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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
18 May 2017 (18.05.2017)

WIPO | PCT

(10) International Publication Number
WO 2017/083656 A1

(51) International Patent Classification:
C12P 5/02 (2006.01) *C12P 11/00* (2006.01)
C12P 7/40 (2006.01) *C12N 9/10* (2006.01)

(21) International Application Number:
PCT/US2016/061539

(22) International Filing Date:
11 November 2016 (11.11.2016)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/255,276 13 November 2015 (13.11.2015) US

(71) Applicant: **INVISTA NORTH AMERICA S.A.R.L** [LU/US]; Three Little Falls Center, 2801 Centerville Road, Wilmington, DE 19808 (US).

(72) Inventors: **BOTES, Adriana, Leonora**; Red House Farm, Daleside Road, Rosedale East Y018 8RL (GB). **KADI, Nadia**; 41 Columbine Close, Marton-in-cleveland, Middlesbrough TS7 8SQ (GB).

(74) Agents: **CHAPMAN, Ernest, F.** et al.; Finnegan, Henderson, Farabow, GARRETT & DUNNER, LLP, 901 New York Avenue, NW, Washington, DC 20001-4413 (US).

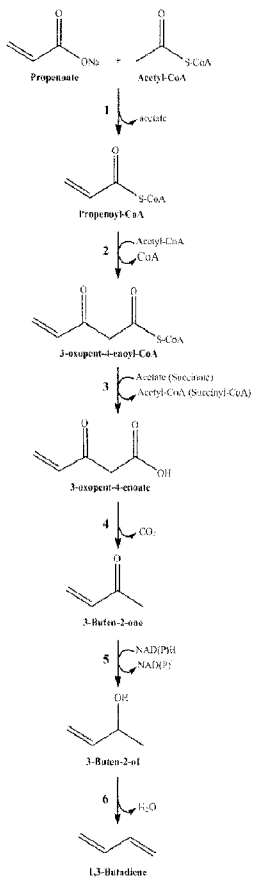
(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM,

[Continued on next page]

(54) Title: POLYPEPTIDES FOR CARBON-CARBON BOND FORMATION AND USES THEREOF

(57) Abstract: This document describes polypeptides with dual CoA transferase and β-ketothioiase activities and variants thereof, use of such polypeptides in biosynthetic methods, and non-naturally occurring hosts comprising such polypeptides.

Figure 1A



WO 2017/083656 A1

DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, 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, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

— *with sequence listing part of description (Rule 5.2(a))*

POLYPEPTIDES FOR CARBON-CARBON BOND FORMATION AND USES THEREOF

[001] This application claims priority to U.S. Provisional Patent Application No. 62/255,276, filed November 13, 2015.

SEQUENCE LISTING

[002] The instant application contains a Sequence Listing which has been submitted electronically in ASCII format and is hereby incorporated by reference in its entirety. Said ASCII copy, created on November 8, 2016, is named 12444_6013-00000_SL.txt and is 113,030 bytes in size.

TECHNICAL FIELD

[003] This disclosure provides multifunctional polypeptides. This disclosure provides polypeptides having a β-ketothiolase and CoA transferase activities and variants thereof. The disclosure provides methods for C-C bond formation useful for biosynthesizing thioesters of 3-keto-acids and derivatives thereof via condensation of acetyl-CoA with alkanoic, (substituted) alkenoic or hydroxy- or haloacids or their CoA esters. These 3-keto-acyl-CoA esters are useful for, for example, the microbial or biocatalytic production of industrially important compounds including straight chain fatty acids, ω-alkenoic acids, ω-hydroxy fatty acids, alkanes, alkenes, ketones, and as precursors for butadiene and isoprene. For example, 3-oxopent-4-enoyl-CoA and 4-methyl-3-oxopent-4-enoyl-CoA can be prepared using such polypeptides. This invention also relates to methods for producing 3-butene-2-ol and 3-methyl-3-butene-2-ol from 3-oxo-pent-4-enoic acid and 4-methyl-3-oxopent-4-enoic acid, respectively using

an alcohol dehydrogenase or a phenylacetaldehyde reductase or using recombinant host cells expressing one or more such polypeptides.

BACKGROUND

[004] Biosynthetic thiolases catalyse carbon-carbon bond formation via a thioester-dependent Claisen-condensation reaction mechanism. This is an essential first step of many biosynthetic pathways relying on the stepwise assembly of carbon backbones from 2- and 3-carbon metabolites, including fatty acids and lipids, polyketides, isoprenoids, cholesterol, steroid hormones and ketone bodies. Many of these compounds are industrially important chemicals due to their biological activity or potential application as building blocks or substrates for the production of bulk chemicals and pharmaceuticals (Klein, M. (2009) *Enzyme and Microbial Technology*, 45, p. 361-366).

[005] Biosynthetic thiolases (EC 2.3.1.9) such as BktB from *Ralstonia eutropha* are also involved in poly-hydroxalkanoate biosynthesis, and serve to condense either two acetyl-CoA's to form acetoacetyl-CoA in polyhydroxybutyrate (PHB) biosynthesis or to condense acetyl-CoA with propionyl-CoA or butyryl-CoA to form valeryl-CoA or hexanoyl-CoA (Kim, Eun-Jung et al., 2014, *Biochemical and Biophysical Research Communications*, 444, 3, p. 365-369).

[006] Sustainable production of advanced biofuels and chemicals from renewable feedstocks requires metabolic engineering of microorganisms to synthesise longer carbon chain length compounds from 2- and 3-carbon metabolic intermediates. Invariably, this requires C-C-bond formation by biosynthetic thiolases. For example, reversal of the β-oxidation cycle provides a platform for the synthesis of fatty acids and

fatty acid-derived chemicals, using biosynthetic thiolases to initiate and reverse the cycle so that the carbon backbone is extended rather than degraded (Clomburg, James et al., 2015, *Metabolic Engineering* 28, p. 202-212). This reverse β-oxidation cycle can be used to generate a diverse range of products (Cintolesi, Angela et al., 2014, *Metabolic engineering* 23, p. 100-115; Dellomonaco, 2011, *Nature*, vol 476, p. 355-359). The intermediates of the reverse β-oxidation cycle can be removed from the cycle to form 3-keto-fatty acids or the corresponding methyl ketones via decarboxylation, medium chain length polyesters, fatty aldehydes, fatty alcohols, fatty acids, alkanes, and alkenes (Yu, Ai-Qun, 2014, *Frontiers in Bioengineering and Biotechnology*, Vol 2, article 78).

[007] The products from the reverse β-oxidation cycle can also be transferred from the reverse β-oxidation cycle to the fatty acid biosynthesis (FAS) cycle by employing enzymes that transfer the acyl-phosphopantetheine group from acyl-CoA to apo-ACP such as *sfp*-type PPTase (phosphopantetheinyl transferases) (*Methods in Enzymology*, Volume 458, Chapter 10; Beld, Joris et al., 2014, *Nat. Prod. Rep.* 31, 61-108). The fatty-acyl-ACP's can similarly be removed from the FAS cycle by a variety of enzymes to produce fatty acids and fatty-acid derived chemicals such as hydroxy fatty acids, fatty aldehydes, fatty alcohols, alkenes, and dicarboxylic acids (Janβen & Steinbüchel, 2014, *Biotechnology for Biofuels*, 7:7).

[008] Supplying 3-keto-acyl-ACP substrates produced by biosynthetic thiolases followed by transfer of the acyl- phosphopantetheine group to *apo*-ACP, also provides an alternative entry of 3-keto-acyl-ACP's into the FAS cycle not relying on KAS III type β-ketoacyl-ACP synthases (EC 2.3.1.180, FabH) that require malonyl-ACP as extender

to improve the overall carbon and energy efficiency of product synthesis (Dellomonaco, 2011, *Nature*, vol 476, p. 355-359).

[009] Enzymes that are capable of C-C bond formation to condense acetyl-CoA with acids or CoA activated acids to form 3-keto-acyl-CoA esters, such as biosynthetic thiolases, are thus essential enzymes, not only in the synthesis of fatty acids and fatty acid-derived chemicals by providing the 3-keto-acyl-CoA or 3-keto-acyl-ACP intermediates to either the reverse β-oxidation or the FAS cycle, but also in polyhydroxyalkanoate biosynthesis and in the production of fermentation products such as butanol, butyric acid, acetone and hydrogen by clostridia (Klein, M. 2009. Enzyme and Microbial Technology, 45, 361-366), and many other biochemical pathways such as isoprenoids and polyketides.

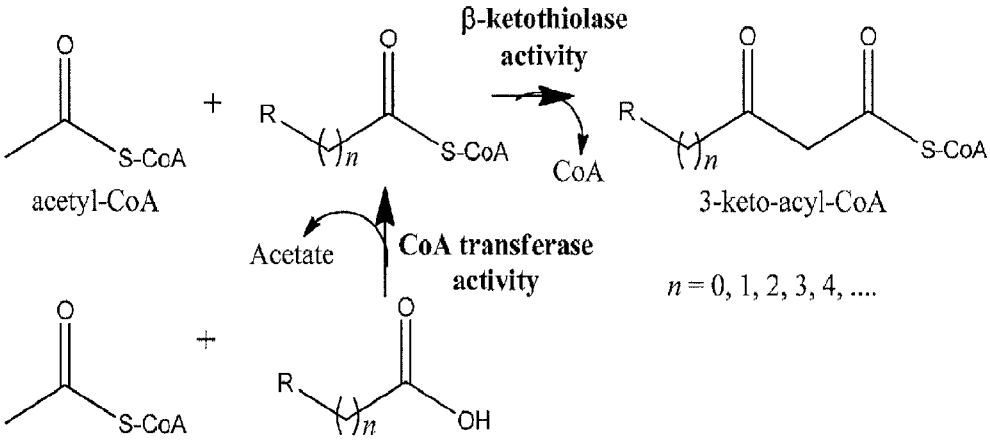
[010] However, all known biosynthetic thiolases have certain limitations. For example, they require two cysteine residues for their catalytic mechanism. In the acyl transfer step, Cys 378 protonates the CoA leaving group, and the acetyl group is transferred to Cys 89. In the subsequent Claisen condensation reaction, the deprotonated Cys378 abstracts the proton of the C2 atom of acetyl-CoA, facilitating its nucleophilic attack on the carbonyl carbon of the acetyl group that is covalently bound to the Cys 89 sulfur atom, which leads to C-C bond formation and release of the acyl group from Cys 89. This two-step "ping-pong" mechanism is also found in the biosynthetic thiolases involved in PHA biosynthesis of haloarchae, but in this case, Cys 89 is replaced by a Ser, leading to a Ser-His-Cys catalytic triad rather than a Cys-His-Cys triad found in other thiolases (Hou, Jing et al., 2013, Applied and Environmental Microbiology, Vol 79, number 17, p5104-5111). Substrates with electrophilic groups,

such as acrylic acid thioesters and haloacetyl-CoA analogs, irreversibly inactivate biosynthetic thiolases through both acylation of Cys 89 and alkylation of Cys 378 (Palmer, M.A. et al., J Biol Chem, 264 (1991), pp. 15293–15297; Palmer, M.A. et al., J Biol Chem, 266 (1991), pp.8369-8375; Davis, Jeffrey T. et al., J Biol Chem, 262 (1987) pp. 90-96). For further example, biosynthetic thiolases are restricted to short chain substrates (C4 or shorter) such as acetyl-CoA, propanoyl-CoA, and butanoyl-CoA. Therefore, the longest acyl chain accepted by biosynthetic thiolases consist of only 4 carbon atoms due to the shape of the substrate binding pocket (Modis, Yorgo and Wierenga, Rik K. 1999, Structure, Vol 7 no. 10 p.1279-1290).

SUMMARY

[011] It is of interest to provide enzymes capable of C-C bond formation to form 3-keto-acyl thioesters, not only to supplement the available biosynthetic thiolases, but also to overcome their limitations: (1) inactivation by electrophilic substrates such as acrylic acid, methacrylic acid, or ω-hydroxy-acid thioesters; and (2) inability to condense acetyl-CoA with substrates of chain length greater than C4. Surprisingly, the inventors have discovered enzymes that can overcome these limitations, which will allow the diversification of chemicals that can be obtained via 3-keto-acyl-CoA intermediates. These chemicals include precursors of bulk chemicals such as butadiene and isoprene, polyhydroxyalkanoate pathways intermediates and products, chemicals derived from fatty acid metabolism (both from reverse β-oxidation and fatty acid biosynthesis), as well as chemicals derived from polyketide and isoprenoid pathways, as well as clostridial fermentation products.

[012] This document is based at least in part on the discovery that it is possible to use CoA transferases from EC 2.8.3.- to catalyze not only CoA transfer, but also Claisen type condensation associated with biosynthetic thiolases, to produce 3-keto-acyl-CoA esters from the condensation of acetyl-CoA with alkanoic, (substituted) alkenoic, hydroxy- or haloacids or their CoA esters.



$R = -CH_3, =CH_2, =CH(CH_3), -CH_2OH, -CH_2X, X = -Cl, -Br, -Fl, -I$

[013] In one aspect, this document provides a method to produce 3-keto acids of carbon chain length *n*+2 by providing cells expressing enzymes from EC 2.8.3.- with straight chain alkanoic acids of carbon chain length *n* (*n* > 2) such as acetate, propionate, butyrate, pentanoic acid, hexanoic acid, and the like, or branched chain alkanoic acids such as isobutyrate, isovaleric acid, and pivalic acid, in addition to a carbon source suitable for growth to provide acetyl-CoA.

[014] In one aspect, this document provides a method to produce 3-keto acids of carbon chain length *n*+2 by providing cells expressing enzymes from EC 2.8.3.- with straight chain alkenoic acids of carbon chain length *n* (*n* > 2) such as acrylic acid, 2-propenoic acid, 3-butenoic acid, 4-pentenoic acid, 5-hexenoic acid, 6-hepteneoic acid,

crotonic acid, and the like, or branched chain alkenoic acids such as methacrylic acid, 3-methyl-3-butenoic acid, 4-methyl-4-pentenoic acid, 5-methyl-5-hexenoic acid, and the like, in addition to a carbon source suitable for growth, to provide acetyl-CoA.

[015] In one aspect, this document provides a method to produce 3-keto acids of carbon chain length n+2 by providing cells expressing enzymes from EC 2.8.3.- with hydroxy- or halo-acids of carbon chain length n ($n > 2$) such as 3-hydroxypropionic acid, 4-hydroxybutyric acid, 5-hydroxyvaleric acid, 6-hydroxy-caproic acid, or the corresponding halogen substituted acid, and the like, in addition to a carbon source suitable for growth to provide acetyl-CoA.

[016] In another aspect, both the acetyl-CoA and the acid partner for the condensation reaction (or its activated CoA ester) are derived from a fermentable carbon source, either via a naturally occurring pathway or via an engineered pathway. As an example, acryloyl-CoA, methacryloyl-CoA, propanoyl-CoA, and butanoyl-CoA are naturally occurring metabolites in many organisms, and their intracellular production can be improved by metabolic engineering strategies. (Pathways leading to acryloyl-CoA and propanoyl-CoA had been described in PCT Application PCT/US2014/048606.)

[017] In another aspect, the 3-keto-acyl-CoA product of the condensation reaction is converted to the free acid by a CoA transferase or a thioesterase.

[018] This document is based at least in part on the discovery that it is possible to construct a biochemical pathway for producing butadiene utilizing a polypeptide having both β-ketothiolase and CoA transferase activity, where that polypeptide can both (i) transfer a Coenzyme A (CoA) moiety from a CoA source such as acetyl-CoA to a short chain alkyl or alkenyl carboxylate such as acrylate, propionate, or butyrate, and

salt forms thereof (e.g., sodium acrylate, sodium propionate, or sodium butyrate) (referred to as CoA transferase activity) and (ii) condense a short-chain acyl-CoA such as acryloyl-CoA, propionyl-CoA, or butyryl-CoA with an acetyl-CoA moiety (referred to as β-ketothiolase activity).

[019] In one aspect, this document provides polypeptides with dual CoA transferase and β-ketothiolase activity. In one embodiment, the polypeptide having both CoA transferase and β-ketothiolase activities is classified under EC 2.8.3.-, for example under EC 2.8.3.8. In one embodiment, the polypeptide having both CoA transferase and β-ketothiolase activities has at least 95% sequence identity to the amino acid sequence set forth in any one of SEQ ID NO: 1, 4, 5, 6, 7, 8, and 9. Also provided are variants of polypeptides having both CoA transferase and β-ketothiolase activities wherein either or both of these activities have been increased, decreased, or abolished via introduction of mutations.

[020] In one aspect, this document provides methods of producing 3-oxopent-4-enoyl-CoA using a polypeptide that has both CoA transferase and β-ketothiolase activities. In one embodiment, said methods can further comprise converting 3-oxopent-4-enoyl-CoA to 3-oxopent-4-enoate. In one embodiment, said methods can further comprise converting 3-oxopent-4-enoate to 3-buten-2-one. In one embodiment, said methods can further comprise converting 3-buten-2-one to 3-buten-2-ol. In one embodiment, said methods can further comprise converting 3-buten-2-ol to 1,3-butadiene.

[021] In one aspect, this document provides methods of producing 3-oxo-acyl-CoA compounds using a polypeptide that has both CoA transferase and β-ketothiolase activities.

[022] In one aspect, this document provides methods of producing 3-oxo-enoyl-CoA compounds using a polypeptide that has both CoA transferase and β-ketothiolase activities.

[023] In one aspect, this document provides methods of producing 3-oxo-hydroxyacyl-CoA compounds using a polypeptide that has both CoA transferase and β-ketothiolase activities. In some embodiments, said 3-oxo-hydroxyacyl-CoA compounds can be further converted to nylon compounds.

[024] In one aspect, methods described in this document may be performed in a non-naturally occurring host.

[025] In one aspect, this document provides hosts capable of producing 3-oxopent-4-enoyl-CoA, said hosts comprising at least one exogenous nucleic acid encoding a polypeptide that has both CoA transferase and β-ketothiolase activities. In one aspect, this document provides hosts capable of producing 1,3-butadiene, said hosts comprising at least one exogenous nucleic acid encoding a polypeptide that has both CoA transferase and β-ketothiolase activities.

[026] In one aspect, this document provides hosts capable of producing 3-oxo-acyl-CoA compounds, said hosts comprising at least one exogenous nucleic acid encoding a polypeptide that has both CoA transferase and β-ketothiolase activities.

[027] In one aspect, this document provides hosts capable of producing 3-oxo-enoyl-CoA compounds, said hosts comprising at least one exogenous nucleic acid encoding a polypeptide that has both CoA transferase and β-ketothiolase activities.

[028] In one aspect, this document provides hosts capable of producing 3-oxo-hydroxyacyl-CoA compounds, said hosts comprising at least one exogenous nucleic acid encoding a polypeptide that has both CoA transferase and β-ketothiolase activities.

[029] In one aspect, the principal carbon source for the methods and hosts described in this document derives from a biological feedstock. In some embodiments, the biological feedstock can be or can derive from, monosaccharides, disaccharides, lignocellulose, hemicellulose, cellulose, lignin, levulinic acid and formic acid, triglycerides, glycerol, fatty acids, agricultural waste, condensed distillers' solubles, or municipal waste.

[030] In one aspect, the principal carbon source for the methods and hosts described in this document derives from a non-biological feedstock. In some embodiments, the non-biological feedstock can be or can derive from natural gas, syngas, CO_2/H_2, methanol, ethanol, benzoate, non-volatile residue (NVR), or a caustic wash waste stream from cyclohexane oxidation processes, or terephthalic acid / isophthalic acid mixture waste streams.

[031] In some embodiments, the host microorganism's tolerance to high concentrations of one or more of 3-oxopent-4-enoyl-CoA, 3-oxopentanoyl-CoA, 3-oxohexanoyl-CoA, 3-buten-2-one, 3-butene-2-ol, or butadiene is improved through continuous cultivation in a selective environment.

[032] In some embodiments, the host microorganism's biochemical network is attenuated or augmented to (1) ensure the intracellular availability of acetyl-CoA, (2) create an NADH or NADPH imbalance that may only be balanced via the formation of one or more of 3-oxopent-4-enoyl-CoA, 3-oxopentanoyl-CoA, 3-oxohexanoyl-CoA, 3-buten-2-one, 3-butene-2-ol, or butadiene, (3) prevent degradation of central metabolites, central precursors leading to and including 3-oxopent-4-enoyl-CoA, 3-oxopentanoyl-CoA, 3-oxohexanoyl-CoA, 3-buten-2-one, 3-butene-2-ol, or butadiene, and (4) ensure efficient efflux from the cell.

[033] In some embodiments, a cultivation strategy is used to achieve anaerobic, micro-aerobic, or aerobic cultivation conditions.

[034] In some embodiments, the cultivation strategy includes limiting nutrients, such as limiting nitrogen, phosphate, or oxygen.

[035] In some embodiments, one or more of 3-oxopent-4-enoyl-CoA, 3-oxopentanoyl-CoA, 3-oxohexanoyl-CoA, 3-buten-2-one, 3-butene-2-ol, or butadiene are produced by a single type of microorganism, e.g., a non-naturally occurring host, such as a recombinant host, containing one or more exogenous nucleic acids, using, for example, a fermentation strategy.

[036] Any of the non-naturally occurring hosts can be a prokaryote such as a prokaryote from a genus selected from the group consisting of *Escherichia*, *Clostridia*, *Corynebacteria*, *Cupriavidus*, *Pseudomonas*, *Delftia*, *Bacilluss*, *Lactobacillus*, *Lactococcus*, and *Rhodococcus*. For example, the prokaryote can be selected from the group consisting of *Escherichia coli*, *Clostridium ljungdahlii*, *Clostridium autoethanogenum*, *Clostridium kluyveri*, *Corynebacterium glutamicum*, *Cupriavidus*

necator, *Cupriavidus metallidurans*, *Pseudomonas fluorescens*, *Pseudomonas putida*, *Pseudomonas oleavorans*, *Delftia acidovorans*, *Bacillus subtillis*, *Lactobacillus delbrueckii*, *Lactococcus lactis*, and *Rhodococcus equi*. Such prokaryotes also can be sources of genes for constructing recombinant host cells described herein that are capable of producing one or more of 3-oxopent-4-enoyl-CoA, 3-oxopentanoyl-CoA, 3-oxohexanoyl-CoA, 3-buten-2-one, 3-butene-2-ol, and butadiene.

[037] Any of the non-naturally occurring hosts can be a eukaryote such as a eukaryote from a genus selected from the group consisting of *Aspergillus*, *Saccharomyces*, *Pichia*, *Yarrowia*, *Issatchenkia*, *Debaryomyces*, *Arxula*, and *Kluyveromyces*. For example, the eukaryote can be selected from the group consisting of *Aspergillus niger*, *Saccharomyces cerevisiae*, *Pichia pastoris*, *Yarrowia lipolytica*, *Issathenkia orientalis*, *Debaryomyces hansenii*, *Arxula adenoinivorans*, and *Kluyveromyces lactis*. Such eukaryotes also can be sources of genes for constructing recombinant host cells described herein that are capable of producing one or more of 3-oxopent-4-enoyl-CoA, 3-oxopentanoyl-CoA, 3-oxohexanoyl-CoA, 3-buten-2-one, 3-butene-2-ol, and butadiene.

[038] This document also features a biochemical network comprising a polypeptide with both β-ketothiolase and CoA transferase activities, for example a polypeptide classified under EC 2.8.3.8, propenoate, acetyl Co-A, propenoyl-CoA, and 3-oxopent-4-enoyl-CoA, wherein the polypeptide enzymatically condenses propenoate and acetyl-CoA to propenoyl-CoA and converts propenoyl-CoA to 3-oxopent-4-enoyl-CoA. The biochemical network further can include a CoA transferase, a decarboxylase,

an alcohol dehydrogenase, a phenylacetaldehyde reductase, a linalool dehydratase, and combinations thereof to enzymatically convert 3-oxopent-4-enoyl-CoA to butadiene.

[039] This document also features a means for producing 3-oxopent-4-enoyl-CoA, wherein the means enzymatically condense propenoate and acetyl-CoA to propenoyl-CoA and convert propenoyl-CoA to 3-oxopent-4-enoyl-CoA. The means can include a polypeptide with both β-ketothiolase and CoA transferase activities, for example a polypeptide classified under EC 2.8.3.8. The means further can include one or more polypeptides for enzymatically converting 3-oxopent-4-enoyl-CoA to butadiene. The means can include a CoA transferase, a decarboxylase, an alcohol dehydrogenase, a phenylacetaldehyde reductase, a linalool dehydratase, and combinations thereof.

[040] This document also features a step for obtaining 3-oxopent-4-enoyl-CoA using a polypeptide with both β-ketothiolase and CoA transferase activities, for example a polypeptide classified under EC 2.8.3.8.

[041] In another aspect, this document features a composition comprising propenoate, acetyl-CoA, propenoyl-CoA, bio-derived 3-oxopent-4-enoyl-CoA, and a polypeptide with both β-ketothiolase and CoA transferase activities and classified under EC 2.8.3.8. The composition can be acellular or cellular.

[042] In another aspect, this document features a composition comprising bio-derived 3-oxopent-4-enoyl-CoA, 3-oxopentanoyl-CoA, 3-oxohexanoyl-CoA, 3-buten-2-one, 3-butene-2-ol, or butadiene. The composition can be acellular or cellular.

[043] In another aspect, this document features bio-derived or fermentation-derived 3-oxopent-4-enoyl-CoA, produced by the method of enzymatically condensing

propenoate and acetyl-CoA to propenoyl-CoA, and converting propenoyl-CoA to 3-oxopent-4-enoyl-CoA using a polypeptide with both β-ketothiolase and CoA transferase activities, for example a polypeptide classified under EC 2.8.3.8.

[044] Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention pertains. Although methods and materials similar or equivalent to those described herein can be used to practice the invention, suitable methods and materials are described below. All publications, patent applications, patents, and other references mentioned herein are incorporated by reference in their entirety. For example, see PCT/US2014/048606 (thiolases from archae), U.S. Patent No. 9,422,580 (e.g., Figures 3, 5, and 9), and U.S. Patent No. 9,422,578 (e.g., Figure 9), teaching, for example, propenoyl-CoA condensation with acetyl-CoA to form 3-oxopent-4-enoyl-CoA; and the thiolase reaction between 4-hydroxybutarate and acetyl-CoA to form 3-oxo-6-hydroxyhexanoyl-CoA, as depicted in various aspects in U.S. Application Serial No. 62/079,903; the thiolase reaction between acetyl-CoA and saturated alkanoic acids depicted in, for example, U.S. Application Serial Nos. 62/255,303 and 62/079,903. Some thiolases such as PaaJ are well known in the art. In case of conflict, the present specification, including definitions, will control. In addition, the materials, methods, and examples are illustrative only and not intended to be limiting. 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 of the disclosure will be apparent from the description and drawings, and from the claims. The word

"comprising" in the claims may be replaced by "consisting essentially of" or with "consisting of," according to standard practice in patent law.

DESCRIPTION OF DRAWINGS

[045] FIG. 1A is a schematic of an exemplary biochemical pathway leading to butadiene from propenoate (also known as acrylate) using (1) a polypeptide with both β-ketothiolase and CoA transferase activities; (2) a CoA transferase; (3) a decarboxylase; (4) an alcohol dehydrogenase or phenylacetaldehyde reductase; and (5) a linalool dehydratase.

[046] FIG. 1B shows four exemplary biochemical pathways to produce 3-keto-acyl-CoA esters from condensation of acetyl-CoA with an (i) alkanoic, (ii) alkenoic, (iii) hydroxy-, or (iv) halo-acid using a polypeptide with both β-ketothiolase and CoA transferase activities.

[047] FIG. 1C shows four exemplary biochemical pathways to produce 3-keto-acyl-CoA esters from condensation of acetyl-CoA with a salt of an (i) alkanoic, (ii) alkenoic, (iii) hydroxy-, or (iv) halo-acid using a polypeptide with both β-ketothiolase and CoA transferase activities.

[048] FIG. 2 contains the amino acid sequences of a *Clostridium propionicum* acetate CoA transferase (YdiF) (see GenBank Accession No. Q9L3F7, SEQ ID NO: 1), a modified *Clostridium propionicum* acetate CoA transferase (SEQ ID NO: 2), a modified *Clostridium propionicum* acetate CoA transferase containing an E324G mutation (SEQ ID NO: 3), a modified *Clostridium propionicum* acetate CoA transferase containing an E324L mutation (SEQ ID NO: 4), a *Peptostreptococcaceae acetate* CoA-transferase (Uniprot Accession No. U2L5C9, SEQ ID NO: 5), a *Firmicutes* bacterium

acetate CoA-transferase (Uniprot Accession No. R5ADR5, SEQ ID NO: 6), a *Megasphaera elsdenii* acetate CoA-transferase (Uniprot Accession No. G0VND6, SEQ ID NO: 7), a *Dyadobacter fermentans* Acetyl-CoA acetyltransferase (Uniprot Accession No. C6VTZ3, SEQ ID NO: 8), a *Salmonella enterica subsp. houtenaeserovar* acetate CoA-transferase (Uniprot Accession No. V1HBS2, SEQ ID NO: 9), an *Escherichia coli* acetyl-CoA:acetoacetyl-CoA transferase encoded by *atoAD* (GenBank Accession Nos. AAC75282.1 (beta subunit) and AAC75281.1 (alpha subunit), SEQ ID NOs:10 and 11, respectively), a *Pseudomonas putida* 3-oxoacid CoA-transferase encoded by *pcaIJ* (GenBank Accession No. ACA73091.1 (A subunit) and ACA73090.1 (B subunit), SEQ ID NO: 12 and 13, respectively), a *Chromobacterium violaceum* acetoacetate decarboxylase (GenBank Accession No. AAQ61181.1, SEQ ID NO: 14), a *Clostridium acetobutylicum* acetoacetate decarboxylase (Genbank Accession No. AAA63761.1, SEQ ID NO: 15), a *Nocardia rhamnosiphila* alcohol dehydrogenase (GenBank Accession No. WP_030525792, SEQ ID NO: 16), a *Rhodococcus* sp. ST-10 phenylacetaldehyde reductase (Uniprot Accession No. Q9ZN85, SEQ ID NO: 17); a *Castellaniella defragrans* linalool dehydratase (Uniprot Accession No. E1XUJ2, SEQ ID NO: 18), a *Clostridium aminobutyricum* CoA-transferase (Uniprot Accession No. Q9RM86, SEQ ID NO: 19), and a *Candida parapsilosis* carbonyl reductase (Uniprot Accession No. B2KJ46, SEQ ID NO: 20).

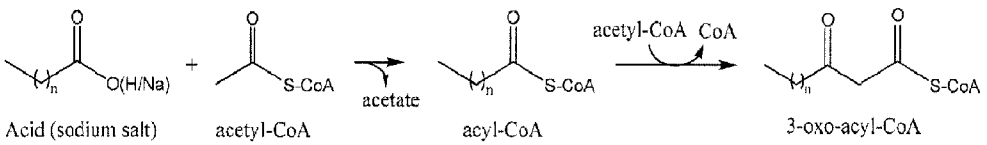
[049] FIG. 3 is a graphical representation of the secondary and tertiary structure of an *E. coli* YdiF and potential CoA transferase active site.

[050] FIG. 4 reports the sequence alignment of *E. coli* 2AHV YdiF and 16 other polypeptides (SEQ ID NOS 23-33, respectively, in order of appearance).

DETAILED DESCRIPTION

[051] In general, this document provides enzymes, non-naturally occurring pathways, cultivation strategies, feedstocks, non-naturally occurring host microorganisms, and attenuations to the host's biochemical network, for producing 3-oxopent-4-enoyl-CoA, 3-butene-2-one, 3-butene-2-ol, and/or butadiene from propenoate (also known as acrylate) or a salt form thereof (e.g., sodium propenoate) and acetyl-CoA using a polypeptide with both β-ketothiolase and CoA transferase activities.

[052] Also provided are enzymes, non-naturally occurring pathways, cultivation strategies, feedstocks, non-naturally occurring host microorganisms, and attenuations to the host's biochemical network, for producing 3-oxo-acyl-CoA compounds of formula (IIa), such as 3-oxo-pentanoyl-CoA or 3-oxo-hexanoyl-CoA, from carboxylic acids of formula (Ia) or salt forms thereof (such as sodium propionate or sodium butyrate), and acetyl-CoA using a polypeptide with both β-ketothiolase and CoA transferase activities.

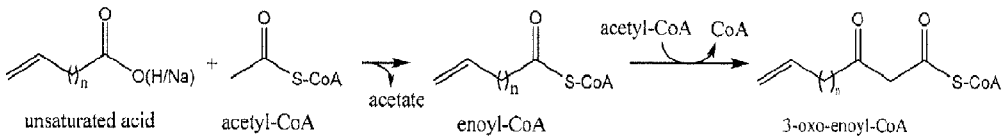


Acid (sodium salt) acetyl-CoA acyl-CoA 3-oxo-acyl-CoA

Formula (Ia) Formula (IIa)

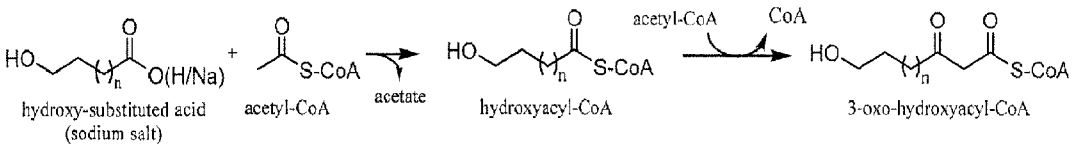
[053] Also provided are enzymes, non-naturally occurring pathways, cultivation strategies, feedstocks, non-naturally occurring host microorganisms, and attenuations to the host's biochemical network, for producing 3-oxo-enoyl-CoA compounds of formula (IIb), such as 3-oxo-pent-4-enoyl-CoA, 3-oxo-hex-5-enoyl-CoA, 3-oxo-hept-6-enoyl-CoA, or 3-oxo-non-8-enoyl-CoA, from unsaturated carboxylic acids of formula (Ib), such

as propenoate, 3-butenoate, 4-pentenoate, or 6-heptenoate, or salt forms thereof, and acetyl-CoA using a polypeptide with both β-ketothiolase and CoA transferase activities.



Formula (Ib) Formula (IIb)

[054] Also provided are enzymes, non-naturally occurring pathways, cultivation strategies, feedstocks, non-naturally occurring host microorganisms, and attenuations to the host's biochemical network, for producing 3-oxo-hydroxyacyl-CoA compounds of formula (IIc), such as 3-oxo-5-hydroxypentanoyl-CoA, 3-oxo-6-hydroxyhexanoyl-CoA, or 3-oxo-7-hydroxyheptanoyl-CoA, from hydroxy-substituted carboxylic acids of formula (Ic) or salt forms thereof (such as 3-hydroxypropionic acid, 4-hydroxybutyric acid, or 5-hydroxypentanoic acid), and acetyl-CoA using a polypeptide with both β-ketothiolase and CoA transferase activities.



Formula (Ic) Formula (IIc)

[055] As used herein, the term "central precursor" is used to denote any metabolite in any metabolic pathway shown herein leading to the synthesis of propenoyl-CoA, 3-oxopent-4-enoyl-CoA, 3-butene-2-one, 3-butene-2-ol, or butadiene. The term "central metabolite" is used herein to denote a metabolite that is produced in all microorganisms to support growth.

[056] Host microorganisms described herein can include endogenous pathways that can be manipulated such that propenoyl-CoA, 3-oxopent-4-enoyl-CoA, 3-butene-2-one, 3-butene-2-ol, or butadiene can be produced. In an endogenous pathway, the host microorganism naturally expresses all of the enzymes catalyzing the reactions within that pathway. A host microorganism containing an engineered pathway does not naturally express all of the enzymes catalyzing the reactions within that pathway but has been engineered such that all of the enzymes within the pathway are expressed in the host.

[057] The term "exogenous" as used herein with reference to a nucleic acid (or a protein) and a host refers to a nucleic acid that does not occur in (and cannot be obtained from) a cell of that particular type as it is found in nature (or a protein encoded by such a nucleic acid). Thus, a non-naturally-occurring nucleic acid is considered to be exogenous to a host once in the host. It is important to note that non-naturally-occurring nucleic acids can contain nucleic acid subsequences or fragments of nucleic acid sequences that are found in nature, provided the nucleic acid as a whole does not exist in nature. For example, a nucleic acid molecule containing a genomic DNA sequence within an expression vector is a non-naturally-occurring nucleic acid, and thus is exogenous to a host cell once introduced into the host, since that nucleic acid molecule as a whole (genomic DNA plus vector DNA) does not exist in nature. Thus, any vector, autonomously replicating plasmid, or virus (e.g., retrovirus, adenovirus, or herpes virus) that as a whole does not exist in nature is considered to be a non-naturally-occurring nucleic acid. It follows that genomic DNA fragments produced by PCR or restriction endonuclease treatment as well as cDNAs are considered to be non-naturally-occurring

nucleic acids since they exist as separate molecules not found in nature. It also follows that any nucleic acid containing a promoter sequence and polypeptide-encoding sequence (e.g., cDNA or genomic DNA) in an arrangement not found in nature is a non-naturally-occurring nucleic acid. A nucleic acid that is naturally-occurring can be exogenous to a particular host microorganism. For example, an entire chromosome isolated from a cell of yeast "x" is an exogenous nucleic acid with respect to a cell of yeast "y" once that chromosome is introduced into a cell of yeast "y."

[058] In contrast, the term "endogenous" as used herein with reference to a nucleic acid (e.g., a gene) (or a protein) and a host refers to a nucleic acid (or protein) that does occur in (and can be obtained from) that particular host as it is found in nature. Moreover, a cell "endogenously expressing" a nucleic acid (or protein) expresses that nucleic acid (or protein) as does a cell of that same particular type as it is found in nature. Moreover, a host "endogenously producing" or that "endogenously produces" a nucleic acid, protein, or other compound produces that nucleic acid, protein, or compound as does a host of the same particular type as it is found in nature.

[059] For example, depending on the host and the compounds produced by the host, one or more of the following enzymes may be expressed in the host: a polypeptide with both β-ketothiolase and CoA transferase activities, a decarboxylase, a CoA transferase, an alcohol dehydrogenase, a carbonyl reductase, a phenylacetaldehyde reductase, and a linalool dehydratase. For example, a recombinant host can include a polypeptide with both CoA transferase and β-ketothiolase activities; a recombinant host can include a polypeptide with both CoA transferase and β-ketothiolase activities and a decarboxylase; a recombinant host can include a polypeptide with both CoA transferase

and β-ketothiolase activities and a CoA transferase; a recombinant host can include a polypeptide with both CoA transferase and β-ketothiolase activities and an alcohol dehydrogenase or a phenylacetaldehyde reductase; a recombinant host can include a CoA transferase with both CoA transferase and β-ketothiolase activities and a linalool dehydratase; a recombinant host can include a polypeptide with both CoA transferase and β-ketothiolase activities, a CoA transferase, and a decarboxylase; a recombinant host can include a polypeptide with both CoA transferase and β-ketothiolase activities, a decarboxylase, and an alcohol dehydrogenase or a phenylacetaldehyde reductase; a recombinant host can include a polypeptide with both CoA transferase and β-ketothiolase activities, a decarboxylase, and a linalool dehydratase; a recombinant host can include a polypeptide with both CoA transferase and β-ketothiolase activities, a CoA transferase, and a linalool dehydratase; a recombinant host can include a polypeptide with both CoA transferase and β-ketothiolase activities, a CoA transferase, a decarboxylase, and a linalool dehydratase; a recombinant host can include a polypeptide with both CoA transferase and β-ketothiolase activities, a CoA transferase, a decarboxylase, an alcohol dehydrogenase, and a linalool dehydratase, or a recombinant host can include a polypeptide with both CoA transferase and β-ketothiolase activities, a CoA transferase, a decarboxylase, a phenylacetaldehyde reductase, and a linalool dehydratase.

[060] For example, a recombinant host can include an exogenous polypeptide with both CoA transferase and β-ketothiolase activities and produce 3-oxopentanoyl-CoA from propionate (e.g., sodium propionate) and acetyl-CoA.

[061] For example, a recombinant host can include an exogenous polypeptide with both CoA transferase and β-ketothiolase activities and produce 3-oxohexanoyl-CoA from butyrate (e.g., sodium butyrate) and acetyl-CoA.

[062] For example, a recombinant host can include an exogenous polypeptide with both CoA transferase and β-ketothiolase activities and produce 3-oxo-hex-5-enoyl-CoA from 3-butenoate (e.g., sodium 3-butenoate) and acetyl-CoA.

[063] For example, a recombinant host can include an exogenous polypeptide with both CoA transferase and β-ketothiolase activities and produce 3-oxo-hept-6-enoyl-CoA from 4-pentenoate (e.g., sodium 4-pentenoate) and acetyl-CoA.

[064] For example, a recombinant host can include an exogenous polypeptide with both CoA transferase and β-ketothiolase activities and produce 3-oxo-non-8-enoyl-CoA from 6-heptenoate (e.g., sodium 6-heptenoate) and acetyl-CoA.

[065] For example, a recombinant host can include an exogenous polypeptide with both CoA transferase and β-ketothiolase activities and produce 3-oxo-5-hydroxypentanoyl-CoA from 3-hydroxypropionic acid (e.g., sodium 3-hydroxypropionate) and acetyl-CoA. 3-oxo-5-hydroxypentanoyl-CoA is a useful intermediate for preparing nylon 5 compounds.

[066] For example, a recombinant host can include an exogenous polypeptide with both CoA transferase and β-ketothiolase activities and produce 3-oxo-6-hydroxyhexanoyl-CoA from 4-hydroxybutyric acid (e.g., sodium 4-hydroxybutyrate) and acetyl-CoA. 3-oxo-6-hydroxyhexanoyl-CoA is a useful intermediate for preparing nylon 6 compounds. For example, 3-oxo-6-hydroxyhexanoyl-CoA can be converted to 6-hydroxyhexanoic acid, which can be converted to one or more of adipic acid, 6-

aminohexanoic acid, hexamethylenediamine, caprolactam, and 1,6-hexanediol using one or more isolated enzymes such as dehydrogenases, reductases, hydratases, thioesterases, monooxygenases, and transaminases or using recombinant host cells expressing one or more such enzymes. For example, 3-oxo-6-hydroxyhexanoyl-CoA can be converted to one or more of adipic acid, 6-aminohexanoic acid, hexamethylenediamine, caprolactam, and 1,6-hexanediol according to methods described in United States Provisional Patent Application No. 62/079,903, whose disclosure is incorporated in its entirety by reference herein.

[067] For example, a recombinant host can include an exogenous polypeptide with both CoA transferase and β-ketothiolase activities and produce 3-oxo-7-hydroxyheptanoyl-CoA from 5-hydroxypentanoic acid (e.g., sodium 5-hydroxypentanoate) and acetyl-CoA. 3-oxo-7-hydroxyheptanoyl-CoA is a useful intermediate for preparing nylon 7 compounds.

[068] For example, a recombinant host can include an exogenous polypeptide with both CoA transferase and β-ketothiolase activities and produce 3-oxopent-4-enoyl-CoA from acrylate (e.g., sodium acrylate) and acetyl-CoA. 3-oxopent-4-enoyl-CoA can be converted to 3-butene-2-one and/or butadiene.

[069] For example, a recombinant host can include an exogenous polypeptide with both CoA transferase and β-ketothiolase activities and an exogenous CoA transferase, and produce 3-oxopent-4-enoate, which can be converted to 3-butene-2-one, 3-buten-2-ol, and/or butadiene.

[070] For example, a recombinant host can include an exogenous polypeptide with both CoA transferase and β-ketothiolase activities, an exogenous CoA transferase,

and an exogenous decarboxylase and produce 3-butene-2-one, which can be converted to 3-butene-2-ol or butadiene.

[071] For example, a recombinant host can include an exogenous polypeptide with both CoA transferase and β-ketothiolase activities, an exogenous CoA transferase, an exogenous decarboxylase, and an exogenous alcohol dehydrogenase or an exogenous phenylacetaldehyde reductase, and produce 3-butene-2-ol, which can be converted to butadiene.

[072] For example, a recombinant host can include an exogenous polypeptide with both CoA transferase and β-ketothiolase activities, an exogenous CoA transferase, an exogenous decarboxylase, an exogenous alcohol dehydrogenase or an exogenous phenylacetaldehyde reductase, and an exogenous linalool dehydratase, and produce butadiene. In some embodiments, a recombinant host can include an exogenous polypeptide with both CoA transferase and β-ketothiolase activities, an exogenous CoA transferase, an exogenous decarboxylase, an exogenous alcohol dehydrogenase, and an exogenous linalool dehydratase, and produce butadiene. In some embodiments, a recombinant host can include an exogenous polypeptide with both CoA transferase and β-ketothiolase activities, an exogenous CoA transferase, an exogenous decarboxylase, an exogenous phenylacetaldehyde reductase, and an exogenous linalool dehydratase, and produce butadiene.

[073] For example, a recombinant host can include an exogenous alcohol dehydrogenase or an exogenous phenylacetaldehyde reductase, and an exogenous linalool dehydratase, and produce butadiene.

[074] Within an engineered pathway, the enzymes can be from a single source, i.e., from one species or genera, or can be from multiple sources, i.e., from different species or genera. Nucleic acids encoding the enzymes described herein have been identified from various organisms and are readily available in publicly available databases such as GenBank or EMBL.

[075] Any of the enzymes described herein that can be used for production of one or more of 3-oxopentanoyl-CoA, 3-oxohexanoyl-CoA, 3-oxopent-4-enoyl-CoA, 3-oxo-hex-5-enoyl-CoA, 3-oxo-hept-6-enoyl-CoA, 3-oxo-non-8-enoyl-CoA, 3-oxo-5-hydroxypentanoyl-CoA, 3-oxo-6-hydroxyhexanoyl-CoA, 3-oxo-7-hydroxyheptanoyl-CoA, 3-butene-2-one, 3-butene-2-ol, and/or butadiene, *in vitro* or in a recombinant host, can have at least 70% sequence identity (homology) (e.g., at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100%) to the amino acid sequence of the corresponding wild-type enzyme. It will be appreciated that the sequence identity can be determined on the basis of the mature enzyme (e.g., with any signal sequence removed) or on the basis of the immature enzyme (e.g., with any signal sequence included). It also will be appreciated that the initial methionine residue may or may not be present on any of the enzyme sequences described herein.

[076] For example, a polypeptide with both CoA transferase and β-ketothiolase activities described herein can have at least 70% sequence identity (homology) (e.g., at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100%) to the amino acid sequence of a *Clostridium propionicum* (YdiF) (see GenBank Accession No. Q9L3F7, SEQ ID NO: 1), a *Clostridium aminobutyricum* (see GenBank Accession No. Q9RM86, SEQ ID NO: 19, a *Peptostreptococcaceae* (see Uniprot

Accession No. U2L5C9, SEQ ID NO: 5); a *Firmicutes* bacterium (see Uniprot Accession No. R5ADR5, SEQ ID NO: 6), a *Megasphaera elsdenii* (see Uniprot Accession No. G0VND6, SEQ ID NO: 7), or a *Salmonella enterica subsp. houtenaeserovar* (see Uniprot Accession No. V1HBS2, SEQ ID NO: 9) acetate CoA transferase, or a modified *Clostridium propionicum* acetate CoA transferase having the amino acid sequence set forth in SEQ ID NO: 2. See FIG. 2.

[077] For example, a CoA transferase described herein can have at least 70% sequence identity (homology) (e.g., at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100%) to the amino acid sequence of an *Escherichia coli* acetyl-CoA:acetoacetyl-CoA transferase encoded by *atoAD* (see GenBank Accession Nos. AAC75282.1 (beta subunit) and AAC75281.1 (alpha subunit), SEQ ID NOs:10 and 11, respectively) or a *Pseudomonas putida* 3-oxoacid CoA-transferase encoded by *pcaIJ* (see Genbank Accession No. ACA73091.1 (A subunit) and ACA73090.1 (B subunit), SEQ ID NO: 12 and 13, respectively).

[078] For example, a *decarboxylase* described herein such as an acetoacetate decarboxylase can have at least 70% sequence identity (homology) (e.g., at least 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, or 100%) to the amino acid sequence of a *Chromobacterium violaceum* acetoacetate decarboxylase (see Genbank Accession No. AAQ61181.1, SEQ ID NO: 14) or a *Clostridium acetobutylicum* acetoacetate decarboxylase (Genbank Accession No. AAA63761.1, SEQ ID NO: 15). See, FIG. 2.

[079] For example, an alcohol dehydrogenase described herein can have at least 70% sequence identity (homology) (e.g., at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100%) to the amino acid sequence of a

Nocardia rhamnosiphila (for example GenBank Accession No. WP_030525792, SEQ ID NO: 16) alcohol dehydrogenase.

[080] For example, a phenylacetaldehyde reductase described herein can have at least 70% sequence identity (homology) (e.g., at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100%) to the amino acid sequence of a *Rhodococcus* sp. ST-10 (for example GenBank Accession No. BAD51480, SEQ ID NO: 17) phenylacetaldehyde reductase.

[081] For example, a linalool dehydratase described herein can have at least 70% sequence identity (homology) (e.g., at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100%) to the amino acid sequence of a *Castellaniella defragrans* (for example GenBank Accession No. CBW30776, SEQ ID NO: 18) linalool dehydratase. In one embodiment, a linalool dehydratase described herein can be a mutant of the linalool dehydratase from *Castellaniella defragrans*, for example one of the mutants described in U.S. Provisional Patent Application No. 62/126,279, whose disclosure is incorporated in its entirety by reference herein, or one of the mutants described in U.S. Provisional Patent Applications No. 62/126,299 or 62/126,315, whose disclosure is incorporated in its entirety by reference herein.

[082] For example, a carbonyl reductase described herein can have at least 70% sequence identity (homology) (e.g., at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100%) to the amino acid sequence of a *Candida parapsilosis* (for example GenBank Accession No. AFD29185.1, SEQ ID NO: 20) carbonyl reductase.

[083] The percent identity (homology) between two amino acid sequences can be determined as follows. First, the amino acid sequences are aligned using the BLAST 2 Sequences (Bl2seq) program from the stand-alone version of BLASTZ containing BLASTP version 2.0.14. This stand-alone version of BLASTZ can be obtained from the U.S. government's National Center for Biotechnology Information web site (www.ncbi.nlm.nih.gov). Instructions explaining how to use the Bl2seq program can be found in the readme file accompanying BLASTZ. Bl2seq performs a comparison between two amino acid sequences using the BLASTP algorithm. To compare two amino acid sequences, the options of Bl2seq are set as follows: -i is set to a file containing the first amino acid sequence to be compared (e.g., C:\seq1.txt); -j is set to a file containing the second amino acid sequence to be compared (e.g., C:\seq2.txt); -p is set to blastp; -o is set to any desired file name (e.g., C:\output.txt); and all other options are left at their default setting. For example, the following command can be used to generate an output file containing a comparison between two amino acid sequences: C:\Bl2seq –i c:\seq1.txt –j c:\seq2.txt –p blastp –o c:\output.txt. If the two compared sequences share homology (identity), then the designated output file will present those regions of homology as aligned sequences. If the two compared sequences do not share homology (identity), then the designated output file will not present aligned sequences. Similar procedures can be followed for nucleic acid sequences except that blastn is used.

[084] Once aligned, the number of matches is determined by counting the number of positions where an identical amino acid residue is presented in both sequences. The percent identity (homology) is determined by dividing the number of

matches by the length of the full-length polypeptide amino acid sequence followed by multiplying the resulting value by 100. It is noted that the percent identity (homology) value is rounded to the nearest tenth. For example, 78.11, 78.12, 78.13, and 78.14 is rounded down to 78.1, while 78.15, 78.16, 78.17, 78.18, and 78.19 is rounded up to 78.2. It also is noted that the length value will always be an integer.

[085] It will be appreciated that a number of nucleic acids can encode a polypeptide having a particular amino acid sequence. The degeneracy of the genetic code is well known to the art; i.e., for many amino acids, there is more than one nucleotide triplet that serves as the codon for the amino acid. For example, codons in the coding sequence for a given enzyme can be modified such that optimal expression in a particular species (e.g., bacteria or fungus) is obtained, using appropriate codon bias tables for that species.

[086] Functional fragments of any of the enzymes described herein can also be used in the methods of the document. The term "functional fragment" as used herein refers to a peptide fragment of a protein that has at least 25% (e.g., at least 30%, 40%, 50%, 60%, 70%, 75%, 80%, 85%, 90%, 95%, 98%, 99%, 100%, or even greater than 100%) of the activity of the corresponding mature, full-length, wild-type protein. The functional fragment can generally, but not always, be comprised of a continuous region of the protein, wherein the region has functional activity.

[087] The secondary structure of a known *E. coli* YdiF was analyzed along with that of other polypeptides known or expected to have dual CoA transferase and β-ketothiolase activities. A sequence alignment chart is provided in Fig. 4.

[088] At least three conserved amino acid sequence motifs are thought to be associated with thiolase activity in polypeptides having dual CoA transferase and β-ketothiolase activities: a EXGXXG motif, a GXGG(A/F) motif, and a I/V/A/LTE motif. *See* Fig. 4. The EXGXXG motif is associated with a turn between two beta sheets in the secondary structure of polypeptides having dual CoA transferase and β-ketothiolase activities. The GXGG(A/F) motif is associated with a strand before an alpha helix in the secondary structure of polypeptides having dual CoA transferase and β-ketothiolase activities. The I/V/A/LTE motif is associated with a strand between two beta sheets in the secondary structure of polypeptides having dual CoA transferase and β-ketothiolase activities.

[089] All CoA transferases identified herein as having thiolase activity share the I/V/A/LTE motif. On this basis CoA transferase from *N. thermophilus* is likely to have ketothiolase activity if expressed in a thermophilic host and tested at elevated reaction temperature.

[090] In at least one embodiment, the polypeptide having both CoA transferase and β-ketothiolase activities is a CoA transferase whose amino acid sequence includes the motifs EXGXXG and GXGG(A/F). In at least one embodiment, the polypeptide having both CoA transferase and β-ketothiolase activities is a CoA transferase whose amino acid sequence includes the motif I/V/A/LTE.

[091] This document also provides (i) functional variants of the enzymes used in the methods of the document and (ii) functional variants of the functional fragments described above. Functional variants of the enzymes and functional fragments can contain additions, deletions, or substitutions relative to the corresponding wild-type

sequences. Enzymes with substitutions will generally have not more than 50 (e.g., not more than one, two, three, four, five, six, seven, eight, nine, ten, 12, 15, 20, 25, 30, 35, 40, or 50) amino acid substitutions (e.g., conservative substitutions). This applies to any of the enzymes described herein and functional fragments. A conservative substitution is a substitution of one amino acid for another with similar characteristics. Conservative substitutions include substitutions within the following groups: valine, alanine, and glycine; leucine, valine, and isoleucine; aspartic acid and glutamic acid; asparagine and glutamine; serine, cysteine, and threonine; lysine and arginine; and phenylalanine and tyrosine. The nonpolar hydrophobic amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan, and methionine. The polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine, and glutamine. The positively charged (basic) amino acids include arginine, lysine, and histidine. The negatively charged (acidic) amino acids include aspartic acid and glutamic acid. Any substitution of one member of the above-mentioned polar, basic, or acidic groups by another member of the same group can be deemed a conservative substitution. By contrast, a nonconservative substitution is a substitution of one amino acid for another with dissimilar characteristics.

[092] In some embodiments, one or more mutations can be introduced into the polypeptide with both CoA transferase and β-ketothiolase activities (e.g., SEQ ID NOs: 1, 2, 5, 6, 7, 8, 9, 19) without impacting either activity. For example, a polypeptide with both CoA transferase and β-ketothiolase activities can have a substitution at one or more of positions 38, 60, 112, 258, and 390 without impacting either the CoA transferase or β-ketothiolase activity. For example, a polypeptide with both CoA

transferase and β-ketothiolase activities can have a substitution of an alanine for any of the following residues: the serine at position 38, the serine at position 60, the cysteine at position 112, the cysteine at position 258, or the cysteine at position 390 of SEQ ID NO: 1, without affecting the β-ketothiolase or CoA transferase activity of the polypeptide. Residues 38, 60, 112, 258, and 390 have homology to the active site of other known β-ketothiolases. For example, a polypeptide with both CoA transferase and β-ketothiolase activities can have a substitution of a glycine for the glutamic acid at position 324 of SEQ ID NO:1 and maintain both activities. Position 324 was identified as the active site of the polypeptide for its CoA transferase activity. See, Selmer *et al., Eur. J. Biochem.*, 269:372-380 (2002).

[093] In some embodiments, one or more mutations can be introduced into the polypeptide with both CoA transferase and β-ketothiolase activities to reduce or abolish one of the activities, as the active sites are independent of each other. For example, a polypeptide with both CoA transferase and β-ketothiolase activities can have a substitution of a leucine for the glutamic acid at position 324 of SEQ ID NO:1 to reduce the CoA transferase activity while maintaining the β-ketothiolase activity.

[094] Deletion variants can lack one, two, three, four, five, six, seven, eight, nine, ten, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 amino acid segments (of two or more amino acids) or non-contiguous single amino acids. Additions (addition variants) include fusion proteins containing: (a) any of the enzymes described herein or a fragment thereof; and (b) internal or terminal (C or N) irrelevant or heterologous amino acid sequences. In the context of such fusion proteins, the term "heterologous amino acid sequences" refers to an amino acid sequence other than (a). A heterologous

sequence can be, for example a sequence used for purification of the recombinant protein (e.g., FLAG, polyhistidine (e.g., hexahistidine (SEQ ID NO: 21), hemagglutinin (HA), glutathione-S-transferase (GST), or maltosebinding protein (MBP)). Heterologous sequences also can be proteins useful as detectable markers, for example, luciferase, green fluorescent protein (GFP), or chloramphenicol acetyl transferase (CAT). In some embodiments, the fusion protein contains a signal sequence from another protein. In certain host cells (e.g., yeast host cells), expression and/or secretion of the target protein can be increased through use of a heterologous signal sequence. In some embodiments, the fusion protein can contain a carrier (e.g., KLH) useful, e.g., in eliciting an immune response for antibody generation, or ER or Golgi apparatus retention signals. Heterologous sequences can be of varying length and in some cases can be a longer sequences than the full-length target proteins to which the heterologous sequences are attached.

[095] Engineered hosts can naturally express none or some (e.g., one or more, two or more, three or more, four or more, five or more, or six or more) of the enzymes of the pathways described herein. Thus, a pathway within an engineered host can include all exogenous enzymes, or can include both endogenous and exogenous enzymes. Endogenous genes of the engineered hosts also can be disrupted to prevent the formation of undesirable metabolites or prevent the loss of intermediates in the pathway through other enzymes acting on such intermediates. Engineered hosts can be referred to as recombinant hosts or recombinant host cells. As described herein, recombinant hosts can include nucleic acids encoding one or more of a polypeptide with both β-ketothiolase and CoA transferase activities, an alcohol dehydrogenase, a

phenylacetaldehyde reductase, a decarboxylase, and a linalool dehydratase as described herein.

[096] In addition, the production of 3-oxopentanoyl-CoA, 3-oxohexanoyl-CoA, 3-oxopent-4-enoyl-CoA, 3-oxo-hex-5-enoyl-CoA, 3-oxo-hept-6-enoyl-CoA, 3-oxo-non-8-enoyl-CoA, 3-oxo-5-hydroxypentanoyl-CoA, 3-oxo-6-hydroxyhexanoyl-CoA , 3-oxo-7-hydroxyheptanoyl-CoA, 3-butene-2-one, 3-butene-2-ol, and/or butadiene can be performed *in vitro* with the appropriate substrates (e.g., a short chain alkyl or alkenyl carboxylate or salt forms thereof such as sodium acrylate, sodium propioanate, or sodium butyrate, and a CoA source such as acetyl CoA) and using one or more of the isolated enzymes described herein, using a lysate (e.g., a cell lysate) from a host microorganism as a source of the enzymes, or using a plurality of lysates from different host microorganisms as the source of the enzymes.

[097] The reactions of the pathways described herein can be performed in one or more host strains (a) naturally expressing one or more relevant enzymes, (b) genetically engineered to express one or more relevant enzymes, or (c) naturally expressing one or more relevant enzymes and genetically engineered to express one or more relevant enzymes. Alternatively, relevant enzymes can be extracted from any of the above types of host cells and used in a purified or semi-purified form. Moreover, such extracts include lysates (e.g. cell lysates) that can be used as sources of relevant enzymes. In the methods provided by the document, all the steps can be performed in host cells, all the steps can be performed using extracted enzymes, or some of the steps can be performed in cells and others can be performed using extracted enzymes.

Enzymes generating Butadiene

[098] As depicted in FIG. 1, butadiene can be biosynthesized from propenoate, for example sodium propenoate. Propenoate and acetyl-CoA can be condensed to propenoyl-CoA using a polypeptide with both CoA transferase and β-ketothiolase activities, for example a polypeptide classified under EC 2.8.3.-, for example a polypeptide classified under 2.8.3.8. For example, a polypeptide with both CoA transferase and β-ketothiolase activities can be classified under EC 2.8.3.8 such as the polypeptides from *Clostridium propionicum*, *Peptostreptococcaceae*, *a Firmicutes* bacterium, or a *Megasphaera elsdenii*, or classified under 2.8.3.-, such as the polypeptide from *Clostridium aminobutyricum*. Propenoyl-CoA can be converted to 3-oxopent-4-enoyl-CoA using a polypeptide with both CoA transferase and β-ketothiolase activities, for example a polypeptide classified under EC 2.8.3.-, for example a polypeptide classified under 2.8.3.8, for example the polypeptides from *Clostridium propionicum, Peptostreptococcaceae, a Firmicutes* bacterium, a *Megasphaera elsdenii*, or a *Salmonella enterica subsp. houtenaeserovar*. See, e.g., SEQ ID NOs: 1 and 5-9. See Table 1 for examples of such transferases. For example, the polypeptide with both CoA transferase and β-ketothiolase activities can be the product of the YdiF gene from *C. propionicum*.

[099] In other embodiments, a polypeptide with both CoA transferase and β-ketothiolase activities can be classified under EC 2.8.3.1, EC 2.8.3.10, EC 2.8.3.11, EC 2.8.3.12, EC 2.8.3.13, EC 2.8.3.15, EC 2.8.3.16, EC 2.8.3.17, EC 2.8.3.18, EC 2.8.3.19, EC 2.8.3.20, EC 2.8.3.21, EC 2.8.3.5, EC 2.8.3.6, or EC 2.8.3.9. See Table 2 for examples of such transferases.

TABLE 1

EC	Uniprot Gene Accession Number	Enzyme	Organism
2.8.3.8	Q9L3F7	Acetate/Propionate CoA-transferase	*Clostridium propionicum*
2.8.3.8	V1HBS2	Acetate CoA-transferase YdiF	*Salmonella enterica subsp. houtenaeserovar*
2.8.3.8	U2L5C9	Acetate CoA-transferase YdiF	*Peptostreptococcaceae bacterium oral taxon*
2.8.3.8	R5ADR5	Acetate CoA-transferase YdiF	*Firmicutes bacterium CAG*
2.8.3.8	G0VND6	Acetate CoA-transferase YdiF	*Megasphaera elsdenii DSM 20460*

TABLE 2

EC	Uniprot Gene Accession Number	Enzyme	Organism
2.8.3.10	J1G510	Acetyl-CoA hydrolase/transferase	*Citrobacter sp. A1*
2.8.3.18	B3EY95	Succinyl-CoA:acetate CoA-transferase	*Acetobacter aceti*
2.8.3.20	A9WGE3	Succinyl-CoA-D-citramalate CoA-transferase	*Chloroflexus aurantiacus*
2.8.3.-	Q9RM86	4-Hydroxybutyrate-CoA transferase	*Clostridium aminobutyricum*

[0101] In some embodiments, a CoA-transferase classified under, for example, EC 2.8.3.- (e.g., EC 2.8.3.8 or EC 2.8.3.6) such as the gene products of *AtoAD* (see SEQ ID NOs: 10 and 11) or *pcalJ* (see SEQ ID NOs: 12 and 13) is used to hydrolyze the CoA moiety and convert 3-oxopent-4-enoyl-CoA to 3-oxopent-4-enoate.

[0102] In some embodiments, a decarboxylase such as an acetoacetate decarboxylase classified, for example, under EC 4.1.1.4 can be used to remove the carboxy group from 3-oxopent-4-enoate to produce 3-buten-2-one. For example, a suitable acetoacetate decarboxylase can have at least 70% sequence identity to the amino acid sequence set forth in SEQ ID NO: 14 or SEQ ID NO: 15. This reaction also can occur spontaneously.

[0103] In some embodiments, an alcohol dehydrogenase classified under EC 1.1.1.- such as an alcohol dehydrogenase from *Nocardia rhamnosiphila* can be used to convert 3-buten-2-one to 3-buten-2-ol. For example, a suitable alcohol

dehydrogenase can have at least 70% sequence identity to the amino acid sequence set forth in SEQ ID NO: 16.

[0104] In some embodiments, a carbonyl reductase classified under EC 1.1.1.184 such as the carbonyl reductase from *Candida parapsilosis* can be used to convert 3-buten-2-one to 3-buten-2-ol. For example, a suitable carbonyl reductase can have at least 70% sequence identity to the amino acid sequence set forth in SEQ ID NO:20.

[0105] In some embodiments, a phenylacetaldehyde reductase classified under EC 1.2.1.39 such as an phenylacetaldehyde reductase from *Rhodococcus* sp. ST-10 can be used to convert 3-buten-2-one to 3-buten-2-ol. For example, a suitable phenylacetaldehyde reductase can have at least 70% sequence identity to the amino acid sequence set forth in SEQ ID NO: 17.

[0106] In some embodiments, a dehydratase enzyme classified in EC 4.2.1.-, such as linalool dehydratase classified, for example, under EC 4.2.1.127, for example a *Castellaniella defragrans* linalool dehydratase, can be used to convert 3-buten-2-ol to butadiene. For example, a suitable linalool dehydratase can have at least 70% sequence identity to the amino acid sequence set forth in SEQ ID NO: 18.

Biochemical pathways to butadiene

[0107] In some embodiments, propenoate, for example sodium propenoate, and acetyl-CoA are condensed to propenoyl-CoA using a polypeptide with both CoA transferase and β-ketothiolase activities, for example an enzyme classified under EC 2.8.3.-, for example an enzyme classified under 2.8.3.8; then propenoyl-CoA is converted to 3-oxopent-4-enoyl-CoA using a polypeptide with both CoA transferase and

β-ketothiolase activities, for example an enzyme classified under EC 2.8.3.-, for example an enzyme classified under 2.8.3.8; followed by conversion of 3-oxopent-4-enoyl-CoA to 3-oxopent-4-enoate by a CoA transferase classified, for example, under 2.8.3.8 or EC 2.8.3.6; followed by conversion of 3-oxopent-4-enoate to 3-butene-2-one either spontaneously or by a *decarboxylase* classified, for example, under EC 4.1.1.4; followed by conversion of 3-butene-2-one to 3-buten-2-ol by an alcohol dehydrogenase classified, for example, under EC 1.1.1-, or a phenylacetaldehyde reductase classified, for example, under EC 1.2.1.39; followed by conversion of 3-buten-2-ol to 1,3-butadiene by a linalool dehydratase classified, for example, under EC 4.2.1.127.

Cultivation strategy

[0108] In some embodiments, one or more of 3-oxopent-4-enoyl-CoA, 3-oxopentanoyl-CoA, 3-oxohexanoyl-CoA, 3-oxo-hex-5-enoyl-CoA, 3-oxo-hept-6-enoyl-CoA, 3-oxo-non-8-enoyl-CoA, 3-oxo-5-hydroxypentanoyl-CoA, 3-oxo-6-hydroxyhexanoyl-CoA, 3-oxo-7-hydroxyheptanoyl-CoA, 3-butene-2-one, 3-buten-2-ol, and butadiene are biosynthesized in a recombinant host using anaerobic, aerobic, or micro-aerobic cultivation conditions. In some embodiments, the cultivation strategy entails nutrient limitation such as nitrogen, phosphate, or oxygen limitation.

[0109] In some embodiments, a cell retention strategy using, for example, ceramic hollow fiber membranes can be employed to achieve and maintain a high cell density during either fed-batch or continuous fermentation.

[0110] In some embodiments, the principal carbon source fed to the fermentation in the synthesis of one or more of 3-oxopent-4-enoyl-CoA, 3-oxopentanoyl-CoA, 3-oxohexanoyl-CoA, 3-oxo-hex-5-enoyl-CoA, 3-oxo-hept-6-enoyl-CoA, 3-oxo-non-8-

enoyl-CoA, 3-oxo-5-hydroxypentanoyl-CoA, 3-oxo-6-hydroxyhexanoyl-CoA, 3-oxo-7-hydroxyheptanoyl-CoA, 3-butene-2-one, 3-butene-2-ol, and butadiene can derive from biological or non-biological feedstocks.

[0111] In some embodiments, the biological feedstock can be, or can derive from, monosaccharides, disaccharides, lignocellulose, hemicellulose, cellulose, lignin, levulinic acid and formic acid, triglycerides, glycerol, fatty acids, agricultural waste, condensed distillers' solubles, or municipal waste.

[0112] The efficient catabolism of crude glycerol stemming from the production of biodiesel has been demonstrated in several microorganisms such as *Escherichia coli*, *Cupriavidus necator*, *Pseudomonas oleavorans*, *Pseudomonas putida*, and *Yarrowia lipolytica* (Lee *et al.*, *Appl. Biochem. Biotechnol.*, 2012, 166:1801 – 1813; Yang *et al.*, *Biotechnology for Biofuels*, 2012, 5:13; Meijnen *et al.*, *Appl. Microbiol. Biotechnol.*, 2011, 90:885 - 893).

[0113] The efficient catabolism of lignocellulosic-derived levulinic acid has been demonstrated in several organisms such as *Cupriavidus necator* and *Pseudomonas putida* in the synthesis of 3-hydroxyvalerate via the precursor propanoyl-CoA (Jaremko and Yu, 2011, *supra*; Martin and Prather, *J. Biotechnol.*, 2009, 139:61 – 67).

[0114] The efficient catabolism of lignin-derived aromatic compounds such as benzoate analogues has been demonstrated in several microorganisms such as *Pseudomonas putida* and *Cupriavidus necator* (Bugg *et al.*, *Current Opinion in Biotechnology*, 2011, 22, 394 – 400; Pérez-Pantoja *et al.*, *FEMS Microbiol. Rev.*, 2008, 32, 736 – 794).

[0115] The efficient utilization of agricultural waste, such as olive mill waste water, has been demonstrated in several microorganisms, including *Yarrowia lipolytica* (Papanikolaou *et al.*, *Bioresour. Technol.*, 2008, 99(7):2419 - 2428).

[0116] The efficient utilization of fermentable sugars such as monosaccharides and disaccharides derived from cellulosic, hemicellulosic, cane and beet molasses, cassava, corn, and other agricultural sources has been demonstrated for several microorganism such as *Escherichia coli*, *Corynebacterium glutamicum*, *Lactobacillus delbrueckii* and *Lactococcus lactis* (see, e.g., Hermann *et al*, *J. Biotechnol.*, 2003, 104:155 – 172; Wee *et al.*, *Food Technol. Biotechnol.*, 2006, 44(2):163 – 172; Ohashi *et al.*, *J. Bioscience and Bioengineering*, 1999, 87(5):647 - 654).

[0117] The efficient utilization of furfural, derived from a variety of agricultural lignocellulosic sources, has been demonstrated for *Cupriavidus necator* (Li *et al.*, *Biodegradation*, 2011, 22:1215 – 1225).

[0118] In some embodiments, the non-biological feedstock can be or can derive from natural gas, syngas, CO_2/H_2, methanol, ethanol, benzoate, non-volatile residue (NVR) caustic wash waste stream from cyclohexane oxidation processes, or terephthalic acid / isophthalic acid mixture waste streams.

[0119] The efficient catabolism of methanol has been demonstrated for the methylotrophic yeast *Pichia pastoris*.

[0120] The efficient catabolism of ethanol has been demonstrated for *Clostridium kluyveri* (Seedorf *et al.*, *Proc. Natl. Acad. Sci. USA*, 2008, 105(6) 2128 - 2133).

[0121] The efficient catabolism of CO_2 and H_2, which may be derived from natural gas and other chemical and petrochemical sources, has been demonstrated for *Cupriavidus necator* (Prybylski *et al.*, *Energy, Sustainability and Society*, 2012, 2:11).

[0122] The efficient catabolism of syngas has been demonstrated for numerous microorganisms, such as *Clostridium ljungdahlii* and *Clostridium autoethanogenum* (Köpke *et al.*, *Applied and Environmental Microbiology*, 2011, 77(15):5467 – 5475).

[0123] The efficient catabolism of the non-volatile residue waste stream from cyclohexane processes has been demonstrated for numerous microorganisms, such as *Delftia acidovorans* and *Cupriavidus necator* (Ramsay *et al.*, *Applied and Environmental Microbiology*, 1986, 52(1):152 – 156).

[0124] In some embodiments, the host microorganism is a prokaryote. For example, the prokaryote can be a bacterium from the genus *Escherichia* such as *Escherichia coli*; from the genus *Clostridia* such as *Clostridium ljungdahlii*, *Clostridium autoethanogenum* or *Clostridium kluyveri*; from the genus *Corynebacteria* such as *Corynebacterium glutamicum*; from the genus *Cupriavidus* such as *Cupriavidus necator*, or *Cupriavidus metallidurans;* from the genus *Pseudomonas* such as *Pseudomonas fluorescens, Pseudomonas putida*, or *Pseudomonas oleavorans*; from the genus *Delftia* such as *Delftia acidovorans;* from the genus *Bacillus* such as *Bacillus subtillis;* from the genus *Lactobacillus* such as *Lactobacillus delbrueckii*; or from the genus *Lactococcus* such as *Lactococcus lactis*. Such prokaryotes also can be a source of genes to construct recombinant host cells described herein that are capable of producing one or more of 3-oxopent-4-enoyl-CoA, 3-oxopentanoyl-CoA, 3-oxohexanoyl-CoA, 3-oxo-hex-5-enoyl-CoA, 3-oxo-hept-6-enoyl-CoA, 3-oxo-non-8-enoyl-CoA, 3-oxo-5-

hydroxypentanoyl-CoA, 3-oxo-6-hydroxyhexanoyl-CoA, 3-oxo-7-hydroxyheptanoyl-CoA, 3-butene-2-one, 3-butene-2-ol, and butadiene.

[0125] In some embodiments, the host microorganism is a eukaryote. For example, the eukaryote can be a filamentous fungus, e.g., one from the genus *Aspergillus* such as *Aspergillus niger*. Alternatively, the eukaryote can be a yeast, e.g., one from the genus *Saccharomyces* such as *Saccharomyces cerevisiae*; from the genus *Pichia* such as *Pichia pastoris*; or from the genus *Yarrowia* such as *Yarrowia lipolytica*; from the genus *Issatchenkia* such as *Issathenkia orientalis*; from the genus *Debaryomyces* such as *Debaryomyces hansenii*; from the genus *Arxula* such as *Arxula adenoinivorans*; or from the genus *Kluyveromyces* such as *Kluyveromyces lactis*. Such eukaryotes also can be a source of genes to construct recombinant host cells described herein that are capable of producing 3-oxopent-4-enoyl-CoA, 3-oxopentanoyl-CoA, 3-oxohexanoyl-CoA, 3-oxo-hex-5-enoyl-CoA, 3-oxo-hept-6-enoyl-CoA, 3-oxo-non-8-enoyl-CoA, 3-oxo-5-hydroxypentanoyl-CoA, 3-oxo-6-hydroxyhexanoyl-CoA, 3-oxo-7-hydroxyheptanoyl-CoA, 3-butene-2-one, 3-butene-2-ol, and/or butadiene.

Metabolic engineering

[0126] The present document provides methods involving less than all the steps described for all the pathways disclosed herein. Such methods can involve, for example, one, two, three, four, five, six, seven, eight, nine, ten, eleven, twelve or more of such steps. Where less than all the steps are included in such a method, the first, and in some embodiments the only, step can be any one of the steps listed.

[0127] Furthermore, recombinant hosts described herein can include any combination of the above enzymes such that one or more of the steps, e.g., one, two,

three, four, five, six, seven, eight, nine, ten, or more of such steps, can be performed within a recombinant host. This document provides host cells of any of the genera and species listed and genetically engineered to express one or more (e.g., two, three, four, five, six, seven, eight, nine, 10, 11, 12, or more) recombinant forms of any of the enzymes recited in the document. Thus, for example, the host cells can contain exogenous nucleic acids encoding enzymes catalyzing one or more of the steps of any of the pathways described herein.

[0128] In addition, this document recognizes that where enzymes have been described as accepting CoA-activated substrates, analogous enzyme activities associated with [acp]-bound substrates exist that are not necessarily in the same enzyme class.

[0129] Also, this document recognizes that where enzymes have been described as accepting (R)-enantiomers of substrate, analogous enzyme activities associated with (S)-enantiomer substrates exist that are not necessarily in the same enzyme class.

[0130] This document also recognizes that where an enzyme is shown to accept a particular co-factor, such as NADPH, or co-substrate, such as acetyl-CoA, many enzymes are promiscuous in terms of accepting a number of different co-factors or co-substrates in catalyzing a particular enzyme activity. Also, this document recognizes that where enzymes have high specificity for e.g., a particular co-factor such as NADH, an enzyme with similar or identical activity that has high specificity for the co-factor NADPH may be in a different enzyme class.

[0131] In some embodiments, the enzymes in the pathways outlined herein are the result of enzyme engineering via non-direct or rational enzyme design approaches

with aims of improving activity, improving specificity, reducing feedback inhibition, reducing repression, improving enzyme solubility, changing stereo-specificity, or changing co-factor specificity.

[0132] In some embodiments, the enzymes in the pathways outlined here can be gene dosed, i.e., overexpressed, into the resulting genetically modified organism via episomal or chromosomal integration approaches.

[0133] In some embodiments, genome-scale system biology techniques such as Flux Balance Analysis can be utilized to devise genome scale attenuation or knockout strategies for directing carbon flux to 3-oxopent-4-enoyl-CoA, 3-oxopentanoyl-CoA, 3-oxohexanoyl-CoA, 3-oxo-hex-5-enoyl-CoA, 3-oxo-hept-6-enoyl-CoA, 3-oxo-non-8-enoyl-CoA, 3-oxo-5-hydroxypentanoyl-CoA, 3-oxo-6-hydroxyhexanoyl-CoA, 3-oxo-7-hydroxyheptanoyl-CoA, 3-butene-2-one, 3-butene-2-ol, and/or butadiene.

[0134] Attenuation strategies include, but are not limited to: the use of transposons, homologous recombination (double cross-over approach), mutagenesis, enzyme inhibitors, and RNAi interference.

[0135] In some embodiments, fluxomic, metabolomic, and transcriptomal data can be utilized to inform or support genome-scale system biology techniques, thereby devising genome scale attenuation or knockout strategies in directing carbon flux to 3-oxopent-4-enoyl-CoA, 3-oxopentanoyl-CoA, 3-oxohexanoyl-CoA, 3-oxo-hex-5-enoyl-CoA, 3-oxo-hept-6-enoyl-CoA, 3-oxo-non-8-enoyl-CoA, 3-oxo-6-hydroxyhexanoyl-CoA, 3-butene-2-one, 3-butene-2-ol, and/or butadiene.

[0136] In some embodiments, the host microorganism's tolerance to high concentrations of 3-oxopent-4-enoyl-CoA, 3-oxopentanoyl-CoA, 3-oxohexanoyl-CoA, 3-

oxo-hex-5-enoyl-CoA, 3-oxo-hept-6-enoyl-CoA, 3-oxo-non-8-enoyl-CoA, 3-oxo-5-hydroxypentanoyl-CoA, 3-oxo-6-hydroxyhexanoyl-CoA, 3-oxo-7-hydroxyheptanoyl-CoA, 3-butene-2-one, and/or butadiene can be improved through continuous cultivation in a selective environment.

[0137] In some embodiments, the host microorganism's endogenous biochemical network can be attenuated or augmented to (1) ensure the intracellular availability of acetyl-CoA and 2-oxoglutarate, (2) create an NADH or NADPH imbalance that may only be balanced via the formation of one or more of 3-oxopent-4-enoyl-CoA, 3-butene-2-one, and butadiene, (3) prevent degradation of central metabolites, or central precursors leading to and including one or more of 3-oxopent-4-enoyl-CoA, 3-oxopentanoyl-CoA, 3-oxohexanoyl-CoA, 3-oxo-hex-5-enoyl-CoA, 3-oxo-hept-6-enoyl-CoA, 3-oxo-non-8-enoyl-CoA, 3-oxo-5-hydroxypentanoyl-CoA, 3-oxo-6-hydroxyhexanoyl-CoA, 3-oxo-7-hydroxyheptanoyl-CoA, 3-butene-2-one, 3-butene-2-ol, and butadiene and/or (4) ensure efficient efflux from the cell.

[0138] In some embodiments requiring intracellular availability of acetyl-CoA for 3-oxopent-4-enoyl-CoA, 3-oxopentanoyl-CoA, 3-oxohexanoyl-CoA, 3-oxo-hex-5-enoyl-CoA, 3-oxo-hept-6-enoyl-CoA, 3-oxo-non-8-enoyl-CoA, 3-oxo-5-hydroxypentanoyl-CoA, 3-oxo-6-hydroxyhexanoyl-CoA, 3-oxo-7-hydroxyheptanoyl-CoA, 3-butene-2-one, 3-butene-2-ol, and/or butadiene synthesis, endogenous enzymes catalyzing the hydrolysis of acetyl-CoA such as short-chain length thioesterases can be attenuated in the host organism.

[0139] In some embodiments requiring the intracellular availability of acetyl-CoA for 3-oxopent-4-enoyl-CoA, 3-oxopentanoyl-CoA, 3-oxohexanoyl-CoA, 3-oxo-hex-5-

enoyl-CoA, 3-oxo-hept-6-enoyl-CoA, 3-oxo-non-8-enoyl-CoA, 3-oxo-5-hydroxypentanoyl-CoA, 3-oxo-6-hydroxyhexanoyl-CoA, 3-oxo-7-hydroxyheptanoyl-CoA, 3-butene-2-one, 3-butene-2-ol, and/or butadiene synthesis, an endogenous phosphotransacetylase generating acetate such as *pta* can be attenuated (Shen *et al.*, *Appl. Environ. Microbiol.*, 2011, 77(9):2905–2915).

[0140] In some embodiments requiring the intracellular availability of acetyl-CoA for 3-oxopent-4-enoyl-CoA, 3-oxopentanoyl-CoA, 3-oxohexanoyl-CoA, 3-oxo-hex-5-enoyl-CoA, 3-oxo-hept-6-enoyl-CoA, 3-oxo-non-8-enoyl-CoA, 3-oxo-5-hydroxypentanoyl-CoA, 3-oxo-6-hydroxyhexanoyl-CoA, 3-oxo-7-hydroxyheptanoyl-CoA, 3-butene-2-one, and/or butadiene synthesis, an endogenous gene in an acetate synthesis pathway encoding an acetate *kinase*, such as *ack*, can be attenuated.

[0141] In some embodiments requiring the intracellular availability of acetyl-CoA and NADH for 3-oxopent-4-enoyl-CoA, 3-butene-2-one, 3-butene-2-ol, and/or butadiene synthesis, an endogenous gene encoding an enzyme that catalyzes the degradation of pyruvate to lactate such as lactate dehydrogenase encoded by *ldhA* can be attenuated (Shen *et al.*, 2011, *supra*).

[0142] In some embodiments, enzymes that catalyze anaplerotic reactions such as PEP carboxylase and/or pyruvate carboxylase can be overexpressed in the host organism.

[0143] In some embodiments requiring the intracellular availability of acetyl-CoA and NADH for 3-oxopent-4-enoyl-CoA, 3-butene-2-one, 3-butene-2-ol, and/or butadiene synthesis, endogenous genes encoding enzymes, such as menaquinol-fumarate

oxidoreductase, that catalyze the degradation of phophoenolpyruvate to succinate such as *frdBC* can be attenuated (see, e.g., Shen *et al.*, 2011, *supra*).

[0144] In some embodiments requiring the intracellular availability of acetyl-CoA and NADH for 3-oxopent-4-enoyl-CoA, 3-butene-2-one, 3-butene-2-ol, and/or butadiene synthesis, an endogenous gene encoding an enzyme that catalyzes the degradation of acetyl-CoA to ethanol such as the alcohol dehydrogenase encoded by *adhE* can be attenuated (Shen *et al.*, 2011, *supra*).

[0145] In some embodiments, where pathways require excess NADH co-factor for 3-oxopent-4-enoyl-CoA, 3-butene-2-one, 3-butene-2-ol, and/or butadiene synthesis, a recombinant formate dehydrogenase gene can be overexpressed in the host organism (Shen *et al.*, 2011, *supra*).

[0146] In some embodiments, where pathways require excess NADH co-factor for 3-oxopent-4-enoyl-CoA, 3-butene-2-one, 3-butene-2-ol, and/or butadiene synthesis, a recombinant NADH-consuming transhydrogenase can be attenuated.

[0147] In some embodiments, an endogenous gene encoding an enzyme that catalyzes the degradation of pyruvate to ethanol such as pyruvate decarboxylase can be attenuated.

[0148] In some embodiments requiring the intracellular availability of acetyl-CoA for 3-oxopent-4-enoyl-CoA, 3-oxopentanoyl-CoA, 3-oxohexanoyl-CoA, 3-oxo-hex-5-enoyl-CoA, 3-oxo-hept-6-enoyl-CoA, 3-oxo-non-8-enoyl-CoA, 3-oxo-5-hydroxypentanoyl-CoA, 3-oxo-6-hydroxyhexanoyl-CoA, 3-oxo-7-hydroxyheptanoyl-CoA, 3-butene-2-one, 3-butene-2-ol, and/or butadiene synthesis, a recombinant acetyl-CoA

synthetase such as the gene product of *acs* can be overexpressed in the microorganism (Satoh et al., *J. Bioscience and Bioengineering*, 2003, 95(4):335 – 341).

[0149] In some embodiments, carbon flux can be directed into the pentose phosphate cycle to increase the supply of NADPH by attenuating an endogenous glucose-6-phosphate isomerase (EC 5.3.1.9).

[0150] In some embodiments, carbon flux can be redirected into the pentose phosphate cycle to increase the supply of NADPH by overexpression a 6-phosphogluconate dehydrogenase and/or a transketolase (Lee *et al.*, 2003, *Biotechnology Progress*, 19(5), 1444 – 1449).

[0151] In some embodiments, where pathways require excess NADPH co-factor in the synthesis of 3-oxopent-4-enoyl-CoA, 3-butene-2-one, 3-butene-2-ol, and/or butadiene, a gene such as *UdhA* encoding a puridine nucleotide transhydrogenase can be overexpressed in the host organisms (Brigham *et al.*, *Advanced Biofuels and Bioproducts*, 2012, Chapter 39, 1065 - 1090).

[0152] In some embodiments, where pathways require excess NADPH co-factor in the synthesis of 3-oxopent-4-enoyl-CoA, 3-butene-2-one, 3-butene-2-ol, and/or butadiene, a recombinant glyceraldehyde-3-phosphate-dehydrogenase gene such as *GapN* can be overexpressed in the host organisms (Brigham *et al.*, 2012, *supra*).

[0153] In some embodiments, where pathways require excess NADPH co-factor in the synthesis of 3-oxopent-4-enoyl-CoA, 3-butene-2-one, 3-butene-2-ol, and/or butadiene, a recombinant malic enzyme gene such as *maeA* or *maeB* can be overexpressed in the host organisms (Brigham *et al.*, 2012, *supra*).

[0154] In some embodiments, where pathways require excess NADPH co-factor in the synthesis of 3-oxopent-4-enoyl-CoA, 3-butene-2-one, 3-butene-2-ol, and/or butadiene, a recombinant glucose-6-phosphate dehydrogenase gene such as *zwf* can be overexpressed in the host organisms (Lim *et al.*, *J. Bioscience and Bioengineering*, 2002, 93(6), 543 - 549).

[0155] In some embodiments, where pathways require excess NADPH co-factor in the synthesis of 3-oxopent-4-enoyl-CoA, 3-butene-2-one, 3-butene-2-ol, and/or butadiene, a recombinant fructose 1,6 diphosphatase gene such as *fbp* can be overexpressed in the host organisms (Becker *et al.*, *J. Biotechnol.*, 2007, 132:99 - 109).

[0156] In some embodiments, where pathways require excess NADPH co-factor in the synthesis of 3-oxopent-4-enoyl-CoA, 3-butene-2-one, 3-butene-2-ol, and/or butadiene, endogenous triose phosphate isomerase (EC 5.3.1.1) can be attenuated.

[0157] In some embodiments, where pathways require excess NADPH co-factor in the synthesis of 3-oxopent-4-enoyl-CoA, 3-butene-2-one, 3-butene-2-ol, and/or butadiene, a recombinant glucose dehydrogenase such as the gene product of *gdh* can be overexpressed in the host organism (Satoh *et al.*, *J. Bioscience and Bioengineering*, 2003, 95(4):335 – 341).

[0158] In some embodiments, endogenous enzymes facilitating the conversion of NADPH to NADH can be attenuated, such as the NADH generation cycle that may be generated via inter-conversion of glutamate dehydrogenases classified under EC 1.4.1.2 (NADH-specific) and EC 1.4.1.4 (NADPH-specific).

[0159] In some embodiments, an endogenous glutamate dehydrogenase (EC 1.4.1.3) that utilizes both NADH and NADPH as co-factors can be attenuated.

[0160] In some embodiments using hosts that naturally accumulate polyhydroxyalkanoates, the endogenous polymer synthase enzymes can be attenuated in the host strain.

[0161] In some embodiments, enzymes such as a pimeloyl-CoA dehydrogenase classified under, EC 1.3.1.62; an acyl-CoA dehydrogenase classified, for example, under EC 1.3.8.7, EC 1.3.8.1, or EC 1.3.99.-; and/or a butyryl-CoA dehydrogenase classified, for example, under EC 1.3.8.6, that degrade central metabolites and central precursors leading to and including 3-oxopent-4-enoyl-CoA, 3-butene-2-one, 3-butene-2-ol, and/or butadiene can be attenuated.

[0162] In some embodiments, endogenous enzymes activating 3-oxopent-4-enoyl-CoA, 3-butene-2-one, 3-butene-2-ol, and/or butadiene precursors via Coenzyme A esterification such as CoA-ligases (e.g., an adipyl-CoA synthetase) classified under, for example, EC 6.2.1.- can be attenuated.

[0163] In some embodiments, the efflux of 3-oxopent-4-enoyl-CoA, 3-butene-2-one, 3-butene-2-ol, and/or butadiene across the cell membrane to the extracellular media can be enhanced or amplified by genetically engineering structural modifications to the cell membrane or increasing any associated transporter activity for 3-oxopent-4-enoyl-CoA, 3-butene-2-one, 3-butene-2-ol, and/or butadiene.

Producing 3-oxopent-4-enoyl-CoA, 3-oxopentanoyl-CoA, 3-oxohexanoyl-CoA, 3-oxo-hex-5-enoyl-CoA, 3-oxo-hept-6-enoyl-CoA, 3-oxo-non-8-enoyl-CoA, 3-oxo-5-hydroxypentanoyl-CoA, 3-oxo-6-hydroxyhexanoyl-CoA, 3-oxo-7-hydroxyheptanoyl-CoA, 3-butene-2-one, 3-butene-2-ol, or Butadiene Using a Recombinant Host

[0164] Typically, one or more of 3-oxopent-4-enoyl-CoA, 3-oxopentanoyl-CoA, 3-oxohexanoyl-CoA, 3-oxo-hex-5-enoyl-CoA, 3-oxo-hept-6-enoyl-CoA, 3-oxo-non-8-enoyl-CoA, 3-oxo-5-hydroxypentanoyl-CoA, 3-oxo-6-hydroxyhexanoyl-CoA, 3-oxo-7-hydroxyheptanoyl-CoA, 3-butene-2-one, 3-butene-2-ol, or butadiene can be produced by providing a host microorganism and culturing the provided microorganism with a culture medium containing a suitable carbon source as described above. In general, the culture media and/or culture conditions can be such that the microorganisms grow to an adequate density and produce 3-oxopent-4-enoyl-CoA, 3-oxopentanoyl-CoA, 3-oxohexanoyl-CoA, 3-oxo-hex-5-enoyl-CoA, 3-oxo-hept-6-enoyl-CoA, 3-oxo-non-8-enoyl-CoA, 3-oxo-5-hydroxypentanoyl-CoA, 3-oxo-6-hydroxyhexanoyl-CoA, 3-oxo-7-hydroxyheptanoyl-CoA, 3-butene-2-one, 3-butene-2-ol, and/or butadiene efficiently. For large-scale production processes, any method can be used such as those described elsewhere (Manual of Industrial Microbiology and Biotechnology, 2nd Edition, Editors: A. L. Demain and J. E. Davies, ASM Press; and Principles of Fermentation Technology, P. F. Stanbury and A. Whitaker, Pergamon). Briefly, a large tank (e.g., a 100 gallon, 200 gallon, 500 gallon, or more tank) containing an appropriate culture medium is inoculated with a particular microorganism. After inoculation, the microorganism is incubated to allow biomass to be produced. Once a desired biomass is reached, the broth containing the microorganisms can be transferred to a second tank. This second tank can be any size. For example, the second tank can be larger, smaller, or the same size as the first tank. Typically, the second tank is larger than the first such that additional culture medium can be added to the broth from the first tank. In addition, the culture

medium within this second tank can be the same as, or different from, that used in the first tank.

[0165] Once transferred, the microorganisms can be incubated to allow for the production of 3-oxopent-4-enoyl-CoA, 3-oxopentanoyl-CoA, 3-oxohexanoyl-CoA, 3-oxo-hex-5-enoyl-CoA, 3-oxo-hept-6-enoyl-CoA, 3-oxo-non-8-enoyl-CoA, 3-oxo-5-hydroxypentanoyl-CoA, 3-oxo-6-hydroxyhexanoyl-CoA, 3-oxo-7-hydroxyheptanoyl-CoA, 3-butene-2-one, 3-butene-2-ol, and/or butadiene. Once produced, any method can be used to isolate 3-oxopent-4-enoyl-CoA, 3-butene-2-one, and/or butadiene. For example, 3-oxopent-4-enoyl-CoA, 3-butene-2-one, 3-butene-2-ol, and/or butadiene can be recovered selectively from the fermentation broth via adsorption processes. The resulting eluate may be further concentrated via evaporation, crystallized via evaporative and/or cooling crystallization, and the crystals recovered via centrifugation. Distillation may be employed to achieve the desired product purity.

[0166] The invention will be further described in the following examples, which do not limit the scope of the invention described in the claims.

EXAMPLES

EXAMPLE 1: Preparation of YdiF mutants

[0167] YdiF mutants were prepared using the relevant primers to introduce the mutated bases via the Quick change lightning kit (Agilent, code product: 210518-5), and cloned into *E. coli*. In addition, all the proteins used in this experiment have the following extra amino acids added in the N-terminal extremity of the original protein sequence (SEQ ID NO:1): MGHHHHHHSSGLVPRGS. Those additional amino acids correspond to: two amino acids (M and G), a 6xHis-tag followed by a 3 amino acids linker (SSG), and a specific proteolytic cleavage site (thrombin site, underlined) (SEQ ID NO: 22). All the sequences were confirmed by DNA sequencing.

[0168] From a fresh LB plate containing the desired clone transformant, one colony (or small scratch) was picked to inoculate 25 mL of LB supplemented with the relevant antibiotic and the pre-culture was incubated overnight at 37 °C, 230 rpm. The following morning, the TB auto-induced medium (Merck / Code product: 71491-5) was prepared by mixing 60g TB / L supplemented with 10mL Glycerol / L of TB and microwaved during 3 + 2 minutes at full power. The TB was cooled down before use and splitting it in sterile flasks. Then, the pre-culture incubated overnight in LB was centrifuged and the supernatant discarded. The obtained pellet was resuspended with 5mL of freshly prepared TB medium and used to inoculate 500 mL of TB dispensed in sterile flasks and supplemented with the appropriate antibiotic. The inoculated medium was incubated at 28° C for at least 20h, 230 rpm.

[0169] The culture was centrifuged at least at 3000g / 20 min / 4 °C and the pellet used immediately or stored at -80°C.

[0170] The pellets (fresh or thawed) were resuspended in 10 to 20 mL of Buffer A (50mM Hepes+ 150 mM NaCl + 40mM Imidazole + 5% Glycerol - pH 7.5). The resuspended cells were then sonicated in ice for ≈ 5 min at 37 % Amplitude with 5" ON and 15" OFF sonication pulse. The sonicated cells were centrifuged for at least 20 min at 15500g and 4 °C. The supernatant containing the soluble fraction of proteins was recovered, filtered using 0.45 or 0.2um filters, and used for His-trap protein purification.

[0171] The filtered soluble fraction of proteins obtained after extraction of proteins by sonication was used for His-tag protein purification. A 1 mL His-trap (GE Healthcare / Code product: 17-5319-01) column was equilibrated with 5-10 volumes column (VC) using Buffer A*. The soluble fraction of proteins was loaded onto the His-trap column manually using a syringe and 5-10 VC of Buffer A were used to wash the His-trap column. 5-10 VC of Buffer B** were used to elute the His-tag protein directly to a 4 or 20 mL centrifugal filtration unit (VWR / Code product: 512-2850) with a relevant cut-off (5kD).

[0172] The centrifugal filtration unit was centrifuged at 3500g / 5 °C to a volume lower than 400uL concentrate. Around 3 mL of Buffer C*** was added to the concentrate and the centrifugal filtration unit was again centrifuged at 3500g / 5 °C to a volume lower than 400uL. This step was made to remove most of the imidazole used in Buffer B to elute the His-tag.

[0173] The concentrate of pure enzyme was recovered and Buffer C was used to top-up to the desired volume.

[0174] The concentration was checked using a Nanodrop spectrophotometer.

[0175] * Buffer A = 50 mM Hepes + 150 mM NaCl + 40 mM Imidazole + 5 % (v/v) Glycerol - pH 7.5

[0176] ** Buffer B = Buffer A + 400mM Imidazole - pH7.5

[0177] *** Buffer C = Buffer A without Imidazole - pH7.5

EXAMPLE 2: Biotransformation using YdiF or mutants

[0178] Acetyl-CoA was used at a final concentration of 1.5 mM and 30 mM of the substrate being studied (e.g. carboxylic acid salts, such as sodium acrylate) or 1.5 mM of propionyl-CoA in 50mM Hepes buffer (pH 7.5). The reaction was started by addition of 0.5 mg/mL of YdiF pure protein or mutants to a final volume of 250 µL and incubated for 1 to 10 hours at 30 °C.

[0179] Two controls were made from similar mixture except that the enzymes were inactivated by heating-up at 95 °C for 20 min or they were replaced by Buffer C.

[0180] All samples were prepared in triplicate and spun down at 15000 g for 20 min at 4 °C. Detection of the thiol product (e.g. acryloyl-CoA) synthesized from CoA-transferase activity and the keto-thiol product (e.g. 3-oxopent-4-enoyl-CoA) due to putative thiolase activity of YdiF (or homologues) was carried out by Liquid Chromatography coupled to a qTof-Mass spectrometry (in positive mode), using a C18 column (Zorbax Eclipse C18, 2.1x50 mm, 1.8 µ, Agilent, code product: 959757-902) and a gradient of 10 mM ammonium acetate and acetonitrile as mobile phase.

[0181] Propionyl-CoA and butyryl-CoA (from Sigma-Aldrich) were used as authentic standards of the some of the reactions and were analyzed by LC-qTof-MS under the same conditions mentioned above.

EXAMPLE 3: Evaluation of Potential Active Sites

[0182] Mutations were introduced as described in Example 1 at the following positions previously found to be potentially involved in the active site of other β-ketothiolases: C258A, S60A, C112A, S38A, C390A. Activity of the mutant enzymes was studied as described in Example 2. None of the mutations abolished the transferase and thiolase activity of YdiF.

[0183] Mutations were introduced as described in Example 1 to replace the glutamate in position 324 of YdiF (*see* Selmer *et al.*, 2002) with glycine (mutant E324G) or leucine (E324L). Activity of the mutant enzymes was studied as described in Example 2. Abolition of transferase activity in these mutants showed that the glutamate at position 324 is present in the active site involved in the transferase activity of YdiF. The activity of mutant E324L was only partially abolished in comparison to mutant E324G, possibly due to a steric effect of the glutamate in the active site. Providing the mutant E324G with propionyl-CoA and acetyl-CoA showed the presence of the keto-thiol product (3-oxopentanoyl-CoA), consistent with the transferase activity and the β-ketothiolase activity of YdiF being independent and the involvement of separate active sites for those activities.

[0184] It was concluded the active site involved in the thiolase activity of YdiF presents new features in comparison to the β-ketothiolases already known. Oligomerisation of YdiF may also create an active site at the interfaces.

EXAMPLE 4

[0185] YdiF from *Clostridium propionicum* belongs to the enzymatic class EC2.8.3.8. Other members of this enzyme class with similar structural features to YdiF,

as well as proteins with a CoA-transferase activity from close-by enzymatic classes, were screened for dual transferase/ketothiolase activity as described in Example 2. Acetyl-CoA and sodium acrylate were provided as substrates. LC-MS was used to detect the thiol-product (propenoyl-CoA) synthesized via the transferase activity and the ketothiol-product (3-oxopent-4-enoyl-CoA) synthesized via the thiolase activity, as described in Example 2. Results are presented in Table 3.

EXAMPLE 5

[0186] To study the substrate specificity of YdiF, several commercially available substrates were provided to YdiF (and homologues) in combination with acetyl-CoA, as described in Example 2. LC-MS was used to detect the thiol-product synthesized via the transferase activity and the ketothiol-product synthesized via the thiolase activity. Results are presented in Table 4.

[0187] Dual transferase/ketothiolase activity was observed with use of C3- and C4-saturated substrates as well as with medium-length unsaturated substrates, without obvious inhibition effects due to any of the products. 4-Hydroxybutyrate was used as a substrate with acetyl-CoA for protein 244 and the corresponding products from the transferase and ketothiolase activities of protein 244 were detected.

Table 3

EC	Gene (Uniprot)	Bdigene	Protein	Organism	Solubility	thiol-product	ketothiol-product
2.8.3.8	Q9L3F7	237	Acetate/Propionate CoA-transferase	*Clostridium propionicum*	✓	✓	✓
2.8.3.8	V1HBS2	271	Acetate CoA-transferaseYdiF	*Salmonella enterica subsp. houtenaeserovar*	✓	✓	✓
2.8.3.8	U2L5C9	272	Acetate CoA-transferaseYdiF	*Peptostreptococcaceae bacterium oral taxon*	✓	✓	✓
2.8.3.8	R5ADR5	273	Acetate CoA-transferaseYdiF	*Firmicutes bacterium CAG*	✓	✓	✓
2.8.3.8	G0VND6	274	Acetate CoA-transferaseYdiF	*Megasphaera elsdenii DSM 20460*	✓	✓	✓
2.8.1.16	C6VTZ3	275	Acetyl-CoA acetyltransferase	*Dyadobacter fermentans*	✓	×	×
2.3.1.16 / 2.3.1.9	S7V863	276	3-ketoacyl-CoA thiolase	*Cyclobacterium qasimii M12-11B*	✓	✓	×
2.8.3-	Q9RM86	0244	4-Hydroxybutyrate-CoA transferase	*Clostridium aminobutyricum*	✓	✓	✓
2.8.3.1	K3RRN6	0336	Propionate CoA-transferase	*Escherichia coli EC1865*	✓	✓	×
2.8.3.10	J1G510	0337	Acetyl-CoA hydrolase/transferase	*Citrobacter sp. A1*	✓	✓	✓
2.8.3.11	T1CV65	0338	CoA-transferase family III	*mine drainage metagenome*	×	×	×
2.8.3.12	Q0AWW8	0339	Glutaconate CoA-transferase	*Syntrophomonas wolfei subsp. wolfei*	×	×	×

EC	Gene (Uniprot)	Bdigene	Protein	Organism	Solubility	thiol-product	ketothiol-product
2.8.3.13	Q7TNE1	0340	Succinate--hydroxymethylglutarate CoA-transferase	*Mus musculus (Mouse)*	✗	✗	✗
2.8.3.16	Q0K0H8	0342	Formyl-CoA:oxalate CoA-transferase	*Cupriavidus necator H16*	✓	✗	✗
2.8.3.17	Q93AM1	0343	E-cinnamoyl-CoA:R-phenyllactate CoA transferase	*Clostridium sporogenes*	✓	✗	✗
2.8.3.18	B3EY95	0344	Succinyl-CoA:acetate CoA-transferase	*Acetobacter aceti*	✓	✓	✓
2.8.3.20	A9WGE3	0345	Succinyl-CoA--D-citramalate CoA-transferase	*Chloroflexus aurantiacus*	✓	✓	✓
2.8.3.5	P56006	0346	Succinyl-CoA:3-ketoacid coenzyme A transferase subunit A	*Helicobacter pylori*	✓	✗	✗
2.8.3.6/-5/-8	B8CRZ0	0347	3-oxoacid CoA-transferase	*Shewanella piezotolerans*	✗	✗	✗
2.8.3.9	B2A667	0348	Butyryl-CoA:acetate CoA transferase	*Natranaerobius thermophilus*	✓	✗	✗

Table 4

enzyme	YdiF (237)						Protein 244
substrate	propionate	butyrate	3-butenoate	4-pentenoate	6-heptenoate	2-hexenoate	4-Hydroxybutyrate
thiol-product	✓	✓	✓	✓	✓	✓	✓
ketothiol-product	✓	✓	✓	✓	✓	✓	✓

ADDITIONAL EXEMPLARY EMBODIMENTS

[0188] In one embodiment is provided a method of producing 3-oxopent-4-enoyl-CoA comprising enzymatically converting propenoate to 3-oxopent-4-enoyl-CoA using a polypeptide that has both CoA transferase and β-ketothiolase activities.

[0189] In one embodiment is provided a method of producing 3-keto-acyl-CoA esters, said method comprising the step of enzymatically condensing acetyl-CoA with any of an alkanoic acid, an alkenoic acid, a hydroxyacid or a haloacid using a polypeptide that has both CoA transferase and β-ketothiolase activities.

[0190] In one embodiment the alkanoic acid may be selected from straight chain alkanoic acids of carbon chain length *n* (*n* > 2) such as acetate, propionate, butyrate, pentanoic acid, hexanoic acid and the like, or branched chain alkanoic acids such as isobutyrate, isovaleric acid or pivalic acid.

[0191] In one embodiment the alkenoic acid may be selected from straight chain alkenoic acids of carbon chain length *n* (wherein *n* > 2) such as acrylic acid, 2-propenoic acid, 3-butenoic acid, 4-pentenoic acid, 5-hexenoic acid, 6-hepteneoic acid, crotonic acid, and the like, or branched chain alkenoic acids such as methacrylic acid, 3-methyl-3-butenoic acid, 4-methyl-4-pentenoic acid, and 5-methyl-5-hexenoic acid, and the like.

[0192] In one embodiment the hydroxyacid may be selected from hydroxyacids of carbon chain length *n* (wherein *n* > 2) such as 3-hydroxypropionic acid, 4-hydroxybutyric acid, 5-hydroxyvaleric acid, and 6-hydroxy-caproic acid.

[0193] In one embodiment the haloacid may be selected from haloacids of carbon chain length *n* (wherein *n* > 2) such as 3-halopropionic acid, 4-halobutyric acid, 5-halovaleric acid, and 6-halocaproic acid.

[0194] In one embodiment the method comprises the step of enzymatically condensing acetyl-CoA with the respective CoA ester of any of said alkanoic acid, said alkenoic acid, said hydroxyacid, or said haloacid using a polypeptide that has both CoA transferase and β-ketothiolase activities.

[0195] In one embodiment the 3-keto-acyl-CoA ester is converted to its respective free acid by a CoA transferase or a thioesterase.

[0196] In one embodiment the polypeptide that has both CoA transferase and β-ketothiolase activities is classified under EC 2.8.3.-.

[0197] In one embodiment the polypeptide that has both CoA transferase and β-ketothiolase activities is classified under EC 2.8.3.8.

[0198] In one embodiment the polypeptide that has both CoA transferase and β-ketothiolase activities is from *Salmonella enterica, Peptostreptococcaceae bacterium, Firmicutes bacterium, Megasphaera elsdenii, Salmonella enterica subsp. houtenaeserovar, Clostridium aminobutyricum*, or *Clostridium propionicum*.

[0199] In one embodiment the polypeptide that has both CoA transferase and β-ketothiolase activities has at least 95% sequence identity to the amino acid sequence set forth in any one of SEQ ID No: 1, 4, 5, 6, 7, 8, and 9.

[0200] In one embodiment the polypeptide that has both CoA transferase and β-ketothiolase activities has an amino acid substitution at one or more of positions 38, 60, 112, 258, and 390 of SEQ ID NO: 1.

[0201] In one embodiment the polypeptide that has both CoA transferase and β-ketothiolase activities is a CoA transferase whose amino acid sequence includes the motif I/V/A/LTE.

[0202] In one embodiment the polypeptide that has both CoA transferase and β-ketothiolase activities is a CoA transferase whose amino acid sequence includes the motifs EXGXXG and GXGG(A/F).

[0203] In one embodiment is provided a method of producing 3-oxopent-4-enoyl-CoA comprising: enzymatically converting propenoate to 3-oxopent-4-enoyl-CoA using a polypeptide that has both CoA transferase and β-ketothiolase activities; and enzymatically converting 3-oxopent-4-enoyl-CoA to 3-oxopent-4-enoate, for example using a CoA-transferase, for example a CoA-transferase classified under EC 2.8.3.6 or EC 2.8.3.8.

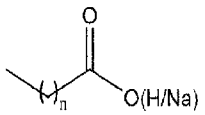
[0204] In one embodiment is provided a method of producing 3-oxopent-4-enoyl-CoA comprising: enzymatically converting propenoate to 3-oxopent-4-enoyl-CoA using a polypeptide that has both CoA transferase and β-ketothiolase activities; enzymatically converting 3-oxopent-4-enoyl-CoA to 3-oxopent-4-enoate; and enzymatically converting 3-oxopent-4-enoate to 3-buten-2-one, for example using a decarboxylase, for example a decarboxylase classified under EC 4.1.1.4.

[0205] In one embodiment is provided a method of producing 3-oxopent-4-enoyl-CoA comprising: enzymatically converting propenoate to 3-oxopent-4-enoyl-CoA using a polypeptide that has both CoA transferase and β-ketothiolase activities; enzymatically converting 3-oxopent-4-enoyl-CoA to 3-oxopent-4-enoate; enzymatically converting 3-oxopent-4-enoate to 3-buten-2-one; and enzymatically converting 3-buten-2-one to 3-buten-2-ol, for example using an alcohol dehydrogenase or a phenylacetaldehyde reductase. In one embodiment the alcohol dehydrogenase has at least 95% sequence identity to the amino acid sequence set forth in SEQ ID NO:16. In one embodiment the

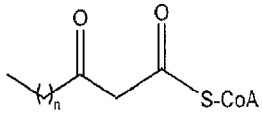
phenylacetaldehyde reductase has at least 95% sequence identity to the amino acid sequence set forth in SEQ ID NO:17.

[0206] In one embodiment is provided a method of producing 3-oxopent-4-enoyl-CoA comprising: enzymatically converting propenoate to 3-oxopent-4-enoyl-CoA using a polypeptide that has both CoA transferase and β-ketothiolase activities; enzymatically converting 3-oxopent-4-enoyl-CoA to 3-oxopent-4-enoate; enzymatically converting 3-oxopent-4-enoate to 3-buten-2-one; enzymatically converting 3-buten-2-one to 3-buten-2-ol; and enzymatically converting 3-buten-2-ol to 1,3 butadiene, for example using a linalool dehydratase, for example a linalool dehydratase classified under EC 4.2.1.127, or using a dehydratase classified under EC 4.2.1.-. In one embodiment the linalool dehydratase has at least 95% sequence identity to the amino acid sequence set forth in SEQ ID NO:18. In one embodiment the dehydratase has greater than 95% sequence identity to the 5-aminovaleryl-CoA dehydratase from C. viride or to a dehydratase classified under EC 4.2.1.- from species such as *Aquincola tertiaricarbonis* or *Methylibium petroleiphilum* PM1.

[0207] In one embodiment is provided a method of producing 3-oxo-acyl-CoA compounds of formula (IIa), said method comprising enzymatically converting a carboxylic acid of formula (Ia) to a 3-oxo-acyl-CoA compound of formula (IIa)



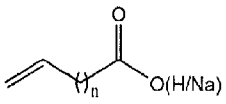
Formula (Ia)



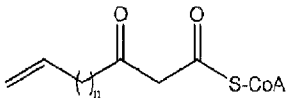
Formula (IIa)

using a polypeptide that has both CoA transferase and β-ketothiolase activities.

[0208] In one embodiment is provided a method of producing 3-oxo-enoyl-CoA compounds of formula (IIb), said method comprising enzymatically converting an unsaturated carboxylic acid of formula (Ib) to a 3-oxo-enoyl-CoA compound of formula (IIb):



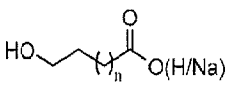
Formula (Ib)



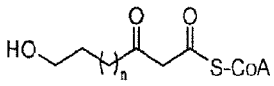
Formula (IIb)

using a polypeptide that has both CoA transferase and β-ketothiolase activities.

[0209] In one embodiment is provided a method of producing 3-oxo-hydroxyacyl-CoA compounds of formula (IIc), said method comprising enzymatically converting a hydroxyl-substituted carboxylic acid of formula (Ic) to a 3-oxo-hydroxyacyl-CoA compound of formula (IIc):



Formula (Ic)



Formula (IIc)

using a polypeptide that has both CoA transferase and β-ketothiolase activities. In one embodiment the 3-oxo-hydroxyacyl-CoA compound of formula (IIc) is converted to a nylon compound.

[0210] In one embodiment 4-hydroxybutyric acid, or a salt thereof, is enzymatically converted to 3-oxo-6-hydroxyhexanoyl-CoA using a polypeptide that has both CoA transferase and β-ketothiolase activities. In one embodiment 3-oxo-6-hydroxyhexanoyl-CoA is enzymatically converted to 6-hydroxyhexanoic acid, and 6-hydroxyhexanoic acid is enzymatically converted to one or more of adipic acid, 6-aminohexanoic acid, hexamethylenediamine, caprolactam, and 1,6-hexanediol using

one or more isolated enzymes selected from dehydrogenases, reductases, hydratases, thioesterases, monooxygenases, and transaminases.

[0211] In one embodiment 3-hydroxypropionic acid, or a salt thereof, is enzymatically converted to 3-oxo-5-hydroxypentanoyl-CoA using a polypeptide that has both CoA transferase and β-ketothiolase activities.

[0212] In one embodiment 5-hydroxypentanoic acid, or a salt thereof, is enzymatically converted to 3-oxo-7-hydroxyheptanoyl-CoA using a polypeptide that has both CoA transferase and β-ketothiolase activities.

[0213] In one embodiment is provided a method as described above wherein said method is performed in a non-naturally occurring host, for example a recombinant host.

[0214] In one embodiment is provided a non-naturally occurring host capable of producing 3-oxopent-4-enoyl-CoA, said host comprising at least one exogenous nucleic acid encoding a polypeptide having both CoA transferase and β-ketothiolase activities.

[0215] In one embodiment is provided a non-naturally occurring host capable of producing 1,3-butadiene, said host comprising at least one exogenous nucleic acid encoding a polypeptide having both CoA transferase and β-ketothiolase activities.

[0216] In one embodiment is provided a non-naturally occurring host capable of producing 3-oxo-acyl-CoA compounds of formula (IIa), said host comprising at least one exogenous nucleic acid encoding a polypeptide having both CoA transferase and β-ketothiolase activities.

[0217] In one embodiment is provided a non-naturally occurring host capable of producing 3-oxo-enoyl-CoA compounds of formula (IIb), said host comprising at least

one exogenous nucleic acid encoding a polypeptide having both CoA transferase and β-ketothiolase activities.

[0218] In one embodiment is provided a non-naturally occurring host capable of producing 3-oxo-hydroxyacyl-CoA compounds of formula (IIc), said host comprising at least one exogenous nucleic acid encoding a polypeptide having both CoA transferase and β-ketothiolase activities.

[0219] In one embodiment is provided a non-naturally occurring host capable of producing nylon compounds, said host comprising at least one exogenous nucleic acid encoding a polypeptide having both CoA transferase and β-ketothiolase activities.

[0220] In one embodiment the host is cultured under aerobic, anaerobic, or micro-aerobic cultivation conditions.

[0221] In one embodiment the host is cultured under conditions of nutrient limitation.

[0222] In one embodiment the host is retained using a ceramic hollow fiber membrane.

[0223] In one embodiment the principal carbon source derives from a biological feedstock, for example a biological feedstock such as, or deriving from, monosaccharides, disaccharides, lignocellulose, hemicellulose, cellulose, lignin, levulinic acid, formic acid, triglycerides, glycerol, fatty acids, agricultural waste, condensed distillers' solubles, or municipal waste.

[0224] In one embodiment the principal carbon source derives from a non-biological feedstock, for example a non-biological feedstock such as, or deriving from, natural gas, syngas, CO_2/H_2, methanol, ethanol, benzoate, non-volatile residue (NVR)

caustic wash waste stream from cyclohexane oxidation processes, or terephthalic acid / isophthalic acid mixture waste streams.

[0225] In one embodiment the host is a prokaryotic host, for example from the genus *Escherichia*, *Clostridia*, *Corynebacteria*, *Cupriavidus*, *Pseudomonas*, *Delftia*, *Bacillus*, *Lactobacillus*, *Lactococcus*, or *Rhodococcus*. In one embodiment the host is *Escherichia coli*, *Clostridium ljungdahlii*, *Clostridium autoethanogenum*, *Clostridium kluyveri*, *Corynebacterium glutamicum*, *Cupriavidus necator*, *Cupriavidus metallidurans*, *Pseudomonas fluorescens*, *Pseudomonas putida*, *Pseudomonas oleavorans*, *Delftia acidovorans*, *Bacillus subtillis*, *Lactobacillus delbrueckii*, *Lactococcus lactis*, or *Rhodococcus equi*.

[0226] In one embodiment the host is a eukaryotic host, for example from the genus *Aspergillus*, *Saccharomyces*, *Pichia*, *Yarrowia*, *Issatchenkia*, *Debaryomyces*, *Arxula*, or *Kluyveromyces*. In one embodiment the host is *Aspergillus niger*, *Saccharomyces cerevisiae*, *Pichia pastoris*, *Yarrowia lipolytica*, *Issathenkia orientalis*, *Debaryomyces hansenii*, *Arxula adenoinivorans*, or *Kluyveromyces lactis*.

[0227] In one embodiment the host is capable of producing 1,3-butadiene and comprises an exogenous alcohol dehydrogenase or an exogenous phenylacetaldehyde reductase, and an exogenous linalool dehydratase. In one embodiment the alcohol dehydrogenase has at least 95% sequence identity to the amino acid sequence set forth in SEQ ID NO:16. In one embodiment the phenylacetaldehyde reductase has at least 95% sequence identity to the amino acid sequence set forth in SEQ ID NO:17. In one embodiment the linalool dehydratase has at least 95% sequence identity to the amino acid sequence set forth in SEQ ID NO:18.

[0228] In one embodiment is provided a composition comprising 3-oxopent-4-enoyl-CoA synthesized by a method described above.

[0229] In one embodiment is provided a composition comprising butadiene synthesized by a method described above.

[0230] In one embodiment is provided method for producing bio-derived 3-oxopent-4-enoyl-CoA, comprising culturing or growing a host described above under conditions and for a sufficient period of time to produce bio-derived 3-oxopent-4-enoyl-CoA.

[0231] In one embodiment is provided a method for producing bio-derived butadiene, comprising culturing or growing a host described above under conditions and for a sufficient period of time to produce bio-derived butadiene.

[0232] In one embodiment is provided a variant polypeptide having at least 95% sequence identity to the amino acid sequence of a wild-type enzyme set forth in any of SEQ NOs: 1, 5, 6, 7, 8, and 9, and comprising at least one mutation to the wild-type enzyme. In one embodiment the at least one mutation results in an increase in the transferase activity, the β-ketothiolase activity, or both activities of the wild-type enzyme. In one embodiment the at least one mutation results in a decrease in or a loss of the transferase activity or the β-ketothiolase activity of the wild-type enzyme. In one embodiment the mutation is a deletion of at least one amino acid. In one embodiment the mutation is an addition of at least one amino acid. In one embodiment the mutation is a substitution of at least one amino acid, for example a conservative substitution.

[0233] In one embodiment is provided a variant polypeptide having the amino acid sequence set forth in SEQ ID NO:1 with an amino acid other than a serine at

position 38, an amino acid other than a serine at position 60, an amino acid other than a cysteine at position 112, an amino acid other than a cysteine at position 258, or an amino acid other than a cysteine at position 390. In one embodiment the variant polypeptide comprises an alanine at position 38, an alanine at position 60, an alanine at position 112, or an alanine at position 390 of SEQ ID NO: 1.

[0234] In one embodiment is provided a variant polypeptide having the amino acid sequence set forth in SEQ ID NO:1 with an amino acid other than glutamic acid at position 324. In one embodiment the variant polypeptide comprises a leucine at position 324. In one embodiment the variant polypeptide comprises a glycine at position 324.

[0235] In one embodiment is provided a variant polypeptide having an amino acid sequence comprising the motifs I/V/A/LTE, EXGXXG, and GXGG(A/F).

[0236] In one embodiment the at least one mutation results in the presence of an XXGXXG motif where an EXGXXG motif was present in the wild-type enzyme.

[0237] In one embodiment the at least one mutation does not affect the β-ketothiolase activity of the wild-type enzyme.

[0238] In one embodiment is provided a bio-derived, bio-based, or fermentation-derived product comprising: (a) a composition comprising at least one bio-derived, bio-based, or fermentation-derived compound prepared (i) using a host described herein, (ii) using a variant polypeptide described herein, or (iii) according to a method described herein, or any combination thereof; (b) a bio-derived, bio-based, or fermentation-derived polymer or resin, for example styrene-butadiene-rubber, poly-butadiene, styrene-butadiene latex, acrylonitrile-butadiene-styrene resin, nitrile rubber, adiponitrile, or nylon

compounds,comprising the bio-derived, bio-based, or fermentation-derived composition or compound of (a), or any combination thereof; (c) a molded substance obtained by molding the bio-derived, bio-based, or fermentation-derived polymer or resin of (b), or any combination thereof; (d) a bio-derived, bio-based, or fermentation-derived formulation comprising the bio-derived, bio-based, or fermentation-derived composition or compound of (a), bio-derived, bio-based, or fermentation-derived polymer or resin of (b), or bio-derived, bio-based, or fermentation-derived molded substance of (c), or any combination thereof, or (e) a bio-derived, bio-based, or fermentation-derived semi-solid or a non-semi-solid stream, comprising the bio-derived, bio-based, or fermentation-derived composition or compound of (a), bio-derived, bio-based, or fermentation-derived polymer or resin of (b), bio-derived, bio-based, or fermentation-derived formulation of (d), or bio-derived, bio-based, or fermentation-derived molded substance of (c), or any combination thereof.

[0239] In one embodiment is provided a method of producing 3-oxopent-4-enoyl-CoA, said method comprising: providing propenoate; and providing a means for enzymatically converting propenoate to 3-oxopent-4-enoyl-CoA, wherein the means comprises both CoA transferase and β-ketothiolase activities.

[0240] In one embodiment is provided a method of producing butadiene, said method comprising: (a) providing propenoate; (b) providing a means for enzymatically converting propenoate to 3-oxopent-4-enoyl-CoA, wherein the means comprises both CoA transferase and β-ketothiolase activities; (c) enzymatically converting 3-oxopent-4-enoyl-CoA to 3-oxopent-4-enoate using a CoA-transferase; (d) enzymatically converting 3-oxopent-4-enoate to 3-buten-2-one using a decarboxylase; (e) enzymatically

converting 3-buten-2-one to 3-buten-2-ol using an alcohol dehydrogenase or a phenylacetaldehyde reductase; and (f) enzymatically converting 3-buten-2-ol to 1,3 butadiene using a linalool dehydratase.

[0241] In one embodiment is provided a non-naturally occurring host capable of producing butadiene, said host comprising: (a) a means for enzymatically converting propenoate to 3-oxopent-4-enoyl-CoA, wherein the means comprises both CoA transferase and β-ketothiolase activities; (b) at least one exogenous nucleic acid encoding a CoA-transferase; (c) at least one exogenous nucleic acid encoding a decarboxylase; (d) at least one exogenous nucleic acid encoding an alcohol dehydrogenase or a phenylacetaldehyde reductase; and (e) at least one exogenous nucleic acid encoding a linalool dehydratase.

[0242] In one embodiment is provided a non-naturally occurring host capable of producing 3-oxo-acyl-CoA compounds of formula (IIa), said host comprising: (a) a carboxylic acid of formula (Ia), or a salt form thereof; and (b) a means for enzymatically converting the carboxylic acid of formula (Ia) to a 3-oxo-acyl-CoA compound of formula (IIa), wherein the means comprises both CoA transferase and β-ketothiolase activities.

[0243] In one embodiment is provided a non-naturally occurring host capable of producing 3-oxo-enoyl-CoA compounds of formula (IIb), said host comprising: (a) an unsaturated carboxylic acid of formula (Ib), or a salt form thereof; and (b) a means for enzymatically converting the unsaturated carboxylic acid of formula (Ib) to a 3-oxo-enoyl-CoA compound of formula (IIb), wherein the means comprises both CoA transferase and β-ketothiolase activities.

[0244] In one embodiment is provided a non-naturally occurring host capable of producing 3-oxo-hydroxyacyl-CoA compounds of formula (IIc), said host comprising: (a) a hydroxy-substituted carboxylic acid of formula (Ic), or a salt form thereof; and (b) a means for enzymatically converting the hydroxy-substituted carboxylic acid of formula (Ic) to a 3-oxo-hydroxyacyl-CoA compound of formula (IIc), wherein the means comprises both CoA transferase and β-ketothiolase activities.

[0245] In one embodiment is provided a polypeptide-substrate complex comprising: (a) a polypeptide having at least 95% sequence identity to the amino acid sequence set forth in any one of SEQ ID No: 1, 3, 4, 5, 6, 7, 8, and 9; and (b) an acyl-CoA compound; wherein said polypeptide has β-ketothiolase activity but no CoA-transferase activity.

[0246] In one embodiment is provided a polypeptide-substrate complex comprising: (a) a polypeptide having the amino acid sequence set forth in any one of SEQ ID No: 3 and 4; and (b) an acyl-CoA compound; wherein said polypeptide has β-ketothiolase activity but no CoA-transferase activity.

[0247] In one embodiment is provided a polypeptide-substrate complex comprising: (a) a polypeptide having the amino acid sequence set forth in any one of SEQ ID No: 3 and 4; and (b) an acyl-CoA compound; wherein said polypeptide has β-ketothiolase activity but reduced CoA-transferase activity compared to a polypeptide having the amino acid sequence set forth in any one of SEQ ID No: 1, 5, 6, 7, 8, and 9.

[0248] In one embodiment is provided a polypeptide having β-ketothiolase activity, wherein said polypeptide does not comprise a ser-his-his and does not comprise a cys-his-cys triad. In one embodiment the polypeptide also has CoA-

transferase activity. In one embodiment the polypeptide has β-ketothiolase activity but no CoA-transferase activity.

[0249] In one embodiment is provided a polypeptide having at least 95% sequence identity to the amino acid sequence set forth in any one of SEQ ID No: 1, 5, 6, 7, 8, and 9, wherein the amino acid sequence of said polypeptide includes the motif I/V/A/LTE. In one embodiment the amino acid sequence of said polypeptide includes at least one of the motifs EXGXXG and GXGG(A/F). In one embodiment the amino acid sequence of said polypeptide includes both of the motifs EXGXXG and GXGG(A/F).

[0250] It is to be understood that while the invention has been described in conjunction with the detailed description thereof, the foregoing description is intended to illustrate and not limit the scope of the invention, which is defined by the scope of the appended claims. Other aspects, advantages, and modifications are within the scope of the following claims.

WHAT IS CLAIMED IS:

1. A method of producing 3-keto-acyl-CoA esters, said method comprising the step of enzymatically condensing acetyl-CoA with any of an alkanoic acid, an alkenoic acid, a hydroxyacid, or a haloacid using a polypeptide that has both CoA transferase and β-ketothiolase activities.

2. The method of claim 1, wherein said alkanoic acid may be selected from straight chain alkanoic acids of carbon chain length *n* (wherein $n > 2$) such as acetate, propionate, butyrate, pentanoic acid, hexanoic acid, and the like, or branched chain alkanoic acids such as isobutyrate, isovaleric acid or pivalic acid.

3. The method of claim 1, wherein said alkenoic acid may be selected from straight chain alkenoic acids of carbon chain length *n* (wherein $n > 2$) such as acrylic acid, 2-propenoic acid, 3-butenoic acid, 4-pentenoic acid, 5-hexenoic acid, 6-hepteneoic acid, crotonic acid, and the like, or branched chain alkenoic acids such as methacrylic acid, 3-methyl-3-butenoic acid, 4-methyl-4-pentenoic acid, and 5-methyl-5-hexenoic acid, and the like.

4. The method of claim 1, wherein said hydroxyacid may be selected from hydroxyacids of carbon chain length *n* (wherein $n > 2$) such as 3-hydroxypropionic acid, 4-hydroxybutyric acid, 5-hydroxyvaleric acid, and 6-hydroxy-caproic acid.

5. The method of claim 1, wherein said haloacid may be selected from haloacids of carbon chain length *n* (wherein $n > 2$) such as 3-halopropionic acid, 4-halobutyric acid, 5-halovaleric acid, and 6-halocaproic acid.

6. The method of claim 1, comprising the step of enzymatically condensing acetyl-CoA with the respective CoA ester of any of said alkanoic acid, said alkenoic acid, said hydroxyacid, or said haloacid using a polypeptide that has both CoA transferase and β-ketothiolase activities.

7. The method of claim 1, wherein the 3-keto-acyl-CoA ester is converted to its respective free acid by a CoA transferase or a thioesterase.

8. A method according to claim 1, wherein said method comprises enzymatically converting propenoate to 3-oxopent-4-enoyl-CoA using a polypeptide that has both CoA transferase and β-ketothiolase activities, resulting in the production of 3-oxopent-4-enoyl-CoA.

9. The method of any of the preceding claims, wherein said polypeptide is classified under EC 2.8.3.-.

10. The method of claim 9, wherein said polypeptide is classified under EC 2.8.3.8.

11. The method of any of the preceding claims, wherein said polypeptide is from *Salmonella enterica, Peptostreptococcaceae bacterium*, *Firmicutes bacterium, Megasphaera elsdenii*, *Salmonella enterica subsp. houtenaeserovar*, *Clostridium aminobutyricum*, or *Clostridium propionicum*.

12. The method of any of the preceding claims, wherein said polypeptide has at least 95% sequence identity to the amino acid sequence set forth in any one of SEQ ID No: 1, 4, 5, 6, 7, 8, and 9.

13. The method of any of the preceding claims, wherein said polypeptide has an amino acid substitution at at least one of position 38, 60, 112, 258, and 390 of SEQ ID NO: 1.

14. The method of any of the preceding claims, said method further comprising enzymatically converting 3-oxopent-4-enoyl-CoA to 3-oxopent-4-enoate.

15. The method of claim 14, wherein 3-oxopent-4-enoyl-CoA is enzymatically converted to 3-oxopent-4-enoate using a CoA-transferase.

16. The method of claim 15, wherein the CoA-transferase is classified under EC 2.8.3.6 or EC 2.8.3.8.

17. The method of any of claims 14-16, said method further comprising enzymatically converting 3-oxopent-4-enoate to 3-buten-2-one.

18. The method of claim 17, wherein 3-oxopent-4-enoate is enzymatically converted to 3-buten-2-one using a decarboxylase.

19. The method of claim 18, wherein the decarboxylase is classified under EC 4.1.1.4.

20. The method of any of claims 17-19, said method further comprising enzymatically converting 3-buten-2-one to 3-buten-2-ol.

21. The method of claim 20, wherein 3-buten-2-one is converted to 3-buten-2-ol using an alcohol dehydrogenase or a phenylacetaldehyde reductase.

22. The method of claim 21, wherein said alcohol dehydrogenase has at least 95% sequence identity to the amino acid sequence set forth in SEQ ID NO:16.

23. The method of claim 21, wherein said phenylacetaldehyde reductase has at least 95% sequence identity to the amino acid sequence set forth in SEQ ID NO:17.

24. The method of any of claims 20-23, said method further comprising enzymatically converting 3-buten-2-ol to 1,3 butadiene.

25. The method of claim 24, wherein 3-buten-2-ol is converted to 1,3 butadiene using a linalool dehydratase.

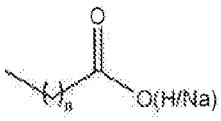
26. The method of claim 25, wherein said linalool dehydratase has at least 95% sequence identity to the amino acid sequence set forth in SEQ ID NO:18.

27. The method of claim 25, wherein said linalool dehydratase is classified under EC 4.2.1.127.

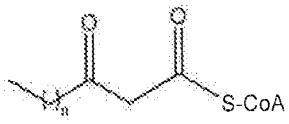
28. The method of claim 24, wherein 3-buten-2-ol is converted to 1,3 butadiene using a dehydratase classified under EC 4.2.1.-.

29. The method of claim 24, wherein 3-buten-2-ol is converted to 1,3 butadiene using a dehydratase having greater than 95% sequence identity to the 5-aminovaleryl-CoA dehydratase from C. viride or to a dehydratase classified under EC 4.2.1.- from species such as *Aquincola tertiaricarbonis* or *Methylibium petroleiphilum* PM1.

30. A method of producing 3-oxo-acyl-CoA compounds of formula (IIa), said method comprising enzymatically converting a carboxylic acid of formula (Ia) to a 3-oxo-acyl-CoA compound of formula (IIa) :



Formula (Ia)

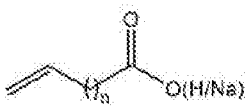


Formula (IIa)

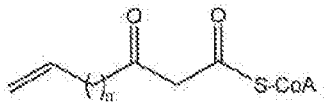
using a polypeptide that has both CoA transferase and β-ketothiolase activities.

31. The method of claim 30, wherein n is 1 or 2.

32. A method of producing 3-oxo-enoyl-CoA compounds of formula (IIb), said method comprising enzymatically converting an unsaturated carboxylic acid of formula (Ib) to a 3-oxo-enoyl-CoA compound of formula (IIb):



Formula (Ib)

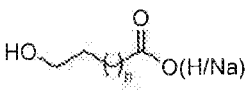


Formula (IIb)

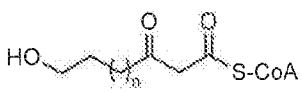
using a polypeptide that has both CoA transferase and β-ketothiolase activities.

33. The method of claim 32, wherein n is 0, 1, 2, 3, or 4.

34. A method of producing 3-oxo-hydroxyacyl-CoA compounds of formula (IIc), said method comprising enzymatically converting a hydroxyl-substituted carboxylic acid of formula (Ic) to a 3-oxo-hydroxyacyl-CoA compound of formula (IIc):



Formula (Ic)



Formula (IIc)

using a polypeptide that has both CoA transferase and β-ketothiolase activities.

35. The method of claim 34, wherein n is 0, 1, or 2.

36. The method of claim 35, said method further comprising converting the 3-oxo-hydroxyacyl-CoA compound of formula (IIc) to a nylon compound.

37. The method of claim 36, said method comprising enzymatically converting 4-hydroxybutyric acid, or a salt thereof, to 3-oxo-6-hydroxyhexanoyl-CoA using a polypeptide that has both CoA transferase and β-ketothiolase activities.

38. The method of claim 37, said method further comprising enzymatically converting 3-oxo-6-hydroxyhexanoyl-CoA to 6-hydroxyhexanoic acid, and enzymatically converting 6-hydroxyhexanoic acid to one or more of adipic acid, 6-aminohexanoic acid, hexamethylenediamine, caprolactam, and 1,6-hexanediol using one or more isolated enzymes selected from dehydrogenases, reductases, hydratases, thioesterases, monooxygenases, and transaminases.

39. The method of claim 34, said method comprising enzymatically converting 3-hydroxypropionic acid, or a salt thereof, to 3-oxo-5-hydroxypentanoyl-CoA using a polypeptide that has both CoA transferase and β-ketothiolase activities.

40. The method of claim 34, said method comprising enzymatically converting 5-hydroxypentanoic acid, or a salt thereof, to 3-oxo-7-hydroxyheptanoyl-CoA using a polypeptide that has both CoA transferase and β-ketothiolase activities.

41. The method of any of claims 30-40, wherein said polypeptide having both CoA transferase and β-ketothiolase activities is classified under EC 2.8.3.-.

42. The method of claim 41, wherein said polypeptide having both CoA transferase and β-ketothiolase activities is classified under EC 2.8.3.8.

43. The method of any of claims 30-40, wherein said polypeptide having both CoA transferase and β-ketothiolase activities is from *Salmonella enterica*, *Peptostreptococcaceae bacterium*, *Firmicutes bacterium*, *Megasphaera elsdenii*,

Salmonella enterica subsp. houtenaeserovar, *Clostridium aminobutyricum*, or *Clostridium propionicum*.

44. The method of any of claims 30-40, wherein said polypeptide having both CoA transferase and β-ketothiolase activities has at least 95% sequence identity to the amino acid sequence set forth in any one of SEQ ID No: 1, 4, 5, 6, 7, 8, and 9.

45. The method of any of claims 30-40, wherein said polypeptide having both CoA transferase and β-ketothiolase activities has an amino acid substitution at at least one of position 38, 60, 112, 258, and 390 of SEQ ID NO: 1.

46. The method of any of claims 30-40, wherein said polypeptide having both CoA transferase and β-ketothiolase activities is a CoA transferase whose amino acid sequence includes the motif I/V/A/LTE.

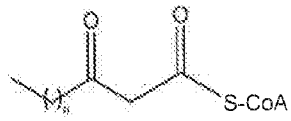
47. The method of any of claims 30-40, wherein said polypeptide having both CoA transferase and β-ketothiolase activities is a CoA transferase whose amino acid sequence includes the motifs EXGXXG and GXGG(A/F).

48. The method of any of the preceding claims, wherein said method is performed in a non-naturally occurring host.

49. A non-naturally occurring host capable of producing 3-oxopent-4-enoyl-CoA, said host comprising at least one exogenous nucleic acid encoding a polypeptide having both CoA transferase and β-ketothiolase activities.

50. A non-naturally occurring host capable of producing butadiene, said host comprising at least one exogenous nucleic acid encoding a polypeptide having both CoA transferase and β-ketothiolase activities.

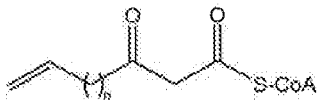
51. A non-naturally occurring host capable of producing 3-oxo-acyl-CoA compounds of formula (IIa):



Formula (IIa) ,

said host comprising at least one exogenous nucleic acid encoding a polypeptide having both CoA transferase and β-ketothiolase activities.

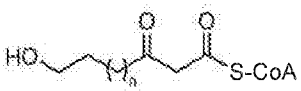
52. A non-naturally occurring host capable of producing 3-oxo-enoyl-CoA compounds of formula (IIb):



Formula (IIb) ,

said host comprising at least one exogenous nucleic acid encoding a polypeptide having both CoA transferase and β-ketothiolase activities.

53. A non-naturally occurring host capable of producing 3-oxo-hydroxyacyl-CoA compounds of formula (IIc)



Formula (IIc) ,

said host comprising at least one exogenous nucleic acid encoding a polypeptide having both CoA transferase and β-ketothiolase activities.

54. A non-naturally occurring host capable of producing nylon compounds, said host comprising at least one exogenous nucleic acid encoding a polypeptide having both CoA transferase and β-ketothiolase activities.

55. The host of any of claims 49-54, wherein said polypeptide having both CoA transferase and β-ketothiolase activities is classified under EC 2.8.3.-.

56. The host of any of claims 49-54, wherein said polypeptide having both CoA transferase and β-ketothiolase activities is classified under EC 2.8.3.8.

57. The host of any of claims 49-54, wherein said polypeptide having both CoA transferase and β-ketothiolase activities is from *Salmonella enterica, Peptostreptococcaceae bacterium, Firmicutes bacterium, Megasphaera elsdenii, Salmonella enterica subsp. houtenaeserovar, Clostridium aminobutyricum*, or *Clostridium propionicum*.

58. The method of claim 41 or the host of any of claims 49-557, wherein said host is cultured under aerobic, anaerobic, or micro-aerobic cultivation conditions.

59. The method of claim 41 or the host of any of claims 49-57, wherein said host is cultured under conditions of nutrient limitation.

60. The method or host of claim 58 or 59, wherein said host is retained using a ceramic hollow fiber membrane.

61. The method or host of any of claims 58-60, wherein the principal carbon source derives from a biological feedstock.

62. The method or host of claim 61, wherein the biological feedstock is, or derives from, monosaccharides, disaccharides, lignocellulose, hemicellulose, cellulose, lignin, levulinic acid, formic acid, triglycerides, glycerol, fatty acids, agricultural waste, condensed distillers' solubles, or municipal waste.

63. The method or host of any of claims 58-60, wherein the principal carbon source derives from a non-biological feedstock.

64. The method or host of claim 63, wherein the non-biological feedstock is, or derives from, natural gas, syngas, CO_2/H_2, methanol, ethanol, benzoate, non-volatile residue (NVR) caustic wash waste stream from cyclohexane oxidation processes, or terephthalic acid / isophthalic acid mixture waste streams.

65. The method or host of any of claims 58-64 wherein the host is a prokaryote.

66. The method or host of claim 65, wherein said prokaryote is from a genus selected from the group consisting of *Escherichia*, *Clostridia*, *Corynebacteria*, *Cupriavidus*, *Pseudomonas*, *Delftia*, *Bacillus*, *Lactobacillus*, *Lactococcus*, and *Rhodococcus*.

67. The method or host of claim 66, wherein said prokaryote is selected from the group consisting of *Escherichia coli*, *Clostridium ljungdahlii*, *Clostridium autoethanogenum*, *Clostridium kluyveri*, *Corynebacterium glutamicum*, *Cupriavidus necator*, *Cupriavidus metallidurans*, *Pseudomonas fluorescens*, *Pseudomonas putida*, *Pseudomonas oleavorans*, *Delftia acidovorans*, *Bacillus subtillis*, *Lactobacillus delbrueckii*, *Lactococcus lactis*, and *Rhodococcus equi*.

68. The method or host of any of claims 58-64 wherein the host is a eukaryote.

69. The method or host of claim 68, wherein said eukaryote is from a genus selected from the group consisting of *Aspergillus*, *Saccharomyces*, *Pichia*, *Yarrowia*, *Issatchenkia*, *Debaryomyces*, *Arxula*, and *Kluyveromyces*.

70. The method or host of claim 69, wherein said eukaryote is selected from the group consisting of *Aspergillus niger*, *Saccharomyces cerevisiae*, *Pichia pastoris*, *Yarrowia lipolytica*, *Issathenkia orientalis*, *Debaryomyces hansenii*, *Arxula adenoinivorans*, and *Kluyveromyces lactis*.

71. The host of any one of claims 49-70, wherein said host further comprises an exogenous alcohol dehydrogenase or an exogenous phenylacetaldehyde reductase, and an exogenous linalool dehydratase, wherein said host is capable of producing 1,3-butadiene.

72. The host of claim 71, wherein said alcohol dehydrogenase has at least 95% sequence identity to the amino acid sequence set forth in SEQ ID NO:16.

73. The host of claim 71, wherein said phenylacetaldehyde reductase has at least 95% sequence identity to the amino acid sequence set forth in SEQ ID NO:17.

74. The host of any one of claims 71-73, wherein said linalool dehydratase has at least 95% sequence identity to the amino acid sequence set forth in SEQ ID NO:18.

75. The host of any one of claims 49-74, wherein said polypeptide having both CoA transferase and β-ketothiolase activities has an amino acid substitution at position 38, 60, 112, 258, or 390 of SEQ ID NO: 1.

76. A composition comprising 3-oxopent-4-enoyl-CoA synthesized by the method of any one of claims 1-29.

77. A composition comprising butadiene synthesized by the method of any one of claims 24-29.

78. A method for producing bio-derived 3-oxopent-4-enoyl-CoA, comprising culturing or growing a host according to any one of claims 41-68 under conditions and for a sufficient period of time to produce bio-derived 3-oxopent-4-enoyl-CoA.

79. A method for producing bio-derived butadiene, comprising culturing or growing a host according to any one of claims 48-75 under conditions and for a sufficient period of time to produce bio-derived butadiene.

80. A variant polypeptide having at least 95% sequence identity to the amino acid sequence of a wild-type enzyme set forth in any of SEQ NOs: 1, 5, 6, 7, 8, and 9, wherein said variant polypeptide comprises at least one mutation to the wild-type enzyme.

81. The variant polypeptide of claim 80, wherein said wild type enzyme has the amino acid sequence set forth in SEQ ID NO:1.

82. The variant polypeptide of claim 80 or 81, wherein said at least one mutation results in an increase in the transferase activity, the β-ketothiolase activity, or both activities of the wild-type enzyme.

83. The variant polypeptide of claim 80 or 81, wherein said at least one mutation results in a decrease in or a loss of the transferase activity or the β-ketothiolase activity of the wild-type enzyme.

84. The variant polypeptide of any one of claims 80-83, wherein said mutation is a deletion of at least one amino acid.

85. The variant polypeptide of any one of claims 80-83, wherein said mutation is an addition of at least one amino acid.

86. The variant polypeptide of any one of claims 80-83, wherein said mutation is a substitution of at least one amino acid.

87. The variant polypeptide of claim 86, wherein said substitution is a conservative substitution.

88. The variant polypeptide of any of claims 80-83, said variant polypeptide having the amino acid sequence set forth in SEQ ID NO:1 with an amino acid other than a serine at position 38, an amino acid other than a serine at position 60, an amino acid other than a cysteine at position 112, an amino acid other than a cysteine at position 258, or an amino acid other than a cysteine at position 390.

89. The variant polypeptide of claim 88, wherein said variant polypeptide comprises an alanine at position 38, an alanine at position 60, an alanine at position 112, or an alanine at position 390 of SEQ ID NO: 1.

90. The variant polypeptide of any of claims 80-83, said variant polypeptide having the amino acid sequence set forth in SEQ ID NO: 1 with an amino acid other than a glutamic acid at position 324.

91. The variant polypeptide of claim 90, wherein said variant polypeptide comprises a leucine at position 324.

92. The variant polypeptide of claim 90, wherein said variant polypeptide comprises a glycine at position 324.

93. The variant polypeptide of any of claims 80-83, said variant polypeptide having an amino acid sequence comprising the motifs I/V/A/LTE, EXGXXG, and GXGG(A/F).

94. The variant polypeptide of any of claims 80-83, wherein said at least one mutation results in the presence of an XXGXXG motif where an EXGXXG motif was present in the wild-type enzyme.

95. The variant polypeptide of claim 94, wherein said at least one mutation results in a decrease in or loss of the transferase activity of the wild-type enzyme.

96. The variant polypeptide of claim 94 or 95, wherein said at least one mutation does not affect the β-ketothiolase activity of the wild-type enzyme.

97. A bio-derived, bio-based, or fermentation-derived product comprising:

(a) a composition comprising at least one bio-derived, bio-based, or fermentation-derived compound prepared (i) using the host of any one of claims 48-75, (ii) using the variant polypeptide of any one of claims 80-96, or (iii) according to the method of any one of claims 1-48, 58-70, or 78-79, or any combination thereof,

(b) a bio-derived, bio-based, or fermentation-derived polymer or resin, for example styrene-butadiene-rubber, poly-butadiene, styrene-butadiene latex, acrylonitrile-butadiene-styrene resin, nitrile rubber, adiponitrile, or nylon compounds,comprising the bio-derived, bio-based, or fermentation-derived composition or compound of (a), or any combination thereof,

(c) a molded substance obtained by molding the bio-derived, bio-based, or fermentation-derived polymer or resin of (b), or any combination thereof,

(d) a bio-derived, bio-based, or fermentation-derived formulation comprising the bio-derived, bio-based, or fermentation-derived composition or compound of (a), bio-derived, bio-based, or fermentation-derived polymer or resin of (b), or bio-derived, bio-based, or fermentation-derived molded substance of (c), or any combination thereof, or

a bio-derived, bio-based, or fermentation-derived semi-solid or a non-semi-solid stream, comprising the bio-derived, bio-based, or fermentation-derived composition or compound of (a), bio-derived, bio-based, or fermentation-derived polymer or resin of (b), bio-derived, bio-based, or fermentation-derived

formulation of (d), or bio-derived, bio-based, or fermentation-derived molded substance of (c), or any combination thereof.

98. A method of producing 3-oxopent-4-enoyl-CoA, said method comprising:
(a) providing propenoate; and
(b) providing a means for enzymatically converting propenoate to 3-oxopent-4-enoyl-CoA,
wherein the means comprises both CoA transferase and β-ketothiolase activities.

99. The method of claim 98, said method further comprising enzymatically converting 3-oxopent-4-enoyl-CoA to 3-oxopent-4-enoate.

100. The method of claim 99, wherein 3-oxopent-4-enoyl-CoA is enzymatically converted to 3-oxopent-4-enoate using a CoA-transferase.

101. The method of claim 100, wherein the CoA-transferase is classified under EC 2.8.3.6 or EC 2.8.3.8.

102. The method of any of claims 99-101, said method further comprising enzymatically converting 3-oxopent-4-enoate to 3-buten-2-one.

103. The method of claim 102, wherein 3-oxopent-4-enoate is enzymatically converted to 3-buten-2-one using a decarboxylase.

104. The method of claim 103, wherein the decarboxylase is classified under EC 4.1.1.4.

105. The method of any of claims 102-104, said method further comprising enzymatically converting 3-buten-2-one to 3-buten-2-ol.

106. The method of claim 105, wherein 3-buten-2-one is converted to 3-buten-2-ol using an alcohol dehydrogenase or a phenylacetaldehyde reductase.

107. The method of claim 106, wherein said alcohol dehydrogenase has at least 95% sequence identity to the amino acid sequence set forth in SEQ ID NO:16.

108. The method of claim 106, wherein said phenylacetaldehyde reductase has at least 95% sequence identity to the amino acid sequence set forth in SEQ ID NO:17.

109. The method of any of claims 105-108, said method further comprising enzymatically converting 3-buten-2-ol to 1,3 butadiene.

110. The method of claim 109, wherein 3-buten-2-ol is converted to 1,3 butadiene using a linalool dehydratase.

111. The method of claim 110, wherein said linalool dehydratase has at least 95% sequence identity to the amino acid sequence set forth in SEQ ID NO:18.

112. A method of producing butadiene, said method comprising:

(a) providing propenoate;

(b) providing a means for enzymatically converting propenoate to 3-oxopent-4-enoyl-CoA, wherein the means comprises both CoA transferase and β-ketothiolase activities;

(c) enzymatically converting 3-oxopent-4-enoyl-CoA to 3-oxopent-4-enoate using a CoA-transferase;

(d) enzymatically converting 3-oxopent-4-enoate to 3-buten-2-one using a decarboxylase;

(e) enzymatically converting 3-buten-2-one to 3-buten-2-ol using an alcohol dehydrogenase or a phenylacetaldehyde reductase; and

(f) enzymatically converting 3-buten-2-ol to 1,3 butadiene using a linalool dehydratase.

113. A non-naturally occurring microorganism comprising at least one exogenous nucleic acid encoding at least one polypeptide having the activity of both a CoA transferase and a β-ketothiolase.

114. A non-naturally occurring host capable of producing butadiene, said host comprising:

(a) a means for enzymatically converting propenoate to 3-oxopent-4-enoyl-CoA, wherein the means comprises both CoA transferase and β-ketothiolase activities;

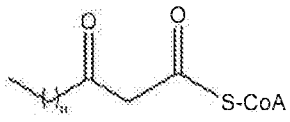
(b) at least one exogenous nucleic acid encoding a CoA-transferase;

(c) at least one exogenous nucleic acid encoding a decarboxylase;

(d) at least one exogenous nucleic acid encoding an alcohol dehydrogenase or a phenylacetaldehyde reductase; and

(e) at least one exogenous nucleic acid encoding a linalool dehydratase.

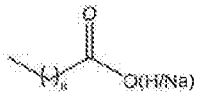
115. A non-naturally occurring host capable of producing 3-oxo-acyl-CoA compounds of formula (IIa):



(IIa),

said host comprising:

(a) a carboxylic acid of formula (Ia):

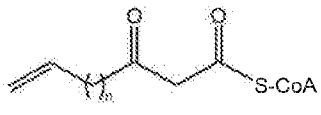


(Ia)

or a salt form thereof; and

(b) a means for enzymatically converting the carboxylic acid of formula (Ia) to a 3-oxo-acyl-CoA compound of formula (IIa), wherein the means comprises both CoA transferase and β-ketothiolase activities.

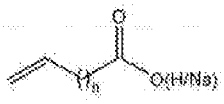
116. A non-naturally occurring host capable of producing 3-oxo-enoyl-CoA compounds of formula (IIb):



(IIb),

said host comprising:

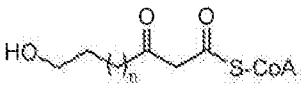
(a) an unsaturated carboxylic acid of formula (Ib):



(Ib), or a salt form thereof; and

(b) a means for enzymatically converting the unsaturated carboxylic acid of formula (Ib) to a 3-oxo-enoyl-CoA compound of formula (IIb), wherein the means comprises both CoA transferase and β-ketothiolase activities.

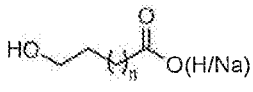
117. A non-naturally occurring host capable of producing 3-oxo-hydroxyacyl-CoA compounds of formula (IIc)



(IIc),

said host comprising:

(a) a hydroxy-substituted carboxylic acid of formula (Ic):



(Ic), or a salt form thereof; and

(b) a means for enzymatically converting the hydroxy-substituted carboxylic acid of formula (Ic) to a 3-oxo-hydroxyacyl-CoA compound of formula (IIc), wherein the means comprises both CoA transferase and β-ketothiolase activities.

118. A polypeptide-substrate complex comprising:
(a) a polypeptide having at least 95% sequence identity to the amino acid sequence set forth in any one of SEQ ID No: 1, 3, 4, 5, 6, 7, 8, and 9; and
(b) an acyl-CoA compound;
wherein said polypeptide has β-ketothiolase activity but no CoA-transferase activity.

119. A polypeptide-substrate complex comprising:
(a) a polypeptide having the amino acid sequence set forth in any one of SEQ ID No: 3 and 4; and
(b) an acyl-CoA compound;
wherein said polypeptide has β-ketothiolase activity but no CoA-transferase activity.

120. A polypeptide-substrate complex comprising:
(a) a polypeptide having the amino acid sequence set forth in any one of SEQ ID No: 3 and 4; and
(b) an acyl-CoA compound;
wherein said polypeptide has β-ketothiolase activity but reduced CoA-transferase activity compared to a polypeptide having the amino acid sequence set forth in any one of SEQ ID No: 1, 5, 6, 7, 8, and 9.

121. A polypeptide having β-ketothiolase activity, wherein said polypeptide does not comprise a ser-his-his and does not comprise a cys-his-cys triad.

122. The polypeptide of claim 121, wherein said polypeptide also has CoA-transferase activity.

123. The polypeptide of claim 121, wherein said polypeptide has β-ketothiolase activity but no CoA-transferase activity.

124. A polypeptide having at least 95% sequence identity to the amino acid sequence set forth in any one of SEQ ID No: 1, 5, 6, 7, 8, and 9, wherein the amino acid sequence of said polypeptide includes the motif I/V/A/LTE.

125. The polypeptide of claim 124, wherein the amino acid sequence of said polypeptide includes at least one of the motifs EXGXXG and GXGG(A/F).

126. The polypeptide of claim 124, wherein the amino acid sequence of said polypeptide includes both of the motifs EXGXXG and GXGG(A/F).

Figure 1A

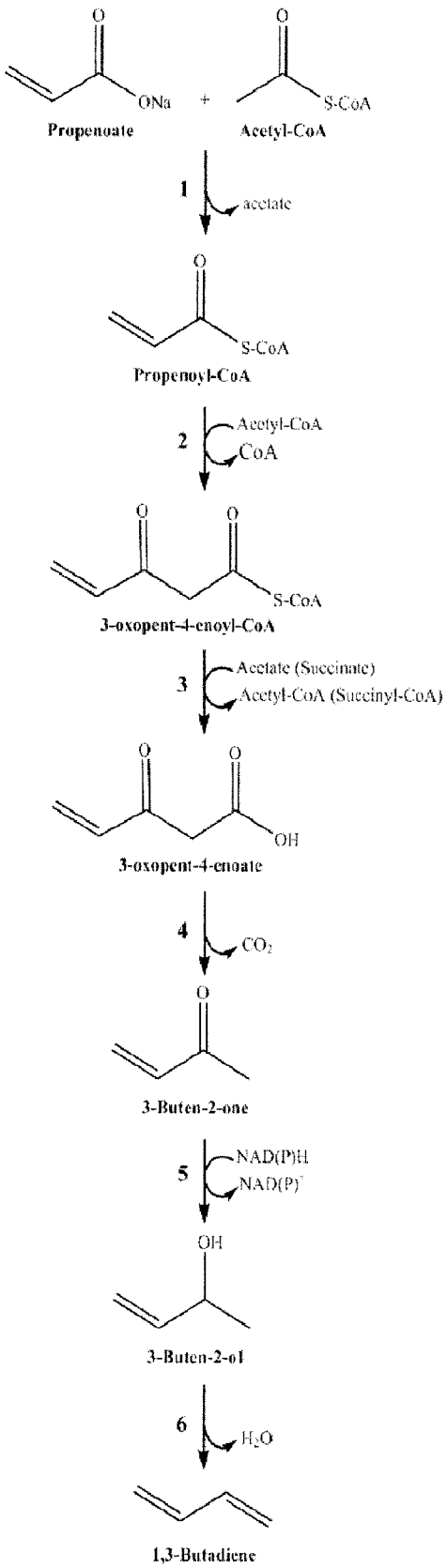
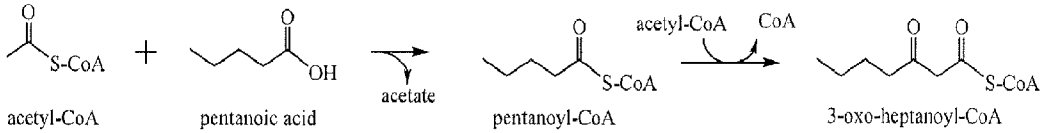

ONa
S-CoA
Propenoate
Acetyl-CoA
1
acetate
S-CoA
Propenoyl-CoA
2
Acetyl-CoA
CoA
S-CoA
3-oxopent-4-enoyl-CoA
3
Acetate (Succinate)
Acetyl-CoA (Succinyl-CoA)
OH
3-oxopent-4-enoate
4
CO2
3-Buten-2-one
5
NAD(P)H
NAD(P)+
OH
3-Buten-2-ol
6
H2O
1,3-Butadiene

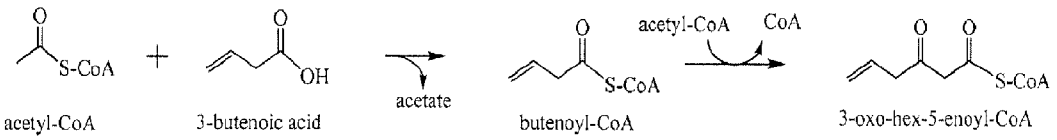
Figure 1B

(i)



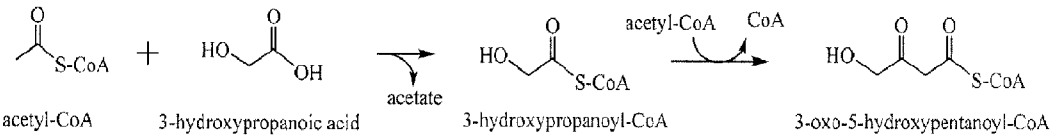
acetyl-CoA + pentanoic acid → pentanoyl-CoA → 3-oxo-heptanoyl-CoA

(ii)



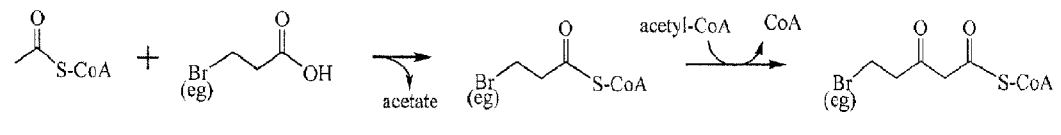
acetyl-CoA + 3-butenoic acid → butenoyl-CoA → 3-oxo-hex-5-enoyl-CoA

(iii)



acetyl-CoA + 3-hydroxypropanoic acid → 3-hydroxypropanoyl-CoA → 3-oxo-5-hydroxypentanoyl-CoA

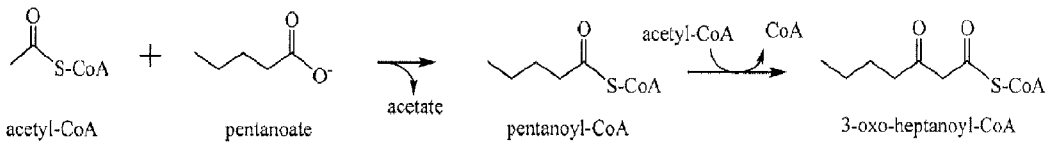
(iv)



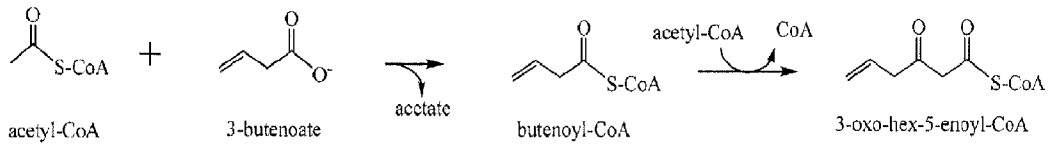
acetyl-CoA + 4-halobutanoic acid → 4-halobutanoyl-CoA → 3-oxo-6-halobutanoyl-CoA

Figure 1C

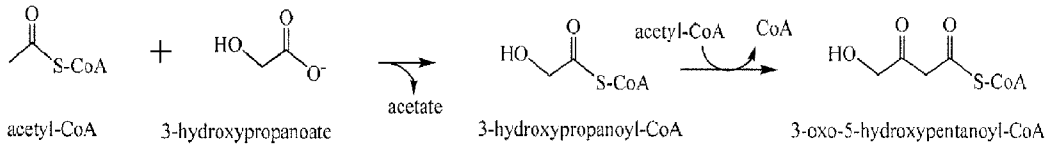
(i)



(ii)



(iii)



(iv)

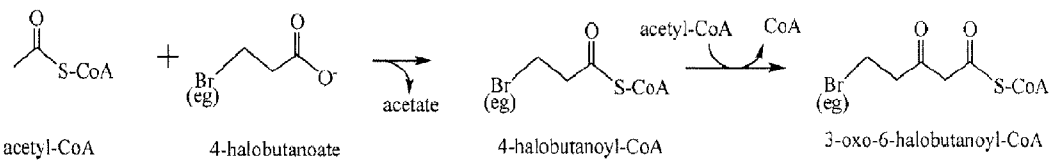


Figure 2

SEQ ID NO	Source	Sequence
1	*Clostridium propionicum*, *YdiF*	MRKVPIITADEAAKLIKDGDTVTTSGFVGNAIPEALDRAVEKRFLETGEPKNITY VYCGSQGNRDGRGAEHFAHEGLLKRYIAGHWATVPALGKMAMENKMEAYNV SQGALCHLFRDIASHKPGVFTKVGIGTFIDPRNGGGKVNDITKEDIVELVEIKGQ EYLFYPAFPIHVALIRGTYADESGNITFEKEVAPLEGTSVCQAVKNSGGIVVVQ VERVVKAGTLDPRHVKVPGIYVDYVVVADPEDHQQSLDCEYDPALSGEHRRP EVVGEPLPLSAKKVIGRRGAIELEKDVAVNLGVGAPEYVASVADEEGIVDFMTL TAESGAIGGVPAGGVRFGASYNADALIDQGYQFDYYDGGGLDLCYLGLAECD EKGNINVSRFGPRIAGCGGFINITQNTPKVFFCGTFTAGGLKVKIEDGKVIIVQE GKQKKFLKAVEQITFNGDVALANKQQVTYITERCVFLLKEDGLHLSEIAPGIDLQ TQILDVMDFAPIIDRDANGQIKLMDAALFAEGLMGLKEMKS

SEQ ID NO	Source	Sequence
2	Modified *Clostridium propionicum*, where 2 amino acids were added to the N-terminal sequence followed by 6xHis-tag, a linker 3 amino acids in length and a specific proteolytic cleavage site (thrombin site, underlined)	MGHHHHHHSSGLVPRGSMRKVPITADEAAKLIKDGDTVTTSGFVGNAIPEAL DRAVEKRFLETGEPKNITYVYCGSQGNRDGRGAEHFAHEGLLKRYIAGHWAT VPALGKMAMENKMEAYNVSQGALCHLFRDIASHKPGVFTKVGIGTFIDPRNG GGKVNDITKEDIVELVEIKGQEYLFYPAFPIHVALIRGTYADESGNITFEKEVAPL EGTSVCQAVKNSGGIVVVQVERVVKAGTLDPRHVKVPGIYVDYVVVAADPEDH QQSLDCEYDPALSGEHRRPEVVGEPLPLSAKKVIGRRGAIELEKDVAVNLGVG APEYVASVADEEGIVDFMTLTAESGAIGGVPAGVRFGASYNADALIDQGYQF DYYDGGGLDLCYLGLAECDEKGNINVSRFGPRIAGCGGFINITQNTPKVFFCG TFTAGGLKVKIEDGKVIIVQEGKQKKFLKAVEQITFNGDVALANKQQVTYITERC VFLLKEDGLHLSEIAPGIDLQTQILDVMDFAPIIDRDANGQIKLMDAALFAEGLM GLKEMKS

SEQ ID NO	Source	Sequence
3	YdiF mutant E324G (numbering as based on SEQ ID NO: 1, position 324 underlined)	MGHHHHHHSSGLVPRGSMRKVPIITADEAAKLIKDGDTVTTSGFVGNAIPEAL DRAVEKRFLETGEPKNITYVYCGSQGNRDGRGAEHFAHEGLLKRYIAGHWAT VPALGKMAMENKMEAYNVSQGALCHLFRDIASHKPGVFTKVGIGTFIDPRNG GGKVNDITKEDIVELVEIKGQEYLFYPAFPIHVALIRGTYADESGNITFEKEVAPL EGTSVCQAVKNSGGIVVVQVERVVKAGTLDPRHVKVPGIYVDYVVVADPEDH QQSLDCEYDPALSGEHRRPEVVGEPLPLSAKKVIGRRGAIELEKDVAVNLGVG APEYVASVADEEGIVDFMTLTAQSGAIGGVPAGGVRFGASYNADALIDQGYQF DYYDGGGLDLCYLGLAECDEKGNINVSRFGPRIAGCGGFINITQNTPKVFFCG TFTAGGLKVKIEDGKVIIVQEGKQKKFLKAVEQITFNGDVALANKQQVTYITERC VFLLKEDGLHLSEIAPGIDLQTQILDVMDFAPIIDRDANGQIKLMDAALFAEGLM GLKEMKS

SEQ ID NO	Source	Sequence
4	YdiF mutant E324L (numbering as based on SEQ ID NO: 1, position 324 underlined)	MGHHHHHHSSGLVPRGSMRKVPITADEAAKLIKDGDTVTTSGFVGNAIPEAL DRAVEKRFLETGEPKNITYVYCGSQGNRDGRGAEHFAHEGLLKRYIAGHWAT VPALGKMAMENKMEAYNVSQGALCHLFRDIASHKPGVFTKVGIGTFIDPRNG GGKVNDITKEDIVELVEIKGQEYLFYPAFPIHVALIRGTYADESGNITFEKEVAPL EGTSVCQAVKNSGGIVVVQVERVVKAGTLDPRHVKVPGIYVDYVVVAADPEDH QQSLDCEYDPALSGEHRRPEVVGEPLPLSAKKVIGRRGAIELEKDVAVNLGVG APEYVASVADEEGIVDFMTLTALSGAIGGVPAGGVRFGASYNADALIDQGYQF DYYDGGGLDLCYLGLAECDEKGNINVSRFGPRIAGCGGFINITQNTPKVFFCG TFTAGGLKVKIEDGKVIIVQEGKQKKFLKAVEQITFNGDVALANKQQVTYITERC VFLLKEDGLHLSEIAPGIDLQTQILDVMDFAPIIDRDANGQIKLMDAALFAEGLM GLKEMKS
5	*Peptostreptococcaceae* bacterium, Uniprot Accession No. U2L5C9	MAKFVTLEEAVSVVKNGDTVATTGFVQVANPEALEWALGKRFEETKEPRDLT LFYCAGQGDGDCRAVNHFAKEGMLKRVVAGHFNMAPLLRQFISDNKCEAYN VPQGVLCNMVRDIAAKKPGVISHVGLNTFADPRIEGCKINLVTKEDIVELMMIN GEEKLFYKTFPLTIAFIKGTYADERGNVTLENEGIPSEATSIAQSVHNCGGKVIV QVEKVVAAGTLDPKLVKVPGIYVDYIVQVDDPSMRQQCYGVDYEPELAGNVYI PLSDIPLKTPLNERKIIARRGAFEIRKGNVGNLGIGVPEVVSEVVSEEGITDWLT LTVEVGPVGGSPQGKNRFGTAINAEAILDQPYQFDFYDGGGLDIAYLGLAQAD AKGNLNVSKFGDRVAGCGGFIDISQNSKAVVFCGSFTAGGLKVEVNDGKLNIV QEGKVKKFVNKVQQITFSGEYARKTGQRVFYVTERAVFQMKPEGLTLIEIAPG VDLEKDVLNQMEFKPLIAKDLKLMDERIFRPGPMGIKNDN

SEQ ID NO	Source	Sequence
6	*Firmicutes* bacterium CAG:102, Uniprot Accession No. R5ADR5	MARQVKVITAAEAAALIKNGDTVTTSGFVASAIPEALDRAVEERFLATGEPRDIT YVYCGSQGNKDGRGAEHFAHEGLLKRYIAGHWATVPALQKMALENKMEAYN VSQGALCHLFRDIAAHRPGCFTKVGLGTFIDPRNGGGKVNDVTKEDIIELVNIK GQDYLFYPAFPINVALIRGTYADESGNISFEKEVSPLEGTSVCQAVKNSGGIVV VQVEKLVKAGTLDPRLVKVPGIYVDYVVVADPKDHQQTLDCDYDPALSGEMR NPDVAPEPLPLSAKKIGRRGAVELEKDVAVNLGVGAPEYVASVANEEGIGDF MTLTVEGGAVGGVPAGGIRFGSAYNADALLDQGYQFDFYDGGGLDLCYLGLA ECDPHGSINVSRFGPRIAGCGGFINITQCTPKVFFCGTFTAGGLKVKVEDGKV VIAQEGKNKKFVKSVEQVTFNGDIANKNGQHVMMYITERCVFVLKEDGLHLTEIA PGIDLQTQILDQMEFEPIIDRNADGSITLMDAKLFADGLMGLKKEMKEGK
7	*Megasphaera elsdenii*, Uniprot Accession No. G0VND6	MRKVEIITAEQAAQLVKDNDTITSIGFVSSAHPEALTKALEKRFLDTNTPQNLTYI YAGSQGKRDGRAAEHLAHTGLLKRAIIGHWQTVPAIGKLAVENKIEAYNFSQG TLVHWFRALAGHKLGVFTDIGLETFLDPRQLGGKLNDVTKEDLVKLIEVDGHE QLFYPTFPVNVAFLRGTYADESGNITMDEEIGPFESTSVAQAVHNCGGKVVVQ VKDVVAHGSLDPRMVKIPGIYVDYVVVAAPEDHQQTYDCEYDPSLSGEHRAP EGATDAALPMSAKKIGRRGALELTENAVVNLGVGAPEYVASVAGEEGIADTIT LTVEGGAIGGVPQGGARFGSSRNADAIIDHTYQFDFYDGGGLDIAYLGLAQCD GSGNINVSKFGTNVAGCGGFPNISQQTPNVYFCGTFTAGGLKIAVEDGKVKIL QEGKAKKFIKAVDQITFNGSYAARNGKHVLYITERCVFELTKEGLKLIEVAPGIDI EKDILAHMDFKPIIDNPKLMDARLFQIDGPMGLKK

SEQ ID NO	Source	Sequence
8	*Dyadobacter fermentans (strain ATCC 700827 / DSM 18053 / NS114)*, Uniprot Accession No. C6VTZ3	MNAYIVAGYRTAVGKAPRGGFRFTRPDDLGAAVIKHLLEKTPQLDPTRVDDVI VGNAVPEAEQGMQMGRYVALLSLPKNVSGITINRYCGSGVEAIAMASAKIHAG MAECIIAGGTESMSLVPTMGWKTALNYEIAHTNPDYYLSMGLTAEQVAQDFKI SREAQDNFSFQSHQKALRAQKEGWFAEGIVPVTVKETYFDQASGKKKTKETV ISQDEGPRADTTLEALNKLKPVFAAGGSVTAGNSSQTSDGAAFVLVMSERLVN ELGLKPIARMMSYATAGVDPRVMGIGPVAAVPLALKQAGLQLKDIQQVELNEA FAAQSLAVIQELGIDPEIVNPNGGAIALGHALGSTGARLSVQLFNEMKRRDQKY GMVTACVGGGQGVAGIYERLN
9	*Salmonella enterica subsp. houtenaeserovar*, Uniprot Accession No. V1HBS2	MLSTKQFTAQQAVELIQDGDKVILGGFIGAVVPEAIEKAIEDKFLAEGHPCNLGL IFAAGQGDAKEKAINRLAHEGLVSSAIGGHWGLIPGLQRLASEGKITGYNLPQG VICHLLRDSAAGKAGTLTHVGLGTFVDPRIEGGKINAKTTEDIVTYNINDVENLL YKKLDANIAILRGTTADTHGNITMEDECLILENLAAAQLVHNQGGKVIVQVKRIV PKGSLDPLQVKIPGIFVDALVVADGEAHMQTFAEAMNENYVGRGEKGIRERKI RPLDVKKVIARRAAMELKKNAIVNYGIGIPEIIAQVADEENVTQELIATVEPGAIG GSPAGGLSFGASAFPEAIITQDQMFDFYDGGGLDQAFLGLAETDAKGDLNVS KFGVKIAGCGGFINITQNAKHVFFCGSFTAGDSDIIVEEGKLIIRRDGQIKKFIKH VQQITFSSDTARKNHKPVLYITERAVFRLAAETIELIEIAPGIDLQHDILDKMEFR PTISPALKEMDKRIFSEALMSLSLK

SEQ ID NO	Source	Sequence
10 and 11	*Escherichia coli*, see GenBank Accession Nos. AAC75282.1 & AAC75281.1, respectively	MDAKQRIARRVAQELRDGDIVNLGIGLPTMVANYLPEGIHITLQSENGFLGLGP VTTAHPDLVNAGGQPCGVLPGAAMFDSAMSFALIRGGHIDACVLGGLQVDEE ANLANWVVPGKMVPGMGGAMDLVTGSRKVIIAMEHCAKDGSAKILRRCTMPL TAQHAVHMLVTELAVFRFIDGKMWLTEIADGCDLATVRAKTEARFEVAADLNT QRGDL (SEQ ID NO:10, beta subunit) MKTKLMTLQDATGFFRDGMTIMVGGFMGIGTPSRLVEALLESGVRDLTLIAND TAFVDTGIGPLIVNGRVRKVIASHIGTNPETGRRMISGEMDVVLVPQGTLIEQIR CGGAGLGGFLTPTGVGTVVEEGKQTLTLDGKTWLLERPLRADLALIRAHRCDT LGNLTYQLSARNFNPLIALAADITLVEPDELVETGELQPDHIVTPGAVIDHIIVSQ ESK (SEQ ID NO: 11, alpha subunit)

SEQ ID NO	Source	Sequence
12 and 13	*Pseudomonas putida* Genbank Accession No. ACA73091.1 (A subunit) and ACA73090.1 (B subunit)	MINKTYESIASAVEGITDGSTIMVGGFGTAGMPSELIDALIDTGTRDLTIISNNAG NGEIGLAALLKAGSVRKVVCSFPRQSDSYVFDELYRAGKIELEVVPQGNLAERI RAAGSGIGAFFSPTGYGTLLSEGKETREIDGRQYVLEMPLHADFALIKAYKGD RWGNLIYRKAARNFGPIMAMAAKTAIAQVDQIVELGELDPEHIITPGIFVQRVVA VTGAASSIANAV (SEQ ID NO:12, A subunit) MTITTKLSRTQMAQRVAADIQEGAYVNLGIGAPTLVANFLGDKEVFLHSENGLL GMGPSPAPGEEDDDLINAGKQHVTLLTGGAFFHHADSFSMMRGGHLDIAVLG AFQVSVKGDLANWHTGAEGSIPAVGGAMDLATGARQVFVMMDHLTKSGESKI VPECTYPLTGIGCVSRIYTDLAVLEVTSDGLKVVEICADIDFDELQKLSGVPLIK (SEQ ID NO: 13, B subunit)
14	*Chromobacterium violaceum*, see GenBank Accession No. AAQ61181.1	MKQQEVRQRAFAMPLTSPAFPPGPYRFVNREYMIITYRTDPAAIEAVLPEPLQ MAEPVVRYEFIRMPDSTGFGDYSESGQVIPVTFRGERGSYTLAMFLDDQPPL AGGRELWGFPKKAGKPRLEVHQDTLVGSLDFGPVRIATGTMGYKYEALDRSA LLASLAEPNFLLKIIPHVDGSPRICELVRYHTTDVAIKGAWSAPGSLELHPHALA PVAALPVLEVLSARHFVCDLTLDLGTVVFDYLRQ
15	*Clostridium acetobutylicum*, see GenBank Accession No. AAA63761.1	MLKDEVIKQISTPLTSPAFPRGPYKFHNREYFNIVYRTDMDALRKVVPEPLEID EPLVRFEIMAMHDTSGLGCYTESGQAIPVSFNGVKGDYLHMMYLDNEPAIAVG RELSAYPKKLGYPKLFVDSDTLVGTLDYGKLRVATATMGYKHKALDANEAKD QICRPNYMLKIPNYDGSPRICELINAKITDVTVHEAWTGPTRLQLFDHAMAPLN DLPVKEIVSSSHILADILPRAEVIYDYLK

SEQ ID NO	Source	Sequence
16	*Nocardia rhamnosiphila*, see GenBank accession No. WP_030525792	MRIRGAVLERIGAPVPYAESAPTISELELADPGPGEILVRIEAAGLCHSDLSVV DGNRVRPVPMLLGHEASGKVVQAGPGVDLPVGRRVAMTFLPRCGECAGCAS GGRTPCIPGSAANNAGELLGGGRRLHRDGAEVQHHLGVSGFATHAVVDRRS VVPVDDDVPPEVAAVLGCAVLTGGGALLNSAKPAATDRVMVVGLGGVGMAA VLVAVSLGVREVIAVDTVPDKLALARELGAGSAHTPAEVADRGVQAEVVVEAV GSARAFESAVAATAPGGVTVTVGLPAPDARATISPLGLVAQGRSIVGSYLGSA VPSRDIPEYVRMWREGRLPVEKLISARIGLADINGAMDELAAGHALRQVIMF
17	*Rhodococcus sp. ST-10*, see Uniprot Accession No. Q9ZN85	MKAIQYTRIGAEPELTEIPKPEPGPGEVLLEVTAAGVCHSDDFIMSLPEEQYTY GLPLTLGHEGAGKVAAVGEGVEGLDIGTNVVVYGPWGCGNCWHCSQGLENY CSRAQELGINPPGLGAPGALAEFMIVDSPRHLVPIGDLDPVKTVPLTDAGLTPY HAIKRSLPKLRGGSYAVVIGTGGLGHVAIQLLRHLSAATVIALDVSADKLELATK VGAHEVVLSDKDAAENVRKITGSQGAALVLDFVGYQPTIDTAMAVAGVGSDVT IVGIGDGQAHAKVGFFQSPYEASVTVPYWGARNELIELIDLAHAGIFDIAVETFS LDNGAEAYRRLAAGTLSGRAVVVPGL

SEQ ID NO	Source	Sequence
18	*Castellaniella defragrans*, see Uniprot Accession No. E1XUJ2	MRFTLKTTAIVSAAALLAGF GPPPRAAELP PGRLATTEDY FAQQAKQAVT PDVMAQLAYM NYIDFISPFY SRGCSFEAWE LKHTPQRVIK YSIAFYAYGL ASVALIDPKL RALAGHDLDI AVSKMKCKRV WGDWEEDGFG TDPIEKENIM YKGHLNLMYG LYQLVTGSRR YEAEHAHLTR IIHDEIAANP FAGIVCEPDN YFVQCNSVAY LSLWVYDRLH GTDYRAATRA WLDFIQKDLI DPERGAFYLS YHPESGAVKP WISAYTTAWT LAMVHGMDPA FSERYYPRFK QTFVEVYDEG RKARVRETAG TDDADGGVGL ASAFTLLLAR EMGDQQLFDQ LLNHLEPPAK PSIVSASLRY EHPGSLLFDE LLFLAKVHAG FGALLRMPPP AAKLAGK
19	*Clostridium aminobutyricum*, see Uniprot Accession No. Q9RM86	MDWKKIYEDRTCTADEAVKSIKSGDRVLFAHCVAEPPVLVEAMVANAAAYKN VTVSHMVTLGKGEYSKPEYKENFTFEGWFTSPSTRGSIAEGHGQFVPVFFHE VPSLIRKDIFHVDVFMVMVSPPDHNGFCCVGVSSDYTMQAIKSAKIVLAEVND QVPVVYGDTFVHVSEIDKFVETSHPLPEIGLPKIGEVEAAIGKHCASLIEDGSTL QLGIGAIPDAVLSQLKDKKHLGIHSEMISDGVVDLYEAGVIDCSQKSIDKGKMAI TFLMGTKRLYDFAANNPKVELKPVDYINHPSVVAQCSKMVCINACLQVDFMG QIVSDSIGTKQFSGVGGQVDFVRGASMSIDGKGKAIIAMPSVAKKKDGSMISKI VPFIDHGAAVTTSRNDADYVVTEYGIAEMKGKSLQDRARALINIAHPDFKDELK AEFEKRFNAAF

SEQ ID NO	Source	Sequence
20	*Candida parapsilosis*, see Uniprot Accession No. B2KJ46	MGEIESYCNK ELGPLPTKAP TLSKNVLDLF SLKGKVASVT GSSGGIGWAV AEAYAQAGAD VAIWYNSHPA DEKAEHLQKT YGVHSKAYKC NISDPKSVEE TISQQEKDFG TIDVFVANAG VTWTQGPEID VDNYDSWNKI ISVDLNGVYY CSHNIGKIFK KNGKGSLIIT SSISGKIVNI PQLQAPYNTA KAACTHLAKS LAIEWAPFAR VNTISPGYID TDITDFASKD MKAKWWQLTP LGREGLTQEL VGGYLYLASN ASTFTTGSDV VIDGGYTCP

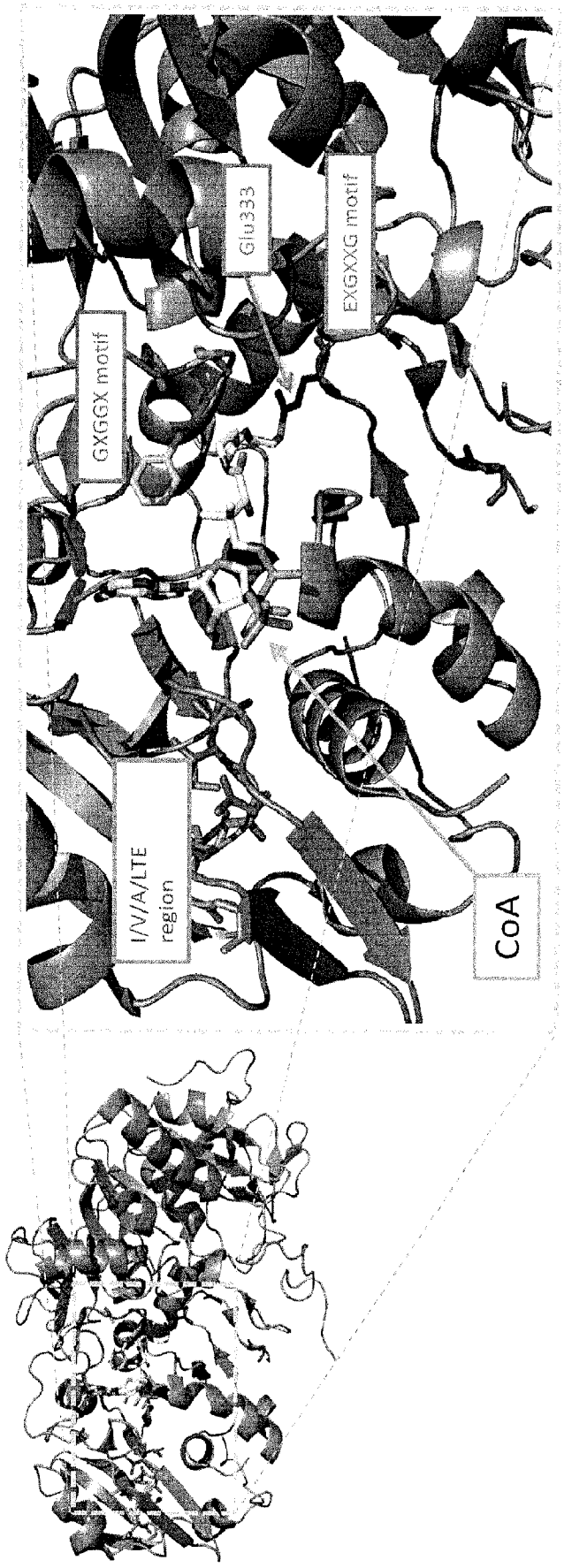
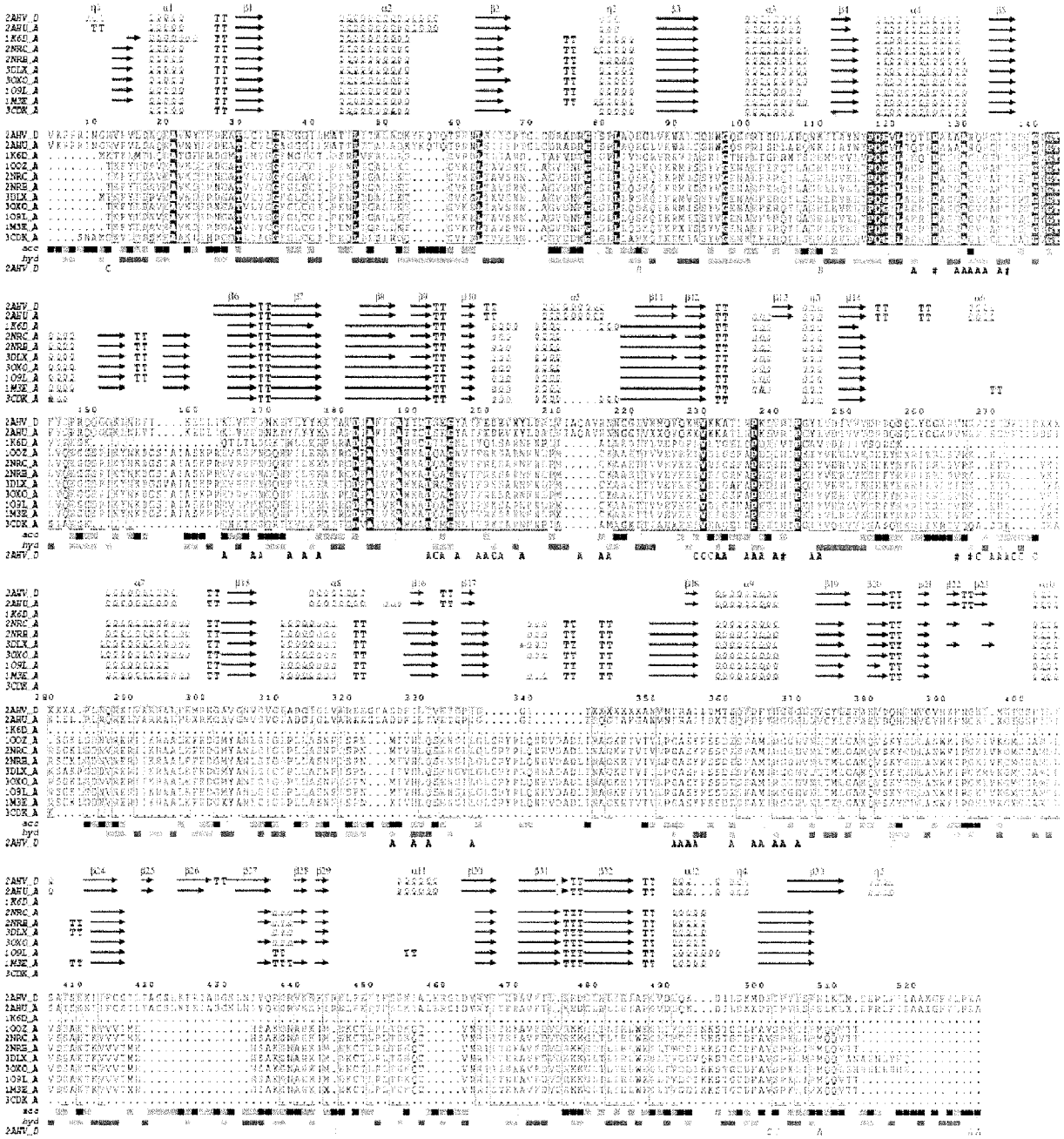


Figure 3

Figure 4



INTERNATIONAL SEARCH REPORT

International application No
PCT/US2016/061539

A. CLASSIFICATION OF SUBJECT MATTER
INV. C12P5/02 C12P7/40 C12P11/00 C12N9/10
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
C12P C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
EPO-Internal, BIOSIS, COMPENDEX, EMBASE, FSTA, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2015/017404 A1 (INVISTA NORTH AMERICA SARL [US]) 5 February 2015 (2015-02-05) cited in the application	76,77,97
A	abstract; figures 3, 4 paragraphs [0012] - [0022]	1-48, 58-70, 78,79, 98-112

A	WO 2014/047407 A1 (BIOAMBER INC [US]) 27 March 2014 (2014-03-27) page 2, paragraph 2; figures 1, 2 page 8, paragraph 3	1-7

A	WO 2015/031653 A2 (INVISTA NORTH AMERICA SARL [US]) 5 March 2015 (2015-03-05) abstract; claims 1, 5, 6; figure 1; sequences 4, 5 page 8, lines 5-12	12

☐ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
17 January 2017	20/03/2017

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Schröder, Gunnar

1

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2016/061539

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:

   a. [X] forming part of the international application as filed:

      [X] in the form of an Annex C/ST.25 text file.

      [ ] on paper or in the form of an image file.

   b. [ ] furnished together with the international application under PCT Rule 13*ter*.1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.

   c. [ ] furnished subsequent to the international filing date for the purposes of international search only:

      [ ] in the form of an Annex C/ST.25 text file (Rule 13*ter*.1(a)).

      [ ] on paper or in the form of an image file (Rule 13*ter*.1(b) and Administrative Instructions, Section 713).

2. [ ] In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that forming part of the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

3. Additional comments:

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2016/061539

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fees, this Authority did not invite payment of additional fees.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1-48, 76, 77, 98-112(completely); 58-70, 78, 79, 97(partially)

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. claims: 1-48, 76, 77, 98-112(completely); 58-70, 78, 79, 97(partially)

Method of producing 3-keto-acyl-CoA esters comprising the use of a polypeptide having both CoA transferase and beta-ketothiolase activities; compositions obtained thereby.

2-8. claims: 49-57, 71-75, 113-117(completely); 58-70, 78, 79, 97(partially)

Non-naturally occurring hosts according to claims 49-54 or 113 comprising at least one exogenous nucleic acid encoding a polypeptide having both CoA transferase and beta-ketothiolase activities

---

9-14. claims: 80-96(completely); 97(partially)

A variant polypeptide according to claim 80, with reference to sequences SEQ ID NO: 1, 5, 6, 7, 8 and 9, respectively

---

15. claims: 118-120

A polypeptide-substrate complex according to claims 118-120

---

16. claims: 121-123

A polypeptide having beta-ketothiolase activity according to claim 121

---

17. claims: 124-126

A polypeptide according to claim 124

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/US2016/061539

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2015017404 A1	05-02-2015	US 2016168610 A1	16-06-2016
		WO 2015017404 A1	05-02-2015
WO 2014047407 A1	27-03-2014	EP 2898082 A1	29-07-2015
		US 2015291987 A1	15-10-2015
		WO 2014047407 A1	27-03-2014
WO 2015031653 A2	05-03-2015	CN 105705650 A	22-06-2016
		EP 3039151 A2	06-07-2016
		US 2016201094 A1	14-07-2016
		US 2016237461 A1	18-08-2016
		WO 2015031653 A2	05-03-2015