error-prone PCR, several mutants were isolated that showed increased production of hexanoic acid and n-hexanol compared to the wild-type strain. It also helped to pinpoint the ketoacyl-CoA reduction step as possible limiting step in the pathway, and, therefore, a target for protein directed evolution. Two rounds of growth rescue were enough to enrich for improved variants of PaaH1 and allow easy isolation of individual variant clones. DNA sequencing analyses of the enriched variants showed that most carry mutations on the first half of the protein on the NAD-binding domain

which could have improved the affinity for the cofactor improving the enzyme activity (Rodriguez-Zavala, 2008). Curiously, one mutant that improved anaerobic cell growth and n-hexanol production contained only a silent mutation which replaced a frequently used leucine codon CTG with a rare codon in *E. coli*, CTA. One possible explanation for this phenotype could be the depletion of frequently used tRNA's due to overproduction of many enzymes in the *E. coli* strain, a phenomenon previously referred as "codon hunger" (Kane, 1995; Kurland and Gallant, 1996). The presence of a rare codon could alleviate the over usage of the CTG codon. Codon optimization by codon usage bias (or codon adaptation index, CAI) is a frequently used method for superexpression of heterologous genes in *E. coli*, but it does not seem to present consistent results. Changes in the mRNA structure could affect the mRNA stability and improve overall expression of the proteins.

[00189] The disclosure thus provides a selection platform for enzymes that favors higher chain alcohols in a reverse β -oxidation pathway. Using this platform a homolog protein that increased n-hexanol production by 10-fold was identified and then evolved to gain an additional 67% increase. The success of this selection and enrichment platform also suggests it can be used for further evolution or bioprospecting new genes from pools of genomic or expression libraries avoiding the cost- and time consuming single colony screening approach.

[00190] A number of embodiments of the invention have been described. Nevertheless, it will be understood that various modifications may be made without departing from the spirit and scope of the invention. Accordingly, other embodiments are within the scope of the following claims.

WHAT IS CLAIMED IS

1. A recombinant microorganism that produces a C6-8 alcohol comprising a recombinant pathway including a heterologous beta-ketothiolase and one or more enzymes selected from the group consisting of an acetyl-CoA acetyltransferase, a 3-hydroxybutyryl-CoA dehydrogenase, a crotonase, a trans-2-enoyl-CoA reductase and a bifunctional aldehyde/alcohol dehydrogenase.

2. The recombinant microorganism of claim 1, comprising a heterologous polypeptide having beta ketothiolase activity and a heterologous polypeptide having acetyl-coA acetyltransferase activity.

3. The recombinant microorganism of claim 1 or 2, comprising heterologous expression or elevated expression of a polypeptide having trans-2-enoyl-CoA reductase activity, as compared to a parental microorganism, wherein the recombinant microorganism produces a metabolite comprising butyryl-CoA from a substrate that includes crotonyl-CoA.

4. The recombinant microorganism of claim 3, wherein the trans-2-enoyl-CoA reductase (ter) is derived from a *Treponema denticola* or *F. succinogenes*.

5. The recombinant microorganism of claim 4, wherein the polypeptide having trans-2-enoyl-CoA (Ter) polypeptide activity is at least 50% identical to a sequence as set forth in SEQ ID NO:68, 69, 70, or 71.

6. The recombinant microorganism of claim 5, wherein the Ter comprises an M11K substitution.

7. The recombinant microorganism of claim 1 or 2, comprising heterologous expression or elevated expression of a polypeptide having acetyl-CoA acetyltransferase activity, as compared to a parental microorganism, wherein the recombinant microorganism

produces a metabolite comprising acetoacetyl-CoA from a substrate comprising acetyl-CoA.

8. The recombinant microorganism of claim 7, wherein the polypeptide having acetyl-CoA acetyltransferase activity is encoded by a polynucleotide having at least about 50% identity to a sequence as set forth in SEQ ID NO:29.

9. The recombinant microorganism of claim 7, wherein the polypeptide having acetyl-CoA acetyltransferase activity is encoded by an *atoB* gene or homolog thereof, or a *fadA* gene or homolog thereof.

10. The recombinant microorganism of claim 9, wherein the *atoB* gene or *fadA* gene is derived from the genus *Escherichia*.

11. The recombinant microorganism of claim 10, wherein the *Escherichia* is *E. coli*.

12. The recombinant microorganism of claim 1 or 2, comprising heterologous expression or elevated expression of a polypeptide having hydroxybutyryl-CoA dehydrogenase activity, as compared to a parental microorganism, wherein the recombinant microorganism produces a metabolite comprising 3-hydroxybutyryl-CoA from a substrate comprising acetoacetyl-CoA.

13. The recombinant microorganism of claim 12, wherein the polypeptide having hydroxybutyryl-CoA activity is encoded by a polynucleotide having at least about 50% identity to a sequence as set forth in SEQ ID NO:35, 72, 73, 74, 75, 76 or 77.

14. The recombinant microorganism of claim 11, wherein the hydroxybutyryl-CoA dehydrogenase is encoded by an *hbd* gene or homolog thereof or a *paah1* gene or homolog thereof.

15. The recombinant microorganism of claim 14, wherein the *hbd* gene is derived from a microorganism selected from the group

consisting of *Clostridium acetobutylicum*, *Clostridium difficile*, *Dastricha ruminatum*, *Butyrivibrio fibrisolvens*, *Treponema phagedemes*, *Acidaminococcus fermentans*, *Clostridium kluyveri*, *Syntrophospora bryanti*, and *Thermoanaerobacterium thermosaccharolyticum*.

16. The recombinant microorganism of claim 15, wherein the microorganism is *Clostridium acetobutylicum*.

17. The recombinant microorganism of claim 1 or 2, comprising heterologous expression or elevated expression of a polypeptide having crotonase activity, as compared to a parental microorganism, wherein the recombinant microorganism produces a metabolite comprising crotonyl-CoA from a substrate comprising 3-hydroxybutyryl-CoA.

18. The recombinant microorganism of claim 17, wherein the polypeptide having crotonase activity is encoded by a polynucleotide having at least about 50% identity to a sequence as set forth in SEQ ID NO:51.

19. The recombinant microorganism of claim 18, wherein the crotonase is encoded by a *crt* gene or homolog thereof.

20. The recombinant microorganism of claim 19, wherein the *crt* gene is derived from a microorganism selected from the group consisting of *Clostridium acetobutylicum*, *Butyrivibrio fibrisolvens*, *Thermoanaerobacterium thermosaccharolyticum*, and *Clostridium difficile*.

21. The recombinant microorganism of claim 20, wherein the microorganism is *Clostridium acetobutylicum*.

22. The recombinant microorganism of claim 1 or 2, comprising heterologous expression or elevated expression of a polypeptide having aldehyde/alcohol dehydrogenase activity, as compared to a parental microorganism, wherein the recombinant microorganism

produces a metabolite comprising buteraldehyde from a substrate comprising butyryl-CoA.

23. The recombinant microorganism of claim 22, wherein the polypeptide having aldehyde/alcohol dehydrogenase activity is encoded by a polynucleotide having at least about 50% identity to a sequence as set forth in SEQ ID NO:67.

24. The recombinant microorganism of claim 22, wherein the polypeptide having aldehyde/alcohol dehydrogenase is encoded by an *aad* gene or homolog thereof, or an *adhE2* gene or homolog thereof.

25. The recombinant microorganism of claim 24, wherein the *aad* gene or *adhE2* gene is derived from *Clostridium acetobutylicum*.

26. The recombinant microorganism of claim 1 or 2, comprising elevated expression of a polypeptide having beta-ketothiolase activity, as compared to a parental microorganism, wherein the recombinant microorganism produces a metabolite comprising 3-ketohexanoyl-CoA from substrates comprising butyryl-CoA and acetyl-CoA.

27. The recombinant microorganism of claim 26, wherein the polypeptide having beta-ketothiolase activity is encoded by a *bktB* gene or homolog thereof.

28. The recombinant microorganism of claim 27, wherein the *bktB* is derived from *R. eutropha*.

29. The recombinant microorganism of claim 27, wherein the *bktB* gene or homolog thereof comprises a sequence having at least 50% identity to SEQ ID NO:37.

30. The recombinant microorganism of any one of the foregoing claims, wherein the microorganism comprises a knockout in a gene selected from the group consisting of *ldhA*, *adhE*, *frdBC* and any combination thereof.

31. The recombinant microorganism of claim 1, wherein the microorganism expresses polypeptides having beta-ketothiolase activity, acetyl-CoA acetyltransferase activity, 3-hydroxybutyryl-CoA dehydrogenase activity, crotonase activity, trans-2-enoyl-CoA reductase activity and aldehyde/alcohol dehydrogenase activity.

32. The recombinant microorganism of claim 1, 2 or 31, wherein the microorganism comprises an *atoB*, *BktB*, *hbd* or *paaH1*, *crt* and a *ter* gene.

33. A method of making a recombinant microorganism of claim 1, comprising transforming a parental cell with one or more vectors for expressing a polypeptide having beta-ketothiolase activity, acetyl-CoA acetyltransferase activity, 3-hydroxybutyryl-CoA dehydrogenase activity, crotonase activity, trans-2-enoyl-CoA reductase activity and/or bifunctional aldehyde/alcohol dehydrogenase activity.

34. A method of making n-hexanol and/or octanol comprising culturing a recombinant microorganism of any one of claims 1-29 with a substrate and under conditions to produce n-hexanol and/or octanol.

35. The recombinant microorganism or method of any of the preceding claims wherein the C6-C8 alcohol are selected from n-hexanol and n-octanol.

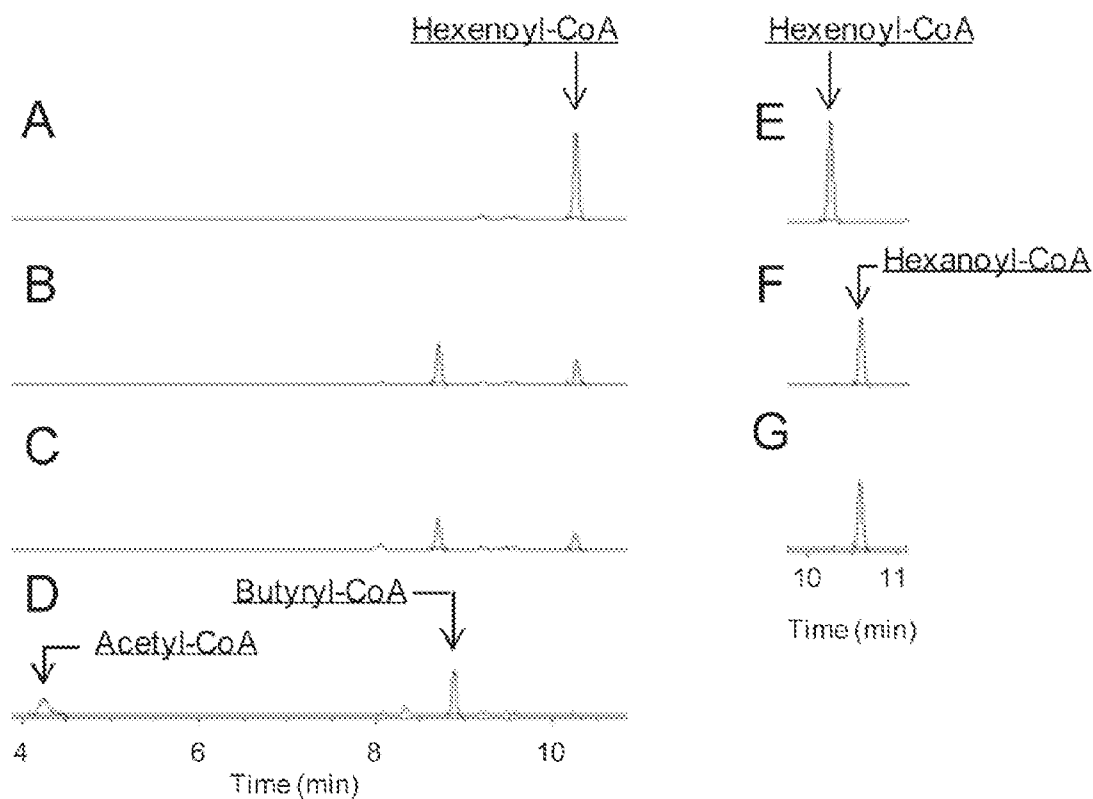


Figure 1A-G

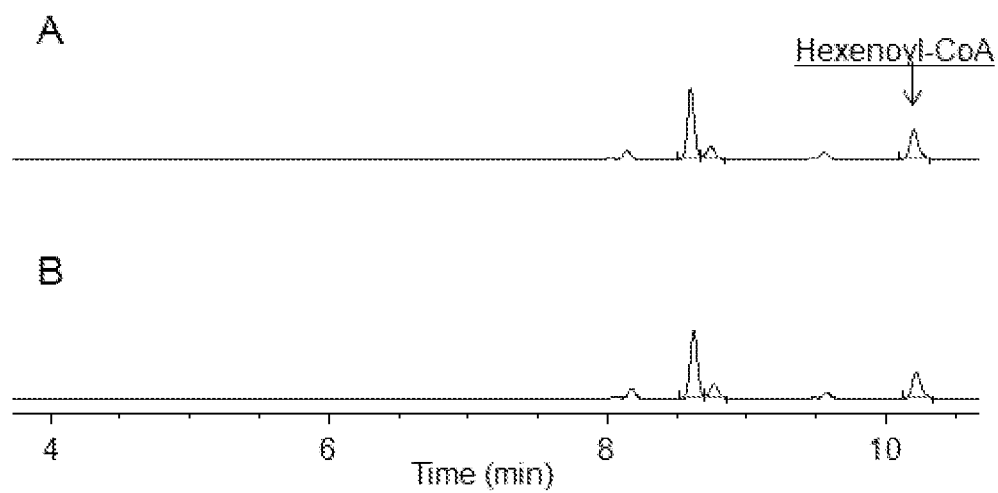


Figure 2A-B

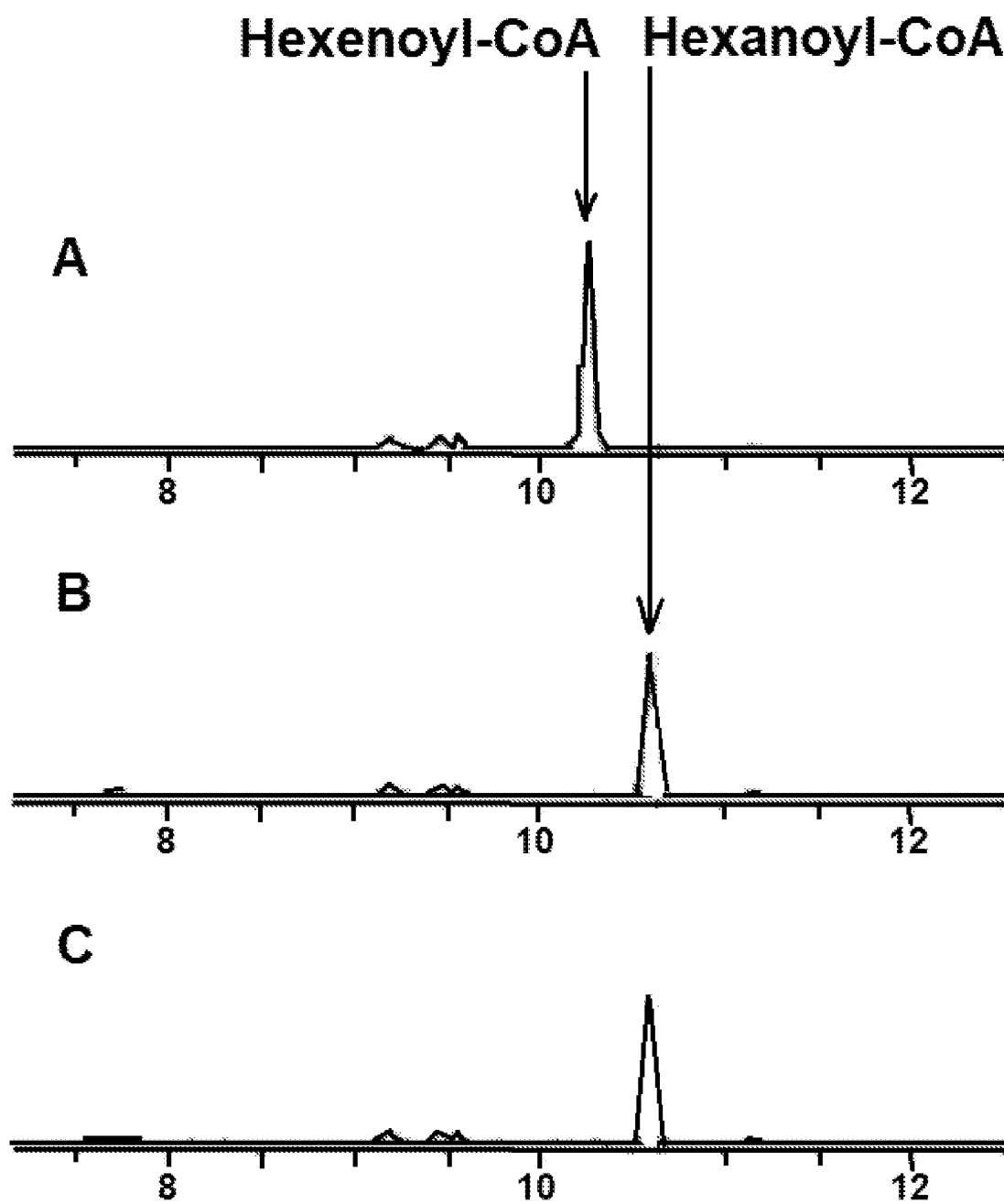
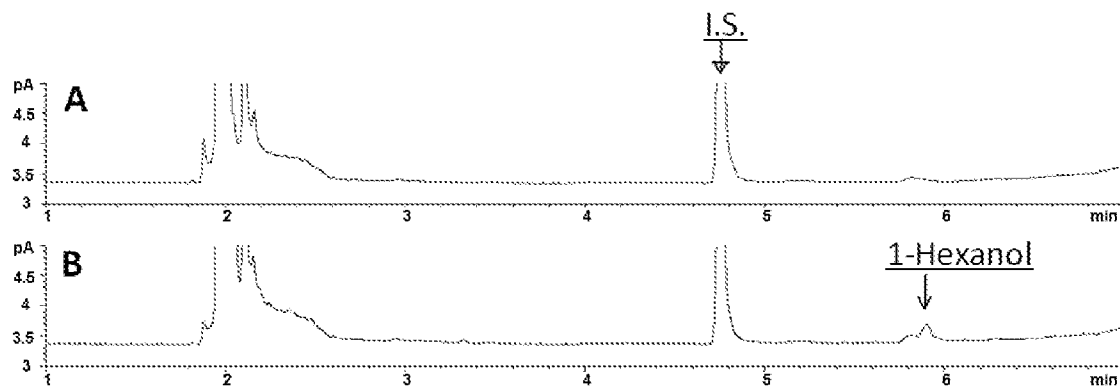


Figure 3A-C



C

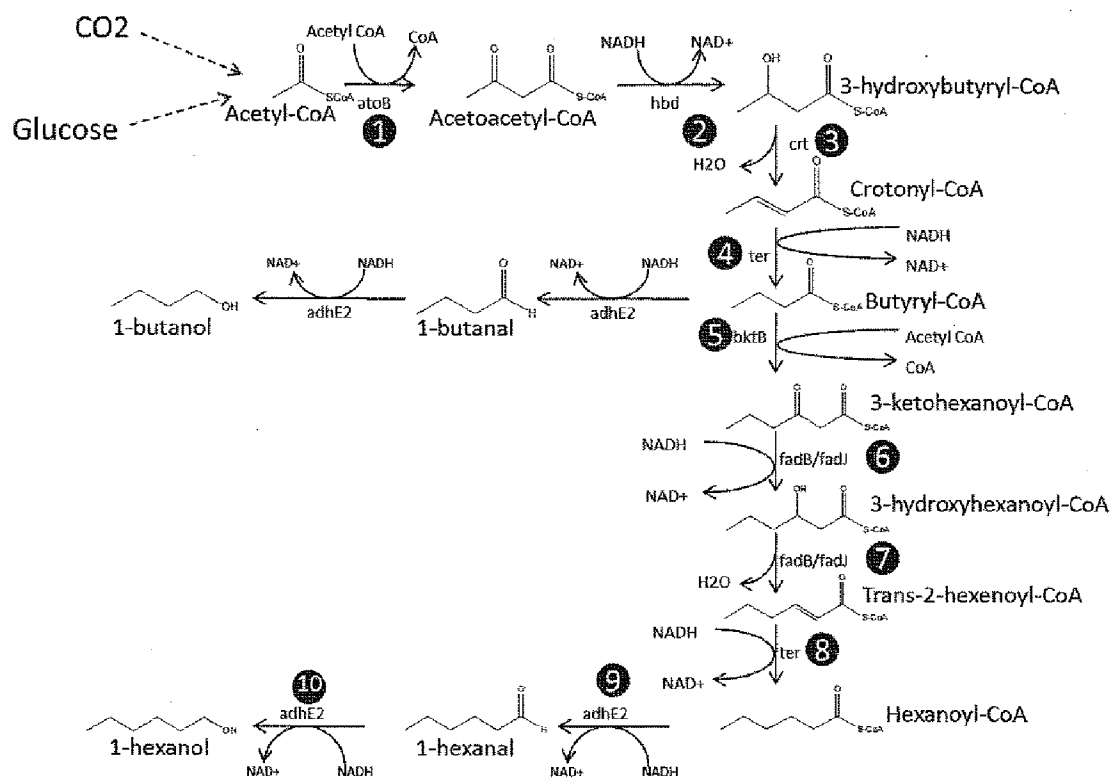


Figure 4A-C

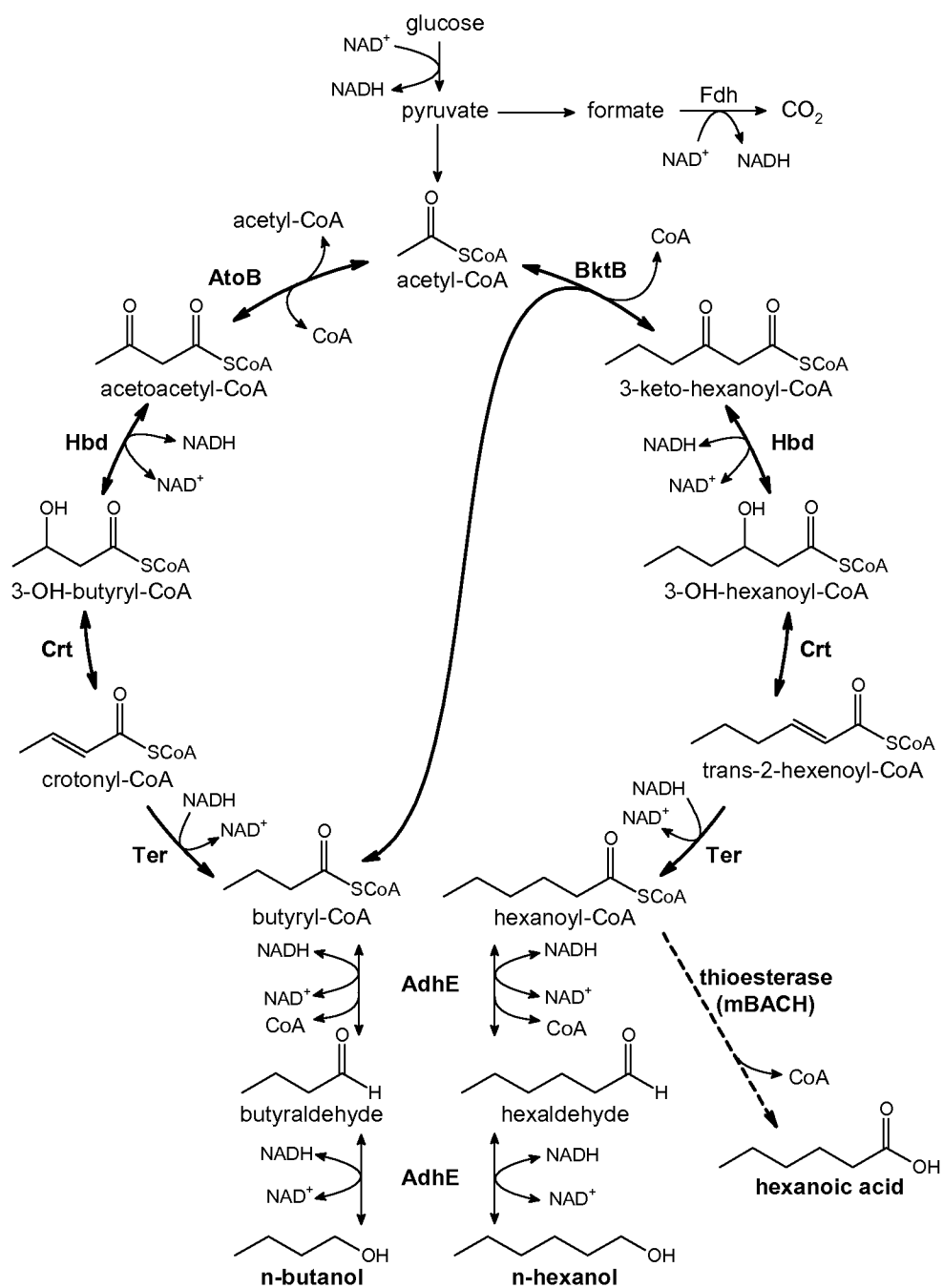


FIGURE 4D

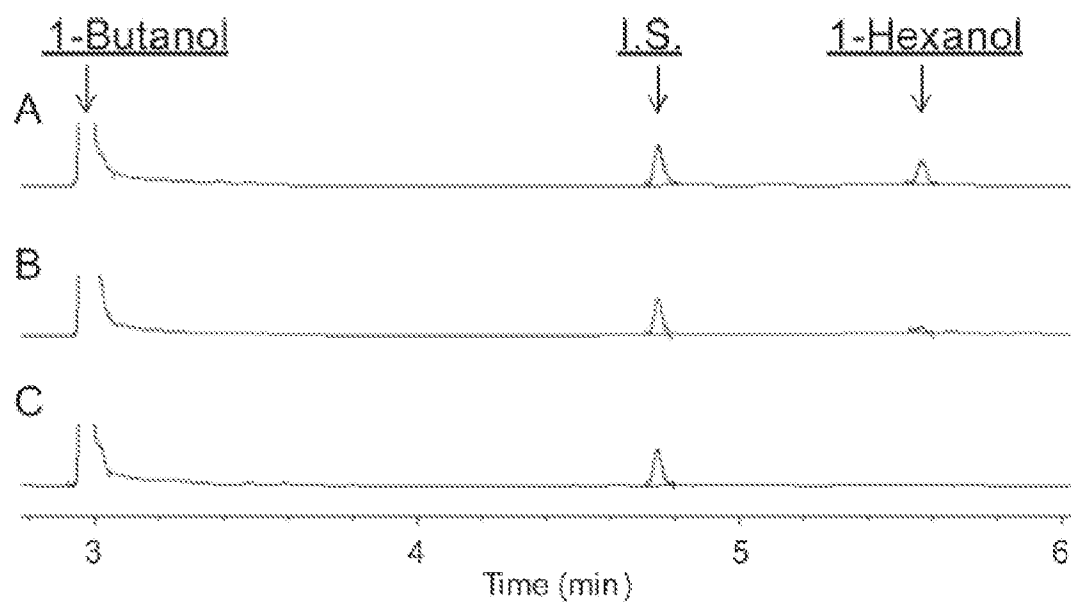


FIGURE 5

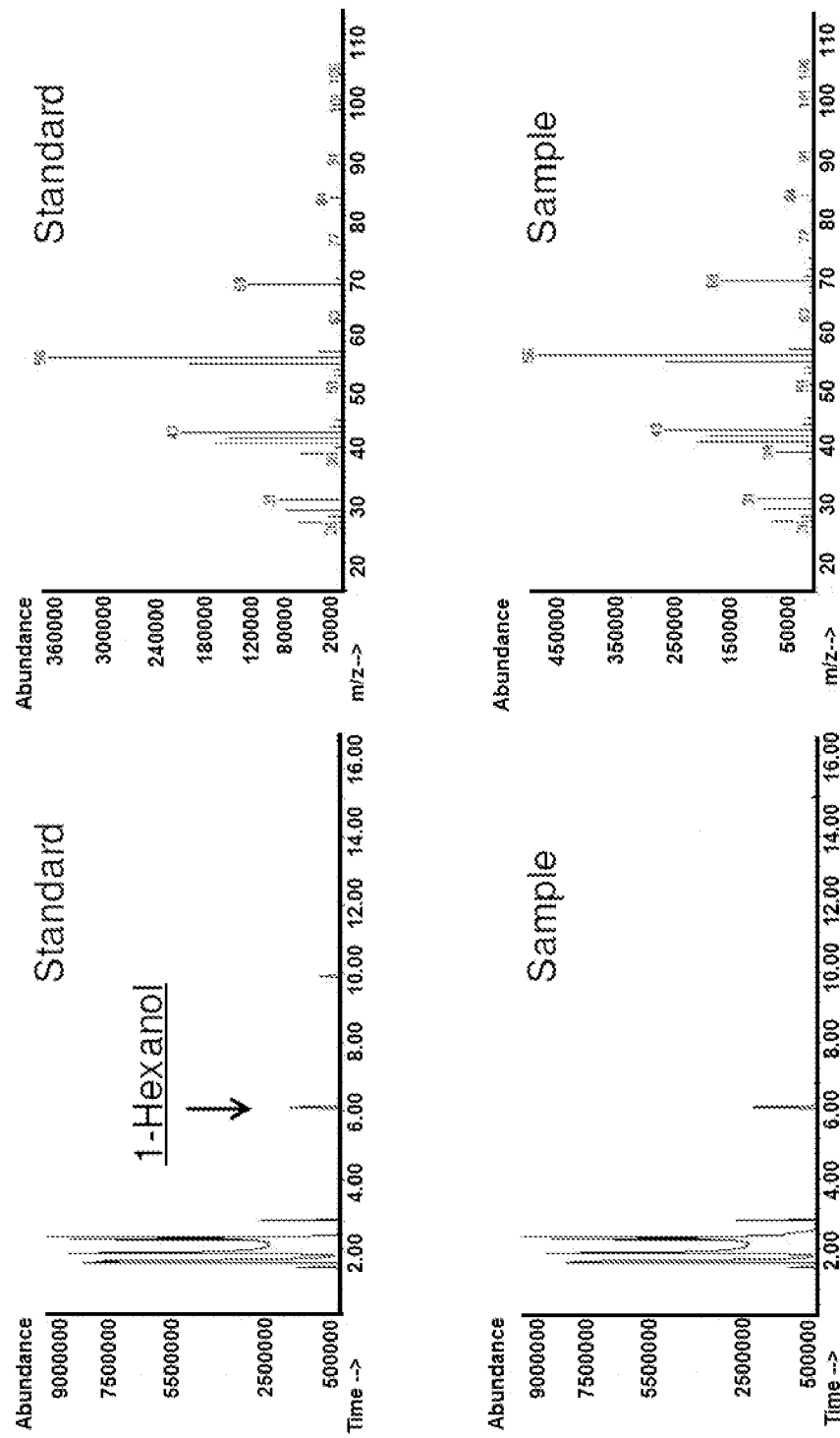


FIGURE 6

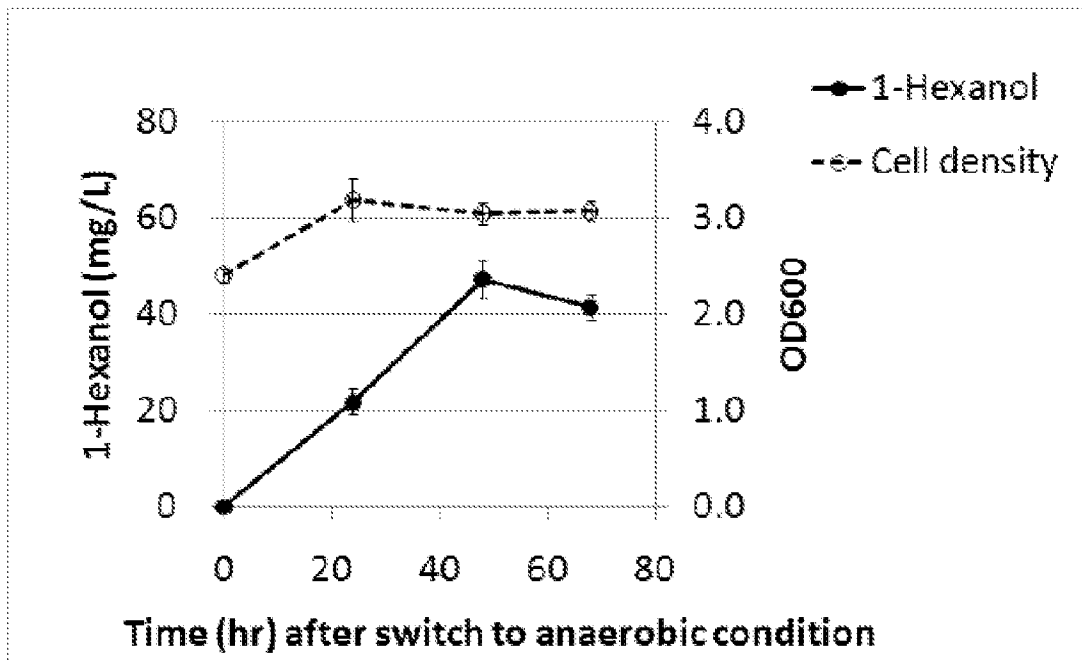


FIGURE 7

atoB (*E. coli*) SEQ ID NO:29

atgaaaaattgtgtcatcgtcagtgcggtacgtactgctatcggtagttttaacggttcaactcgcttccaccagc
gccatcgacctgggggacagtaattaaagccgccattgaacgtgcaaaaatcgattcacaacacggttgatgaa
gtgattatgggtaacgtgttacaagccgggctggggcaaaaatccggcgcgctcaggcactgttaaaagcgggctg
gcagaaacgggtgtgcggttacacgggtcaataaagtatgtggttcgggtcttaaaagtgtggcgcttgccgcccag
gccattcaggcaggtcaggcgcagagcattgtggcggggggtatggaaaaatagagtttagccccctacttactc
gatgcaaaagcacgctctggttatcgtcttggagacggacaggtttatgacgtaatcctgcgcgatggcctgatg
tgcgccacccatgggttatcatatggggattaccgcccgaacagtggttaaagagtagcgaattaccggtgaaatg
caggatgaactggcgctacattcacagcgtaaagcggcagccgaattgagtcgggtgcttttacagccgaaatc
gtcccggtaaatgttgtcactcgaaagaaaaccttcgtcttcagtcaagacgaattcccgaagcgaattcaacg
gctgaagcggttaggtgcattgcgccccgcttcgataaagcaggaacagtcaccgctgggaacgcgtctggtatt
aacgacgggtgctgccgctctggtgattatggaagaatctgcggcgctggcagcaggccttacccttggctcgc
attaaaagtattgccagcggtggcggtgccccccgcatatgatgggtatggggccagtagcctgccacgcaaaaagcg
ttacaactggcggggctgcaactggcggtatattgatctcattgaggctaataagcatttgctgcacagttcctt
gccgttgggaaaaacctgggcttggattctgagaaagtgaatgtcaacggcgggggccatcgcgctcgggcacacct
atcggtgccagtggtgctcgtattctggtcacactattacatgccatgcaggcacgcgataaaacgctggggctg
gcaacactgtgcattggcgcggtcagggaattgcgatggtgattgaacggttgaattaa

Figure 8

thlA (*Clostridium acetobutylicum*) SEQ ID NO:31

atgaagaagttgtaatagctagtgcagtaagaacagcgattggatcttatggaaagtctcttaaggatgtacca
gcagtagatttaggagctacagctataaaggaagcagttaaaaaagcaggaataaaaccagaggatgttaatgaa
gtcatttttaggaaatgttcttcaagcagggttaggacagaatccagcaagacaggcatcttttaagcaggatta
ccagttgaaattccagctatgactattaataaggtttgtggttcaggacttagaacagtttagcttagcagcacia
attataaaagcaggagatgctgacgtaataatagcagggtggtatggaaaaatgtctagagctccttacttagcg
aataacgctagatggggatatagaatgggaaacgctaaatttgttgatgaaatgatcactgacggattgtgggat
gcatttaattgattaccacatgggaataacagcagaaaaacatagctgagagatggaacatttcaagagaagaacia
gatgagtttgctcttgcacacaaaaaaagctgaagaagctataaaatcagggtcaatttaagatgaaatagtt
cctgtagtaattaaaggcagaaagggagaaactgtagttgatacagatgagcaccctagatttggatcaactata
gaaggacttgcaaaattaaaacctgccttcaaaaaagatggaacagttacagctggttaatgcatcaggattaaat
gactgtgcagcagtagttgtaatcatgagtgcagaaaaagctaaagagcttgagtaaaaccacttgctaagata
gtttcttatggttcagcaggagttgaccagcaataatgggatatggaccttctatgcaacaaaagcagctatt
gaaaaagcagggttgacagttgatgaattagatttaataagaatcaaatgaagcttttgacgctcaaagtttagca
gtagcaaaagatttaaaatttgatatgaataaagtaaatgtaaatggaggagctattgcccttggctcatccaatt
ggagcatcagggtgcaagaatactcgttactctgtacacgcaatgcaaaaaagagatgcaaaaaagggcttagca
actttatgtataggtggcgagacaaggaacagcaatattgctagaaaagtgtag

Figure 9

crt (*Clostridium acetobutylicum*) SEQ ID NO:33

atggaactaaacaatgtcatccttgaaaaggaaggttaaagttgctgtagttaccattaacagacctaagcatta
aatgcgttaaatagtgatacactaaaagaaatggattatgttataggtgaaattgaaaatgatagcgaagtactt
gcagtaattttaactggagcaggagaaaaatcattttagcaggagcagatatttctgagatgaaggaaatgaat
accattgaaggtagaaaaattcgggatacttgaaaataaagtgtttagaagattagaacttcttgaaaagcctgta
atagcagctgttaatggttttgctttaggaggcggtgcgaaatagctatgtcttgtgatataagaatagcttca
agcaacgcaagatttggtaaccagaagtaggtctcggaataacacctggttttgggtgtacacaaagactttca
agattagttggaatgggcatggcaaagcagcttatatttactgcacaaaatataaaggcagatgaagcattaaga
atcggacttgtaataaaggtagtagaacctagtgaattaatgaatacagcaaaagaaattgcaaacaaaattgtg
agcaatgctccagtagctgttaagtttaagcaaacaggctattaatagaggaatgcagtgtagattgatactgct
ttagcatttgaatcagaagcatttggagaatgcttttcaacagaggatcaaaaggatgcaatgacagctttcata
gagaaaagaaaaattgaaggcttcaaaaatagatag

Figure 10

hbd (*Clostridium acetobutylicum*) SEQ ID NO:35

atgaaaaaggtatgtgttataggtgcaggtactatgggttcaggaattgctcaggcatttgcagctaaaggattt
gaagtagtattaagagatatataagatgaatttggtagatagaggattagattttatcaataaaaatctttctaaa
ttagttaaaaaaggaaagatagaagaagctactaaagttgaaatcttaactagaatttccggaacagttgacctt
aatatggcagctgattgcgatttagttatagaagcagctgttgaaagaatggatattaaaaagcagatttttgc
gacttagacaatatatgcaagccagaaacaattcttgcatcaaatacatcatcactttcaataacagaagtggca
tcagcaactaaaagacctgataaggttataggtatgcatttctttaatccagctcctgttatgaagctttagag
gtaataagaggaatagctacatcacagaactttttagtcaggttaaagagacatctatagcaataggaaaagat
cctgtagaagtagcagaagcaccaggatttgttgtaataagaatattaataccaatgattaatgaagcagttgg
atattagcagaaggaatagcttcagtagaagacatagataaaagctatgaaacttgagctaatcacccaatggga
ccattagaattaggtgattttataggtcttgatataatgcttgcataatggatgttttatactcagaaactgga
gattctaagtatagaccacatacattacttaagaagtatgtaagagcaggatggcttggaagaaatcaggaaaa
ggtttctacgattattcaaaataa

BktB (beta-ketothiolase) (*Ralstonia eutropha*) (SEQ ID NO:37)

atgacgcgtgaagtggtagtgtaagcgggtgtccgtaccgcgatcgggacctttggcggcagcctgaaggatgtg
gcaccggcgagctgggcgcactgggtgggtgcgcgagcgctggcgcgcgcgaggtgtcgggcgacgatgtcggc
cacgtgggtattcggcaacgtgatccagaccgagcgcgcgacatgtatctgggcgcgctcgcggccgtcaacggc
ggggtgacgatcaacgccccgcgctgacgtgaaccgcctgtgcggctcgggcctgcaggccattgtcagcgcc
gcgcgacccatcctgctgggcgataccgacgtcgccatcggcggcgggcgcggaagcatgagccgcgcaccgtac
ctggcgccggcagcgcgctggggcgacgcacgtgggcgacgcggcctgggtcgacatgatgctgggtgcgctgcac
gatcccttccatcgcatccacatgggcgtgaccgcccagaaatgtcgccaaggaatacgcacatctcgcgcgcgag
caggacgaggccgcgctggaatcgcaaccgcgcgcttcggcagcgatcaaggccggctacttcaaggaccagatc
gtcccgggtggtgagcaaggccgcaaggcgacgtgaccttcgacacccgacgagcacgtgcgccatgacgccacc
atcgacgacatgaccaagctcaggccgggtcttcgtcaaggaaaaacggcacgggtcacggccggcaatgcctcgggc
cgctggtgtcgtacggccatgccggcggtggaccgaaggccatgggcatcggcccggtgcggcgacgaagatc
gcgctggagcgcgccggcctgcaggtgtcgacactggacgtgatcgaagccaacgaagcctttgccgcacaggcg
tgcgccgtgaccaaggcgctcggtctggaccggccaaaggttaaccgaacggctcgggcatctcgctgggccac
ccgatcggcgccaccgggtgccctgatcacgggtgaaggcgctgcagctgaaccgcgtgcagggccgctacgcg
ctgggtgacgatgtgcatcgcgcgggcaggccattgccgccatcttcgagcgtatctga

Figure 11

bcd (*Clostridium acetobutylicum*) SEQ ID NO:39

atggatttttaatttaacaagagaacaagaattagtaagacagatgggttagagaatttgctgaaaatgaagttaaa
cctatagcagcagaaaattgatgaaacagaaagatttccaatggaaaatgtaaagaaaatgggtcagtatgggatg
atgggaattccattttcaaaagagtatgggtggcgaggtggagatgtattatcttatataatcgcggttgaggaa
ttatcaaaggtttgcggtactacaggagttattctttcagcacatacatcactttgtgcttcattaataaatgaa
catggtacagaagaacaaaaacaaaaatatttagtacctttagctaaaggtgaaaaaataggtgcttatggattg
actgagccaaatgcaggaaacagattctggagcacacaacacagtagctgtacttgaaggagatcattatgtaatt
aatgggttcaaaaatattcataactaatggaggagttgcagatacttttggttatatttgcaatgactgacagaact
aaaggaacaaaaggtatatcagcatttataatagaaaaaggcttcaaaggtttctctattggtaaagttgaacaa
aagcttggataagagcttcatcaacaactgaacttgtatttgaagatatgatagtaccagtagaaaaacatgatt
ggtaaagaaggaaaaggcttccctatagcaatgaaaactcttgatggaggaagaattgggtatagcagctcaagct
ttaggtatagctgaaggtgctttcaacgaagcaagagcttacatgaaggagagaaaaacaatttggagaagcctt
gacaaattccaaggtcttgcatggatgatggcagatatggatgtagctatagaatcagctagatatatttagtatat
aaagcagcatatcttaaacaaagcaggacttccatacacagttgatgctgcaagagctaagcttcatgctgcaaat
gtagcaatggatgtaacaactaaggcagtagacaattatttgggtggatagcgatatacaaaaagattatccagttgaa
agaatgatgagagatgctaagataactgaaatatatgaaggaacttcagaagttcagaaaattagttatttcagga
aaaatttttagataa

Figure 12

etfA (*Clostridium acetobutylicum*) SEQ ID NO:41

atgaataaagcagattacaaggcgctatgggtgtttgctgaacaaagagacggagaattacaaaagggtatcattg
gaattattaggtaaaggtaaggaaatggctgagaaattaggcggttgaattaacagctgttttacttggacataat
actgaaaaaatgtcaaaggatttattatctcatggagcagataagggttttagcagcagataatgaacttttagca
cattttttcaacagatggatagctaaagttatatgtgatttagttaatgaaagaaagccagaaatattattcata
ggagctactttcataggaagagatttaggaccaagaatagcagcaagactttctactggtttaactgctgattgt
acatcacttgacatagatgtagaaaaatagagatttattggctacaagaccagcggttgggtggaatttgatagct
acaatagtttgttcagaccacagaccacaaatggctacagtaagacctgggtgtgtttgaaaaattacctgttaat
gatgcaaatgtttctgatgataaaatagaaaaagttgcaattaaattaacagcatcagacataagaacaaaagtt
tcaaaagttgttaagcttgctaaagatatcagcagatatcgagagaagctaaggattagttgctggtggtagagga
gttggaaagcaaaagaaaactttgaaaaacttgaagagttagcaagtttacttgggtggaacaatagccgcttcaaga
gcagcaatagaaaaagaatgggttgataaggaccttcaagtaggtcaaaactggtaaaactgtaagaccaactctt
tatattgcatgtggtatatcaggagctatccagcatttagcaggtatgcaagattcagattacataattgctata
aataaagatgtagaagcccccaataatgaaggtagcagatttggtctatagttggtgatgtaataaagttgtacca
gaattaatagctcaagttaaagctgctaataattaa

Figure 13

etfB (*Clostridium acetobutylicum*) SEQ ID NO:43

atgaatatagttgtttgttttaaaacaagttccagatacagcggaagttagaatagatccagttaagggaaacactt
ataagagaaggagttccatcaataataaatccagatgataaaaacgcacttgaggaagcttttagtattaaaagat
aattatgggtgcacatgtaacagttataagtatgggacctccacaagctaaaaatgcttttagtagaagctttggct
atgggtgctgatgaagctgtacttttaacagatagagcatttggaggagcagatacacttgcgacttcacataca
attgcagcaggaattaagaagctaaaaatgatataagtttttgcgtggaaggcaggctatagatggagatacagct
caggttggaccagaaatagctgagcatcttggaaacacccaagtaacttatgttgagaaagttgaagttgatgga
gatacttttaagattagaaaagcttgggaagatggatatgaagttggtgaagttaagacaccagttccttttaaca
gcaattaaagaattaaatgttccaagatatatgagtgtagaaaaaatattcggagcatttgataaagaagtaaaa
atgtggactgccgatgatataagatgtagataaggctaatttaggtcttaaagggtcaccaactaaagttaagaag
tcatcaactaaagaagttaaaggacagggagaagttattgataagcctgttaaggaagcagctgcatatgttgctc
tcaaaattaaaagaagaacactatatttaa

Figure 14

bcd (*Megasphaera elsdenii*) SEQ ID NO:45

atggatttttaacttaacagatattcaacaggacttcttaaaactcgctcatgatttcggcgaaaagaaattagca
ccgaccgttacggaacgcgaccacaaaggtatttatgacaaagaactcatcgacgaattgctcagcctcggtatt
accggcgcttacttcgaagaaaaatacggcggttccggcgatgacggcggcgacgtttttgagctacatcctcgct
gttgaagaattggctaaatacgcagctgggtgttgctatcaccttgctggcaacgggtttccctttgctgaacccg
atttggcagttcggtacagaagctcagaagaaaaatcctcgcttccctttggttgaaggcactaaactcggcgct
ttcggttgaccgaaccgaacgcaggtactgatgcttccggccagcagaccattgctacgaagaacgatgacggc
acttacacgttgaacggctccaagatcttcatcaccaacggcggcgctgctgacatctacattgtcttcgctatg
accgataagagcaaaggcaaccacggcattacagccttcatcctcgaagacggtaactccgggctttacttacggc
aagaaagaagacaagatgggcatccatacttcgcagaccatggaactcgtattccaggacgtcaaagttccggct
gaaaacatgctcggcgaagaaggcaaggcttcaagattgctatgatgaccttgacggcggcgctatcggcgctt
gctgctcaggctctcggcattgcagaagctgctttggcagatgctggtgaatactccaaacagcgtgtacagttc
ggcaaacgctctgcaaattccagtcatttcttcaaactggctgacatgaagatgcagatcgaagctgctcgt
aacctcgtttacaaagctgcttgcagaacaggaaggcaaaccttcaccggttgacgctgctatcgcaaacgc
gttgcttccgacgtcgctatgcgcgtaacgaccgaagctgtccagatcttcggcggtatggctacagcgaagaa
tatccgggttgctcgtcacatgcgcgatgctaagattactcagatctacgaaggcacgaacgaagttcagctcatg
gttacaggcggtgctctgttaagataa

Figure 15

etfA (*Megasphaera elsdenii*) SEQ ID NO:47

atggatttagcagaatataaaagccatttatgttaattgctgaacagttcgaaggcaaattacgtgatgtatctttc
gaattgttggggccaggctcgcacatttgggtgacaccatcggcgacgaagtcggtgcaatcctcattggtaaagac
gtaaaaccgttgggtcaggaaattatcgctcacgggtgctcataaagtatacgtttatgatgatcctcagctcgaa
cattacaatacgacggttatgcaaaagttatttgcgatttcttccatgaagaaaaaccgaacgtattcctcgtt
ggcgctaccaacatcgccgtgacctcgcccgcggtgtcgctaaactccttgaagactggcctcacggctgactgc
acgcagctcggcggttgacgacgacaaaaagaccatcgatggaccgcgtcgggtctcggcggaacatcatggct
gaaatcatctgcccggacaaccgtccgcagatgggtactgtccgtccgcagtgcttcaaaaaaccggaagcagat
ccttctgcaactggcgaagttatcgaaaagaaagctaaccctcgcgatgctgacttcgatgaccaaattcgtcgaa
ctcatcaaattggcgccgaaggcggttaaaatcgaagacgctgacgttatcggtgctggcgccgtggcatgaac
agtgaagaaccgttcaagaccgggtatcctcaaagaatgtgcagacgctcctcgccggcggtgttgggtgcatccgt
gcagctgttgacgctgggtggatcgatgctctccatcaggttggccagactggtaaaacagttgggtccgaagatc
tacattgcatgcgctatttccggtgctatccagccattggcagggcatgactgggtctgactgcatcattgctatc
aacaagacgaagatgctccgatcttcaaagtcgagctatgggtatcgtaggcgatgtcttcaaagttctcccg
ctcctcacggaagccatcaagaaacagaaaggcattgcataa

Figure 16

etfB (*Megasphaera elsdenii*) SEQ ID NO:49

atggaaatattgggtatgtgtcaaacagggttccggacactgcagaagttaagattgaccccgtaaaacatacggctc
atccgcgctgggtgttccctaacatttttaaccccttcgaccagaacgcttttgaagcagctctcgcattgaaagat
gctgacaaagacgtaaaaatcacacttctctcgatgggtcctgatcaggcaaaagacgttcttcgtgaaggcctc
gcaatggcgctgacgatgcttatcttctgtccgaccgcaaactcgggtgggttccgatacgttagctacgggctat
gctttggcacaggctatcaaaaaattggctgctgacaaaggatcgaacagttcgatatcatcctctgcccga
caggctattgacggcgataccgcacagggttggcccgacatcgcttgcgaactcgggtattcctcagattacgtat
gcccgcgacatcaaagtcgaaggcgacaaagtactgttcagcaggaaaacgaagaaggctacatcgtaacggaa
gctcagttccctgttttgatcacggctgtttaaagacttgaaacgaaccggtttcccgaccattcgtggcacgatg
aaagcaaaacgcccgcgaatcccgaacttgacgctgctgctgttgcagctgacgacgctcagatcggtttgtct
ggctctccgactaaagtcgtaagattttcacaccgcctcagagatccgggtgggtctcgttctcaaagttgaagat
gacaacgaacaggcaatcgtcgaccagggtcatggaaaaactgggttggccagaaaatcatttaa

crt (*Clostridium acetobutylicum*) (SEQ ID NO:51)

atggaaactaaacaatgtcatccttgaaaaggaaggtaaagttgctgtagttaccattaacagacctaaagcatta
aatgcgttaaatagtgtacactaaaagaaatggattatgttataggtgaaattgaaaatgatagcgaagtactt
gcagtaatttttaactggagcaggagaaaaatcattttagcaggagcagatatttctgagatgaaggaaatgaat
accattgaaggtagaaaaattcgggatacttggaaataaagtgtttagaagattagaacttcttgaaaagcctgta
atagcagctgttaatgggtttgcttttaggagggcgatgcgaaatagctatgtcttgtgatataagaatagcttca
agcaacgcaagatttgggtcaaccagaagtaggtctcggaataacacctgggttttgggtgtacacaaagactttca
agattagttggaatgggcatggcaagcagcttatatttactgcacaaaatataaaggcagatgaagcattaaga
atcggacttgtaaataaggtagtagaacctagtgaattaatgaataacagcaaaagaaattgcaaacaaaattgtg
agcaatgctccagtagctgttaagttaagcaaacaggctattaatagagggaatgcagtgtgatattgatactgct
ttagcatttgaatcagaagcatttggagaatgcttttcaacagaggatcaaaaggatgcaatgacagcgtttcata
gagaaaagaaaaattgaaggcttcaaaaatagatag

Figure 17

ccr (*Streptomyces coelicolor*) SEQ ID NO:53

gtgaccgtgaaggacatcctggacgcatccagtcgcccgaactccacgcggccgacatcgccgaactgcccgtc
cccgaagtcgtaccgcgcgatcacctgcacaaggacgagaccgagatgttcgcgggcctcgagaccgcgacaag
gacccccgcaagtcgatccacctggacgacgtgccggtgccgagctgggccccggcgaggccctgggtggccgtc
atggcctcctcggtcaactacaactcggtgtggacctcgatcttcgagccgctgtccaccttcgggttcctggag
cgctacggccgggtcagcgacctcgccaagcggcacgacctgccgtaccacgtcatcggtccgacctcgccggt
gtcgtcctgcgccaccggtccggggtcaacgcctggcaggcgggagcagaggtcgtcgcgcaactgacctctcgtc
gagctggagtcctccgacggccacaacgacacgatgctcgaccccgagcagcgcatctggggcttcgagaccaac
ttcgcgggcctcgcgagatcgcgctgggtcaagtccaaccagctgatgcgaagccggaccacctgagctgggag
gaggccgcgctcccggtgggtcaactccacgcgtaccgccagctcgtctcccgcaacggcgccggcatgaag
caggggcgacaacgtgctcatctggggcgagcggcggaactcggtcgtacgccacccagttcgccctcgccggc
ggcgccaacccgatctgctcgtctcctcgccgcagaaggcggagatctgcccgcgatggggcgccgagggcgtc
atcgaccgcaacgcggaggggtaccggttctggaaggacgagaacacccaggaccgcaaggagtggaagcgcttc
ggcaagcgcatccgcgaactgaccggcgggcagagacatcgacatcgtcttcgagcaccggcgccgagaccttc
ggcgctcctcgtcttcgtcaccggcaaggggcgccaccatcaccacctgcccctcgacctcggggtacatgcacgag
tacgacaaccgctacctgtggatgtccctgaagcgcatcatcggtcgcacttcgccaaactaccgcgagggcctgg
gaggccaaccgcctcatcgccaagggcaggatccacccacgctctccaaggtgactccctcgaggacaccggc
caggcgccctacgacgtccaccgcaacctccaccagggaagggtcggtcgtgtgctgctggcgcccgaggaggc
ctggggtgctgcgaccgggagaagcgcgcgagcacctcgacgccatcaaccgcttcgggaacatctga

Figure 18

ccr (*Pseudomonas syringae*) SEQ ID NO:55

atgaatcaagcactgactgaaaccatgcaggcctttctgatccgccccgagcgctatggcgaaccgcagcaggcc
atccagctcgaacaggtccagatccccaccctgggtccgcatcaggctcctcatcgaagtgtggcagccggactc
aactacaacaacgtctgggcccggccagggttaagccggtgacatcatcgccgcgcggcgcaagcggaaaccgtgac
gccgaacccttccacatcgagggtcggaagcctccggttacgtgaaagccgtgggcgacgctgtcaccacgctc
aagggtgggcataccgtggtggtgtcctgctcggtctacgacgccacggccatcgaatcgcgctcgccccgac
cccatgttctgcagcaaccaggaaatctacgggtacgagaccagctacgggtccttcgccgaatacacccctcgtc
gaagactaccaatgttcccaaaaccaaagttcctgagctgggaggaaagtgccaccctgatgctcaatgggtccg
accgcctacaagcagctcacgcattgggcacccaataaccgtcaagcctggagacgcagtcctgatctggggcgcg
gcaggtggcctgggtctatgtctatccagttgaccgcgcgctcggggggctgccggtggccgtgggtgtccagt
ccagacagggggcgctacgcctgcgaactcggcgccgtgggtacttgctcagaaccgactatccgcacctggga
cgtctgccggacttgaactccgacgctcacagcgcttgaccaaagcttcgcgagtttcgctcgacattcttc
atgacgtggggaaaaaggagctgccccaaagtgggtgatcgagcactccggccaagccaccttccccacctcgtg
cagatctgcgaccgctccggcatgggtgggtcatcggtgggtggcacgtccggctacaactgcgacttcgatgtccgc
cacctgtggatgcaccagaagcgcatccagggtcccaactacgccaaacatccgcgagtgccaggaattcctgcaa
ctagtcgaacaacgcgggtagtgccgacctgaacacctctatcgcttcgaggagacacctagggcgcatcag
gcgctactgagtggaagtcgtaggcaatgcgcctgctggtcaaggccgagcgacccggcctaggggtcggt
tggtga

Figure 19

ccr (*Rhodobacter sphaerides*) SEQ ID NO:57

atggccctcgacgtgcagagcgatatcgctgcctacgacgcgccccagaaggacctctacgagatcggcgagatg
ccgctctcggccatgtgccgaaggagatgtatgcttgggccatccggcgcgagcgtcatggcgagccggatcag
gccatgcagatcgaggtggtcgagacgcctcgcacgacgacgaggtgctcgttctcgtgatggcgggcgggc
gtgaactacaacggcatctgggcccgcctcggcgtgcccgctctcgccgttcgacgggtcacaagcagccctatcac
atcgcggggtccgacgcgtcgggcatcgtctgggcggtggcgacaaaggtcaagcgctggaaggtggcgacgag
gtcgtgatccactgcaaccaggacgacggcgacgacgaggaatgcaacggcgggcgaccgatgttctcgcccacc
cagcggatctggggctacgagacgcggacggctccttcgcccagttcaccgcgtgcaggcgacgacgtgatg
aagcgctccgaagcacctgacctgggaagaggcgccctgctacacgctgacctcgccaccgcctaccggatgctc
ttcgggccacaagccgcacgacctgaagcggggcgagaacgtgctggtctggggcgccctcggcgccctcggtcc
tacgcatccagctcatcaacacggcgggcgccaatgccatcgcgctcatctcagaggaagacaagcgcgacttc
gtcatggggctggggcgccaaggcgctcatcaaccgcaaggacttcaagtgtggggccagctgcccagggtgaac
tcgcccgaatataacgagtggctgaaggaggcgcgcaagtccggcaaggccatctgggacatcaccggcaagggc
atcaacgtcgacatggtgttcgaacatccggcgagggcgaccttcccggtctcgtcgctggtggtgaagaaggc
ggcatggtcgtgatctgcgcgggcaccacgggcttcaactgcaccttcgacgtccgctacatgtggatgcaccag
aagcgctgcaggggcagccatttcgccaacctcaagcaggcctccggggccaaccagctgatgatcgagcgccgc
ctcgatccctgcatgtccgaggtcttccctgggcccagatcccggtgcccatacgaagatgtataagaaccag
cacaagcccggcaacatggcggtgctggtgcaggccccgcgcacggggttgcgcaccttcggcgacgtgctcgag
gccggccgcaaggcctga

Figure 20

ccr (*Rhodospirillum rubrum*) SEQ ID NO:59

atgaccacgtcggcggaagtcatagaactcaatccccggcactggccggaaggatctttacgaactcgggtgaatt
ccgcccgtcggccacgttcccaagtctatgtacgcctgggtcatccgcccggatcgccatggcgaaaccgagaag
tctttccaggttgaagtcgttgaaacgccaactcttgacagccacgacgtcttggtgatggtgatggcgggccggc
gtcaactacaacgggatctgggcccggattggggccagccgatcagcgttttcgactcgcataaggccgcttatcac
atcgccgggttcggatgcggcgggcatcgtctgggcccgtcggcgccaagggtcaagcgctggaaggtcggcgacgag
gtggtcgtccactgcaatcagaccgacggcgacgacgaggaatgcaatggtggcgatccgatgttctcgccgacc
cagcgcatctggggctatgagacccccgatggctccttcgcccagttcaccgcgtgcagtcaccagcaggtgatg
gcccgtccgcgccatctgacctgggaggaaagtgcacgtacgtgctggttctggccaccgcctatcgcatgctg
ttcgggccaccgcccccatgtgctgcgcgccgggtcacaacgtgctgatctggggcgccctcggcgggcctgggatcg
atggcgatccagctgtgcgccacggcgggcgccaatgccatcgcgctcatctccgatgagaccaagcgcgatttc
gtcatgagcctggggcgccaaggcggtgatcaaccgcaaggatttcaattgctggggccaattgcccacgggtcaat
ggcgaggggttcgacgcctatatgaaagaggtgcgcaagtccggcaaggcgatctgggacatcaccggcaagggc
aacgacgttgatttcgtgttcgaacatccggcgagcagaccttcccggtctcgtgcaatgtggtcaagcgcggt
ggcatggtggtgttttgcccggcaccacgggcttcaacctgaccttcgacgcccgcctttgtgtggatgcgccag
aagcgcatccaggcagccacttcgccaatctgctccaggcctcgcaagccaaccagttggtcatcgagcgggcg
atcgatccgtgcatgagcgaagtgttttccctgggacgatattcccaaggcccacaccaagatgtggaagaatcag
cataagcccgggaatatggcggtgctggtccaggcccatcgcccgggcccgcgcaccttgagagattgcgcgagag
gaagggtga

Figure 21

ccr (*Streptomyces avermitilis*) SEQ ID NO:61

gtgaaggaaatcctggacgcgattcagtcacagacggccacgtctgcgcacttcgcgcactgcgcgtccccgac
tcgtaccgcgcgatcaccgtgcacaaggacgagacggagatgttcgcggggtcagcaccgcgcacaaggacccc
cgcaagtcgatccacctggacgacgtgcccgtgcccggagctcgccccggcgaggccctggtggcgctcatggcg
tcctcctgtaactacaactcggctctggacgtcgatcttcgagccgggtgtcgaccttcaacttcctggagcgctac
ggcgcggtcagcgatctcagcaagcgccacgacctgccgtaccacatcatcggttctgacctcgccggggcgctcgtc
ctgcgcacccggccccgggagtgcaactcctggaagccccggcgacgaggtcgctgcgcactgtctctcggtcgagctg
gagtcgtccgacggccacaacgacacgatgctcgaccccgagcagcgcatctggggcttcgagaccaacttcggc
gggctcgcgcgagatcgcgctcgtcaagtccaaccagctgatgccgaagccggaccacctcagctgggaggaggcc
ggcgctccgggctggtgaactcgaccgcgtaccggcagctcgtctcccgcaacggcgccggcgatgaagcagggc
gacaacgtcctcatctggggcgcgagcggtggactgggctcgtaacgcaacgcagttcgcgctcgcggggcgcc
aaccgcatctgcgtcgtctccagcgagcagaaggcgacatctgcgcgtcgatggggcgccgaggcgatcatcgac
cgcaacgcgcgagggtacaaagtctggaaggacgagaccacccaggacccgaaggagtgaagcgcttcggcaag
cgcatccgcgagttcaccgcggcgaggacatcgacatcgtcttcgagcaccgcggcgagacacctcggcgcgc
tcggtctacgtcaccgcgaaggcgccaccatcaccacctgcgcctcgacctcgggctacatgcacgagtagcac
aaccgctacctgtggatgtcgtcgaagcggatcatcggtcgcacttcgcgaactaccgcgaggcctgggaggcc
aaccgctcgtcgccaagggaagatccacccacgctctccaaggtctactccctggaggacaccggggcaggcc
gcctacgacgtgcaccgcaacctccaccagggaagggtcgcgctgctcgccctcgcgccccgcgagggcctgggc
gtgcgcgacgaggagaagcgcgcgcagcacatcgacgccatcaaccgcttcggaacatctga

Figure 22

ccr (*Silicibacter pomeroyi*) SEQ ID NO:63

atggctttggacaccgacagcggatcgcgctctacgcggcgccccgagaaagacctctatgagatgggtgaaatc
ccccgatgggattcgtgcccaagaagatgtatgcgtggcgatccgcaaagagcgccacggatgacccgatacc
gogatgcaggtcgaaagtgggtgacgtgcccagcgtcgacagccacgaggtgctggttctggtgatggcgcgtggc
gtcaactacaatggcgctctgggctccaaagggtgttcggatttcccccttcgatggccacggacagccctatcac
atcgccgggttcgatgcttcgggtatcgtctgggcccgtgggggacaagggtcaagcgctggaaggctggcgacgag
gtcgtgatccactgcaatcaggatgatgggtgacgacgagcactgcaatggcggtgacccgatgtattcgcccagt
cagcggatctgggggttacgagacgcggacggatcctttgctcagttcaccaatgtgcaggcgacgagctgatg
ccgcggccccaaagcacctgacctgggaagaagcggcatgttacacgctgacgctggcgaccgcctaccggatgctg
tttggccatgagccgatgatctcaagcccggtcagaacgttctgggtctggggtgcgtccggtgggtctggggtcc
tatgcgatccagcttatcaatacggcggggtgcgaacgcgattggcgctcatctcgatgaaagcaagcgccagttt
gtcatggaccttggcgcaaaagggtgtcatcaaccgcaaggatttcaactgctggggtcaactgcccacgggtgaac
acccccgaatatgccgagtggttcaaggaaagccgcaagttcggcaaggcgatctgggacattaccggcaagggc
gtgaacgtggacatggctcttcgagcaccgcggcgagagcaggttcccggtctcgaccttcgtggtgaagaagggc
ggatgggttgatctgcgcgggcaccagcggctacaacctgacctttgacgtgcgctatatgtggatgcaccag
aagcgcttcaggggcagccacttcgcccctctcaagcaggcaatggccgcgaaccagctgatggtcgagcgccgg
ctcgacccgtgcatgtccgaggtgttcacctgggcccgatctgcccaggcgcatatgaagatgatgcgcaacgag
cacaagccgggcaacatgtcgggtgctggtgcaatcgccccgacccgggtgcgcaccctcgaagaggttctggac
gcccgcggttaa

Figure 23

ccr (*Xanthobacter autotrophicus*) SEQ ID NO:65

atggcccagacggcagccgccaacgcgaacgagggaccggtgaaggacctttatgagctgggagaggttcccccc
ctcggtcacgtccccgccaagatgtacgcctgggccatccgccgcgagcgccatggggccgcggaagagtcgttc
cagctggaagtgggtgcccacctgggagctggggcgagaacgacgtgctggtctacgtcatggccgcccgcgtcaac
tacaacggcatctgggcggtcctcgccagccgatctcgccgttcgacgtgcacaaggcgccctccacatcgcc
ggctccgatgcctcgggtatcgtctgggcggtgggctccaaggtgaagcgctggaaggtgggagcagaggtggtc
gtgcactgtaaccaggacgacggcgacgagaggtgcaacggcgggcgaccccatgttctccccgtcccagcgc
atctggggctatgagacggcgacggctcgttcgcccagttctgcccgggtgcaggcgcgccagctgatgcccgcgc
cccaagcacctgacctgggaagagagcgctgctacaccctaccatggccaccgcctaccgcatgctgttcggc
catccgcgcacacgggtgaagcggcgactacgtgctggtgtggggcgccctcgggcggtcctcggcgtgttcggc
gtgcagctcgccgcgcctcggcgccccatgtgatcggtgatctccgacgagaccaagcgcgactatgtcctc
ggcctcggcgccaagggcggtgatcaaccgcaaggatttcaagtgtggggccagctgcccagggtcaactcgccg
gaatacaatgagtggaaccaaggaagcccgaagttcggcaaggccatttgggacatcagcggcaagcgcgacgtg
gacatcggttcgagcatcctggcgagcagaccttccgggtctcgaccctcgtgggcaagcgcgggcgcatgatc
gtgttctgcccggcaccacggcttcaacatcaccttcgacgcccgtacgtgtggatgcccagaagcgcatc
cagggctcccacttcgctcacctcaagcaggcctcggccgcaatcagttcatcgcaccggcgcggtggacccc
tgcatgtcggaagtgttccgtgggacggcatccccgagggcgacaccaagatgtggaagaaccagcagccct
ggcaacatggcggtgctggtcaacacccccgcaccggcctgctgctaccctcgaggacgtgatcgaggccggcgcg
aagaagtga

Figure 24

adhE2 (*Clostridium acetobutylicum*) SEQ ID NO:67

atgaaagttacaaatcaaaaagaactaaaacaaaagctaaatgaattgagagaagcgcaaaaagaagtttgcaacc
tatactcaagagcaagttgataaaatttttaacaatgtgccatagccgcagctaaagaaagaataaacttagct
aaattagcagtagaagaaacaggaataggtccttgtagaagataaaattataaaaaatcattttgcagcagaatat
atatacaataaatataaaaaatgaaaaaacttgtggcataatagacatgacgattccttaggcataacaaaggtt
gctgaaccaatttgaattgttgcagccatagttcctactactaatccaacttccacagcaattttcaaatcatta
atttcttttaaaaacaagaacgcaatattcttttcaccacatccacgtgcaaaaaaatctacaattgctgcagca
aaattaatttttagatgcagctgttaaagcaggagcacctaaaaatataataggctggatagatgagccatcaata
gaactttctcaagattttagtgagtgaagctgatatataattagcaacaggaggtccttcaatgggttaaagcggcc
tattcatctggaaaacctgcaattgggtgttggagcaggaataacaccagcaataatagatgagagtgcagatata
gatattggcagtaagctccataattttatcaaagacttatgacaatggagtaatatgcgcttctgaacaatcaata
ttagttatgaattcaatatacgaaaaagttaaagaggaaatttgtaaaacgaggatcatatatactcaatcaaaat
gaaatagctaaaaataaaaagaactatgtttaaaaatggagctattaatgctgacatagttggaaaatctgcttat
ataattgctaaaaatggcaggaattgaagttcctcaaaactacaaagatacttataggcgaagtacaatctgttgaa
aaaagcgagctgttctcacatgaaaaactatcaccagtaacttgcaatgtataaagttaaaggattttgatgaagct
ctaaaaaaggcacaaaaggctaataagaattaggtggaagtggacacacgctcatctttatatatagattcacaaaac
aataaggataaagttaaagaatttggattagcaatgaaaacttcaaggacatttattaacatgccttcttcacag
ggagcaagcggagattttatacaattttgcgatagcaccatcatttactcttggatgcggcacttggggaggaac
tctgtatcgcaaatgtagagcctaacaattttatataatataaaaagtgttgctgaaagaagggaaaaatatgctt
tggtttaaagtgccacaaaaataatattttaaatatggatgtcttagatttgcattaaaaagaattaaaagatatg
aataagaaaagagcctttatagtaacagataaagatctttttaaaacttggatatgttaataaaaataacaaaggta
ctagatgagatagatattaaatacagtatatttacagatatataatctgatccaactattgattcagtaaaaaaa
ggtgctaaagaaatgcttaactttgaacctgatactataatctctatttggtggtggatcgccaatggatgcagca
aaggttatgcacttggtatatgaatatccagaagcagaaattgaaaaatctagctataaaactttatggatataaga
aagagaatatgcaatttccctaaattaggtacaaaggcgatttcagtagctatttctacaactgctggtaccggt
tcagagggcaacaccttttgcagttataactaatatgaaacaggaatgaaataccctttaaacttcttatgaattg
accccaaacatggcaataatagatactgaattaatgtttaaatatgctagaaaatttaacagcagcaactggaata
gatgcatttagttcatgctatagaagcatatgtttcgggttatggctacggattatactgatgaattagccttaaga
gcaataaaaaatgatatttaaatatttgcctagagcctataaaaaatgggactaacgacattgaagcaagagaaaaa
atggcacatgcctctaataattgcggggatggcatttgc aaatgctttcttaggtgtatgccattcaatgggtcat
aaacttggggcaatgcatcagttccacatggaattgctgtgctgtattaatagaagaagttattaaatataac
gctacagactgtccaacaaagcaaacagcattccctcaatataaatctcctaattgctaagagaaaatatgctgaa
attgcagagtatttgaatttaaagggtactagcgataccgaaaaggtaacagccttaatagaagctatttcaaag
ttaagatagattttagtattccacaaaatataagtccgctggaataaataaaaaagattttttataatacgtata
gataaaatgtcagagcttgcctttgatgaccaatgtacaacagctaatacctaggtatccacttataagtgaactt
aaggatatctatataaaaatcattttaa

Figure 25

(SEQ ID NO:)

```

TD      MIVKPMVRNNICLNAHPQGCKKGVEDQIEYTKKRITAIEVKAGAKAPKNVLVLGCSNGYGL (68)
TV      MSMKPMRLSNNICLNAHPQGCKKAVEDQIAYTRKRAASHPAG-TATPKNVLVIGCSGGYGL (69)
FS      MIIKPLIRSNMCLNAHPKGCAADVKKHIEFIKKKFTTRSIP-ADAPKTVLVLCSTGYGL (70)
FJ      MIIIEPRMRGFTCLTAHPAGCEQNVKNQIEYIKS-----KGAIAGAKKVLVIGASTGFGL (71)
      * : : * : * . : * . * * * * * * : * . * * : : . * . * * : * . * : * *

TD      ASRITAAFGYGAATIGVSFEKAGSE-----TKYGTTPGWYNNLAFDEAAKREGLYSVTI
TV      ASRITAAFGYGAATIGVSFEKAGSE-----KKWGTTPGWYNNLAVDAAAKEAGLISVTI
FS      ASRIVAAFGYKAATIGVSFEKEGSDGGIGESREKTGTPGWYNNMAFDKFAKEAGLDAVTF
FJ      ASRITSAFGSDAATIGVFFEKPPVEG-----KTASPGWYNSAAFEKEAHKAGLYAKSI
      * * * * . : * * * * * * * : * * : * . : * * * * . * . : * . * * : :

TD      DGDAFSDEIKAQVIEEAKKKGIKFDLIVYSLASPVRTDPDTGIMHKSVLKPFKGKFTTGKT
TV      NGDAFSDAIKAQVIDEAKKLNKIFDLIVYSVASSVRTDPDTGVTYRSALKPFGKPFKGKFTTGKT
FS      NGDAFSHEMRQNVIDTLKKMGRKVDLLVYSVASSVRVDPDNGTIYRSVLKPIDKVFVTGAT
FJ      NGDAFSNEIKRETLDLIKADLGQVDLVIYSLASPVRTNPNTGVTTHRSVLKPIGQFTFNKT
      : * * * * . : : : : * : . * * : * * : * * : * * : * * : * * : * * . *

TD      VDPFTGELKEISAEAPANDEEAAATVKVMGGEDWERWIKQLSKEGLLEEGCITLAYSYIGP
TV      LDPFTGALTEITAEPATDEEAAATVKVMGGEDWQRWIEKLGAADVLAQGCITVAYSYIGP
FS      IDCLSGKISTISAEPATAEAAANTVKVMGGEDWALWVRKLKEAGVLAEGVKTVAYSYIGP
FJ      VDFHTGNVSEVSIAPANEEDIENTVAVMGGEDWAMWIDALKNENLLAEGATTIAYSYIGP
      : * : * : : : * . * : * * * * * * * : * : * : * * : * * : * * : * *

TD      EATQALYRKGTIGKAKEHLEATAHRLNKENP-SIR--AFVSVNKLVTTRASAVIPVIPLY
TV      EATQALYRKGTIGKAKEHLEATAHALNAKLA-ALKGKAFVSVNKLVTTRASAVIPVIPLY
FS      KLSHAIYRDGTIGGAKKHLEATALELNKELQNDLHGEAYVSVNKLVTTRSSAVIPIIPMY
FJ      ELTEAVYRKGTIGRAKDHLEATAFTITDTLK-SLGGKAYVSVNKALVTQASSAIPVIPLY
      : : . : * . * * * * * * . * * * * * : . : * : * * * . * * : : * . * : * : *

TD      LASLFLKVMKEKGNHEGCIEQITRLYAERLYRKDGTIPVDEENRIRIDDWELEEDVQKAVS
TV      LASLFLKVMKEKGTHEGCIEQINRLFDSRLYTADGVIPTDNENRIRIDDWELDEGVQSAVA
FS      ISVLFLKVMKEMGNHEGCIEQMERLMTERTLYTG-SKVPTDENHLIRIDDYELDPKVQAEVD
FJ      ISLLYKIMKEEGIHEGCIEQIQRLFQDRLYNG-SEVPVDEKGRIRIDDWEMREDVQAKVA
      : : * . * : * * * * * : * * . * * . : * . : : * * * : * : * * *

TD      ALMEKVTGENAESLTDLAGYRHDFLASNGFDVEGINYEAEVERFDRIX----
TV      KIMATVTDETSCKLTDVDEYRHDFLAINGFDIAGIDYDAEIDRFDRIX----
FS      KRMATVTQENLAEVGDLEGYRHDFLATNGFDIDGVDYADVQTLTSIX----
FJ      ALWKEATTETLPSIGDLAGYRNDFLNLFGFEFAGVDYKADTNEVVNIESIKX
      . * * . : : * : * * : * * * * * : * : . * : : * . : : . *

```

Figure 26A

```

PaaH1-WT      1 10 20 30 40 50
PaaH1-P01     MSIRTVGIVGAGTMGNGIAQACAVVGLNVVMVDISSDAAVCKGVATVASSL
PaaH1-P042    MSIRTVGIVGAGTMGNGIAQACAVVGLNVVMVDISSDAAVCKGVATVASSL
PaaH1-P06     MSIRTVGIVGAGTMGNGIAQACAVVGLNVVMVDISSDAAVCKGVATVASSL
PaaH1-P09     MSIRTVGIVGAGTMGNGIAQACAVVGLNVVMVDISSDAAVCKGVATVASSL
PaaH1-P13     MSIRTVGIVGAGTMGNGIAQACAVVGLNVVMVDISSDAAVCKGVATVASSL

PaaH1-WT      60 70 80 90 100
PaaH1-P01     DRLIKKEKLEADKASALARIKGSTSYDOLKATDIVIEAATENYDLKVKI
PaaH1-P042    DRLIKKEKLEADKASALARIKGSTSYDOLKATDIVIEAATENYDLKVKI
PaaH1-P06     DRLIKKEKLEADKASALARIKGSTSYDOLKATDIVIEAATENYDLKVKI
PaaH1-P09     DRLIKKEKLEADKASALARIKGSTSYDOLKATDIVIEAATENYDLKVKI
PaaH1-P13     DRLIKKEKLEADKASALARIKGSTSYDOLKATDIVIEAATENYDLKVKI

PaaH1-WT      110 120 130 140 150
PaaH1-P01     LKQIDGIVGENVILASNTSSGISITKLAAVTSRADRFIGMHEFFNFPVPVMAI
PaaH1-P042    LKQIDGIVGENVILASNTSSGISITKLAAVTSRADRFIGMHEFFNFPVPVMAI
PaaH1-P06     LKQIDGIVGENVILASNTSSGISITKLAAVTSRADRFIGMHEFFNFPVPVMAI
PaaH1-P09     LKQIDGIVGENVILASNTSSGISITKLAAVTSRADRFIGMHEFFNFPVPVMAI
PaaH1-P13     LKQIDGIVGENVILASNTSSGISITKLAAVTSRADRFIGMHEFFNFPVPVMAI

PaaH1-WT      160 170 180 190 200
PaaH1-P01     VELIPGLOQTSDDTHAAVEALSKQLGKYPITVKNSPGFVVNRIILCEMINEA
PaaH1-P042    VELIPGLOQTSDDTHAAVEALSKQLGKYPITVKNSPGFVVNRIILCEMINEA
PaaH1-P06     VELIPGLOQTSDDTHAAVEALSKQLGKYPITVKNSPGFVVNRIILCEMINEA
PaaH1-P09     VELIPGLOQTSDDTHAAVEALSKQLGKYPITVKNSPGFVVNRIILCEMINEA
PaaH1-P13     VELIPGLOQTSDDTHAAVEALSKQLGKYPITVKNSPGFVVNRIILCEMINEA

PaaH1-WT      210 220 230 240 250
PaaH1-P01     FCVLGEGLASPEEIDEEMKILGCHNPIGFLALADMIGLDTMLAVMEVLYTE
PaaH1-P042    FCVLGEGLASPEEIDEEMKILGCHNPIGFLALADMIGLDTMLAVMEVLYTE
PaaH1-P06     FCVLGEGLASPEEIDEEMKILGCHNPIGFLALADMIGLDTMLAVMEVLYTE
PaaH1-P09     FCVLGEGLASPEEIDEEMKILGCHNPIGFLALADMIGLDTMLAVMEVLYTE
PaaH1-P13     FCVLGEGLASPEEIDEEMKILGCHNPIGFLALADMIGLDTMLAVMEVLYTE

PaaH1-WT      260 270 280 285
PaaH1-P01     FADPKYRFPAMLMREMVAAGYLGRKKTGRGVYVYSK*
PaaH1-P042    FADPKYRFPAMLMREMVAAGYLGRKKTGRGVYVYSK*
PaaH1-P06     FADPKYRFPAMLMREMVAAGYLGRKKTGRGVYVYSK*
PaaH1-P09     FADPKYRFPAMLMREMVAAGYLGRKKTGRGVYVYSK*
PaaH1-P13     FADPKYRFPAMLMREMVAAGYLGRKKTGRGVYVYSK*

```

FIGURE 26B

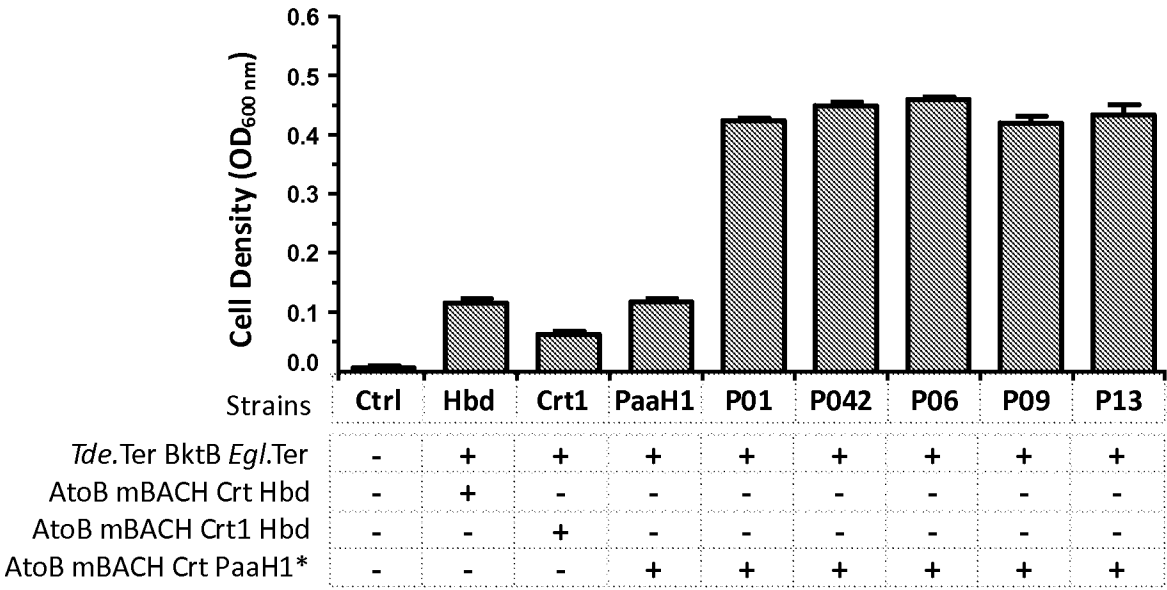


FIGURE 27

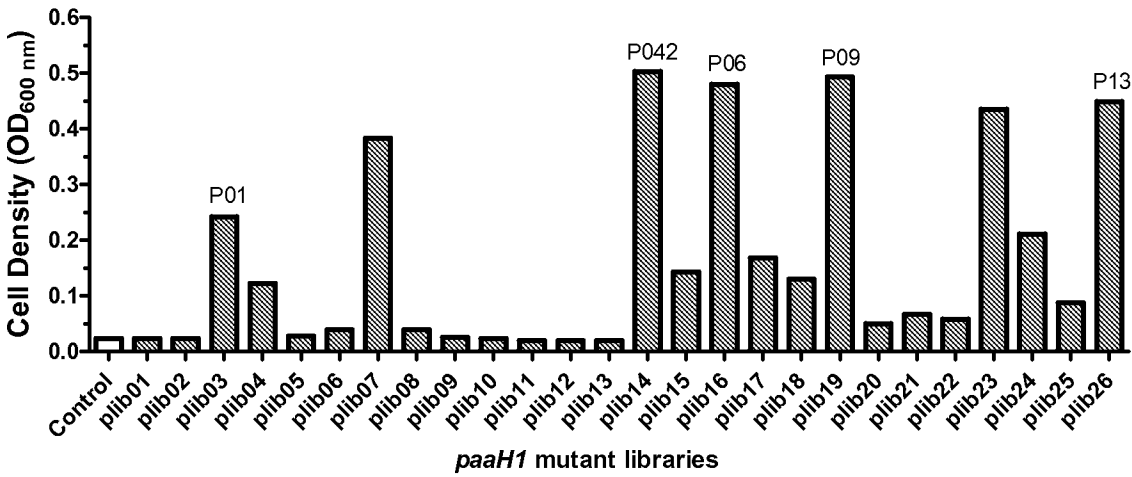


FIGURE 28

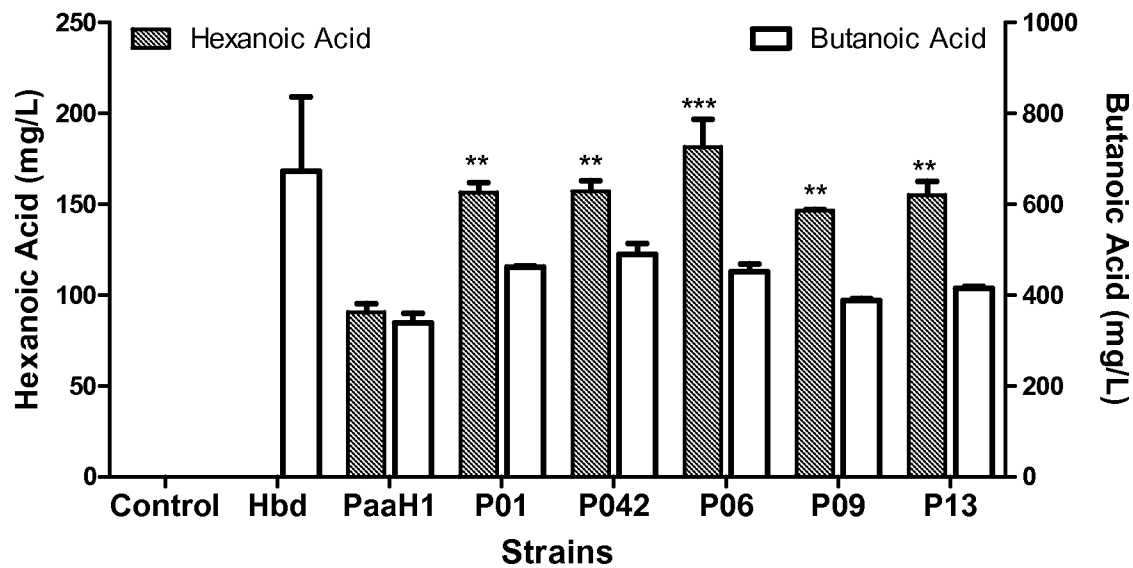


FIGURE 29

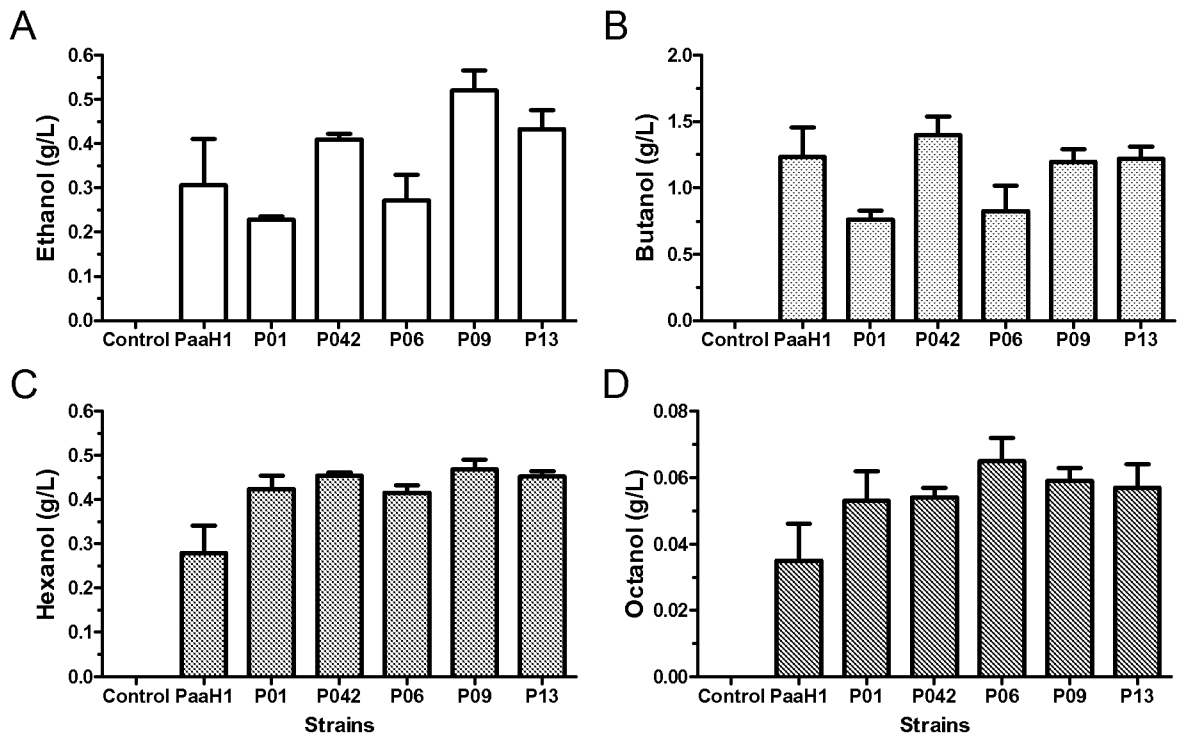


FIGURE 30