



(51) International Patent Classification:

C12P 7/40 (2006.01) *C12N 9/10* (2006.01)
C12P 7/26 (2006.01)

(21) International Application Number:

PCT/US2014/048606

(22) International Filing Date:

29 July 2014 (29.07.2014)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

61/859,439 29 July 2013 (29.07.2013) US

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

(72) Inventors: **CONRADIE, Alex Van Eck**; 4 Aberdovey
Drive, Eaglescliffe Cleveland TS16 9EZ (GB). **BOTES,**
Adriana Leonora; Red House Farm, Daleside Road,
Rosedale East Cleveland Y018 8RL (GB).

(74) Agents: **MCCORMICK GRAHAM, Monica** et al.; Fish
& Richardson P.C., P.O. Box 1022, Minneapolis, Min-
nesota 55440-1022 (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, DK, DM,
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,
KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME,
MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ,
OM, 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.

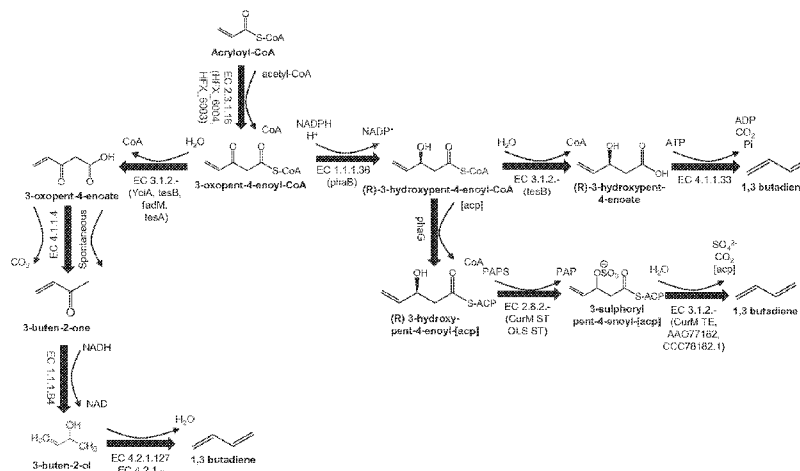
(84) Designated States (unless otherwise indicated, for every
kind of regional protection available): ARIPO (BW, GH,
GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ,
UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ,
TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK,
EE, ES, FI, FR, GB, GR, HR, HU, IE, 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))
- with sequence listing part of description (Rule 5.2(a))

(54) Title: METHODS FOR BIOSYNTHESIZING 3-OXOPENT-4-ENOYL-CoA

FIGURE 4



(57) Abstract: This document describes biochemical pathways that include the production of 3-oxopent-4-enoyl-CoA by condensation of acryloyl-CoA and acetyl-CoA using a β -ketothiolase with a SER-HIS-HIS catalytic triad. These pathways described herein rely on enzymes such as, *inter alia*, dehydrogenases, dehydratases and β -ketothiolases.

METHODS FOR BIOSYNTHESIZING 3-OXOPENT-4-ENOYL-COA

CROSS-REFERENCE TO PRIOR APPLICATION

[1] This application claims priority to U.S. Provisional Application No. 61/859,439 filed July 29, 2013, the entire contents of which is incorporated herein by reference in its entirety.

TECHNICAL FIELD

[2] This invention relates to methods for biosynthesizing 3-oxopent-4-enoyl-CoA, and more particularly to synthesizing 3-oxopent-4-enoyl-CoA as a bio-based precursor to 1,3-butadiene, using one or more isolated enzymes such as *dehydrogenases*, *dehydratases* and *β -ketothiolases*, or using recombinant host cells expressing one or more of such enzymes.

BACKGROUND

[3] 1,3-Butadiene (sometimes referred to herein as “butadiene”) is an important monomer for the production of synthetic rubbers including styrene-butadiene-rubber (SBR), polybutadiene (PB), styrene-butadiene latex (SBL), acrylonitrile-butadiene-styrene resins (ABS), nitrile rubber, and adiponitrile. Adiponitrile is used in the manufacture of Nylon-66 (White, *Chemico-Biological Interactions*, 2007, 166, 10-14).

[4] Butadiene is typically produced as a co-product from the steam cracking process, distilled to a crude butadiene stream, and purified via extractive distillation (White, *Chemico-Biological Interactions*, 2007, 166, 10-14).

[5] On-purpose butadiene has been prepared among other methods by dehydrogenation of n-butane and n-butene (Houdry process); and oxidative dehydrogenation of n-butene (Oxo-D or O-X-D process) (White, *Chemico-Biological Interactions*, 2007, 166, 10-14).

[6] Industrially, 95% of global butadiene production is undertaken via the steam cracking process using petrochemical-based feedstocks such as naphtha. Production

of on-purpose butadiene is not significant, given the high cost of production and low process yield (White, *Chemico-Biological Interactions*, 2007, 166, 10-14).

[7] Given reliance on petrochemical feedstocks and, for on-purpose butadiene, energy intensive catalytic steps; biotechnology offers an alternative approach via biocatalysis. Biocatalysis is the use of biological catalysts, such as enzymes, to perform biochemical transformations of organic compounds.

[8] Accordingly, against this background, it is clear that there is a need for sustainable methods for producing intermediates, in particular butadiene, wherein the methods are biocatalyst based (Jang et al., *Biotechnology & Bioengineering*, 2012, 109(10), 2437 – 2459).

[9] Both bioderived feedstocks and petrochemical feedstocks are viable starting materials for the biocatalysis processes.

[10] There are no known enzyme pathways leading to the synthesis of butadiene in prokaryotes or eukaryotes. Three potential pathways have been suggested for producing 1,3-butadiene from biomass-sugar: (1) from acetyl-CoA via crotonyl-CoA; (2) from erythrose-4-phosphate; and (3) via a condensation reaction with malonyl-CoA and acetyl-CoA. However, no performance information using these strategies has been reported (Jang *et al.*, *Biotechnology & Bioengineering*, 2012, 109(10), 2437 – 2459).

[11] An alternate pathway strategy using 3-oxopent-4-enoyl-CoA as a precursor to 1,3-butadiene is a plausible route to the synthesis of 1,3-butadiene via biocatalysis processes (See WO2013082542 A2).

SUMMARY

[12] Synthesis of 3-oxopent-4-enoyl-CoA via β -ketothiolase activity, such as the gene product of *bktB* classified under EC 2.3.1.16, would entail condensation of acryloyl-CoA and acetyl-CoA. However, acryloyl-CoA is an inactivator of β -ketothiolases reliant on a CYS-HIS-CYS catalytic triad, as is typical for polyhydroxybutyrate producing bacterial species (Palmer et al., *Journal of Biological Chemistry*, 1991, 266(13), 8369 – 8375; Modis & Wierenga, *Structure*, 1999, 7, 1279 – 1290).

[13] This document is based at least in part on the conclusion by the inventors that the SER-HIS-HIS catalytic triad associated with β -ketothiolase activity from *haloarchaea* (Hou et al., Appl. Environ. Microbiol., 2013, 79(14)), as opposed to the CYS-HIS-CYS catalytic triad of other β -ketothiolases, would not be inactivated by acryloyl-CoA and that it is possible to construct biochemical pathways producing 3-oxopent-4-enoyl-CoA as a precursor leading to the synthesis of butadiene. These pathways that are described herein rely on enzymes such as *dehydrogenases*, *dehydratases* and β -ketothiolases.

[14] Enzymes capable of condensing acryloyl-CoA and acetyl-CoA have not previously been described or suggested.

[15] Thus, in one aspect, this document provides enzymes that can produce 3-oxopent-4-enoyl-CoA, serving as a precursor in the synthesis of butadiene (See FIG. 4).

[16] In some embodiments, the condensation of acryloyl-CoA and acetyl-CoA to form 3-oxopent-4-enoyl-CoA is catalysed via β -ketothiolase activity not reliant on cysteine residues for catalysis such as the CYS-HIS-CYS catalytic triad associated with particularly polyhydroxybutyrate producing bacterial species.

[17] In some embodiments, the condensation of acryloyl-CoA and acetyl-CoA to form 3-oxopent-4-enoyl-CoA is catalysed via β -ketothiolase activity, such as encoded by HFX_6004 (*bktBa*) and HFX_6003 (*bktB β*), associated with *haloarchaea*. The gene product of the *bktBa* endows the β -ketothiolase with its substrate specificity.

[18] In some embodiments, the condensation of acryloyl-CoA and acetyl-CoA to form 3-oxopent-4-enoyl-CoA is catalysed via β -ketothiolase activity associated with a SER-HIS-HIS catalytic triad.

[19] In some embodiments, acryloyl-CoA can be produced from propanoyl-CoA by using a short-chain *acyl-CoA dehydrogenase* such as classified under EC 1.3.8.1 or EC 1.3.8.7 or a *2-methylacyl-CoA dehydrogenase* such as classified under EC 1.3.99.12 (See FIG. 3). Propanoyl-CoA can be produced via numerous pathways as described in FIG. 1 and FIG. 2.

[20] In some embodiments, acryloyl-CoA can be produced from 3-hydroxypropionoyl-CoA using a *3-hydroxypropionyl-CoA dehydratase* such as classified under EC 4.2.1.116 (See FIG. 2). The *3-hydroxypropionyl-CoA* can be produced from 3-hydroxypropionate using a *CoA-ligase* such as classified under EC 6.2.1.-. The 3-hydroxypropionate can be produced from malonate semialdehyde using a *3-hydroxypropionate dehydrogenase* such as classified under EC 1.1.1.59 (See FIG. 3). The malonate semialdehyde can be produced from the central metabolite malonyl-CoA using a *malonyl-CoA reductase* such as classified under EC 1.2.1.75.

[21] In some embodiments, acryloyl-CoA can be produced from lactoyl-CoA using a *lactoyl-CoA dehydratase* such as classified under EC 4.2.1.54 (See FIG. 3). The lactoyl-CoA can be produced from lactate using a *propionate CoA-transferase* such as classified under EC 2.8.3.1. The lactate can be produced from pyruvate using a *lactate dehydrogenase* such as classified under EC 1.1.1.27.

[22] In any of the methods described herein, the method can be performed using isolated enzymes, using cell lysates comprising the enzymes, or using a recombinant host.

[23] The recombinant host can be anaerobically, micro-aerobically or aerobically cultivated.

[24] Recombinant host cells can be retained in ceramic hollow fiber membranes to maintain a high cell density during fermentation.

[25] The principal carbon source fed to the fermentation can derive from biological or non-biological feedstocks. For example, the biological feedstock can be, or can be derived from, monosaccharides, disaccharides, lignocellulose, hemicellulose, cellulose, levulinic acid, furfural, lignin, triglycerides such as glycerol and fatty acids, agricultural waste or municipal waste. The non-biological feedstock can be, or can be derived from, either natural gas, syngas, CO₂/H₂, methanol, ethanol, non-volatile residue (NVR) or caustic wash waste stream from cyclohexane oxidation processes.

[26] The host microorganism can be a prokaryote 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 oleovorans*; from the genus *Delftia* such as *Delftia acidovorans*; from the genus *Bacillus* such as *Bacillus subtilis*; from the genus *Lactobacillus* such as *Lactobacillus delbrueckii*; or from the genus *Lactococcus* such as *Lactococcus lactis*.

[27] The host microorganism can be a eukaryote from the genus *Aspergillus* such as *Aspergillus niger*; from the genus *Saccharomyces* such as *Saccharomyces cerevisiae*; from the genus *Pichia* such as *Pichia pastoris*; from the genus *Yarrowia* such as *Yarrowia lipolytica*; from the genus *Issatchenkia* such as *Issatchenkia 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*.

[28] In any of the recombinant hosts described herein, the enzymes catalyzing the hydrolysis of propionyl-CoA and acetyl-CoA can be attenuated; the enzymes consuming propanoyl-CoA via the methyl-citrate cycle can be attenuated; the enzymes consuming propanoyl-CoA to pyruvate can be attenuated; the enzymes consuming propanoyl-CoA to malonyl-CoA can be attenuated; a feedback-resistant *threonine deaminase* can be genetically engineered into the host organism; the β -*ketothiolases* catalyzing the condensation of acetyl-CoA to acetoacetyl-CoA such as the gene products of *AtoB* or *phaA* can be attenuated; the *polymer synthase* enzymes in a host strain that naturally accumulates polyhydroxyalkanoates can be attenuated; a gene encoding a *phosphotransacetylase*, such as *pta*, can be attenuated; a gene encoding an *acetate kinase* degrading propanoate, such as *ack*, can be attenuated; a gene encoding the degradation of pyruvate to lactate can be attenuated; a gene encoding the degradation of phosphoenolpyruvate to succinate such as *frdBC* can be attenuated; a gene encoding the degradation of acetyl-CoA to ethanol such as *adhE* can be attenuated; the enzymes catalyzing anaplerotic reactions supplementing the citric acid cycle intermediates can be amplified; a *puridine nucleotide transhydrogenase* gene such as *UdhA* can be overexpressed; a *glyceraldehyde-3P-dehydrogenase* gene such as *GapN* can be overexpressed in the host organisms; a *malic enzyme* gene such as *maeA* or *maeB* can be overexpressed in the host organism; a *glucose-6-phosphate dehydrogenase* gene such as *zwf* can be overexpressed in the

host organism; a *fructose 1,6 diphosphatase* gene such as *fbp* can be overexpressed in the host organism.

[29] More specifically, the present document provides a method of 3-oxopent-4-enoyl-CoA synthesis; the method can include the condensation of acryloyl-CoA and acetyl-CoA using a protein with β -ketothiolase activity. At least one subunit of the protein can be a polypeptide, or can be derived from a polypeptide, encoded by an archaeon gene, e.g., a haloarchaeon gene. The protein can contain a SER-HIS-HIS catalytic triad associated with its catalytic mechanism.

[30] In addition, the protein can be a heterodimer in which one or both of the two polypeptides of the heterodimer are: full-length gene products of the *Haloferax mediterranei* HFX_6004 (*bktBa*) and HFX_6003 (*bktB β*) genes; biologically active fragments of the gene products; or biologically active variants of the full-length gene products or the biologically active fragments. One or both of the polypeptides can be an engineered polypeptide having greater than 50% homology to the corresponding polypeptide chain of the BktB β -ketothiolase of *Haloferax mediterranei*.

[31] In any of the methods, the acryloyl-CoA can be produced using a *butyryl-CoA dehydrogenase* (EC 1.3.8.1), a *medium-chain acyl-CoA dehydrogenase* (EC 1.3.8.7), a *2-methylacyl-CoA dehydrogenase* (EC 1.3.99.12), a *lactoyl-CoA dehydratase* (EC 4.2.1.54), or a *3-hydroxypropionyl-CoA dehydratase* (EC 4.2.1.116).

[32] Any of the above methods can include a fermentation process using a host cell expressing the protein. The host cell can be a prokaryote or a eukaryote. The prokaryote can be a bacterium of the genus *Escherichia* such as *Escherichia coli*; of the genus *Clostridia* such as *Clostridium ljungdahlii*, *Clostridium autoethanogenum* or *Clostridium kluyveri*; of the genus *Corynebacteria* such as *Corynebacterium glutamicum*; of the genus *Cupriavidus* such as *Cupriavidus necator* or *Cupriavidus metallidurans*; of the genus *Pseudomonas* such as *Pseudomonas fluorescens* or *Pseudomonas putida*; of the genus *Bacillus* such as *Bacillus subtilis*; or of the genus *Rhodococcus* such as *Rhodococcus equi*. Alternatively, the prokaryote can be an archaeon of the genus *Methanocaldococcus* such as *Methanocaldococcus jannaschii*; of the genus *Haloferax* such as *Haloferax mediterranei*; of the genus *Metallosphaera* such as *Metallosphaera sedula*; of the genus *Methanobacterium* such as *Methanobacterium thermoautotrophicum*; or of the genus *Pyrococcus* such as

Pyrococcus abyssi. The eukaryote can be a fungus of the genus *Aspergillus* such as *Aspergillus niger*; of the genus *Saccharomyces* such as *Saccharomyces cerevisiae*; of the genus *Pichia* such as *Pichia pastoris*; of the genus *Yarrowia* such as *Yarrowia lipolytica*; of the genus *Issatchenkia* such as *Issatchenkia orientalis*; of the genus *Debaryomyces* such as *Debaryomyces hansenii*; of the genus *Arxula* such as *Arxula adeninivorans*; or of the genus *Kluyveromyces* such as *Kluyveromyces lactis*. The fermentation process of any of the above methods can include anaerobic, micro-aerobic or aerobic cell cultivation. Moreover, cell retention strategies using, for example, ceramic hollow fiber membranes can be employed to achieve and maintain a high cell density during fermentation. Furthermore, the principal carbon source fed to the fermentation can derive from biological or non-biological feedstocks. The biological feedstock can be, or can be derived from, monosaccharides, disaccharides, hemicellulose such as levulinic acid and furfural, cellulose, lignocellulose, lignin, triglycerides such as glycerol and fatty acids, agricultural waste or municipal waste. In addition, the non-biological feedstock can be, or can be derived from, natural gas, syngas, CO₂/H₂, methanol, ethanol, non-volatile residue (NVR), caustic wash from a cyclohexane oxidation processes, or other waste stream from the chemical or petrochemical industries.

[33] 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. 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.

[34] The details of one or more embodiments of the invention are set forth in the accompanying drawings and the description below. Other features, objects, and advantages of the invention 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

[35] FIG. 1 is a schematic of biochemical pathways leading to the production of propanoyl-CoA from central metabolites.

[36] FIG. 2 is a schematic of biochemical pathways leading to the production of propanoyl-CoA from central metabolites.

[37] FIG. 3 is a schematic of biochemical pathways leading to the production of acryloyl-CoA.

[38] FIG. 4 is a schematic of biochemical pathways leading to the production of 1,3-butadiene using 3-oxopent-4-enoyl-CoA as central precursor.

[39] FIG. 5 is a depiction of the amino acid sequence of the *Haloferax mediterranei* β -ketothiolase α -chain (BktB α) (Genbank Accession No. AFK21132.1, SEQ ID NO: 1). The amino acids of the SER-HIS-HIS catalytic triad are shown in bold and underlined.

[40] FIG. 6 is a depiction of the amino acid sequence of the *Haloferax mediterranei* β -ketothiolase β -chain (BktB β) (Genbank accession No. AFK21131.1, SEQ ID NO: 2).

[41] Images with KEGG Entry ID numbers in FIG. 1 to FIG. 4 originate from the world wide web at genome.jp/kegg (Kanehisa and Goto, *Nucleic Acids Res.*, 2012, 40, D109 – D114; Kanehisa and Goto, *Nucleic Acids Res.*, 2000, 28, 27-30).

DETAILED DESCRIPTION

[42] In particular, this document provides enzymes, non-natural pathways, cultivation strategies, feedstocks, host microorganisms and attenuations to the host's biochemical network, which allow the production of 3-oxopent-4-enoyl-CoA as central precursor leading to the synthesis of 1,3 butadiene (sometimes referred to as “butadiene” herein) from central metabolites. As used herein, the term “central precursor” is used to denote any metabolite in any metabolic pathway shown herein leading to the synthesis of 3-oxopent-4-enoyl-CoA and/or butadiene. The term

"central metabolite" is used herein to denote a metabolite that is produced in all microorganisms to support growth.

[43] As such, host microorganisms described herein can include endogenous pathways that can be manipulated such that 3-oxopent-4-enoyl-CoA can be produced. In an endogenous pathway, the host microorganism naturally expresses all of the enzymes catalyzing the reactions within the pathway. A host microorganism containing an engineered pathway does not naturally express all of the enzymes catalyzing the reactions within the pathway but has been engineered such that all of the enzymes within the pathway are expressed in the host. Within an engineered pathway, the enzymes can be from a single source, i.e., from one species, or can be from multiple sources, i.e., different species. 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.

[44] Where the enzyme is one having the β -ketothiolase activity described herein, the enzyme will most commonly be one that is, or is derived from an archaeon, e.g., the haloarchaeon *Haloferax mediterranei*. The active form of this enzyme is a heterodimer composed of an α -chain and a β -chain, referred to as BktB α and BktB β , respectively. The genes encoding these two polypeptides in *H. mediterranei* are referred to as HFX_6004 (or *bktBa*) and HFX_6003 (or *bktB β*), respectively. Contrary to other β -ketothiolases, which are homotetramers and contain a CYS-HIS-CYS catalytic triad, the *H. mediterranei* enzyme contains a SER-HIS-HIS catalytic triad (Hou *et al.*, *Appl. Environ. Microbiol.*, 2013, 79(14)). It is noted that the amino acids of neither the CYS-HIS-CYS nor the SER-HIS-HIS catalytic triads are consecutively located on the relevant polypeptide subunit but are separated by different numbers of amino acids. FIG. 5 shows the amino acid sequence of the *H. mediterranei* BktB α subunit (SEQ ID NO: 1) containing the SER-HIS-HIS triad and the relevant SER and HIS residues of the catalytic triad are shown in bold and underlined; the three residues are located at amino acid positions 83, 281, and 331, respectively. The catalytic function of the three amino acids was predicted by bioinformatics analyses (amino acid sequence alignments (BLAST (see below for more details) and the GenDoc programs) and prediction by the NCBI conserved domain search (CDD) service) and confirmed by genetic and functional studies (Hou

et al., *Appl. Environ. Microbiol.*, 2013, 79(14)). While the BktBa subunit endows the substrate specificity of the enzyme, the BktB β subunit (amino acid sequence of the *H. mediterranei* BktB β subunit shown in FIG. 6, SEQ ID NO:2) appears also to be essential for its activity and it seems that an OB fold (oligonucleotide/oligosaccharide-binding fold) in the BktB β subunit (at amino acid residues 32-94 in the *H. mediterranei* BktB β subunit) is required for the activity (Hou *et al.*, *Appl. Environ. Microbiol.*, 2013, 79(14)). Bioinformatic analyses (such as those mentioned above) together with genetic and functional studies well known in the art can be used to identify the amino acid residues and regions of the SER-HIS-HIS catalytic triad (e.g., the region corresponding to amino acids 83-331 of the *H. mediterranei* BktBa subunit) and the OB fold (e.g., the region corresponding to amino acids 32-94 of the *H. mediterranei* BktB β subunit) in the other β -ketothiolase polypeptides.

[45] Functional fragments of any of the enzymes, or enzyme subunits, 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. It will be understood that, in the case of BktBa chains, such fragments should preferably be, or contain, the region containing the SER-HIS-HIS catalytic triad (such as amino acids 83-331 of the *H. mediterranei* BktBa chain). Moreover, in the case of BktB β chains, such fragments should preferably be, or contain, the region containing the OB fold (such as amino acids 32-94 of the *H. mediterranei* BktB β subunit).

[46] 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 non-conservative substitution is a substitution of one amino acid for another with dissimilar characteristics.

[47] 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), hemagglutinin (HA), glutathione-S-transferase (GST), or maltose binding 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.

[48] All the enzymes, or enzyme subunits, described herein that can be used for 3-oxopent-4-enoyl-CoA and/or butadiene production can have at least 50% sequence identity (homology) (e.g., at least 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, or 100%) to the amino acid sequence of the corresponding wild-type enzyme or enzyme subunit. 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 (BI2seq) program from the stand-alone version of BLASTZ containing BLASTP version 2.0.14. This stand-alone version of BLASTZ can be obtained from Fish & Richardson's web site (e.g., www.fr.com/blast/) or the U.S. government's National Center for Biotechnology Information web site (www.ncbi.nlm.nih.gov). Instructions explaining how to use the BI2seq program can be found in the readme file accompanying BLASTZ. BI2seq performs a comparison between two amino acid sequences using the BLASTP algorithm. To compare two amino acid sequences, the options of BI2seq 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:\BI2seq -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 used for nucleic acid sequences except that blastn is used.

[49] 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.

[50] It is understood that substitutions and deletions will not be of the three amino acids of the SER-HIS-HIS catalytic triad of the BtkB α subunit. Moreover, substitutions, deletions, and additions will preferably be outside the region of the BtkB α chain containing the catalytic triad (see above) and outside the region of BtkB β subunit polypeptide containing the OB fold (see above).

[51] 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.

[52] 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. 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. Thus, as described herein recombinant hosts can include nucleic acids encoding one or more of a *dehydrogenase*, a *dehydratase* or a β -*ketothiolase* as described in more detail below.

[53] In addition, the production of 3-oxopent-4-enoyl-CoA can be performed *in vitro* using 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.

Enzymes and pathways generating acryloyl-CoA (e.g., to be used for the biosynthesis of 3-oxopent-4-enoyl-CoA)

[54] As depicted in FIG. 3, acryloyl-CoA can be produced from 3-hydroxypropionyl-CoA, propanoyl-CoA or lactoyl-CoA.

[55] In some embodiments, acryloyl-CoA can be produced from: (a) propanoyl-CoA by a *butyryl-CoA dehydrogenase* (EC 1.3.8.1), a *medium-chain acyl-CoA dehydrogenase* (EC 1.3.8.7), or a *2-methylacyl-CoA dehydrogenase* (EC 1.3.99.12); (b) lactoyl-CoA by a *lactoyl-CoA dehydratase* (EC 4.2.1.54); or (c) 3-hydroxypropionyl-CoA by a *3-hydroxypropionyl-CoA dehydratase* (EC 4.2.1.116) (See, FIG. 3).

[56] Propanoyl-CoA can be produced via several pathways as described in FIG. 1 and FIG. 2.

[57] In some embodiments, *3-hydroxypropionyl-CoA* can be produced from 3-hydroxypropionate using a *CoA-ligase* such as that classified under EC 6.2.1.- (FIG. 3). The 3-hydroxypropionate can be produced from malonate semialdehyde using a *3-hydroxypropionate dehydrogenase* such as that classified under EC 1.1.1.59 (FIG. 3). The malonate semialdehyde can be produced from the central metabolite malonyl-CoA using a *malonyl-CoA reductase* such as that classified under EC 1.2.1.75 (FIG. 3).

[58] In some embodiments, lactoyl-CoA can be produced from lactate using a *propionate CoA-transferase* such as classified under EC 2.8.3.1 (FIG. 3). The lactate can be produced from pyruvate using a *lactate dehydrogenase* such as classified under EC 1.1.1.27 (FIG. 3).

Enzymes generating the 3-oxopent-4-enoyl-CoA

[59] In some embodiments, the condensation of acryloyl-CoA and acetyl-CoA to form 3-oxopent-4-enoyl-CoA is catalysed via β -ketothiolase activity not reliant on cysteine residues for catalysis such as the CYS-HIS-CYS catalytic triad associated with, for example, polyhydroxybutyrate producing bacterial species.

[60] In some embodiments, the condensation of acryloyl-CoA and acetyl-CoA to form 3-oxopent-4-enoyl-CoA is catalysed via β -ketothiolase activity, such as that encoded by HFX_6004 (*bktBa*) and HFX_6003 (*bktB β*), associated with *haloarchaea*.

[61] In some embodiments, the condensation of acryloyl-CoA and acetyl-CoA to form 3-oxopent-4-enoyl-CoA is catalysed via β -ketothiolase activity associated with a SER-HIS-HIS catalytic triad.

Enzymes generating 1,3-butadiene from 3-oxopent-4-enoyl-CoA

[62] In some embodiments, 1,3-butadiene can be synthesized by, for example, the enzymes and metabolic pathways shown in FIG. 4.

Cultivation strategy

[63] In some embodiments, 3-oxopent-4-enoyl-CoA is biosynthesized in a recombinant host using a fermentation strategy that can include anaerobic, micro-aerobic or aerobic cultivation of the recombinant host.

[64] In some embodiments, a cell retention strategy using, for example, ceramic hollow fiber membranes is employed to achieve and maintain a high cell density during either fed-batch or continuous fermentation in the synthesis of 3-oxopent-4-enoyl-CoA as central precursor to butadiene.

[65] In some embodiments, the principal carbon source fed to the fermentation in the synthesis of 3-oxopent-4-enoyl-CoA derives from biological or non-biological feedstocks.

[66] In some embodiments, the biological feedstock is, includes, or derives from, monosaccharides, disaccharides, lignocellulose, hemicellulose, cellulose, lignin such as levulinic acid and furfural, lignin, triglycerides such as glycerol and fatty acids, agricultural waste or municipal waste.

[67] 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 oleovorans*, *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).

[68] 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, *Journal of Biotechnology*, 2011, 155, 293 – 298; Martin and Prather, *Journal of Biotechnology*, 2009, 139, 61 – 67).

[69] 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).

[70] 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).

[71] 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* and *Lactobacillus delbrueckii* and *Lactococcus lactis* (see, e.g., Hermann *et al.*, *Journal of Biotechnology*, 2003, 104, 155 – 172; Wee *et al.*, *Food Technol. Biotechnol.*, 2006, 44(2), 163 – 172; Ohashi *et al.*, *Journal of Bioscience and Bioengineering*, 1999, 87(5), 647 - 654).

[72] 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).

[73] In some embodiments, the non-biological feedstock can be, or can be derived from, natural gas, syngas, CO₂/H₂, methanol, ethanol, non-volatile residue (NVR) or a caustic wash waste stream from cyclohexane oxidation processes.

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

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

[76] The efficient catabolism of CO₂ and H₂, 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).

[77] 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).

[78] 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).

[79] In some embodiments, the host microorganism is a prokaryote. For example, the prokaryote can be 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 oleovorans*; from the genus *Delftia* such as *Delftia acidovorans*; from the genus *Bacillus* such as *Bacillus subtilis*; from the genus *Lactobacillus* such as *Lactobacillus delbrueckii*; or from the genus *Lactococcus* such as *Lactococcus lactis*; from the genus *Synechococcus* such as *Synechococcus elongatus*. Such prokaryotes also can be sources of genes to construct recombinant host cells described herein that are capable of producing butadiene.

[80] In some embodiments, the prokaryotic host microorganism is an *archaeon*. For example, the *archaeon* can be from the genus *Methanocaldococcus* such as *Methanocaldococcus jannaschii*; from the genus *Haloferax* such as *Haloferax mediterranei*; from the genus *Metallosphaera* such as *Metallosphaera sedula*; from the genus *Methanobacterium* such as *Methanobacterium thermoautotrophicum*; or from the genus *Pyrococcus* such as *Pyrococcus abyssi*.

[81] In some embodiments, the host microorganism is a eukaryote. Eukaryotes can be, for example, fungi (e.g., filamentous fungi or yeasts). For example, the eukaryote can be from the genus *Aspergillus* such as *Aspergillus niger*; 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 *Issatchenkia orientalis*; from the genus *Debaryomyces* such as *Debaryomyces hansenii*; from the genus *Arxula* such as *Arxula adeninivorans*; 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 butadiene.

Metabolic engineering

[82] The present document provides methods involving less than all the steps described for all the above pathways. Such methods can involve, for example, one, two, three, four, five, six, seven, eight, nine, ten, or more of such steps. Where less than all the steps are included in such a method, the first step can be any one of the steps listed. Again, all these methods can be *in vitro* (see above) or in recombinant host cells.

[83] 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.

[84] 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.

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

[86] 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.

[87] In some embodiments, the enzymes in the pathways outlined in section 4.1 and section 4.2 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.

[88] In some embodiments, the enzymes in the pathways outlined in section 4.1 and section 4.2 are gene dosed, i.e., overexpressed, into the resulting genetically modified organism via episomal or chromosomal integration approaches.

[89] In some embodiments, genome-scale system biology techniques such as Flux Balance Analysis are utilized to devise genome scale attenuation or knockout strategies for directing carbon flux to butadiene.

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

[91] In some embodiments, fluxomic, metabolomic and transcriptomal data are utilized to inform or support genome-scale system biology techniques, thereby devising genome scale attenuation or knockout strategies in directing carbon flux to butadiene.

[92] In some embodiments requiring intracellular availability of propanoyl-CoA or acryloyl-CoA for 3-oxopent-4-enoyl-CoA synthesis, genes (e.g., endogenous genes) encoding enzymes catalyzing the hydrolysis of propionyl-CoA and acetyl-CoA can be attenuated in the host organism.

[93] In some embodiments requiring the intracellular availability of propanoyl-CoA or acryloyl-CoA for 3-oxopent-4-enoyl-CoA synthesis, genes (e.g., endogenous genes) encoding enzymes consuming propanoyl-CoA via the methyl-citrate cycle can be attenuated in the host organism (Upton and McKinney, *Microbiology*, 2007, 153, 3973 - 3982).

[94] In some embodiments requiring the intracellular availability of propanoyl-CoA or acryloyl-CoA for 3-oxopent-4-enoyl-CoA synthesis, genes (e.g., endogenous genes) encoding enzymes consuming propanoyl-CoA to pyruvate can be attenuated in the host organism.

[95] In some embodiments requiring the intracellular availability of propanoyl-CoA or acryloyl-CoA for 3-oxopent-4-enoyl-CoA synthesis, genes (e.g., endogenous genes) encoding enzymes consuming propanoyl-CoA to malonyl-CoA can be attenuated in the host organism.

[96] In some embodiments requiring the intracellular availability of propanoyl-CoA or acryloyl-CoA via L-threonine as central metabolite for 3-oxopent-4-enoyl-CoA synthesis, a feedback-resistant *threonine deaminase* is genetically engineered into the host organism (Tseng *et al.*, *Microbial Cell Factories*, 2010, 9:96).

[97] In some embodiments requiring condensation of acetyl-CoA and acryloyl-CoA for 3-oxopent-4-enoyl-CoA synthesis, the genes (e.g., endogenous genes) encoding β -ketothiolases catalyzing the condensation of acetyl-CoA to acetoacetyl-CoA (such as the *AtoB* or *phaA* genes) can be attenuated.

[98] In some embodiments using hosts that naturally accumulate polyhydroxyalkanoates, the genes (e.g., endogenous genes) encoding *polymer synthase* enzymes can be attenuated in the host strain.

[99] In some embodiments requiring the intracellular availability of acetyl-CoA for 3-oxopent-4-enoyl-CoA synthesis, a host that is deficient (e.g., attenuated level of activity) in one or more enzymes in the acetate synthesis pathway can be used. For example, a host that is deficient in a *phosphotransacetylase* (encoded by the *pta* gene) can be used (Shen *et al.*, *Appl. Environ. Microbio.*, 2011, 77(9), 2905 – 2915).

[100] In some embodiments requiring the intracellular availability of acetyl-CoA for 3-oxopent-4-enoyl-CoA synthesis, a gene (e.g., an endogenous gene) in an acetate synthesis pathway encoding an acetate kinase, such as *ack*, can be attenuated.

[101] In some embodiments requiring the intracellular availability of acetyl-CoA for 3-oxopent-4-enoyl-CoA synthesis, a gene (e.g., an endogenous gene) encoding an enzyme catalyzing the degradation of pyruvate to lactate, such as *ldhA*, can be attenuated (Shen *et al.*, *Appl. Environ. Microbio.*, 2011, 77(9), 2905 - 2915).

[102] In some embodiments requiring the intracellular availability of acetyl-CoA for 3-oxopent-4-enoyl-CoA synthesis, a gene (an endogenous gene) encoding an enzyme catalyzing the degradation of phosphoenolpyruvate to succinate, such as *frdBC*, can be attenuated (see, e.g., Shen *et al.*, 2011, *supra*).

[103] In some embodiments requiring the intracellular availability of acetyl-CoA for 3-oxopent-4-enoyl-CoA synthesis, a gene (e.g., an endogenous gene) encoding an enzyme catalyzing the degradation of acetyl-CoA to ethanol, such as *adhE*, can be attenuated (Shen *et al.*, 2011, *supra*).

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

[105] In some embodiments, where pathways require excess NADPH co-factor in the synthesis of 3-oxopent-4-enoyl-CoA, a glyceraldehyde-3P-dehydrogenase gene such as *GapN* can be overexpressed in the host organisms (Brigham *et al.*, 2012, *supra*).

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

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

[108] In some embodiments, where pathways require excess NADPH co-factor in the synthesis of 3-oxopent-4-enoyl-CoA, a fructose 1,6 diphosphatase gene such as *fbp* can be overexpressed in the host organisms (Becker *et al.*, *Journal of Biotechnology*, 2007, 132, 99 - 109).

OTHER EMBODIMENTS

[109] 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 3-oxopent-4-enoyl-CoA synthesis, the method comprising condensing acryloyl-CoA and acetyl-CoA using a protein with β -ketothiolase activity.
2. The method of claim 1, wherein at least one subunit of the protein is a polypeptide, or is derived from a polypeptide, encoded by an archaeon gene.
3. The method of claim 2, wherein the archaeon gene is a haloarchaeon gene.
4. The method according to any one of claims 1-3, wherein the protein comprises a SER-HIS-HIS catalytic triad associated with its catalytic mechanism.
5. The method according to any one of claims 1-4, wherein the protein is a heterodimer and wherein one or both of the two polypeptides of the heterodimer are:
full-length gene products of the *Haloferax mediterranei* HFX_6004 (*bktBa*) and HFX_6003 (*bktB β*) genes;
biologically active fragments of the gene products; or
biologically active variants of the full-length gene products or the biologically active fragments.
6. The method according to claim 5, wherein one or both of the polypeptides is an engineered polypeptide having greater than 50% homology to the corresponding polypeptide chain of the BktB β -ketothiolase of *Haloferax mediterranei*.
7. The method according to any one of claims 1-6, wherein the acryloyl-CoA is produced using a *butyryl-CoA dehydrogenase* (EC 1.3.8.1), a *medium-chain acyl-CoA dehydrogenase* (EC 1.3.8.7), a *2-methylacyl-CoA dehydrogenase* (EC 1.3.99.12), *lactoyl-CoA dehydratase* (EC 4.2.1.54) or a *3-hydroxypropionyl-CoA dehydratase* (EC 4.2.1.116).
8. The method according to any one of claims 1 - 7, wherein the method comprises a fermentation process using a host cell expressing the protein.
9. The method according to claim 8, wherein the host cell is a prokaryote or a eukaryote.
10. The method according to claim 9, wherein the prokaryote is a bacterium of the genus *Escherichia* such as *Escherichia coli*; of the genus *Clostridia* such as *Clostridium ljungdahlii*, *Clostridium autoethanogenum* or *Clostridium kluyveri*; of the genus *Corynebacteria* such as *Corynebacterium glutamicum*; of the genus *Cupriavidus* such as *Cupriavidus necator* or *Cupriavidus metallidurans*; of the genus

Pseudomonas such as *Pseudomonas fluorescens* or *Pseudomonas putida*; of the genus *Bacillus* such as *Bacillus subtilis*; or of the genus *Rhodococcus* such as *Rhodococcus equi*.

11. The method according to claim 9, wherein the eukaryote is a fungus of the genus *Aspergillus* such as *Aspergillus niger*; of the genus *Saccharomyces* such as *Saccharomyces cerevisiae*; of the genus *Pichia* such as *Pichia pastoris*; of the genus *Yarrowia* such as *Yarrowia lipolytica*; of the genus *Issatchenkia* such as *Issatchenkia orientalis*; of the genus *Debaryomyces* such as *Debaryomyces hansenii*; of the genus *Arxula* such as *Arxula adeninivorans*; or of the genus *Kluyveromyces* such as *Kluyveromyces lactis*.

12. The method according to claim 9, wherein the prokaryote is an archaeon of the genus *Methanocaldococcus* such as *Methanocaldococcus jannaschii*; of the genus *Haloferax* such as *Haloferax mediterranei*; of the genus *Metallosphaera* such as *Metallosphaera sedula*; of the genus *Methanobacterium* such as *Methanobacterium thermoautotrophicum*; or of the genus *Pyrococcus* such as *Pyrococcus abyssi*.

13. The method according to any one of claims 8 - 12, wherein the fermentation process comprises anaerobic, micro-aerobic or aerobic cell cultivation.

14. The method according to any one of claims 8 - 13, wherein cell retention strategies using, for example, ceramic hollow fibre membranes are employed to achieve and maintain a high cell density during fermentation.

15. The method according to any one of claims 8 - 14, wherein the principal carbon source fed to the fermentation derives from biological or non-biological feedstocks.

16. The method according to claim 15, where the biological feedstock is, or derives from, monosaccharides, disaccharides, hemicellulose such as levulinic acid and furfural, cellulose, lignocellulose, lignin, triglycerides such as glycerol and fatty acids, agricultural waste or municipal waste.

17. The method according to claim 15, where the non-biological feedstock is, or derives from, natural gas, syngas, CO₂/H₂, methanol, ethanol, non-volatile residue (NVR), caustic wash from a cyclohexane oxidation processes, or other waste stream from the chemical or petrochemical industries.

FIGURE 3

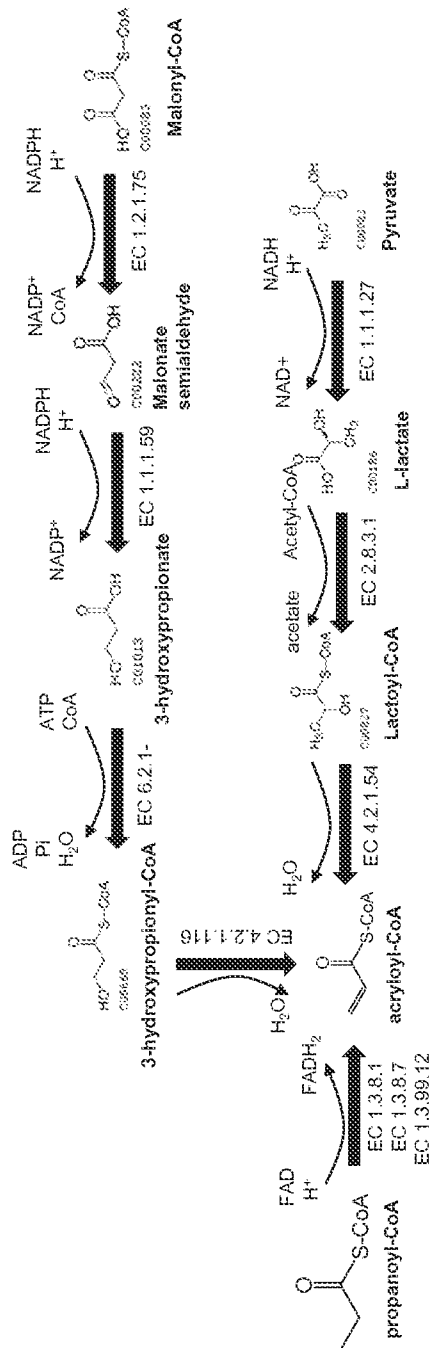


FIGURE 4

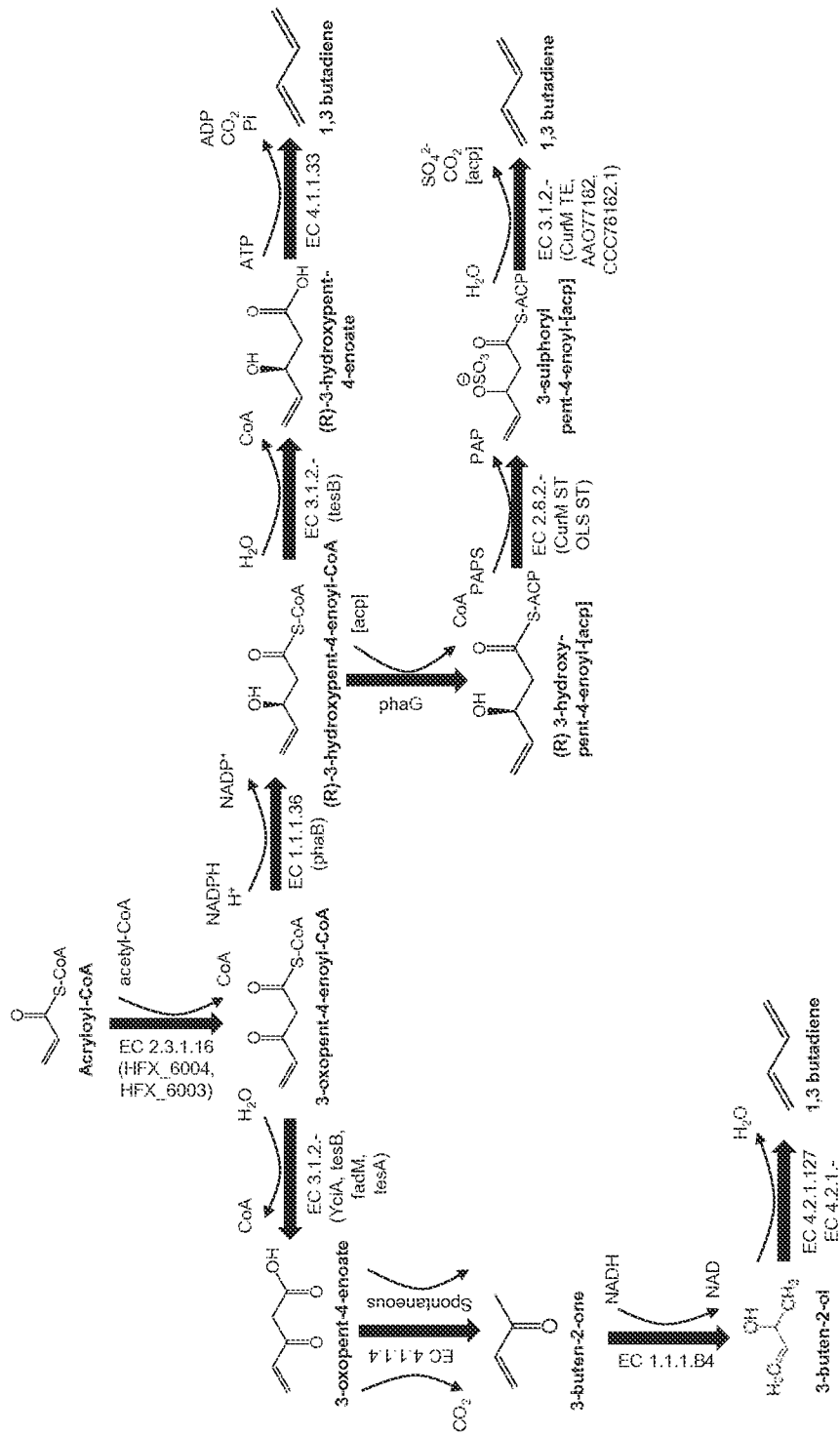


FIGURE 5

10 MEVAVIGSSM TKFGQRSAMI RELISEAGQA CLEDAGVAPA SVDHLYVSNM ASGEFEGQTG 60
 70 VMNALAHDLG VIPAYTQRID QTSSSGGAGI YEAWQSIASG VSEMTLIVGG EKMTHTTGE 120
 130 STDI IASCTH PEEYKHGVTI PSFAGMTARN YLERFDAPRE SLARVAVKNH RGVNDNPKAQ 180
 190 FQKEIDIETA LESPIIADPL RLYDFCPITD GSAAMFETTE ERAQEITDEY AIVSGVGGAT 240
 250 DTHVVHERDD PTVMGGVVES SKQAYEMAGV GPDDLDAEL HDMFTILEFL QLEGIGVADH 300
 310 GAAWELAMDG VTAKDGGGLPI NTSGGLKSKG HPLIGASGVAQ GVEIYEQLVG EAGPRQVEAD 360
 370 TALACNVGGF GNCVITTIME AAK 380

FIGURE 6

10 20 30 40 50 60
MAIVEFEVDG QAVRALGQVT TDDIETGDVV EPVYVEELRD PEVGIKAPES QSWDGYRWDP V

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2014/048606

A. CLASSIFICATION OF SUBJECT MATTER
INV. C12P7/40 C12P7/26 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, WPI Data, BIOSIS, Sequence Search, EMBASE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2013/082542 A2 (INVISTA NORTH AMERICA SARL [US]) 6 June 2013 (2013-06-06)	1-4,7-17
Y	paragraphs [0068], [0097] -----	5,6
Y	DATABASE UniProt [Online] 5 September 2012 (2012-09-05), "SubName: Full=3-ketoacyl-CoA thiolase; SubName: Full=PHA-specific beta-ketothiolase alpha subunit; EC=2.3.1.9;" XP002730873, retrieved from EBI accession no. UNIPROT:I3RA72 Database accession no. I3RA72 the whole document ----- -/-	5,6



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

9 October 2014

Date of mailing of the international search report

31/10/2014

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Schneider, Patrick

INTERNATIONAL SEARCH REPORT

International application No

PCT/US2014/048606

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2012/021478 A1 (OSTERHOUT ROBIN E [US] ET AL) 26 January 2012 (2012-01-26) paragraph [0327] -----	1-17

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/US2014/048606

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2013082542 A2	06-06-2013	EP 2785848 A2	08-10-2014
		WO 2013082542 A2	06-06-2013

US 2012021478 A1	26-01-2012	AU 2011286199 A1	14-02-2013
		CA 2806230 A1	09-02-2012
		CN 103025877 A	03-04-2013
		CO 6680664 A2	31-05-2013
		EP 2607340 A1	26-06-2013
		JP 2013535203 A	12-09-2013
		KR 20130043170 A	29-04-2013
		SG 187174 A1	28-02-2013
		US 2012021478 A1	26-01-2012
		WO 2012018624 A2	09-02-2012
