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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
01 February 2018 (01.02.2018)

WIPO | PCT

(10) International Publication Number
WO 2018/022595 A1

(51) International Patent Classification:
C12N 15/52 (2006.01) *C12N 15/90* (2006.01)

(21) International Application Number:
PCT/US2017/043682

(22) International Filing Date:
25 July 2017 (25.07.2017)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/366,565 25 July 2016 (25.07.2016) US

(71) Applicants: **INVISTA TEXTILES (U.K.) LIMITED** [GB/GB]; One St. Peter's Square, Manchester M2 3DE (GB). **INVISTA NORTH AMERICA S.A.R.L.** [LU/US]; Three Little Falls Centre, 2801 Centerville Road, Wilmington, DE 19808 (US).

(72) Inventors: **FOSTER, Alexander, Brett**; 4 Westholme, Hutton Ruddy, Yarm TS15 0EA (GB). **CARTMAN, Stephen, Thomas**; 14 Glenluce Close, Eaglescliffe, Stockton-on-Tees TS16 9HR (GB). **KENNEDY, Jonathan**; 49 Cleveland Avenue, Stokesley, North Yorkshire TS9 5HB (GB).

(74) Agent: **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, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, 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,

(54) Title: MATERIALS AND METHODS UTILIZING BIOTIN PRODUCING MUTANT HOSTS FOR THE PRODUCTION OF 7-CARBON CHEMICALS

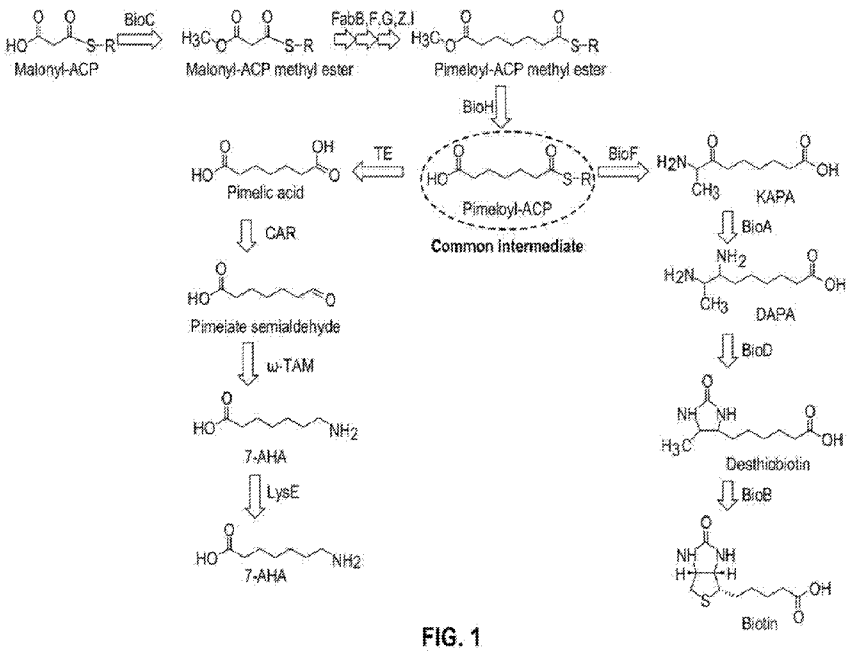

FIG. 1

(57) Abstract: Disclosed are methods for regulating biosynthesis of at least one of pimelic acid, 7-aminoheptanoic acid, 7-hydroxyheptanoic acid, heptamethylenediamine, 7-arninoheiptanol and 1,7-heptanediol (C7 building blocks) using a pathway having a pimeloyl-ACP intermediate, the method including the step of downregulating the activity of BioF. Also disclosed are recombinant hosts by fermentation in which the above methods are performed. Further disclosed are recombinant hosts for producing pimeloyl-ACP, the recombinant host including a deletion of a *bioF* gene.

WO 2018/022595 A1

[Continued on next page]

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))*

MATERIALS AND METHODS UTILIZING BIOTIN PRODUCING MUTANT HOSTS FOR THE PRODUCTION OF 7-CARBON CHEMICALS

CROSS-REFERENCE TO RELATED APPLICATIONS

[1] The present application claims the benefit of U.S. Provisional Application No. 62/366,565, filed on July 25, 2016, which is incorporated herein by reference in its entirety.

SEQUENCE LISTING

[2] 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 July 24, 2017, is named 12444_0684-00304_SL.txt and is 111,898 bytes in size.

BACKGROUND

[3] This disclosure relates to methods for regulating biosynthesis of at least one seven carbon compounds. This disclosure relates to materials and methods for regulating biosynthesis of at least one of pimelic acid, 7-aminoheptanoic acid, 7-hydroxyheptanoic acid, heptamethylenediamine, 7-aminohelptanol and 1,7-heptanediol (hereafter "C7 building blocks") using a pathway having a pimelyl-ACP intermediate. The methods include the step of downregulating the activity of BioF, an 8-amino-7-oxononanoate synthase. The disclosure also relates to recombinant hosts by fermentation in which the above-mentioned method is performed. Also disclosed are pimeloyl-ACP producing recombinant hosts that include a deletion of a *bioF* gene.

[4] Nylons are synthetic polyamides which are sometimes synthesized by the condensation polymerisation of a diamine with a dicarboxylic acid. Similarly, Nylons may be produced by the condensation polymerisation of lactams. A ubiquitous Nylon is Nylon 6,6, which is produced by reaction of hexamethylenediamine (HMD) and adipic acid. Nylon 6 is produced by a ring opening polymerisation of caprolactam (Anton & Baird, Polyamides Fibers, Encyclopedia of Polymer Science and Technology, 2001).

[5] Nylon 7 and Nylon 7,7 represent novel polyamides with value-added characteristics compared to Nylon 6 and Nylon 6,6. Nylon 7 is produced by polymerisation of 7-aminoheptanoic acid, whereas Nylon 7,7 is produced by condensation polymerisation of

pimelic acid and heptamethylenediamine. No economically viable petrochemical routes exist to producing the monomers for Nylon 7 and Nylon 7,7.

[6] Given no economically cost competitive petrochemical monomer feedstocks, biotechnology offers an alternative approach via biocatalysis. Biocatalysis is the use of biological catalysts, such as enzymes, to perform biochemical transformations of organic compounds.

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

[8] However, no wild-type prokaryote or eukaryote naturally overproduces or excretes C7 building blocks to the extracellular environment. Nevertheless, the metabolism of pimelic acid has been reported.

[9] The dicarboxylic acid, pimelic acid, is converted efficiently as a carbon source by a number of bacteria and yeasts via β-oxidation into central metabolites. β-oxidation of CoEnzyme A (CoA) activated pimelate to CoA-activated 3-oxopimelate facilitates further catabolism via, for example, pathways associated with aromatic substrate degradation. The catabolism of 3-oxopimeloyl-CoA to acetyl-CoA and glutaryl-CoA by several bacteria has been characterized comprehensively (Harwood and Parales, *Annual Review of Microbiology*, 1996, 50, 553 – 590).

[10] The optimality principle states that microorganisms regulate their biochemical networks to support maximum biomass growth. Beyond the need to express heterologous pathways in a host organism, directing carbon flux towards C7 building blocks that serve as carbon sources rather than to biomass growth constituents, contradicts the optimality principle. For example, transferring the 1-butanol pathway from *Clostridium* species into other production strains has often fallen short by an order of magnitude compared to the production performance of native producers (Shen *et al.*, *Appl. Environ. Microbiol.*, 2011, 77(9), 2905 – 2915).

[11] The efficient synthesis of the C7 aliphatic backbone precursor is a key consideration in synthesizing C7 building blocks prior to forming terminal functional groups, such as carboxyl, amine or hydroxyl groups, on the C7 aliphatic backbone. Accordingly, against this background, it is clear that there is a need for methods for producing pimelic acid, 7-aminoheptanoic acid, heptamethylenediamine, 7-hydroxyheptanoic acid, 7-aminohelptanol and 1,7-heptanediol (hereafter "C7 building blocks") wherein the methods are biocatalyst-based.

SUMMARY

[12] Described herein are methods and genetically modified hosts that allow for more efficient use of seven carbon aliphatic backbone precursors and production of C7 building blocks, for example, by use of a recombinant host in a *bioF* deficient background.

[13] The present disclosure relates to regulating biosynthesis for certain seven-carbon building blocks ("C7 building blocks") by downregulating the activity of a BioF. In one embodiment, downregulating the activity of BioF comprises downregulating the expression of a BioF protein. In another embodiment, downregulating the activity of BioF comprises downregulating the activity of the protein itself. The pathway for producing the C7 building blocks has a pimelyl-ACP intermediate. The C7 building blocks may include, for example, pimelic acid, 7-aminoheptanoate ("7-AHA"), 7-hydroxyheptanoate, heptamethylenediamine, 7-aminoheptanol and 1,7-heptanediol.

[14] In some embodiments, the biochemical pathways for the production of 7-AHA currently being pursued are based on the production of pimelic acid, followed by the conversion of pimelic acid to 7-AHA. *See* FIG. 1. For example, E. coli may act as a host for the pathways producing pimelic acid its subsequent conversion to 7-AHA.

[15] Pimeloyl-ACP, or pimelyl-ACP, is an intermediate of the biotin pathway in host, such as *E. coli*. For example, some methods in this disclosure for producing pimelic acid in *E. coli* are focused on metabolic engineering approaches, such as increasing the flux through this pathway, and the effective interception of the pimelic acid intermediate. *See* FIG. 1.

[16] Pimelic acid and pimelate, 7-hydroxyheptanoic acid and 7-hydroxyheptanoate, and 7-aminoheptanoic acid and 7-aminoheptanoate are used interchangeably herein to refer to the compound in any of its neutral or ionized forms, including any salt forms thereof. It is understood by those skilled in the art that the specific form will depend on pH. These pathways, metabolic engineering and cultivation strategies described herein rely on fatty acid elongation and synthesis enzymes or homologs accepting methyl-ester shielded dicarboxylic acids as substrates.

[17] For compounds containing carboxylic acid groups such as organic monoacids, hydroxyacids, aminoacids and dicarboxylic acids, these compounds may be formed or converted to their ionic salt form when an acidic proton present in the parent compound either is replaced by a metal ion, e.g., an alkali metal ion, an alkaline earth ion, or an aluminum ion; or coordinates with an organic base. Acceptable organic bases include

ethanolamine, diethanolamine, triethanolamine, tromethamine, N-methylglucamine, and the like. Acceptable inorganic bases include aluminum hydroxide, calcium hydroxide, potassium hydroxide, sodium carbonate, sodium hydroxide, and the like. The salt can be isolated as is from the system as the salt or converted to the free acid by reducing the pH to below the pKa through addition of acid or treatment with an acidic ion exchange resin.

[18] For compounds containing amine groups such as but not limited to organic amines, aminoacids and diamine, these compounds may be formed or converted to their ionic salt form by addition of an acidic proton to the amine to form the ammonium salt, formed with inorganic acids such as hydrochloric acid, hydrobromic acid, sulfuric acid, nitric acid, phosphoric acid, and the like; or formed with organic acids such as acetic acid, propionic acid, hexanoic acid, cyclopentanepropionic acid, glycolic acid, pyruvic acid, lactic acid, malonic acid, succinic acid, malic acid, maleic acid, fumaric acid, tartaric acid, citric acid, benzoic acid, 3-(4-hydroxybenzoyl)benzoic acid, cinnamic acid, mandelic acid, methanesulfonic acid, ethanesulfonic acid, 1,2-ethanedisulfonic acid, 2-hydroxyethanesulfonic acid, benzenesulfonic acid, 2-naphthalenesulfonic acid, 4-methylbicyclo-[2.2.2]oct-2-ene-1-carboxylic acid, glucoheptonic acid, 4,4'-methylenebis-(3-hydroxy-2-ene-1-carboxylic acid), 3-phenylpropionic acid, trimethylacetic acid, tertiary butylacetic acid, lauryl sulfuric acid, gluconic acid, glutamic acid, hydroxynaphthoic acid, salicylic acid, stearic acid or muconic acid. Acceptable inorganic bases are known in the art and include aluminum hydroxide, calcium hydroxide, potassium hydroxide, sodium carbonate, sodium hydroxide, and the like. The salt can be isolated as is from the system as a salt or converted to the free amine by raising the pH to above the pKb through addition of base or treatment with a basic ion exchange resin.

[19] For compounds containing both amine groups and carboxylic acid groups such as but not limited to aminoacids, these compounds may be formed or converted to their ionic salt form by either 1) acid addition salts, formed with inorganic acids such as hydrochloric acid, hydrobromic acid, sulfuric acid, nitric acid, phosphoric acid, and the like; or formed with organic acids such as acetic acid, propionic acid, hexanoic acid, cyclopentanepropionic acid, glycolic acid, pyruvic acid, lactic acid, malonic acid, succinic acid, malic acid, maleic acid, fumaric acid, tartaric acid, citric acid, benzoic acid, 3-(4-hydroxybenzoyl)benzoic acid, cinnamic acid, mandelic acid, methanesulfonic acid, ethanesulfonic acid, 1,2-ethanedisulfonic acid, 2-hydroxyethanesulfonic acid, benzenesulfonic acid, 2-naphthalenesulfonic acid, 4-methylbicyclo-[2.2.2]oct-2-ene-1-carboxylic acid, glucoheptonic acid, 4,4'-methylenebis-(3-hydroxy-2-ene-1-carboxylic acid),

3-phenylpropionic acid, trimethylacetic acid, tertiary butylacetic acid, lauryl sulfuric acid, gluconic acid, glutamic acid, hydroxynaphthoic acid, salicylic acid, stearic acid, muconic acid Acceptable inorganic bases include aluminum hydroxide, calcium hydroxide, potassium hydroxide, sodium carbonate, sodium hydroxide, and the like or 2) when an acidic proton present in the parent compound either is replaced by a metal ion, e.g., an alkali metal ion, an alkaline earth ion, or an aluminum ion; or coordinates with an organic base. Acceptable organic bases are known in the art and include ethanolamine, diethanolamine, triethanolamine, tromethamine, N-methylglucamine, and the like. Acceptable inorganic bases are known in the art and include aluminum hydroxide, calcium hydroxide, potassium hydroxide, sodium carbonate, sodium hydroxide, and the like. The salt can be isolated as is from the system or converted to the free acid by reducing the pH to below the pKa through addition of acid or treatment with an acidic ion exchange resin.

[20] In the face of the optimality principle, it surprisingly has been discovered that appropriate non-natural pathways, feedstocks, host microorganisms, attenuation strategies to the host's biochemical network and cultivation strategies may be combined to efficiently produce one or more C7 building blocks. It has also been discovered that the pathway having pimelyl-ACP as intermediate can be regulated by downregulating the activity BioF. The BioF is ACP-dependent enzyme. It converts, for example, the pimeloyl-ACP to 8-amino-7-oxononanoic acid. The BioF is a 8-amino-7-oxononanoate synthase having at least 70% sequence identity or homology to the amino acid sequence set forth in SEQ ID NO: 23.

[21] In some embodiments, a terminal amine group can be enzymatically formed using a *ω-transaminase* or a *deacetylase*. See FIG. 5 and FIGs. 6A-6C.

[22] In some embodiments, a terminal hydroxyl group can be enzymatically formed using a *4-hydroxybutyrate dehydrogenase*, *5-hydroxypentanoate dehydrogenase* or a *6-hydroxyhexanoate dehydrogenase* or an *alcohol dehydrogenase*. See FIG. 7 and FIG. 8.

[23] In one aspect, this disclosure features a method for biosynthesizing a product selected from the group consisting of pimelic acid, 7-aminoheptanoate, 7-hydroxyheptanoate, heptamethylenediamine, 7-aminoheptanol and 1,7-heptanediol. The method includes enzymatically synthesizing a C7 aliphatic backbone from (i) acetyl-CoA and malonyl-CoA via two cycles of methyl ester shielded carbon chain elongation or (ii) malonyl-ACP via two cycles of methyl-ester shielded carbon chain elongation, and enzymatically forming one or two terminal functional groups selected from the group consisting of carboxyl, amine, and hydroxyl groups in the backbone, thereby forming the product.

[24] The present disclosure relates to a method for regulating biosynthesis of at least one C7 building block using a pathway having a pimelyl-ACP intermediate. The method may comprise the step of downregulating the activity of BioF.

[25] The at least one C7 building block is selected from the group consisting of pimelic acid, 7-aminoheptanoate, 7-hydroxyheptanoate, heptamethylenediamine, 7-aminoheptanol and 1,7-heptanediol.

[26] In some embodiments, the step of downregulating the activity of BioF may comprise reducing the activity to a level below that of a wild type BioF, for example, reducing the activity to zero. In one embodiment, downregulating the activity of BioF comprises downregulating the expression of BioF protein. In another embodiment, downregulating the activity of BioF comprises downregulating the activity of the protein itself.

[27] In some embodiments, the method for regulating biosynthesis of at least one C7 building block may further comprise the step of overexpressing BioW and a CoA-specific BioF. BioW enzymatically converts pimelic acid to pimeloyl-CoA. The BioW is a pimeloyl-CoA ligase having at least 70% sequence identity or homology to the amino acid sequence set forth in SEQ ID NO: 18.

[28] CoA-specific BioF enzymatically converts pimeloyl-CoA to 8-amino-7-oxo-nonanoic acid.

[29] The above-mentioned CoA-specific BioF is a 8-amino-7-oxononanoate synthase having at least 70% sequence identity or homology to the amino acid sequence set forth in SEQ ID NO: 19.

[30] The above-mentioned pathway comprises enzymatically synthesizing a C7 aliphatic backbone from malonyl-ACP via two cycles of methyl-ester shielded carbon chain elongation, and enzymatically forming two terminal functional groups selected from the group consisting of carboxyl, amine, and hydroxyl groups in said backbone, thereby forming the C7 building block.

[31] A *S-adenosyl-L-methionine (SAM)-dependent methyltransferase*, or BioC, converts malonyl-ACP to malonyl-ACP methyl ester. The *S-adenosyl-L-methionine (SAM)-dependent methyltransferase* has at least 70% sequence identity or homology to the amino acid sequence set forth in SEQ ID NO: 20.

[32] The C7 aliphatic backbone can be pimeloyl-ACP. A *malonyl-ACP O-methyltransferase* can covert malonyl-CoA to a malonyl-CoA methyl ester or can convert malonyl-ACP to a malonyl-ACP methyl ester. Each of the two cycles of carbon chain

elongation can include using (i) a *β-ketoacyl-ACP synthase* or a *β-ketothiolase*, (ii) a *3-oxoacyl-ACP reductase*, an *acetoacetyl-CoA reductase*, a *3-hydroxyacyl-CoA dehydrogenase* or a *3-hydroxybutyryl-CoA dehydrogenase*, (iii) an *enoyl-CoA hydratase* or a *3-hydroxyacyl-ACP dehydratase*, and (iv) an *enoyl-ACP reductase*. Or the two cycles of carbon chain elongation may involve a *trans-2-enoyl-CoA reductase* to produce pimeloyl-ACP methyl ester from malonyl-ACP methyl ester.

[33] The two terminal functional groups can be the same (e.g., amine or hydroxyl) or can be different (e.g., a terminal amine and a terminal carboxyl group; or a terminal hydroxyl group and a terminal carboxyl group). For example, the two terminal functional groups can be both amine, or the two terminal functional groups can be both hydroxyl groups. In other embodiments, the C7 building block may comprise a terminal amine and a terminal carboxyl group, or the C7 building block may comprise a terminal hydroxyl group and a terminal carboxyl group.

[34] A 6-hydroxyhexanoate dehydrogenase, a 5-hydroxypentanoate dehydrogenase, a 4-hydroxybutyrate dehydratase, or an alcohol dehydrogenase can enzymatically form a hydroxyl group.

[35] A thioesterase, an aldehyde dehydrogenase, a 7-oxoheptanoate dehydrogenase, a 6-oxohexanoate dehydrogenase, a CoA-transferase (e.g. a glutaconate CoA transferase), or a reversible CoA-ligase (e.g., a reversible succinate-CoA ligase) can enzymatically forms a terminal carboxyl group.

[36] The thioesterase ("TE") can have at least 70% sequence identity or homology to the amino acid sequence set forth in SEQ ID NO: 1.

[37] A *ω-transaminase* or a *deacetylase* can enzymatically form an amine group. The *ω-transaminase* can have at least 70% sequence identity or homology to any one of the amino acid sequences set forth in SEQ ID NO. 8 – 13.

[38] A *carboxylate reductase* ("CAR") and a *phosphopantetheinyl transferase* can form a terminal aldehyde group as an intermediate in forming the product. The *carboxylate reductase* can have at least 70% sequence identity or homology to any one of the amino acid sequences set forth in SEQ ID NO. 2 – 7.

[39] Any of the methods can be performed in a recombinant host by fermentation. The host can be subjected to a cultivation strategy under aerobic, anaerobic, micro-aerobic or mixed oxygen/denitrification cultivation conditions. The host can be cultured under conditions of nutrient limitation. The host can be retained using a ceramic hollow fiber membrane to maintain a high cell density during fermentation.

[40] The principal carbon source fed to the fermentation can derive from biological or non-biological feedstocks. In some embodiments, the biological feedstock is, includes, or derives from, at least one chosen from monosaccharides, disaccharides, lignocellulose, hemicellulose, cellulose, lignin, levulinic acid and formic acid, triglycerides, glycerol, fatty acids, agricultural waste, condensed distillers' solubles, or municipal waste.

[41] In some embodiments, the non-biological feedstock is, or derives from, at least one chosen 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 a terephthalic acid/isophthalic acid mixture waste stream.

[42] The recombinant host can be a prokaryote, e.g., 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 acidovorans*, from the genus *Bacillus* such as *Bacillus subtillis*; from the genes *Lactobacillus* such as *Lactobacillus delbrueckii*; from the genus *Lactococcus* such as *Lactococcus lactis* or from the genus *Rhodococcus* such as *Rhodococcus equi*.

[43] The recombinant host can be a eukaryote, e.g., 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 *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*.

[44] In any of the methods, the host's tolerance to high concentrations of at least one C7 building block can be improved through continuous cultivation in a selective environment.

[45] Any of the recombinant hosts described herein further can include one or more of the following attenuated enzymes: *polyhydroxyalkanoate synthase*, an *acetyl-CoA thioesterase*, an *acetyl-CoA specific β-ketothiolase*, a *phosphotransacetylase forming acetate*, an *acetate kinase*, a *lactate dehydrogenase*, a *menaquinol-fumarate oxidoreductase*, a *2-oxoacid decarboxylase* producing isobutanol, an *alcohol dehydrogenase* forming ethanol, a *triose phosphate isomerase*, a *pyruvate decarboxylase*, a *glucose-6-phosphate isomerase*, a *transhydrogenase* dissipating the NADH or NADPH imbalance, an *glutamate dehydrogenase*

dissipating the NADH or NADPH imbalance, a NADH/NADPH-utilizing *glutamate dehydrogenase*, a *pimeloyl-CoA dehydrogenase*; an *acyl-CoA dehydrogenase* accepting C7 building blocks and central precursors as substrates; a *glutaryl-CoA dehydrogenase*; or a *pimeloyl-CoA synthetase*.

[46] Any of the recombinant hosts described herein further can overexpress one or more genes encoding: an *acetyl-CoA synthetase*, a *6-phosphogluconate dehydrogenase*; a *transketolase*; a *puridine nucleotide transhydrogenase*; a *formate dehydrogenase*; a *glyceraldehyde-3P-dehydrogenase*; a *malic enzyme*; a *glucose-6-phosphate dehydrogenase*; a *fructose 1,6 diphosphatase*; a *L-alanine dehydrogenase*; a *L-glutamate dehydrogenase* specific to the NADH or NADPH used to generate a co-factor imbalance; a *methanol dehydrogenase*, a *formaldehyde dehydrogenase*, a *diamine transporter; a dicarboxylate transporter*; an *S-adenosylmethionine synthetase*, and/or a *multidrug transporter*.

[47] The disclosure also relates to a recombinant host for producing for pimeloyl-ACP. The recombinant host may comprise a deletion of a *bioF* gene. In some embodiments, the recombinant host expresses BioF activity at a level below that of a wild type. In some embodiments, the recombinant host does not express BioF activity.

[48] In some embodiments, the recombinant host comprises at least one exogenous nucleic acid encoding *bioW* and a CoA-specific *bioF*.

[49] Also diclosed herein is a non-naturally occurring organism comprising at least one exogenous nucleic acid encoding at least one polypeptide having the activity of at least one enzyme depicted in any one of Figs. 1 to 21.

[50] In some embodiments, a nucleic acid construct or expression vector comprises a polynucleotide encoding a polypeptide having *8-amino-7-oxononanoate synthase* activity, wherein the polynucleotide is operably linked to one or more heterologous control sequences that direct production of the polypeptide and wherein the polypeptide having *8-amino-7-oxononanoate synthase* activity is a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 23.

[51] In some embodiments, a nucleic acid construct or expression vector comprises a polynucleotide encoding a polypeptide having *thioesterase* activity, wherein the polynucleotide is operably linked to one or more heterologous control sequences that direct production of the polypeptide and wherein the polypeptide having *thioesterase* activity is selected from the group consisting of: (a) a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 1; (b) a polypeptide having at least 70%

sequence identity or homology to the polypeptide of SEQ ID NO: 21; and (c) a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 22.

[52] In some embodiments, a nucleic acid construct or expression vector comprises a polynucleotide encoding a polypeptide having *carboxylate reductase* activity, wherein the polynucleotide is operably linked to one or more heterologous control sequences that direct production of the polypeptide and wherein the polypeptide having *carboxylate reductase* activity is selected from the group consisting of: (a) a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 2; (b) a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 3; (c) a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 4; (d) a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 5; (e) a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 6; and (f) a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 7.

[53] In some embodiments, a nucleic acid construct or expression vector comprises a polynucleotide encoding a polypeptide having *ω-transaminase* activity, wherein the polynucleotide is operably linked to one or more heterologous control sequences that direct production of the polypeptide and wherein the polypeptide having *ω-transaminase* activity is selected from the group consisting of: (a) a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 8; (b) a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 9; (c) a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 10 (d) a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 11; (e) a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 12; and (f) a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 13.

[54] In some embodiments, a nucleic acid construct or expression vector comprises a polynucleotide encoding a polypeptide having *phosphopantetheinyl transferase* activity, wherein the polynucleotide is operably linked to one or more heterologous control sequences that direct production of the polypeptide and wherein the polypeptide having *phosphopantetheinyl transferase* activity is selected from the group consisting of: (a) a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 14 and (b) a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 15.

[55] In some embodiments, a nucleic acid construct or expression vector comprises a polynucleotide encoding a polypeptide having *pimeloyl-ACP methyl ester esterase* activity, wherein the polynucleotide is operably linked to one or more heterologous control sequences that direct production of the polypeptide and wherein the polypeptide having *pimeloyl-ACP methyl ester esterase* activity is a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 17.

[56] In some embodiments, a nucleic acid construct or expression vector comprises a polynucleotide encoding a polypeptide having *6-carboxyhexanoate-CoA ligase* activity, wherein the polynucleotide is operably linked to one or more heterologous control sequences that direct production of the polypeptide and wherein the polypeptide having *6-carboxyhexanoate-CoA ligase* activity is a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 18.

[57] In some embodiments, a nucleic acid construct or expression vector comprises a polynucleotide encoding a polypeptide having *8-amino-7-oxononanoate synthase* activity, wherein the polynucleotide is operably linked to one or more heterologous control sequences that direct production of the polypeptide and wherein the polypeptide having *8-amino-7-oxononanoate synthase* activity is a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 19.

[58] In some embodiments, a nucleic acid construct or expression vector comprises a polynucleotide encoding a polypeptide having *malonyl-ACP O-methyltransferase* activity, wherein the polynucleotide is operably linked to one or more heterologous control sequences that direct production of the polypeptide and wherein the polypeptide having *malonyl-ACP O-methyltransferase* activity is a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 20.

[59] In some embodiments, a composition comprises any of the above-described nucleic acid construct or expression vector.

[60] Also disclosed is a non-naturally occurring biochemical network comprising a 5-hydroxypentanoyl-CoA, an exogenous nucleic acid encoding a polypeptide having the activity of a β-ketothiolase classified under EC. 2.3.1, and a 3-oxo-7-hydroxyheptanoyl-CoA.

[61] In some embodiments, a bio-derived, bio-based or fermentation-derived product comprises:

i. a composition comprising at least one bio-derived, bio-based or fermentation-derived compound produced or biosynthesized according to any one of claims 1-75 or or any one of FIGS. 1-21, or any combination thereof;

ii. a bio-derived, bio-based or fermentation-derived polymer comprising the bio-derived, bio-based or fermentation-derived composition or compound of i., or any combination thereof;

iii. a bio-derived, bio-based or fermentation-derived resin comprising the bio-derived, bio-based or fermentation-derived compound or bio-derived, bio-based or fermentation-derived composition of i. or any combination thereof or the bio-derived, bio-based or fermentation-derived polymer of ii. or any combination thereof;

iv. a molded substance obtained by molding the bio-derived, bio-based or fermentation-derived polymer of ii. or the bio-derived, bio-based or fermentation-derived resin of iii., or any combination thereof;

v. a bio-derived, bio-based or fermentation-derived formulation comprising the bio-derived, bio-based or fermentation-derived composition of i., bio-derived, bio-based or fermentation-derived compound of i., bio-derived, bio-based or fermentation-derived polymer of ii., bio-derived, bio-based or fermentation-derived resin of iii., or bio-derived, bio-based or fermentation-derived molded substance of iv, or any combination thereof; or

vi. 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 of i., bio-derived, bio-based or fermentation-derived compound of i., bio-derived, bio-based or fermentation-derived polymer of ii., bio-derived, bio-based or fermentation-derived resin of iii., bio-derived, bio-based or fermentation-derived formulation of v., or bio-derived, bio-based or fermentation-derived molded substance of iv., or any combination thereof.

[62] The reactions of the pathways described herein can be performed in one or more cell (e.g., host cell) 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. Extracted enzymes can optionally be immobilized to a solid substrate such as the floors and/or walls of appropriate reaction vessels. Moreover, such extracts include lysates (e.g. cell lysates) that can be used as sources of relevant enzymes. In the methods provided by the disclosure, all the steps can be performed in cells (e.g., 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.

[63] Many of the enzymes described herein catalyze reversible reactions, and the reaction of interest may be the reverse of the described reaction. The schematic pathways shown in FIGs. 1-3 and 5-8 illustrate the reaction of interest for each of the intermediates.

[64] The disclosure also provides the following additional embodiments:

[65] Embodiment 1. A method for regulating biosynthesis of at least one C7 building block using a pathway having a pimeloyl-ACP intermediate, said method comprising the step of downregulating the activity of a BioF enzyme.

[66] Embodiment 2. The method of embodiment 1, wherein the BioF has at least 70% sequence identity or homology to the amino acid sequence set forth in SEQ ID NO: 23.

[67] Embodiment 3. The method of embodiment 1, wherein the at least one C7 building block is selected from the group consisting of pimelic acid, 7-aminoheptanoate, 7-hydroxyheptanoate, heptamethylenediamine, 7-aminoheptanol and 1,7-heptanediol.

[68] Embodiment 4. The method of embodiment 1, wherein the step of downregulating the activity of BioF comprises reducing the activity to a level below that of a wild type BioF protein.

[69] Embodiment 5. The method of embodiment 1, wherein the step of downregulating the activity of BioF comprises reducing the activity to zero.

[70] Embodiment 6. A method for regulating biosynthesis of at least one C7 building block using a pathway having a pimeloyl-ACP intermediate, said method comprising

the step of downregulating the activity of BioF; and

the step of overexpressing BioW and a CoA-specific BioF.

[71] Embodiment 7. The method of embodiment 6, wherein said BioW enzymatically converts pimelic acid to pimeloyl-CoA.

[72] Embodiment 8. The method of embodiment 6 or 7, wherein said BioW is a pimeloyl-CoA ligase having at least 70% sequence identity or homology to the amino acid sequence set forth in SEQ ID NO: 14.

[73] Embodiment 9. The method of any one of embodiments 6-8, wherein CoA-specific BioF enzymatically converts pimeloyl-CoA to 8-amino-7-oxo-nonanoic acid.

[74] Embodiment 10. The method of any one of embodiments 6-9, wherein said CoA-specific BioF is a 8-amino-7-oxononanoate synthase having at least 70% sequence identity or homology to the amino acid sequence set forth in SEQ ID NO: 15.

[75] Embodiment 11. The method of embodiment 1, wherein said pathway comprises enzymatically synthesizing a C7 aliphatic backbone from malonyl-ACP via two

cycles of methyl-ester shielded carbon chain elongation, and enzymatically forming two terminal functional groups independently selected from the group consisting of carboxyl, amine, and hydroxyl groups in said backbone, thereby forming the C7 building block.

[76] Embodiment 12. The method of embodiment 11, wherein a *S-adenosyl-L-methionine (SAM)-dependent methyltransferase* converts malonyl-ACP to malonyl-ACP methyl ester.

[77] Embodiment 13. The method of embodiment 11, wherein the *S-adenosyl-L-methionine (SAM)-dependent methyltransferase* has at least 70% sequence identity or homology to the amino acid sequence set forth in SEQ ID NO: 16.

[78] Embodiment 14. The method of embodiment 10, wherein each of said two cycles of carbon chain elongation comprises using (i) a *β-ketoacyl-ACP synthase*, (ii) a *3-oxoacyl-ACP reductase*, (iii) a *3-hydroxyacyl-ACP dehydratase*, and (iv) an *enoyl-ACP reductase*.

[79] Embodiment 15. The method of embodiment 10, wherein said two cycles of carbon chain elongation produce pimeloyl-ACP methyl ester from malonyl-ACP methyl ester using a *trans-2-enoyl-CoA reductase*.

[80] Embodiment 16. The method of any one of embodiments 1-15, wherein said two terminal functional groups are the same.

[81] Embodiment 17. The method of any one of embodiments 1-16, wherein said two terminal functional groups are amine.

[82] Embodiment 18. The method of any one of embodiments 1-16, wherein said two terminal functional groups are hydroxyl groups.

[83] Embodiment 19. The method of any one of embodiments 1-15, wherein said two terminal functional groups are different.

[84] Embodiment 20. The method of any one of embodiments 1-15 and 19, wherein said at least one C7 building block comprises a terminal amine and a terminal carboxyl group.

[85] Embodiment 21. The method of any one of embodiments 1-15 and 19, wherein said at least one C7 building block comprises a terminal hydroxyl group and a terminal carboxyl group.

[86] Embodiment 22. The method of any one of embodiments 1-21, wherein the hydroxyl group is enzymatically formed by a *6-hydroxyhexanoate dehydrogenase*, a *5-*

hydroxypentanoate dehydrogenase, a *4-hydroxybutyrate dehydratase*, or an *alcohol dehydrogenase*.

[87] Embodiment 23. The method of any one of embodiments 1-22, wherein the carboxyl group is enzymatically formed by a *thioesterase*, an *aldehyde dehydrogenase*, a *7-oxoheptanoate dehydrogenase*, a *6-oxohexanoate dehydrogenase*, a *glutaconate CoA-transferase*, or a *reversible succinyl-CoA ligase*.

[88] Embodiment 24. The method of embodiment 23, wherein said *thioesterase* has at least 70% sequence identity or homology to the amino acid sequence set forth in SEQ ID NO: 21 or 22.

[89] Embodiment 25. The method of any one of embodiments 1-24, wherein the amine group is enzymatically formed by a *ω-transaminase* or a *deacetylase*.

[90] Embodiment 26. The method of embodiment 25, wherein said *ω-transaminase* has at least 70% sequence identity or homology to any one of the amino acid sequences set forth in any of SEQ ID NOs: 8 – 13.

[91] Embodiment 27. The method of any one of embodiments 1-26, wherein an intermediate (in forming the at least one product) with a terminal aldehyde group is formed by a *carboxylate reductase* and enhanced by a *phosphopantetheinyl transferase*.

[92] Embodiment 28. The method of embodiment 27, wherein said *carboxylate reductase* has at least 70% sequence identity or homology to any one of the amino acid sequences set forth in any of SEQ ID NOs: 2 – 7.

[93] Embodiment 29. The method of any of the preceding embodiments, wherein said method is performed in a recombinant host by fermentation.

[94] Embodiment 30. The method of embodiment 29, wherein said recombinant host is subjected to a cultivation condition under aerobic, anaerobic, micro-aerobic or mixed oxygen/denitrification cultivation conditions.

[95] Embodiment 31. The method of embodiment 29 or embodiment 30, wherein said recombinant host is cultured under conditions of nutrient limitation.

[96] Embodiment 32. The method according to any one of embodiments 29-31, wherein said recombinant host is retained using a ceramic hollow fiber membrane to maintain a high cell density during fermentation.

[97] Embodiment 33. The method of any one of embodiments 29-32, wherein the principal carbon source fed to the fermentation derives from biological or non-biological feedstocks.

[98] Embodiment 34. The method of embodiment 33, wherein the biological feedstock is, or derives from, at least one chosen from monosaccharides, disaccharides, lignocellulose, hemicellulose, cellulose, lignin, levulinic acid, formic acid, triglycerides, glycerol, fatty acids, agricultural waste, condensed distillers' solubles, or municipal waste.

[99] Embodiment 35. The method of embodiment 33, wherein the non-biological feedstock is, or derives from, at least one chosen 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.

[100] Embodiment 36. The method of any one of embodiments 29-35, wherein the recombinant host is a prokaryote.

[101] Embodiment 37. The method of embodiment 36, wherein said prokaryote is chosen from the genii *Escherichia*, *Clostridia*, *Corynebacteria*, *Cupriavidus*, *Pseudomonas*, *Delftia acidovorans*, *Bacillus*, *Lactobacillus*, *Lactococcus*, or *Rhodococcus*.

[102] Embodiment 38. The method of embodiment 37, wherein said prokaryote is *Escherichia coli*.

[103] Embodiment 39. The method of embodiment 37, wherein said prokaryote is *Clostridium ljungdahlii*, *Clostridium autoethanogenum*, or *Clostridium kluyveri*.

[104] Embodiment 40. The method of embodiment 37, wherein said prokaryote is *Corynebacterium glutamicum*.

[105] Embodiment 41. The method of embodiment 37, wherein said prokaryote is *Cupriavidus necator* or *Cupriavidus metallidurans*.

[106] Embodiment 42. The method of embodiment 37, wherein said prokaryote is *Pseudomonas fluorescens*, *Pseudomonas putida* or *Pseudomonas oleavorans*.

[107] Embodiment 43. The method of embodiment 37, wherein said prokaryote is *Bacillus subtillis*.

[108] Embodiment 44. The method of embodiment 37, wherein said prokaryote is *Lactobacillus delbrueckii*.

[109] Embodiment 45. The method of embodiment 37, wherein said prokaryote is *Lactococcus lactis*.

[110] Embodiment 46. The method of embodiment 37, wherein said prokaryote is *Rhodococcus equi*.

[111] Embodiment 47. The method of any one of embodiments 29-35, wherein the host is a eukaryote.

[112] Embodiment 48. The method of embodiment 47, wherein said eukaryote is from the genii *Aspergillus*, *Saccharomyces*,*Pichia, Yarrowia, Issatchenkia, Debaryomyces, Arxula,* or *Kluyveromyces.*

[113] Embodiment 49. The method of embodiment 47, wherein said eukaryote is *Aspergillus niger.*

[114] Embodiment 50. The method of embodiment 47, wherein said eukaryote is *Saccharomyces cerevisiae.*

[115] Embodiment 51. The method of embodiment 47, wherein said eukaryote is *Pichia pastoris.*

[116] Embodiment 52. The method of embodiment 47, wherein said eukaryote is *Yarrowia lipolytica.*

[117] Embodiment 53. The method of embodiment 47, wherein said eukaryote is *Issathenkia orientalis.*

[118] Embodiment 54. The method of embodiment 47, wherein said eukaryote is *Debaryomyces hansenii.*

[119] Embodiment 55. The method of embodiment 47, wherein said eukaryote is *Arxula adenoinivorans.*

[120] Embodiment 56. The method of embodiment 47, wherein said eukaryote is *Kluyveromyces lactis.*

[121] Embodiment 57. The method of embodiment 29, wherein the recombinant host's tolerance to high concentrations of at least one C7 building block is improved through continuous cultivation in a selective environment.

[122] Embodiment 58. The method of any one of embodiments 29-57, wherein said recombinant host comprises one or more of attenuated enzymes chosen from *polyhydroxyalkanoate synthase*, an *acetyl-CoA thioesterase*, an *acetyl-CoA specific β-ketothiolase*, a *phosphotransacetylase forming acetate*, an *acetate kinase*, a *lactate dehydrogenase*, a *menaquinol-fumarate oxidoreductase*, a *2-oxoacid decarboxylase* producing isobutanol, an *alcohol dehydrogenase* forming ethanol, a *triose phosphate isomerase*, a *pyruvate decarboxylase*, a *glucose-6-phosphate isomerase*, a *transhydrogenase* dissipating the NADH or NADPH imbalance, an *glutamate dehydrogenase* dissipating the NADH or NADPH imbalance, a NADH/NADPH-utilizing *glutamate dehydrogenase*, a *pimeloyl-CoA dehydrogenase*; an *acyl-CoA dehydrogenase* accepting C7 building blocks and central precursors as substrates; a *glutaryl-CoA dehydrogenase*; or a *pimeloyl-CoA synthetase.*

[123] Embodiment 59. The method of any one of embodiments 29-58, wherein said recombinant host overexpresses one or more genes encoding an *acetyl-CoA synthetase*, a *6-phosphogluconate dehydrogenase*; a *transketolase*; a *puridine nucleotide transhydrogenase*; a *formate dehydrogenase*; a *glyceraldehyde-3P-dehydrogenase*; a *malic enzyme*; a *glucose-6-phosphate dehydrogenase*; a *fructose 1,6 diphosphatase*; a *L-alanine dehydrogenase*; a *L-glutamate dehydrogenase* specific to the NADH or NADPH used to generate a co-factor imbalance; a *methanol dehydrogenase*, a *formaldehyde dehydrogenase*, a *diamine transporter*; a *dicarboxylate transporter*; an *S-adenosylmethionine synthetase* and/or a *multidrug transporter*.

[124] Embodiment 60. A recombinant host for producing for pimeloyl-ACP, said recombinant host comprising a deletion of a *bioF* gene.

[125] Embodiment 61. A recombinant host for producing for pimeloyl-ACP, wherein said recombinant host expresses a BioF at a level below that of a wild type host.

[126] Embodiment 62. The recombinant host of embodiment 1, wherein said recombinant host does not express BioF activity.

[127] Embodiment 63. The recombinant host of embodiment [47], wherein said recombinant host comprises at least one exogenous nucleic acid encoding *bioW* and a CoA-specific *bioF*.

[128] Embodiment 64. A non-naturally occurring organism comprising at least one exogenous nucleic acid encoding at least one polypeptide having the activity of at least one enzyme depicted in any one of Figs. 1 to 21.

[129] Embodiment 65. A nucleic acid construct or expression vector comprising a polynucleotide encoding a polypeptide having *8-amino-7-oxononanoate synthase* activity, wherein the polynucleotide is operably linked to one or more heterologous control sequences that direct production of the polypeptide and wherein the polypeptide having *8-amino-7-oxononanoate synthase* activity is a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 23.

[130] Embodiment 66. A nucleic acid construct or expression vector comprising a polynucleotide encoding a polypeptide having *thioesterase* activity, wherein the polynucleotide is operably linked to one or more heterologous control sequences that direct production of the polypeptide and wherein the polypeptide having *thioesterase* activity is selected from the group consisting of: (a) a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 1; (b) a polypeptide having at least 70%

sequence identity or homology to the polypeptide of SEQ ID NO: 21; and (c) a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 22.

[131] Embodiment 67. A nucleic acid construct or expression vector comprising a polynucleotide encoding a polypeptide having *carboxylate reductase* activity, wherein the polynucleotide is operably linked to one or more heterologous control sequences that direct production of the polypeptide and wherein the polypeptide having *carboxylate reductase* activity is selected from the group consisting of: (a) a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 2; (b) a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 3; (c) a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 4; (d) a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 5; (e) a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 6; and (f) a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 7.

[132] Embodiment 68. A nucleic acid construct or expression vector comprising a polynucleotide encoding a polypeptide having *ω-transaminase* activity, wherein the polynucleotide is operably linked to one or more heterologous control sequences that direct production of the polypeptide and wherein the polypeptide having *ω-transaminase* activity is selected from the group consisting of: (a) a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 8; (b) a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 9; (c) a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 10 (d) a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 11; (e) a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 12; and (f) a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 13.

[133] Embodiment 69. A nucleic acid construct or expression vector comprising a polynucleotide encoding a polypeptide having *phosphopantetheinyl transferase* activity, wherein the polynucleotide is operably linked to one or more heterologous control sequences that direct production of the polypeptide and wherein the polypeptide having *phosphopantetheinyl transferase* activity is selected from the group consisting of: (a) a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 14 and (b) a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 15.

[134] Embodiment 70. A nucleic acid construct or expression vector comprising a polynucleotide encoding a polypeptide having *pimeloyl-ACP methyl ester esterase* activity, wherein the polynucleotide is operably linked to one or more heterologous control sequences that direct production of the polypeptide and wherein the polypeptide having *pimeloyl-ACP methyl ester esterase* activity is a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 17.

[135] Embodiment 71. A nucleic acid construct or expression vector comprising a polynucleotide encoding a polypeptide having *6-carboxyhexanoate-CoA ligase* activity, wherein the polynucleotide is operably linked to one or more heterologous control sequences that direct production of the polypeptide and wherein the polypeptide having *6-carboxyhexanoate-CoA ligase* activity is a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 18.

[136] Embodiment 72. A nucleic acid construct or expression vector comprising a polynucleotide encoding a polypeptide having *8-amino-7-oxononanoate synthase* activity, wherein the polynucleotide is operably linked to one or more heterologous control sequences that direct production of the polypeptide and wherein the polypeptide having *8-amino-7-oxononanoate synthase* activity is a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 19.

[137] Embodiment 73. A nucleic acid construct or expression vector comprising a polynucleotide encoding a polypeptide having *malonyl-ACP O-methyltransferase* activity, wherein the polynucleotide is operably linked to one or more heterologous control sequences that direct production of the polypeptide and wherein the polypeptide having *malonyl-ACP O-methyltransferase* activity is a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 20.

[138] Embodiment 74. A composition comprising the nucleic acid construct or expression vector of any one of embodiments 65-73.

[139] Embodiment 75. A non-naturally occurring biochemical network comprising a 5-hydroxypentanoyl-CoA, an exogenous nucleic acid encoding a polypeptide having the activity of a β-ketothiolase classified under EC. 2.3.1, and a 3-oxo-7-hydroxyheptanoyl-CoA.

[140] Embodiment 76. A bio-derived, bio-based or fermentation-derived product, wherein said product comprises:

i. a composition comprising at least one bio-derived, bio-based or fermentation-derived compound produced or biosynthesized according to any one of embodiments 1-75 or or any one of FIGS. 1-21, or any combination thereof;

ii. a bio-derived, bio-based or fermentation-derived polymer comprising the bio-derived, bio-based or fermentation-derived composition or compound of i., or any combination thereof;

iii. a bio-derived, bio-based or fermentation-derived resin comprising the bio-derived, bio-based or fermentation-derived compound or bio-derived, bio-based or fermentation-derived composition of i. or any combination thereof or the bio-derived, bio-based or fermentation-derived polymer of ii. or any combination thereof;

iv. a molded substance obtained by molding the bio-derived, bio-based or fermentation-derived polymer of ii. or the bio-derived, bio-based or fermentation-derived resin of iii., or any combination thereof;

v. a bio-derived, bio-based or fermentation-derived formulation comprising the bio-derived, bio-based or fermentation-derived composition of i., bio-derived, bio-based or fermentation-derived compound of i., bio-derived, bio-based or fermentation-derived polymer of ii., bio-derived, bio-based or fermentation-derived resin of iii., or bio-derived, bio-based or fermentation-derived molded substance of iv, or any combination thereof; or

vi. 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 of i., bio-derived, bio-based or fermentation-derived compound of i., bio-derived, bio-based or fermentation-derived polymer of ii., bio-derived, bio-based or fermentation-derived resin of iii., bio-derived, bio-based or fermentation-derived formulation of v., or bio-derived, bio-based or fermentation-derived molded substance of iv., or any combination thereof.

[141] 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.

[142] 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 the 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.

BRIEF DESCRIPTION OF DRAWING(S)

[143] FIG 1 illustrates an exemplary biotin biosynthesis in E. coli through intermediates 8-amino-7-oxononanoate (KAPA) and 7,8-diaminopelargonic acid (DAPA), and the regulation of the biosynthesis to divert pimelic acid for 7-AHA production.

[144] FIG. 2 is a schematic of exemplary biochemical pathways leading to 8-amino-7-oxononanoate (KAPA) starting from pimelic acid.

[145] FIG. 3 is a schematic of an exemplary biochemical pathway leading to pimelic acid from malonyl-ACP through two cycles of methyl-ester shielded carbon chain elongation.

[146] FIG. 4 illustrates the constructs for expression of pathway for conversion of pimelic acid to 7-AHA in the *ΔbioF* E. coli strain.

[147] FIG. 5 is a schematic of exemplary biochemical pathways leading to 7-aminoheptanoate using pimelate or pimelate semialdehyde as precursors.

[148] FIGs. 6A-6C are schematics of exemplary biochemical pathways leading to heptamethylenediamine using 7-aminoheptanoate, 7-hydroxyheptanoate or pimelate semialdehyde as central precursors.

[149] FIG. 7 is a schematic of exemplary biochemical pathways leading to 7-hydroxyheptanoate using pimelate, pimeloyl-CoA, or pimelate semialdehyde as precursors.

[150] FIG. 8 is a schematic of an exemplary biochemical pathway leading to 1,7-heptanediol using 7-hydroxyheptanoate as a precursor.

[151] FIGs. 9A-9D contain the amino acid sequences of an *Escherichia coli thioesterase* encoded by *tesB* (see GenBank Accession No. AAA24665.1, SEQ ID NO: 1), a *Mycobacterium marinum carboxylate reductase* (see Genbank Accession No. ACC40567.1, SEQ ID NO: 2), a *Mycobacterium smegmatis carboxylate reductase* (see Genbank Accession No. ABK71854.1, SEQ ID NO: 3), a *Segniliparus rugosus carboxylate reductase* (see Genbank Accession No. EFV11917.1, SEQ ID NO: 4), a *Mycobacterium smegmatis carboxylate reductase* (see Genbank Accession No. ABK75684.1, SEQ ID NO: 5), a *Mycobacterium massiliense carboxylate reductase* (see Genbank Accession No. EIV11143.1, SEQ ID NO: 6), a *Segniliparus rotundus carboxylate reductase* (see Genbank Accession No. ADG98140.1, SEQ ID NO: 7), a *Chromobacterium violaceum ω-transaminase* (see Genbank Accession No. AAQ59697.1, SEQ ID NO: 8), a *Pseudomonas aeruginosa ω-transaminase* (see Genbank Accession No. AAG08191.1, SEQ ID NO: 9), a *Pseudomonas syringae ω-transaminase* (see Genbank Accession No. AAY39893.1, SEQ ID NO: 10), a *Rhodobacter*

sphaeroides ω-transaminase (see Genbank Accession No. ABA81135.1, SEQ ID NO: 11), an *Escherichia coli ω-transaminase* (see RefSeq Accession No. NP_417544.5, SEQ ID NO: 12), a *Vibrio fluvialis ω-transaminase* (see Genbank Accession No. AEA39183.1, SEQ ID NO: 13), a *Bacillus subtilis phosphopantetheinyl transferase* (see RefSeq Accession No. WP_003234549.1, SEQ ID NO:14), a *Nocardia sp.* NRRL 5646 *phosphopantetheinyl transferase* (see Genbank Accession No. ABI83656.1, SEQ ID NO:15), a *Bacillus cereus malonyl-CoA methyltransferase* (see GenBank Accession No. AAS43086.1, SEQ ID NO: 16), an *Escherichia coli pimelyl-ACP methyl ester esterase* (see GenBank Accession No. AAC76437.1, SEQ ID NO:17), a *6-carboxyhexanoate-CoA ligase* (BioW) (EC:6.2.1.14, see GenBank Accession No. AAB17457.1, SEQ ID NO:18), a *8-amino-7-oxononanoate synthase* (BioF) (EC:2.3.1.47, see GenBank Accession No. AAB17459.1, SEQ ID NO: 19), a *SAM-dependent malonyl-ACP O-methyltransferas* (BioC) (EC:2.1.1.197, see RefSeq Accession No. NP_415298.1, SEQ ID NO:20), a *thioesterase* encoded by *tesA* or *fatB* (see GenBank Accession No. ABJ63754.1, SEQ ID NO:21; see GenBank Accession No. CCC78182.1, SEQ ID NO:22), and a *8-amino-7-oxononanoate synthase* (BioF, or ACP-dependent BioF) (EC:2.3.1.47, see GenBank Accession No. AAA23516.1, SEQ ID NO: 23).

[152] FIG. 10 is a bar graph summarizing the change in absorbance at 340 nm after 20 minutes, which is a measure of the consumption of NADPH and activity of *carboxylate reductases* relative to the enzyme only controls (no substrate).

[153] FIG. 11 is a bar graph of the change in absorbance at 340 nm after 20 minutes, which is a measure of the consumption of NADPH and the activity of *carboxylate reductases* for converting pimelate to pimelate semialdehyde relative to the empty vector control.

[154] FIG. 12 is a bar graph of the change in absorbance at 340 nm after 20 minutes, which is a measure of the consumption of NADPH and the activity of *carboxylate reductases* for converting 7-hydroxyheptanoate to 7-hydroxyheptanal relative to the empty vector control.

[155] FIG. 13 is a bar graph of the change in absorbance at 340 nm after 20 minutes, which is a measure of the consumption of NADPH and the activity of *carboxylate reductases* for converting N7-acetyl-7-aminoheptanoate to N7-acetyl-7-aminoheptanal relative to the empty vector control.

[156] FIG. 14 is a bar graph of the change in absorbance at 340 nm after 20 minutes, which is a measure of the consumption of NADPH and activity of *carboxylate*

reductases for converting pimelate semialdehyde to heptanedial relative to the empty vector control.

[157] FIG. 15 is a bar graph summarizing the percent conversion after 4 hours of pyruvate to L-alanine (mol/mol) as a measure of the *ω-transaminase* activity of the enzyme only controls (no substrate).

[158] FIG. 16 is a bar graph of the percent conversion after 4 hours of pyruvate to L-alanine (mol/mol) as a measure of the *ω-transaminase* activity for converting 7-aminoheptanoate to pimelate semialdehyde relative to the empty vector control.

[159] FIG. 17 is a bar graph of the percent conversion after 4 hours of L-alanine to pyruvate (mol/mol) as a measure of the *ω-transaminase* activity for converting pimelate semialdehyde to 7-aminoheptanoate relative to the empty vector control.

[160] FIG. 18 is a bar graph of the percent conversion after 4 hours of pyruvate to L-alanine (mol/mol) as a measure of the *ω-transaminase* activity for converting heptamethylenediamine to 7-aminoheptanal relative to the empty vector control.

[161] FIG. 19 is a bar graph of the percent conversion after 4 hours of pyruvate to L-alanine (mol/mol) as a measure of the *ω-transaminase* activity for converting N7-acetyl-1,7-diaminoheptane to N7-acetyl-7-aminoheptanal relative to the empty vector control.

[162] FIG. 20 is a bar graph of the percent conversion after 4 hours of pyruvate to L-alanine (mol/mol) as a measure of the *ω-transaminase* activity for converting 7-aminoheptanol to 7-oxoheptanol relative to the empty vector control.

[163] FIG. 21 is a table of the conversion after 1 hour of pimeloyl-CoA methyl ester to pimeloyl-CoA by a pimeloyl-[acp] methyl ester methylesterase.

DETAILED DESCRIPTION

[164] This disclosure provides enzymes, non-natural pathways, cultivation strategies, feedstocks, host microorganisms and attenuations to the host's biochemical network, which generates a C7 aliphatic backbone from central metabolites in which two terminal functional groups may be formed leading to the synthesis of pimelic acid, 7-aminoheptanoic acid, heptamethylenediamine, 7-aminohelptanol or 1,7-heptanediol (referred to as "C7 building blocks" herein). As used herein, the term "central precursor" is used to denote any metabolite in any metabolic pathway shown herein leading to the synthesis of a C7 building block. The term "central metabolite" is used herein to denote a metabolite that is produced in all microorganisms to support growth.

[165] The term "*bioF*" refers to any one of a number of *bioF* genes well known in the art and present in multiple organisms. The term "BioF" refers to a protein encoded by any one of such genes. Numerous *bioF* genes and BioF proteins from different organisms are well known in the art and can be easily identified in public databases such as GenBank, ExPASy, and via Enzyme Commission numbers.

[166] Host microorganisms described herein can include endogenous pathways that can be manipulated such that one or more C7 building blocks 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.

[167] 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 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 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 acid 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 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.

[168] 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 host of the 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.

[169] 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 in addition to a *malonyl-ACP O-methyltransferase*; a *β-ketoacyl-[acp] synthase*, a *β-ketothiolase*, a *3-oxoacyl-[acp] reductase*, *acetoacetyl-CoA reductase*, a *3-hydroxyacyl-CoA dehydrogenase*, a *3-hydroxybutyryl-CoA dehydrogenase*, an *enoyl-CoA hydratase*, *3-hydroxyacyl-ACP dehydratase*, an *enoyl-ACP reductase*, a *trans-2-enoyl-CoA reductase*, a *thioesterase*, a reversible *CoA ligase*, a *CoA-transferase*, an *acetylating aldehyde dehydrogenase*, a *6-oxohexanoate dehydrogenase*, a *7-oxoheptanoate dehydrogenase*, an *aldehyde dehydrogenase*, a *carboxylate reductase*, a *ω-transaminase*, a *N-acetyl transferase*, an *alcohol dehydrogenase*, a *deacetylase*, a *6-hydroxyhexanoate dehydrogenase*, a *5-hydroxypentanoate dehydrogenase*, or a *4-hydroxybutyrate dehydrogenase*. In recombinant hosts expressing a *carboxylate reductase*, a *phosphopantetheinyl transferase* also can be expressed as it enhances activity of the *carboxylate reductase*.

[170] For example, a recombinant host can include at least one exogenous nucleic acid encoding (i) a malonyl-ACP O-methyltransferase, (ii) a β-ketoacyl-ACP synthase or a β-ketothiolase, (iii) a 3-oxoacyl-ACP reductase, acetoacetyl-CoA reductase, a 3-hydroxyacyl-CoA dehydrogenase or a 3-hydroxybutyryl-CoA dehydrogenase, (iv) an enoyl-CoA hydratase or 3-hydroxyacyl-ACP dehydratase, (v) an enoyl-ACP reductase or a trans-2-enoyl-CoA reductase and and produce pimeloyl-ACP or pimeloyl-CoA.

[171] Such recombinant hosts producing pimeloyl-ACP or pimeloyl-CoA further can include at least one exogenous nucleic acid encoding one or more of a *thioesterase*, an *aldehyde dehydrogenase*, a *7-oxoheptanoate dehydrogenase*, a *6-oxohexanoate dehydrogenase*, a *glutaconate CoA-transferase*, a reversible *succinyl-CoA ligase*, an *acetylating aldehyde dehydrogenase*, or a *carboxylate reductase* and produce pimelic acid or pimelate semialdehyde. For example, a recombinant host producing pimeloyl-ACP or pimeloyl-CoA further can include a *thioesterase*, a reversible *Co-ligase* (e.g., a reversible

succinyl-CoA ligase), or a *CoA transferase* (e.g., a *glutaconate CoA-transferase*) and produce pimelic acid. For example, a recombinant host producing pimeloyl-CoA further can include an *acetylating aldehyde dehydrogenase* and produce pimelate semilaldehyde. For example, a recombinant host producing pimelate further can include a *carboxylate reductase* and produce pimelate semialdehyde.

[172] In some embodiments, the recombinant host can comprise at least one exogenous nucleic acid encoding *bioW* and a CoA-specific *bioF*.

[173] A recombinant hosts producing pimelate semialdehyde further can include at least one exogenous nucleic acid encoding a *ω-transaminase* and produce 7-aminoheptanoate. In some embodiments, a recombinant host producing pimeloyl-CoA includes a *carboxylate reductase* and a *ω-transaminase* to produce 7-aminoheptanoate.

[174] A recombinant host producing pimelate or pimelate semialdehyde further can include at least one exogenous nucleic acid encoding a *6-hydroxyhexanoate dehydrogenase*, a *5-hydroxypentanoate dehydrogenase* or a *4-hydroxybutyrate dehydrogenase*, and produce 7-hydroxyheptanoic acid. In some embodiments, a recombinant host producing pimeloyl-CoA includes an *acetylating aldehyde dehydrogenase*, and a *6-hydroxyhexanoate dehydrogenase*, a *5-hydroxypentanoate dehydrogenase* or a *4-hydroxybutyrate dehydrogenase* to produce 7-hydroxyheptanoate. In some embodiments, a recombinant host producing pimelate includes a *carboxylate reductase* and a *6-hydroxyhexanoate dehydrogenase*, a *5-hydroxypentanoate dehydrogenase* or a *4-hydroxybutyrate dehydrogenase* to produce 7-hydroxyheptanoate.

[175] A recombinant hosts producing 7-aminoheptanoate, 7-hydroxyheptanoate or pimelate semialdehyde further can include at least one exogenous nucleic acid encoding a *ω-transaminase*, a *deacetylase*, a *N-acetyl transferase*, or an *alcohol dehydrogenase*, and produce heptamethylenediamine. For example, a recombinant host producing 7-hydroxyheptanoate can include a *carboxylate reductase* with a *phosphopantetheine transferase* enhancer, a *ω-transaminase* and an *alcohol dehydrogenase*.

[176] A recombinant host producing 7-hydroxyheptanoic acid further can include one or more of a *carboxylate reductase* with a *phosphopantetheine transferase* enhancer and an *alcohol dehydrogenase*, and produce 1,7-heptanediol.

[177] Within an engineered pathway, the enzymes can be from a single source, i.e., from one species or genus, or can be from multiple sources, i.e., 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.

[178] Any of the enzymes described herein that can be used for production of one or more C7 building blocks can have at least 50%, 60% or 70% sequence identity or homology (e.g., at least 50%, 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. 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).

[179] The percent identity and 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 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 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 following for nucleic acid sequences except that blastn is used.

[180] 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 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 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.

[181] When percentage of sequence identity is used in reference to proteins it is recognized that residue positions which are not identical often differ by conservative amino

acid substitutions, where amino acid residues are substituted for other amino acid residues with similar chemical properties (e.g., charge or hydrophobicity) and therefore do not change the functional properties of the molecule. Where sequences differ in conservative substitutions, the percent sequence identity may be adjusted upwards to correct for the conservative nature of the substitution and this process results in "sequence homology" of, e.g., at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99%. Means for making this adjustment are well known to those of skill in the art. Typically this involves scoring a conservative substitution as a partial rather than a full mismatch, thereby increasing the percentage sequence identity. Thus, for example, where an identical amino acid is given a score of 1 and a non-conservative substitution is given a score of zero, a conservative substitution is given a score between zero and 1. The scoring of conservative substitutions is calculated, e.g., according to the algorithm of Meyers and Miller, Computer. Applic. Biol. Sci., 1988, 4, 11-17, e.g., as implemented in the program PC/GENE (Intelligenetics, Mountain View, Cailf., USA). This alignment and the percent homology or identity can be determined using any suitable software program known in the art, for example those described in CURRENT PROTOCOLS IN MOLECULAR BIOLOGY (F. M. Ausubel et al. (eds) 1987, Supplement 30, section 7.7.18). Such programs may include the GCG Pileup program, FASTA (Pearson et al., Proc. Natl. Acad. Sci. USA, 1988, 85, 2444-2448), and BLAST (BLAST Manual, Altschul et al., Nat'l Cent. Biotechnol. Inf., Nat'l Lib. Med. (NCIB NLM NIH), Bethesda, Md., and Altschul et al., NAR, 1997, 25, 3389-3402). Another alignment program is ALIGN Plus (Scientific and Educational Software, Pa.), using default parameters. Another sequence software program that finds use is the TFASTA Data Searching Program available in the Sequence Software Package Version 6.0 (Genetics Computer Group, University of Wisconsin, Madison, Wis.).

[182] 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.

[183] For example, a *8-amino-7-oxononanoate synthase* (BioF, or ACP-dependent BioF) described herein can have at least 70% sequence identity or homology (e.g., at least 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, or 100%) to the amino acid sequence of an *Escherichia coli 8-amino-7-oxononanoate synthase* (BioF) (EC:2.3.1.47, see GenBank Accession No. AAA23516.1, SEQ ID NO: 23), See FIG. 9D.

[184] For example, a *thioesterase* (TE) described herein can have at least 70% sequence identity or homology (e.g., at least 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, or 100%) to the amino acid sequence of an *Escherichia coli thioesterase* encoded by *tesB* (see GenBank Accession No. AAA24665.1, SEQ ID NO: 1), or the gene products encoded by *tesA* or *fatB* (see GenBank Accession No. ABJ63754.1, SEQ ID NO:21; see GenBank Accession No. CCC78182.1, SEQ ID NO:22). See FIGs. 9A and 9D.

[185] For example, a *carboxylate reductase* described herein can have at least 70% sequence identity or homology (e.g., at least 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, or 100%) to the amino acid sequence of a *Mycobacterium marinum* (see Genbank Accession No. ACC40567.1, SEQ ID NO: 2), a *Mycobacterium smegmatis* (see Genbank Accession No. ABK71854.1, SEQ ID NO: 3), a *Segniliparus rugosus* (see Genbank Accession No. EFV11917.1, SEQ ID NO: 4), a *Mycobacterium smegmatis* (see Genbank Accession No. ABK75684.1, SEQ ID NO: 5), a *Mycobacterium massiliense* (see Genbank Accession No. EIV11143.1, SEQ ID NO: 6), or a *Segniliparus rotundus* (see Genbank Accession No. ADG98140.1, SEQ ID NO: 7) *carboxylate reductase*. See, FIGs. 9A-9C.

[186] For example, a *ω-transaminase* described herein can have at least 70% sequence identity or homology (e.g., at least 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, or 100%) to the amino acid sequence of a *Chromobacterium violaceum* (see Genbank Accession No. AAQ59697.1, SEQ ID NO: 8), a *Pseudomonas aeruginosa* (see Genbank Accession No. AAG08191.1, SEQ ID NO: 9), a *Pseudomonas syringae* (see Genbank Accession No. AAY39893.1, SEQ ID NO: 10), a *Rhodobacter sphaeroides* (see Genbank Accession No. ABA81135.1, SEQ ID NO: 11), an *Escherichia coli* (see RefSeq Accession No. NP_417544.5, SEQ ID NO: 12), or a *Vibrio fluvialis* (see Genbank Accession No. AEA39183.1, SEQ ID NO: 13) *ω-transaminase*. Some of these *ω-transaminases* are *diamine ω-transaminases*. See, FIGs. 9C and 9D.

[187] For example, a *phosphopantetheinyl transferase* described herein can have at least 70% sequence identity or homology (e.g., at least 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, or 100%) to the amino acid sequence of a *Bacillus subtilis phosphopantetheinyl transferase* (see RefSeq Accession No. WP_003234549.1, SEQ ID NO:14) or a *Nocardia sp.* NRRL 5646 *phosphopantetheinyl transferase* (see Genbank Accession No. ABI83656.1, SEQ ID NO:15). See FIG. 7.

[188] For example, a *pimeloyl-ACP methyl ester esterase* described herein can have at least 70% sequence identity or homology (e.g., at least 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, or 100%) to the amino acid sequence of an *Escherichia coli pimeloyl-ACP methyl ester esterase* (BioH) (see GenBank Accession No. AAC76437.1, SEQ ID NO:17). See, FIG. 9D.

[189] For example, a *6-carboxyhexanoate-CoA ligase* described herein can have at least 70% sequence identity or homology (e.g., at least 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, or 100%) to the amino acid sequence of an *Bacillus subtilis* (strain 168) *6-carboxyhexanoate-CoA ligase* (BioW) (EC:6.2.1.14, see GenBank Accession No. AAB17457.1, SEQ ID NO:18). See, FIG. 9D.

[190] For example, a *8-amino-7-oxononanoate synthase* described herein can have at least 70% sequence identity or homology (e.g., at least 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, or 100%) to the amino acid sequence of an *Bacillus subtilis* (strain 168) *8-amino-7-oxononanoate synthase* (BioF) (EC:2.3.1.47, see GenBank Accession No. AAB17459.1, SEQ ID NO: 19). See, FIG. 9D. Other *bioF* genes and BioF proteins from different hosts are well known in the art and can be easily identified in public databases such as GenBank, ExPASy, and via Enzyme Commission numbers.

[191] For example, a *SAM-dependent malonyl-ACP O-methyltransferase* described herein can have at least 70% sequence identity or homology (e.g., at least 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, or 100%) to the amino acid sequence of a *Escherichia coli malonyl-ACP O-methyltransferase* (BioC) (EC:2.1.1.197, see RefSeq Accession No. NP_415298.1, SEQ ID NO: 20). See, FIG. 9D.

[192] 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.

[193] Functional fragments of any of the enzymes described herein can also be used in the methods of the disclosure. 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.

[194] This disclosure also provides (i) functional variants of the enzymes used in the methods of the disclosure 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.

[195] 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: 24)), 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.

[196] 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 *methyltransferase*, a *synthase*, *β-ketothiolase*, a *dehydratase*, a *hydratase*, a *dehydrogenase*, a *methylesterase*, a *thioesterase*, a *reversible CoA-ligase*, a *CoA-transferase*, a *reductase*, *deacetylase*, *N-acetyl transferase* or a *ω-transaminase* as described in more detail below.

[197] In addition, the production of one or more C7 building blocks 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 generating the C7 aliphatic backbone for conversion to C7 Building Blocks

[198] As depicted in FIG. 3, the first step in biotin biosynthesis is the methylation of malonyl-ACP by BioC (a *S-adenosyl-L-methionine (SAM)-dependent methyltransferase*). The malonyl-ACP methyl ester thus generated then serves as a starter unit for the fatty acid biosynthesis pathway, two rounds of fatty acid elongation and reduction then occur (catalysed

by FabB, F, G, Z and I) to generate pimelyl-ACP methyl ester. Pimelyl-ACP is then generated by the removal of the methyl group by the esterase BioH. Biotin biosynthesis then continues via a series of reactions catalysed by the enzymes BioF, BioA, BioB and BioD. One of the intermediates in biotin biosynthesis, pimelyl-ACP, can also be converted to 7-aminoheptanoic acid (7-AHA) in a synthetic metabolic pathway. In this pathway a thioesterase (TE) is used to release pimelic acid from pimelyl-ACP. Then carboxylic acid reductase (CAR) reduces the free pimelic acid to the cognate semi-aldehyde, with the final catalytic step being the amination of the semi-aldehyde by the ω-transaminase (ω-TAM) to produce 7-AHA. The export of 7-AHA from the cell then may be facilitated by the transport protein LysE.

[199] The native biotin pathway thus provides a potential source of the pimelic acid intermediate that can be subsequently metabolized to 7-AHA. As biotin is only required in trace amounts the natural flux through this pathway is very low. Initial attempts to increase the flux to pimelic acid in E. coli have yielded only trace amounts of pimelic acid (<0.1 ppm), due in part to complex regulation of the biotin pathway and the toxicity associated with overexpression of pathway enzymes. In order to bypass these issues strains of E. coli with increased flux through the biotin pathway were sourced. A series of strains, generated by multiple rounds of random mutagenesis, have been reported to produce almost 1 g/L of biotin (US patent 6,284,500 B1). These strains were used in order to provide a genetic background in which flux through the native biotin pathway is high with feedback and other negative regulatory mechanisms non-functional or much reduced.

[200] In some embodiments, a methyl ester shielded carbon chain elongation associated with biotin biosynthesis route comprises using a *malonyl-ACP O-methyltransferase* to form a malonyl-ACP methyl ester, and then performing two cycles of carbon chain elongation using a *β-ketoacyl-ACP synthase*, a *3-oxoacyl-ACP reductase*, a *3-hydroxyacyl-ACP dehydratase*, and an *enoyl-ACP reductase*. A *pimeloyl-ACP methyl ester esterase* can be used to cleave the resulting pimeloyl-ACP methyl ester.

[201] In some embodiments, a *methyltransferase* can be a *malonyl-ACP O-methyltransferase* classified, for example, under EC 2.1.1.197 such as the gene product of *bioC* from *Bacillus cereus* (see Genbank Accession No. AAS43086.1, SEQ ID NO:16) (see, for example, Lin, 2012, *Biotin Synthesis in Escherichia coli*, Ph.D. Dissertation, University of Illinois at Urbana-Champaign).

[202] In some embodiments, a *β-ketoacyl-ACP synthase* may be classified, for example, under EC 2.3.1.- (e.g., EC 2.3.1.41, EC 2.3.1.179 or EC 2.3.1.180) such as the gene product of *fabB*, *fabF*, or *fabH*.

[203] In some embodiments, a *3-oxoacyl-ACP reductase* may be classified under EC 1.1.1.100, such as the gene product of *fabG*.

[204] In some embodiments, an *enoyl-ACP dehydratase* such as a *3-hydroxyacyl-ACP dehydratase* may be classified under EC 4.2.1.59, such as the gene product of *fabZ*.

[205] In some embodiments, a *trans-2-enoyl-CoA reductase* may be classified under EC 1.3.1.- (e.g., EC 1.3.1.38, EC 1.3.1.8, EC 1.3.1.44), such as the gene product of *ter* (Nishimaki *et al.*, *J. Biochem.*, 1984, 95, 1315 – 1321; Shen *et al.*, 2011, *supra*) or *tdter* (Bond-Watts *et al.*, *Biochemistry*, 2012, 51, 6827 – 6837).

[206] In some embodiments, an *enoyl-ACP reductase* may be classified under EC 1.3.1.10 such as the gene product of *fabI*.

[207] In some embodiments, a *pimeloyl-ACP methyl ester esterase* may be classified, for example, under EC 3.1.1.85 such as the gene product of *bioH* from *E. coli*. See Genbank Accession No. AAC76437.1, SEQ ID NO:17.

Enzymes generating the terminal carboxyl groups in the biosynthesis of C7 Building Blocks

[208] As depicted in FIG. 1, a terminal carboxyl group can be enzymatically formed using a *thioesterase* (TE).

[209] In some embodiments, the second terminal carboxyl group leading to the synthesis of a C7 building block is enzymatically formed by a *thioesterase* classified, for example, under EC 3.1.2.-, such as the gene product of *YciA*, *tesB* (Genbank Accession No. AAA24665.1, SEQ ID NO: 1) or the gene products encoded by *tesA* or *fatB* (see GenBank Accession No. ABJ63754.1, SEQ ID NO:21; see GenBank Accession No. CCC78182.1, SEQ ID NO:22) (see, for example, Cantu *et al.*, *Protein Science*, 2010, 19, 1281 – 1295; Zhuang *et al.*, *Biochemistry*, 2008, 47(9), 2789 – 2796; or Naggert *et al.*, *J. Biol. Chem.*, 1991, 266(17), 11044 – 11050).

Enzymes generating the terminal amine groups in the biosynthesis of C7 Building Blocks

[210] As depicted in FIG. 5 and FIGs. 6A-6C, terminal amine groups can be enzymatically formed using a *ω-transaminase* or a *deacetylase*.

[211] In some embodiments, the first or second terminal amine group leading to the synthesis of 7-aminoheptanoic acid is enzymatically formed by a *ω-transaminase* classified, for example, under EC 2.6.1.18, EC 2.6.1.19, EC 2.6.1.29, EC 2.6.1.48, or EC 2.6.1.82 such as that obtained from *Chromobacterium violaceum* (Genbank Accession No. AAQ59697.1, SEQ ID NO: 8), *Pseudomonas aeruginosa* (Genbank Accession No. AAG08191.1, SEQ ID NO: 9), *Pseudomonas syringae* (Genbank Accession No. AAY39893.1, SEQ ID NO: 10), *Rhodobacter sphaeroides* (Genbank Accession No. ABA81135.1, SEQ ID NO: 11), *Escherichia coli* (RefSeq Accession No. NP_417544.5, SEQ ID NO: 12), *Vibrio Fluvialis* (Genbank Accession No. AEA39183.1, SEQ ID NO: 13), *Streptomyces griseus*, or *Clostridium viride*. Some of these *ω-transaminases* are *diamine ω-transaminases* (e.g., SEQ ID NO: 12). For example, the *ω-transaminases* classified, for example, under EC 2.6.1.29 or EC 2.6.1.82 may be *diamine ω-transaminases*.

[212] The reversible *ω-transaminase* from *Chromobacterium violaceum* (Genbank Accession No. AAQ59697.1, SEQ ID NO: 8) has demonstrated analogous activity accepting 6-aminohexanoic acid as amino donor, thus forming the first terminal amine group in adipate semialdehyde (Kaulmann *et al.*, *Enzyme and Microbial Technology*, 2007, 41, 628 – 637).

[213] The reversible *4-aminobubyrate:2-oxoglutarate transaminase* from *Streptomyces griseus* has demonstrated analogous activity for the conversion of 6-aminohexanoate to adipate semialdehyde (Yonaha *et al.*, *Eur. J. Biochem.*, 1985, 146:101 - 106).

[214] The reversible *5-aminovalerate transaminase* from *Clostridium viride* has demonstrated analogous activity for the conversion of 6-aminohexanoate to adipate semialdehyde (Barker *et al.*, *J. Biol. Chem.*, 1987, 262(19), 8994 – 9003).

[215] In some embodiments, a terminal amine group leading to the synthesis of 7-aminoheptanoate or heptamethylenediamine is enzymatically formed by a *diamine ω-transaminase*. For example, the second terminal amino group can be enzymatically formed by a *diamine ω-transaminase* classified, for example, under EC 2.6.1.29 or classified, for example, under EC 2.6.1.82, such as the gene product of *YgjG* from *E. coli* (RefSeq Accession No. NP_417544.5, SEQ ID NO: 12).

[216] The gene product of *ygjG* accepts a broad range of diamine carbon chain length substrates, such as putrescine, cadaverine and spermidine (see, for example, Samsonova *et al.*, *BMC Microbiology*, 2003, 3:2).

[217] The *diamine ω-transaminase* from *E.coli* strain B has demonstrated activity for 1,7 diaminoheptane (Kim, *The Journal of Chemistry*, 1964, 239(3), 783 – 786).

[218] In some embodiments, the second terminal amine group leading to the synthesis of heptamethylenediamine is enzymatically formed by a *deacetylase* such as *acetylputrescine deacetylase* classified, for example, under EC 3.5.1.62. The *acetylputrescine deacetylase* from *Micrococcus luteus* K-11 accepts a broad range of carbon chain length substrates, such as acetylputrescine, acetylcadaverine and N^8.acetylspermidine (see, for example, Suzuki *et al.*, 1986, *BBA – General Subjects*, 882(1):140 – 142).

Enzymes generating the terminal hydroxyl groups in the biosynthesis of C7 Building Blocks

[219] As depicted in FIG. 7 and FIG. 8, a terminal hydroxyl group can be enzymatically formed using an *alcohol dehydrogenase*.

[220] In some embodiments, a terminal hydroxyl group leading to the synthesis of 1,7 heptanediol is enzymatically formed by an *alcohol dehydrogenase* classified, for example, under EC 1.1.1.- (e.g., 1, 2, 21, or 184) such as the gene product of *YMR318C* (classified, for example, under EC 1.1.1.2, see Genbank Accession No. CAA90836.1) (Larroy *et al.*, 2002, *Biochem J.*, 361(Pt 1), 163 – 172), the gene product of *YghD*, the gene product of *cpnD* (Iwaki et al., 2002, *Appl. Environ. Microbiol.*, 68(11):5671 – 5684), the gene product of *gbd*, or a *6-hydroxyhexanoate dehydrogenase* classified, for example, under EC 1.1.1.258 such as the gene product of *ChnD* (Iwaki *et al.*, *Appl. Environ. Microbiol.*, 1999, *supra*).

Biochemical pathways

Pathways from malonyl-ACP to pimeloyl-ACP as central precursor leading to C7 Building Block

[221] As depicted in FIG. 1 and FIG. 3, the first step in biotin biosynthesis is the methylation of malonyl-ACP by BioC (a *S-adenosyl-L-methionine (SAM)-dependent methyltransferase*). The malonyl-ACP methyl ester thus generated then serves as a starter unit for the fatty acid biosynthesis pathway, two rounds of fatty acid elongation and reduction then occur (catalysed by FabB, F, G, Z and I) to generate pimelyl-ACP methyl ester. Pimelyl-ACP is then generated by the removal of the methyl group by the esterase BioH. Biotin biosynthesis then continues via a series of reactions catalysed by the enzymes BioF, BioA, BioB and BioD. One of the intermediates in biotin biosynthesis, pimelyl-ACP, can also be converted to 7-aminoheptanoic acid (7-AHA) in a synthetic metabolic pathway. In this pathway a thioesterase (TE) is used to release pimelic acid from pimelyl-ACP. Then

carboxylic acid reductase (CAR) reduces the free pimelic acid to the cognate semi-aldehyde, with the final catalytic step being the amination of the semi-aldehyde by the ω-transaminase (ω-TAM) to produce 7-AHA.The export of 7-AHA from the cell then may be facilitated by the transport protein LysE.

[222] The native biotin pathway thus provides a potential source of the pimelic acid intermediate that can be subsequently metabolized to 7-AHA. As biotin is only required in trace amounts the natural flux through this pathway is very low. Initial attempts to increase the flux to pimelic acid in E. coli have yielded only trace amounts of pimelic acid (<0.1 ppm), due in part to complex regulation of the biotin pathway and the toxicity associated with overexpression of pathway enzymes. In order to bypass these issues strains of E. coli with increased flux through the biotin pathway were sourced. A series of strains, generated by multiple rounds of random mutagenesis, have been reported to produce almost 1 g/L of biotin (US patent 6,284,500 B1). These strains were used in order to provide a genetic background in which flux through the native biotin pathway is high with feedback and other negative regulatory mechanisms non-functional or much reduced.

Pathways using pimeloyl-ACP as central precursor to pimelic acid

[223] In some embodiments, pimelic acid is synthesized from the central precursor, pimeloyl-ACP, by conversion of pimeloyl-ACP to pimelate by a *thioesterase* classified, for example, under EC 3.1.2.- such as the gene products encoded by *tesA* or *fatB* (Genbank Accession No. ABJ63754.1, SEQ ID NO:21; Genbank Accession No. CCC78182.1, SEQ ID NO:22). See FIG. 3.

Pathways using pimelate semialdehyde as precursor to 7-aminoheptanoate

[224] In some embodiments, 7-aminoheptanoate is synthesized from the central precursor, pimelate, by conversion of pimelate to pimelate semialdehyde by a *carboxylate reductase* classified, for example, under EC 1.2.99.6 such as the gene product of *car* in combination with a *phosphopantetheine transferase* enhancer (e.g., encoded by a *sfp* (RefSeq Accession No. WP_003234549.1, SEQ ID NO:14) gene from *Bacillus subtilis* or *npt* (Genbank Accession No. ABI83656.1, SEQ ID NO:15) gene from *Nocardia*) or the gene products of *GriC* and *GriD* from *Streptomyces griseus* (Suzuki *et al.*, *J. Antibiot.*, 2007, 60(6), 380 – 387); followed by conversion of pimelate semialdehyde to 7-aminoheptanoate by a *ω-transaminase* (e.g., EC 2.6.1.18, EC 2.6.1.19, EC 2.6.1.48, EC 2.6.1.29, EC 2.6.1.82 such as SEQ ID NOs:8-13). The *carboxylate reductase* can be obtained, for example, from

Mycobacterium marinum (Genbank Accession No. ACC40567.1, SEQ ID NO: 2), *Mycobacterium smegmatis* (Genbank Accession No. ABK71854.1, SEQ ID NO: 3), *Segniliparus rugosus* (Genbank Accession No. EFV11917.1, SEQ ID NO: 4), *Mycobacterium smegmatis* (Genbank Accession No. ABK75684.1, SEQ ID NO: 5), *Mycobacterium massiliense* (Genbank Accession No. EIV11143.1, SEQ ID NO: 6), or *Segniliparus rotundus* (Genbank Accession No. ADG98140.1, SEQ ID NO: 7). See FIG. 5.

Pathway using 7-aminoheptanoate, 7-hydroxyheptanoate or pimelate semialdehyde as a precursor to heptamethylenediamine

[225] In some embodiments, heptamethylenediamine is synthesized from the central precursor, 7-aminoheptanoate, by conversion of 7-aminoheptanoate to 7-aminoheptanal by a *carboxylate reductase* classified, for example, under EC 1.2.99.6 such as the gene product of *car* (see above) in combination with a *phosphopantetheine transferase* enhancer (e.g., encoded by a *sfp* (RefSeq Accession No. WP_003234549.1, SEQ ID NO:14) gene from *Bacillus subtilis* or *npt* (Genbank Accession No. ABI83656.1, SEQ ID NO:15) gene from *Nocardia*) or the gene product of *GriC & GriD*; followed by conversion of 7-aminoheptanal to heptamethylenediamine by a *ω-transaminase* (e.g., classified, for example, under EC 2.6.1.18, EC 2.6.1.19, EC 2.6.1.29, EC 2.6.1.48, or EC 2.6.1.82 such as SEQ ID NOs:8-13, see above). See FIG. 6A.

[226] The *carboxylate reductase* encoded by the gene product of *car* and the *phosphopantetheine transferase* enhancer *npt* or *sfp* has broad substrate specificity, including terminal difunctional C4 and C5 carboxylic acids (Venkitasubramanian *et al., Enzyme and Microbial Technology*, 2008, 42, 130 – 137).

[227] In some embodiments, heptamethylenediamine is synthesized from the central precursor, 7-hydroxyheptanoate (which can be produced as described in FIG. 5), by conversion of 7-hydroxyheptanoate to 7-hydroxyheptanal by a *carboxylate reductase* classified, for example, under EC 1.2.99.6 such as the gene product of *car* (see above) in combination with a *phosphopantetheine transferase* enhancer (e.g., encoded by a *sfp* gene from *Bacillus subtilis* or *npt* gene from *Nocardia*) or the gene product of *GriC & GriD* (Suzuki *et al., J. Antibiot.*, 2007, 60(6), 380 – 387); followed by conversion of 7-aminoheptanal to 7-aminoheptanol by a *ω-transaminase* classified, for example, under EC 2.6.1.18, EC 2.6.1.19, EC 2.6.1.29, EC 2.6.1.48, or EC 2.6.1.82 such as SEQ ID NOs:8-13, see above; followed by conversion to 7-aminoheptanal by an *alcohol dehydrogenase* classified, for example, under EC 1.1.1.- (e.g., EC 1.1.1.1, EC 1.1.1.2, EC 1.1.1.21, or EC

1.1.1.184) such as the gene product of *YMR318C* (classified, for example, under EC 1.1.1.2, see Genbank Accession No. CAA90836.1) or *YqhD* (from *E. coli*, GenBank Accession No. AAA69178.1) (Liu *et al.*, *Microbiology*, 2009, 155, 2078 – 2085; Larroy *et al.*, 2002, *Biochem J.*, 361(Pt 1), 163 – 172; Jarboe, 2011, *Appl. Microbiol. Biotechnol.*, 89(2), 249 - 257) or the protein having GenBank Accession No. CAA81612.1 (from *Geobacillus stearothermophilus*); followed by conversion to heptamethylenediamine by a *ω-transaminase* classified, for example, under EC 2.6.1.18, EC 2.6.1.19, EC 2.6.1.29, EC 2.6.1.48, or EC 2.6.1.82 such as SEQ ID NOs:8-13, see above. See FIG. 6A.

[228] In some embodiments, heptamethylenediamine is synthesized from the central precursor, 7-aminoheptanoate, by conversion of 7-aminoheptanoate to N7-acetyl-7-aminoheptanoate by a *N-acetyltransferase* such as a *lysine N-acetyltransferase* classified, for example, under EC 2.3.1.32; followed by conversion to N7-acetyl-7-aminoheptanal by a *carboxylate reductase* classified, for example, under EC 1.2.99.6 such as the gene product of *car* (see above) in combination with a *phosphopantetheine transferase* enhancer (e.g., encoded by a *sfp* gene from *Bacillus subtilis* or *npt* gene from *Nocardia*) or the gene product of *GriC* & *GriD*; followed by conversion to N7-acetyl-1,7-diaminoheptane by a *ω-transaminase* classified, for example, under EC 2.6.1.18, EC 2.6.1.19, EC 2.6.1.29, EC 2.6.1.48, EC 2.6.1.46, or EC 2.6.1.82 such as SEQ ID NOs:8-13, see above; followed by conversion to heptamethylenediamine by an *acetylputrescine deacylase* classified, for example, under EC 3.5.1.62. See, FIG. 6B.

[229] In some embodiments, heptamethylenediamine is synthesized from the central precursor, pimelate semialdehyde, by conversion of pimelate semialdehyde to heptanedial by a *carboxylate reductase* classified, for example, under EC 1.2.99.6 such as the gene product of *car* (see above) in combination with a *phosphopantetheine transferase* enhancer (e.g., encoded by a *sfp* gene from *Bacillus subtilis* or *npt* gene from *Nocardia*) or the gene product of *GriC* & *GriD*; followed by conversion to 7-aminoheptanal by a *ω-transaminase* classified, for example, under EC 2.6.1.18, EC 2.6.1.19, or EC 2.6.1.48; followed by conversion to heptamethylenediamine by a *ω-transaminase* classified, for example, under EC 2.6.1.18, EC 2.6.1.19, EC 2.6.1.29, EC 2.6.1.48, EC 2.6.1.46, or EC 2.6.1.82 such as SEQ ID NOs:8-13, see above. See FIG. 6C.

Pathways using pimelate or pimelate semialdehyde as central precursor to 1,7-heptanediol

[230] In some embodiments, 7-hydroxyheptanoate is synthesized from the central precursor, pimelate, by conversion of pimelate to pimelate semialdehyde by a *carboxylate reductase* classified, for example, under EC 1.2.99.6 such as the gene product of *car* (see above) in combination with a *phosphopantetheine transferase* enhancer (e.g., encoded by a *sfp* gene from *Bacillus subtilis* or *npt* gene from *Nocardia*) or the gene product of *GriC & GriD*; followed by conversion to 7-hydroxyheptanoate by a *dehydrogenase* classified, for example, under EC 1.1.1.- such as a *6-hydroxyhexanoate dehydrogenase* classified, for example, under EC 1.1.1.258 such as the gene from of *ChnD* or a *5-hydroxypentanoate dehydrogenase* classified, for example, under EC 1.1.1.- such as the gene product of *CpnD* (see, for example, Iwaki et al., 2002, *Appl. Environ. Microbiol.*, 68(11):5671 – 5684) or a *4-hydroxybutyrate dehydrogenase* such as *gbd*. See FIG. 7. Pimelate semialdehyde also can be produced from pimeloyl-CoA using an *acetylating aldehyde dehydrogenase* as described above. See, also FIG. 7.

[231] In some embodiments, 1,7 heptanediol is synthesized from the central precursor, 7-hydroxyheptanoate, by conversion of 7-hydroxyheptanoate to 7-hydroxyheptanal by a *carboxylate reductase* classified, for example, under EC 1.2.99.6 such as the gene product of *car* (see above) in combination with a *phosphopantetheine transferase* enhancer (e.g., encoded by a *sfp* gene from *Bacillus subtilis* or *npt* gene from *Nocardia*) or the gene product of *GriC & GriD*; followed by conversion of 7-hydroxyheptanal to 1,7 heptanediol by an *alcohol dehydrogenase* classified, for example, under EC 1.1.1.- such as EC 1.1.1.1, EC 1.1.1.2, EC 1.1.1.21, or EC 1.1.1.184) such as the gene product of *YMR318C* or *YqhD* (see, e.g., Liu *et al.*, *Microbiology*, 2009, 155, 2078 – 2085; Larroy *et al.*, 2002, *Biochem J.*, 361(Pt 1), 163 – 172; or Jarboe, 2011, *Appl. Microbiol. Biotechnol.*, 89(2), 249 - 257) or the protein having GenBank Accession No. CAA81612.1 (from *Geobacillus stearothermophilus*). See, FIG. 8.

Cultivation strategy

[232] In some embodiments, the cultivation strategy entails achieving an aerobic, anaerobic or micro-aerobic cultivation condition.

[233] In some embodiments, the cultivation strategy entails nutrient limitation such as nitrogen, phosphate or oxygen limitation.

[234] 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.

[235] In some embodiments, the principal carbon source fed to the fermentation in the synthesis of one or more C7 building blocks can derive from biological or non-biological feedstocks.

[236] In some embodiments, the biological feedstock can be, can include, 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.

[237] 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).

[238] 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, 2011, 293 – 298; Martin and Prather, *Journal of Biotechnology*, 2009, 139, 61 – 67).

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

[240] 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).

[241] 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).

[242] 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).

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

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

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

[246] 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).

[247] 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).

[248] 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).

[249] 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 C7 building blocks.

[250] 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 one or more C7 building blocks.

Metabolic engineering

[251] The present disclosure 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, 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.

[252] 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 disclosure 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 disclosure. 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.

[253] In addition, this disclosure 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.

[254] Also, this disclosure 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.

[255] This disclosure also recognizes that where an enzyme is shown to accept a particular co-factor, such as NADPH, or a 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 disclosure 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.

[256] 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.

[257] In some embodiments, the enzymes in the pathways outlined herein can be gene dosed (i.e., overexpressed by having a plurality of copies of the gene in the host organism), into the resulting genetically modified organism via episomal or chromosomal integration approaches.

[258] 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 a C7 building block.

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

[260] 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 a C7 building block.

[261] In some embodiments, the host microorganism's tolerance to high concentrations of at least one C7 building block can be improved through continuous cultivation in a selective environment.

[262] In some embodiments, where pathways require excess NADPH co-factor in the synthesis of a C7 building block, 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).

[263] 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). For example, avoiding dissipation of an NADH imbalance towards C7 building blocks, a NADPH-specific *glutamate dehydrogenase* can be attenuated.

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

[265] In some embodiments, a membrane-bound *enoyl-CoA reductases* can be solubilized via expression as a fusion protein to a small soluble protein such as a *maltose binding protein* (Gloerich et al., FEBS Letters, 2006, 580, 2092 – 2096).

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

[267] In some embodiments, a *L-alanine dehydrogenase* can be overexpressed in the host to regenerate L-alanine from pyruvate as an amino donor for *ω-transaminase* reactions.

[268] In some embodiments, a *L-glutamate dehydrogenase* specific for the co-factor used to achieve co-factor imbalance can be overexpressed in the host to regenerate L-glutamate from 2-oxoglutarate as an amino donor for *ω-transaminase* reactions. For example, promoting dissipation of the NADH imbalance towards C7 building blocks, a NADH-specific *glutamate dehydrogenase* can be overexpressed.

[269] In some embodiments, enzymes such as *pimeloyl-CoA dehydrogenase* classified under, EC 1.3.1.62; an *acyl-CoA dehydrogenase* classified, for example, under EC 1.3.8.7 or EC 1.3.8.1; and/or a *glutaryl-CoA dehydrogenase* classified, for example, under EC 1.3.8.6 that degrade central metabolites and central precursors leading to and including C7 building blocks can be attenuated.

[270] In some embodiments, endogenous enzymes activating C7 building blocks via Coenzyme A esterification such as *CoA-ligases* (e.g., a *pimeloyl-CoA synthetase*) classified under, for example, EC 6.2.1.14 can be attenuated.

[271] In some embodiments, a *methanol dehydrogenase* and a *formaldehyde dehydrogenase* can be overexpressed in the host to allow methanol catabolism via formate.

[272] In some embodiments, a *S-adenosylmethionine synthetase* can be overexpressed in the host to generate S-Adenosyl-L-methionine as a co-factor for *malonyl-ACP O-methyltransferase*.

[273] In some embodiments, the efflux of a C7 building block 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 a C7 building block.

[274] The efflux of heptamethylenediamine can be enhanced or amplified by overexpressing broad substrate range multidrug transporters such as *Blt* from *Bacillus subtilis* (Woolridge *et al.*, 1997, *J. Biol. Chem.*, 272(14):8864 – 8866); *AcrB* and *AcrD* from *Escherichia coli* (Elkins & Nikaido, 2002, *J. Bacteriol.*, 184(23), 6490 – 6499), NorA from *Staphylococcus aereus* (Ng et al., 1994, *Antimicrob Agents Chemother*, 38(6), 1345 – 1355), or *Bmr* from *Bacillus subtilis* (Neyfakh, 1992, *Antimicrob Agents Chemother*, 36(2), 484 – 485).

[275] The efflux of 7-aminoheptanoate and heptamethylenediamine can be enhanced or amplified by overexpressing the solute transporters such as the *lysE* transporter from *Corynebacterium glutamicum* (Bellmann et al., 2001, *Microbiology*, 147, 1765 – 1774).

[276] The efflux of pimelic acid can be enhanced or amplified by overexpressing a dicarboxylate transporter such as the *SucE* transporter from *Corynebacterium glutamicum* (Huhn et al., *Appl. Microbiol. & Biotech.*, 89(2), 327 – 335).

Producing C7 Building Blocks Using a Recombinant Host

[277] Typically, one or more C7 building blocks 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 a C7 building block 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.

[278] Once transferred, the microorganisms can be incubated to allow for the production of a C7 building block. Once produced, any method can be used to isolate C7 building blocks. For example, C7 building blocks can be recovered selectively from the

fermentation broth via adsorption processes. In the case of pimelic acid and 7-aminoheptanoic acid, the resulting eluate can be further concentrated via evaporation, crystallized via evaporative and/or cooling crystallization, and the crystals recovered via centrifugation. In the case of heptamethylenediamine and 1,7-heptanediol, distillation may be employed to achieve the desired product purity.

[279] The present disclosure is further described in the following examples, which do not limit the scope of the invention described in the claims.

EXAMPLES

EXAMPLE 1

Strain Selection

[280] A series of six biotin overproducing strains of *E. coli* were obtained, these were assessed for biotin productivity in small scale experiments and, based on initial biotin productivity, two of the strains were selected for further analysis, FERM BP 4928 and FERM BP 5667. These two strains were also analysed in high density fermentations with titres of biotin metabolites being >200mg/L in chemically defined growth medium.

[281] The biotin overproducing strains were modified to facilitate further manipulation and analysis; the plasmids bearing extra copies of the biotin operon were removed and; the T7 RNA polymerase gene was introduced to allow high level expression from the T7 promoter. This generated strains FERM BP-4928-P2(DE3)-4 and FERM BP-5667-P2(DE3)-A7.

EXAMPLE 2

Expression of Thioesterases

[282] The release of pimelic acid from pimelyl-ACP is catalysed by acyl-ACP thioesterases. The expression of a suitable thioesterase would be predicted to result in the release of pimelic acid. Nine different thioesterases were selected on the basis of diversity and *in vitro* activity and, following induction of expression, culture supernatants were analysed for pimelic acid. Pimelic acid levels were in all cases below the level of quantification and in some cases induction of gene expression also resulted in the cessation of cell growth. *See* Table 1.

EXAMPLE 3

Deletion of Biotin Biosynthesis Genes

[283] It was sought to increase the potential pool of pimelyl-ACP by deleting specific biotin biosynthesis genes from the chromosome. The bioF gene was the major target as the product of this gene catalyses the conversion of pimelyl-ACP to 8-amino-7-oxo-nonanoic acid ("KAPA") (FIG. 4), and the removal of this activity from the strain would be predicted to result in the accumulation of pimeloyl-ACP. In addition to the deletion of the bioF gene, the bioH gene was also deleted which would be expected to result in a strain that accumulates pimelyl-ACP methyl ester. *See* Table 1.

[284] The strains carrying deletions of bioF and bioH were then analysed, and all were biotin auxotrophs. The bioF mutant strains were found to produce detectable levels of pimelic acid in shake flask experiments (~2.5 mg/L), while the bioH mutant strains were found to produce both pimelic acid methyl ester (~0.6 mg/L) and pimelic acid (~0.2 mg/L) in the culture supernatant. The production of both pimelic acid and pimelic acid methyl ester in the absence of an additional thioesterase indicates that native *E. coli* thioesterase activities are able to release ACP bound pimelic acid (and methyl ester). Additional thioesterase was expressed in the *ΔbioF* strains, however it did not resulted in an increase in pimelic acid levels above those found without an additional thioesterase. *See* Table 1.

[285] In Table 1, pimelate production was analyzed in biotin mutant strains. Strains were cultured in chemically defined medium and were grown for 24 to 48 hours following induction. Culture supernatants were then analysed by LC-MS for pimelic acid and pimelic acid methyl ester.

TABLE 1

Strain	*bioF*	*bioH*	thioesterase	pimelic acid (ppm)	pimelic acid methyl ester (ppm)
FERM BP-4928-P2(DE3)-4	WT	WT	None	0	0
FERM BP-4928-P2(DE3)-4	WT	WT	9 different TEs[1]	0, trace amount - TE2, TE5 (<0.1 ppm)	0, trace amount - TE11
FERM BP-4928-P2(DE3)-4	Δ	WT	None	2.5	0
FERM BP-4928-P2(DE3)-4	WT	Δ	None	0.2	0.6
FERM BP-4928-P2(DE3)-4	Δ	Δ	None	0.5	1.5
FERM BP-4928-P2(DE3)-4	Δ	WT	7 different TEs[2]	0.1-0.8	0, 0.2 with TE11

[1] TEs utilised were TE2, TE5, TE8, TE11, TE14, TE17, YciA, AA077182, 'tesA.

[2] TEs utilised were TE2, TE5, TE8, TE11, TE14, TE17, 'tesA.

EXAMPLE 4

Production of 7-AHA in *E. coli*

[286] Conversion of pimelic acid to 7-AHA requires the activity of two enzymes, CAR and ω-TAM (FIG. 1). Five plasmids were constructed to enable the testing of two different promoters and three different versions of the CAR gene (FIG. 4). These plasmids were introduced into the *ΔbioF* strain and strains were analysed for 7-AHA production. All the strains with the full pathway present were found to produce 7-AHA (maximum level detected 0.03 mg/L) with the identity of the 7-AHA being confirmed by LC-MS/MS comparison to an authentic standard.

[287] As shown in FIG. 4, two different promoters were used PBAD and PT7. Three variants of the CAR gene were used, the S. *rugosus* CAR, the S. *rugosus* improved CAR (CAR-69) and the *S. rotundus* CAR. The negative control plasmid

pCDF_T7_ωta_lysE_sfp lacks a CAR gene. Plasmids were introduced into FERM BP-4928-P2(DE3)-4 ΔbioF. Strains were cultured in chemically defined medium and were grown for 24 to 72 hours following induction. Culture supernatants were analysed by LC-MS for 7-AHA. Titres of 7-AHA indicate the highest titre observed for each strain.

EXAMPLE 5

Enzyme activity of *ω-transaminase* using pimelate semialdehyde as substrate and forming 7-aminoheptanoate

[288] A sequence encoding an N-terminal His-tag was added to the genes from *Chromobacterium violaceum*, *Pseudomonas syringae*, *Rhodobacter sphaeroides*, and *Vibrio Fluvialis* encoding the *ω-transaminases* of SEQ ID NOs: 8, 10, 11 and 13, respectively (see FIGs. 9C and 9D) such that N-terminal HIS tagged *ω-transaminases* could be produced. Each of the resulting modified genes was cloned into a pET21a expression vector under control of the T7 promoter and each expression vector was transformed into a BL21[DE3] *E. coli* host. The resulting recombinant *E. coli* strains were cultivated at 37°C in a 250mL shake flask culture containing 50 mL LB media and antibiotic selection pressure, with shaking at 230 rpm. Each culture was induced overnight at 16 °C using 1 mM IPTG.

[289] The pellet from each induced shake flask culture was harvested via centrifugation. Each pellet was resuspended and lysed via sonication. The cell debris was separated from the supernatant via centrifugation and the cell free extract was used immediately in enzyme activity assays.

[290] Enzyme activity assays in the reverse direction (i.e., 7-aminoheptanoate to pimelate semialdehyde) were performed in a buffer composed of a final concentration of 50 mM HEPES buffer (pH = 7.5), 10 mM 7-aminoheptanoate, 10 mM pyruvate and 100 µM pyridoxyl 5' phosphate. Each enzyme activity assay reaction was initiated by adding cell free extract of the *ω-transaminase* gene product or the empty vector control to the assay buffer containing the 7-aminoheptanoate and incubated at 25°C for 4 h, with shaking at 250 rpm. The formation of L-alanine from pyruvate was quantified via RP-HPLC.

[291] Each enzyme only control without 7-aminoheptanoate demonstrated low base line conversion of pyruvate to L-alanine. See FIG. 15. The gene product of SEQ ID NO 8, SEQ ID NO 10, SEQ ID NO 11 and SEQ ID NO 13 accepted 7-aminoheptanote as substrate as confirmed against the empty vector control. See FIG. 16.

[292] Enzyme activity in the forward direction (i.e., pimelate semialdehyde to 7-aminoheptanoate) was confirmed for the *transaminases* of SEQ ID NO 10, SEQ ID NO 11

and SEQ ID NO 13. Enzyme activity assays were performed in a buffer composed of a final concentration of 50 mM HEPES buffer (pH = 7.5), 10 mM pimelate semialdehyde, 10 mM L-alanine and 100 µM pyridoxyl 5' phosphate. Each enzyme activity assay reaction was initiated by adding a cell free extract of the *ω-transaminase* gene product or the empty vector control to the assay buffer containing the pimelate semialdehyde and incubated at 25°C for 4 h, with shaking at 250 rpm. The formation of pyruvate was quantified via RP-HPLC.

[293] The gene product of SEQ ID NO 10, SEQ ID NO 11 and SEQ ID NO 13 accepted pimelate semialdehyde as substrate as confirmed against the empty vector control. See FIG. 17. The reversibility of the *ω-transaminase* activity was confirmed, demonstrating that the *ω-transaminases* of SEQ ID NO 8, SEQ ID NO 10, SEQ ID NO 11, and SEQ ID NO 13 accepted pimelate semialdehyde as substrate and synthesized 7-aminoheptanoate as a reaction product.

EXAMPLE 6

Enzyme activity of *carboxylate reductase* using pimelate as substrate and forming pimelate semialdehyde

[294] A sequence encoding a HIS-tag was added to the genes from *Segniliparus rugosus* and *Segniliparus rotundus* that encode the *carboxylate reductases* of SEQ ID NOs: 4 and 7, respectively (see FIGs. 9B and 9C), such that N-terminal HIS tagged *carboxylate reductases* could be produced. Each of the modified genes was cloned into a pET Duet expression vector along with a *sfp* gene encoding a HIS-tagged *phosphopantetheine transferase* from *Bacillus subtilis,* both under the T7 promoter. Each expression vector was transformed into a BL21[DE3] *E. coli* host and the resulting recombinant *E. coli* strains were cultivated at 37°C in a 250mL shake flask culture containing 50 mL LB media and antibiotic selection pressure, with shaking at 230 rpm. Each culture was induced overnight at 37 °C using an auto-induction media.

[295] The pellet from each induced shake flask culture was harvested via centrifugation. Each pellet was resuspended and lysed via sonication, and the cell debris was separated from the supernatant via centrifugation. The *carboxylate reductases* and *phosphopantetheine transferases* were purified from the supernatant using Ni-affinity chromatography, diluted 10-fold into 50mM HEPES buffer (pH = 7.5), and concentrated via ultrafiltration.

[296] Enzyme activity assays (i.e., from pimelate to pimelate semialdehyde) were performed in triplicate in a buffer composed of a final concentration of 50 mM HEPES buffer

(pH = 7.5), 2 mM pimelate, 10 mM $MgCl_2$, 1 mM ATP and 1 mM NADPH. Each enzyme activity assay reaction was initiated by adding purified *carboxylate reductase* and *phosphopantetheine transferase* gene products or the empty vector control to the assay buffer containing the pimelate and then incubated at room temperature for 20 min. The consumption of NADPH was monitored by absorbance at 340 nm. Each enzyme only control without pimelate demonstrated low base line consumption of NADPH. See FIG. 10.

[297] The gene products of SEQ ID NO 4 and SEQ ID NO 7, enhanced by the gene product of *sfp*, accepted pimelate as substrate, as confirmed against the empty vector control (see FIG. 11), and synthesized pimelate semialdehyde.

EXAMPLE 7

Enzyme Activity of *Carboxylate Reductase* using 7-Hydroxyheptanoate as Substrate and Forming 7-hydroxyheptanal

[298] A sequence encoding a His-tag was added to the genes from Mycobacterium marinum, Mycobacterium smegmatis, Segniliparus rugosus, Mycobacterium smegmatis, Mycobacterium massiliense, and Segniliparus rotundus that encode the carboxylate reductases of SEQ ID NOs: 2 - 7- respectively (see FIGs. 9A-9C) such that N-terminal HIS tagged carboxylate reductases could be produced. Each of the modified genes was cloned into a pET Duet expression vector alongside a sfp gene encoding a His-tagged phosphopantetheine transferase from Bacillus subtilis, both under control of the T7 promoter.

[299] Each expression vector was transformed into a BL21[DE3] *E. coli* host and the resulting recombinant *E. coli* strains were cultivated at 37°C in a 250mL shake flask culture containing 50 mL LB media and antibiotic selection pressure, with shaking at 230 rpm. Each culture was induced overnight at 37 °C using an auto-induction media.

[300] The pellet from each induced shake flask culture was harvested via centrifugation. Each pellet was resuspended and lysed via sonication. The cell debris was separated from the supernatant via centrifugation. The *carboxylate reductases* and *phosphopantetheine transferase* were purified from the supernatant using Ni-affinity chromatography, diluted 10-fold into 50 mM HEPES buffer (pH = 7.5) and concentrated via ultrafiltration.

[301] Enzyme activity (i.e., 7-hydroxyheptanoate to 7-hydroxyheptanal) assays were performed in triplicate in a buffer composed of a final concentration of 50 mM HEPES buffer (pH = 7.5), 2 mM 7-hydroxyheptanal, 10 mM $MgCl_2$, 1 mM ATP, and 1 mM NADPH. Each enzyme activity assay reaction was initiated by adding purified *carboxylate reductase*

and *phosphopantetheine transferase* or the empty vector control to the assay buffer containing the 7-hydroxyheptanoate and then incubated at room temperature for 20 min. The consumption of NADPH was monitored by absorbance at 340 nm. Each enzyme only control without 7-hydroxyheptanoate demonstrated low base line consumption of NADPH. See FIG. 10.

[302] The gene products of SEQ ID NO 2 - 7, enhanced by the gene product of *sfp*, accepted 7-hydroxyheptanoate as substrate as confirmed against the empty vector control (see FIG. 12), and synthesized 7-hydroxyheptanal.

EXAMPLE 8

Enzyme activity of *ω-transaminase* for 7-aminoheptanol, forming 7-oxoheptanol

[303] A nucleotide sequence encoding an N-terminal His-tag was added to the *Chromobacterium violaceum*, *Pseudomonas syringae* and *Rhodobacter sphaeroides* genes encoding the *ω-transaminases* of SEQ ID NOs: 8, 10 and 11, respectively (see FIG. 9C) such that N-terminal HIS tagged *ω-transaminases* could be produced. The modified genes were cloned into a pET21a expression vector under the T7 promoter. Each expression vector was transformed into a BL21[DE3] *E. coli* host. Each resulting recombinant *E. coli* strain were cultivated at 37°C in a 250mL shake flask culture containing 50 mL LB media and antibiotic selection pressure, with shaking at 230 rpm. Each culture was induced overnight at 16°C using 1 mM IPTG.

[304] The pellet from each induced shake flask culture was harvested via centrifugation. Each pellet was resuspended and lysed via sonication. The cell debris was separated from the supernatant via centrifugation and the cell free extract was used immediately in enzyme activity assays.

[305] Enzyme activity assays in the reverse direction (i.e., 7-aminoheptanol to 7-oxoheptanol) were performed in a buffer composed of a final concentration of 50 mM HEPES buffer (pH = 7.5), 10 mM 7-aminoheptanol, 10 mM pyruvate, and 100 µM pyridoxyl 5' phosphate. Each enzyme activity assay reaction was initiated by adding cell free extract of the *ω-transaminase* gene product or the empty vector control to the assay buffer containing the 7-aminoheptanol and then incubated at 25 °C for 4 h, with shaking at 250 rpm. The formation of L-alanine was quantified via RP-HPLC.

[306] Each enzyme only control without 7-aminoheptanol had low base line conversion of pyruvate to L-alanine. See FIG.15.

[307] The gene products of SEQ ID NO 8, 10 & 11 accepted 7-aminoheptanol as substrate as confirmed against the empty vector control (see FIG. 20) and synthesized 7-oxoheptanol as reaction product. Given the reversibility of the *ω-transaminase* activity (see Example 2), it can be concluded that the gene products of SEQ ID 8, 10 & 11 accept 7-oxoheptanol as substrate and form 7-aminoheptanol.

EXAMPLE 9

Enzyme activity of *ω-transaminase* using heptamethylenediamine as substrate and forming 7-aminoheptanal

[308] A sequence encoding an N-terminal His-tag was added to the Chromobacterium violaceum, Pseudomonas aeruginosa, Pseudomonas syringae, Rhodobacter sphaeroides, Escherichia coli, and Vibrio fluvialis genes encoding the ω-transaminases of SEQ ID NOs: 8 - 13, respectively (see FIGs. 9C and 9D) such that N-terminal HIS tagged ω-transaminases could be produced. The modified genes were cloned into a pET21a expression vector under the T7 promoter. Each expression vector was transformed into a BL21[DE3] E. coli host. Each resulting recombinant E. coli strain were cultivated at 37°C in a 250mL shake flask culture containing 50 mL LB media and antibiotic selection pressure, with shaking at 230 rpm. Each culture was induced overnight at 16°C using 1 mM IPTG.

[309] The pellet from each induced shake flask culture was harvested via centrifugation. Each pellet was resuspended and lysed via sonication. The cell debris was separated from the supernatant via centrifugation and the cell free extract was used immediately in enzyme activity assays.

[310] Enzyme activity assays in the reverse direction (i.e., heptamethylenediamine to 7-aminoheptanal) were performed in a buffer composed of a final concentration of 50 mM HEPES buffer (pH = 7.5), 10 mM heptamethylenediamine, 10 mM pyruvate, and 100 µM pyridoxyl 5' phosphate. Each enzyme activity assay reaction was initiated by adding cell free extract of the *ω-transaminase* gene product or the empty vector control to the assay buffer containing the heptamethylenediamine and then incubated at 25 °C for 4 h, with shaking at 250 rpm. The formation of L-alanine was quantified via RP-HPLC.

[311] Each enzyme only control without heptamethylenediamine had low base line conversion of pyruvate to L-alanine. See FIG. 15.

[312] The gene products of SEQ ID NO 8-13 accepted heptamethylenediamine as substrate as confirmed against the empty vector control (see FIG. 18) and synthesized 7-aminoheptanal as reaction product. Given the reversibility of the ω-transaminase activity

(see Example 5), it can be concluded that the gene products of SEQ ID 8 – 13 accept 7-aminoheptanal as substrate and form heptamethylenediamine.

EXAMPLE 10

Enzyme activity of *carboxylate reductase* for N7-acetyl-7-aminoheptanoate, forming N7-acetyl-7-aminoheptanal

[313] The activity of each of the N-terminal His-tagged *carboxylate reductases* of SEQ ID NOs: 3, 6, and 7 (see Examples 7, and FIGs. 9A, 9B, and 9C) for converting N7-acetyl-7-aminoheptanoate to N7-acetyl-7-aminoheptanal was assayed in triplicate in a buffer composed of a final concentration of 50 mM HEPES buffer (pH = 7.5), 2 mM N7-acetyl-7-aminoheptanoate, 10 mM $MgCl_2$, 1 mM ATP, and 1 mM NADPH. The assays were initiated by adding purified *carboxylate reductase* and *phosphopantetheine transferase* or the empty vector control to the assay buffer containing the N7-acetyl-7-aminoheptanoate then incubated at room temperature for 20 min. The consumption of NADPH was monitored by absorbance at 340 nm. Each enzyme only control without N7-acetyl-7-aminoheptanoate demonstrated low base line consumption of NADPH. See FIG. 10.

[314] The gene products of SEQ ID NO 3, 6, and 7, enhanced by the gene product of *sfp*, accepted N7-acetyl-7-aminoheptanoate as substrate as confirmed against the empty vector control (see FIG. 13), and synthesized N7-acetyl-7-aminoheptanal.

EXAMPLE 11

Enzyme activity of *ω-transaminase* using N7-acetyl-1,7-diaminoheptane, and forming N7-acetyl-7-aminoheptanal

[315] The activity of the N-terminal His-tagged *ω-transaminases* of SEQ ID NOs: 8 – 13 (see Example 9, and FIGs. 9C and 9D) for converting N7-acetyl-1,7-diaminoheptane to N7-acetyl-7-aminoheptanal was assayed using a buffer composed of a final concentration of 50 mM HEPES buffer (pH = 7.5), 10 mM N7-acetyl-1,7-diaminoheptane, 10 mM pyruvate and 100 µM pyridoxyl 5' phosphate. Each enzyme activity assay reaction was initiated by adding a cell free extract of the *ω-transaminase* or the empty vector control to the assay buffer containing the N7-acetyl-1,7-diaminoheptane then incubated at 25 °C for 4 h, with shaking at 250 rpm. The formation of L-alanine was quantified via RP-HPLC.

[316] Each enzyme only control without N7-acetyl-1,7-diaminoheptane demonstrated low base line conversion of pyruvate to L-alanine. See FIG. 15.

[317] The gene product of SEQ ID NO 8 - 13 accepted N7-acetyl-1,7-diaminoheptane as substrate as confirmed against the empty vector control (see FIG. 19) and synthesized N7-acetyl-7-aminoheptanal as reaction product.

[318] Given the reversibility of the ω-transaminase activity (see example 2), the gene products of SEQ ID 8 – 13 accept N7-acetyl-7-aminoheptanal as substrate forming N7-acetyl-1,7-diaminoheptane.

EXAMPLE 12

Enzyme activity of *carboxylate reductase* using pimelate semialdehyde as substrate and forming heptanedial

[319] The N-terminal His-tagged *carboxylate reductase* of SEQ ID NO 7 (see Example 7 and FIG. 9C) was assayed using pimelate semialdehyde as substrate. The enzyme activity assay was performed in triplicate in a buffer composed of a final concentration of 50 mM HEPES buffer (pH = 7.5), 2 mM pimelate semialdehyde, 10 mM $MgCl_2$, 1 mM ATP and 1 mM NADPH. The enzyme activity assay reaction was initiated by adding purified *carboxylate reductase* and *phosphopantetheine transferase* or the empty vector control to the assay buffer containing the pimelate semialdehyde and then incubated at room temperature for 20 min. The consumption of NADPH was monitored by absorbance at 340 nm. The enzyme only control without pimelate semialdehyde demonstrated low base line consumption of NADPH. See FIG. 10.

[320] The gene product of SEQ ID NO 7, enhanced by the gene product of *sfp*, accepted pimelate semialdehyde as substrate as confirmed against the empty vector control (see FIG.14) and synthesized heptanedial.

OTHER EMBODIMENTS

[321] 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.

CLAIMS

What is claimed is:

1. A method for regulating biosynthesis of at least one C7 building block using a pathway having a pimeloyl-ACP intermediate, said method comprising the step of downregulating the activity of a BioF enzyme.

2. The method of claim 1, wherein the BioF has at least 70% sequence identity or homology to the amino acid sequence set forth in SEQ ID NO: 23.

3. The method of claim 1, wherein the at least one C7 building block is selected from the group consisting of pimelic acid, 7-aminoheptanoate, 7-hydroxyheptanoate, heptamethylenediamine, 7-aminoheptanol and 1,7-heptanediol.

4. The method of claim 1, wherein the step of downregulating the activity of BioF comprises reducing the activity to a level below that of a wild type BioF.

5. The method of claim 1, wherein the step of downregulating the activity of BioF comprises reducing the activity to zero.

6. A method for regulating biosynthesis of at least one C7 building block using a pathway having a pimeloyl-ACP intermediate, said method comprising

the step of downregulating the activity of BioF; and

the step of overexpressing BioW and a CoA-specific BioF.

7. The method of claim 6, wherein said BioW enzymatically converts pimelic acid to pimeloyl-CoA.

8. The method of claim 6 or 7, wherein said BioW is a pimeloyl-CoA ligase having at least 70% sequence identity or homology to the amino acid sequence set forth in SEQ ID NO: 14.

9. The method of any one of claims 6-8, wherein CoA-specific BioF enzymatically converts pimeloyl-CoA to 8-amino-7-oxo-nonanoic acid.

10. The method of any one of claims6-9, wherein said CoA-specific BioF is a 8-amino-7-oxononanoate synthase having at least 70% sequence identity or homology to the amino acid sequence set forth in SEQ ID NO: 15.

11. The method of claim 1, wherein said pathway comprises enzymatically synthesizing a C7 aliphatic backbone from malonyl-ACP via two cycles of methyl-ester shielded carbon chain elongation, and enzymatically forming two terminal functional groups independently selected from the group consisting of carboxyl, amine, and hydroxyl groups in said backbone, thereby forming the C7 building block.

12. The method of claim 11, wherein a *S-adenosyl-L-methionine (SAM)-dependent methyltransferase* converts malonyl-ACP to malonyl-ACP methyl ester.

13. The method of claim 11, wherein the *S-adenosyl-L-methionine (SAM)-dependent methyltransferase* has at least 70% sequence identity or homology to the amino acid sequence set forth in SEQ ID NO: 16.

14. The method of claim 10, wherein each of said two cycles of carbon chain elongation comprises using (i) a *β-ketoacyl-ACP synthase*, (ii) a *3-oxoacyl-ACP reductase*, (iii) a *3-hydroxyacyl-ACP dehydratase*, and (iv) an *enoyl-ACP reductase*.

15. The method of claim 10, wherein said two cycles of carbon chain elongation produce pimeloyl-ACP methyl ester from malonyl-ACP methyl ester using a *trans-2-enoyl-CoA reductase*.

16. The method of any one of claims 1-15, wherein said two terminal functional groups are the same.

17. The method of any one of claims 1-16, wherein said two terminal functional groups are amine.

18. The method of any one of claims 1-16, wherein said two terminal functional groups are hydroxyl groups.

19. The method of any one of claims 1-15, wherein said two terminal functional groups are different.

20. The method of any one of claims 1-15 and 19, wherein said at least one C7 building block comprises a terminal amine and a terminal carboxyl group.

21. The method of any one of claims 1-15 and 19, wherein said at least one C7 building block comprises a terminal hydroxyl group and a terminal carboxyl group.

22. The method of any one of claims 1-21, wherein the hydroxyl group is enzymatically formed by a *6-hydroxyhexanoate dehydrogenase*, a *5-hydroxypentanoate dehydrogenase*, a *4-hydroxybutyrate dehydratase*, or an *alcohol dehydrogenase*.

23. The method of any one of claims 1-22, wherein the carboxyl group is enzymatically formed by a *thioesterase*, an *aldehyde dehydrogenase*, a *7-oxoheptanoate dehydrogenase*, a *6-oxohexanoate dehydrogenase*, a *glutaconate CoA-transferase*, or a *reversible succinyl-CoA ligase*.

24. The method of claim 23, wherein said *thioesterase* has at least 70% sequence identity or homology to the amino acid sequence set forth in SEQ ID NO: 21 or 22.

25. The method of any one of claims 1-24, wherein the amine group is enzymatically formed by a *ω-transaminase* or a *deacetylase*.

26. The method of claim 25, wherein said *ω-transaminase* has at least 70% sequence identity or homology to any one of the amino acid sequences set forth in any of SEQ ID NOs: 8 – 13.

27. The method of any one of claims 1-26, wherein an intermediate (in forming the at least one product) with a terminal aldehyde group is formed by a *carboxylate reductase* and enhanced by a *phosphopantetheinyl transferase*.

28. The method of claim 27, wherein said *carboxylate reductase* has at least 70% sequence identity or homology to any one of the amino acid sequences set forth in any of SEQ ID NOs: 2 – 7.

29. The method of any of the preceding claims, wherein said method is performed in a recombinant host by fermentation.

30. The method of claim 29, wherein said recombinant host is subjected to a cultivation condition under aerobic, anaerobic, micro-aerobic or mixed oxygen/denitrification cultivation conditions.

31. The method of claim 29 or claim 30, wherein said recombinant host is cultured under conditions of nutrient limitation.

32. The method according to any one of claims 29-31, wherein said recombinant host is retained using a ceramic hollow fiber membrane to maintain a high cell density during fermentation.

33. The method of any one of claims 29-32, wherein the principal carbon source fed to the fermentation derives from biological or non-biological feedstocks.

34. The method of claim 33, wherein the biological feedstock is, or derives from, at least one chosen from monosaccharides, disaccharides, lignocellulose, hemicellulose, cellulose, lignin, levulinic acid, formic acid, triglycerides, glycerol, fatty acids, agricultural waste, condensed distillers' solubles, or municipal waste.

35. The method of claim 33, wherein the non-biological feedstock is, or derives from, at least one chosen 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.

36. The method of any one of claims 29-35, wherein the recombinant host is a prokaryote.

37. The method of claim 36, wherein said prokaryote is chosen from the genii *Escherichia, Clostridia, Corynebacteria, Cupriavidus, Pseudomonas, Delftia acidovorans, Bacillus, Lactobacillus, Lactococcus,* or *Rhodococcus.*

38. The method of claim 37, wherein said prokaryote is *Escherichia coli.*

39. The method of claim 37, wherein said prokaryote is *Clostridium ljungdahlii, Clostridium autoethanogenum,* or *Clostridium kluyveri.*

40. The method of claim 37, wherein said prokaryote is *Corynebacterium glutamicum.*

41. The method of claim 37, wherein said prokaryote is *Cupriavidus necator* or *Cupriavidus metallidurans.*

42. The method of claim 37, wherein said prokaryote is *Pseudomonas fluorescens, Pseudomonas putida* or *Pseudomonas oleavorans.*

43. The method of claim 37, wherein said prokaryote is *Bacillus subtillis.*

44. The method of claim 37, wherein said prokaryote is *Lactobacillus delbrueckii.*

45. The method of claim 37, wherein said prokaryote is *Lactococcus lactis.*

46. The method of claim 37, wherein said prokaryote is *Rhodococcus equi.*

47. The method of any one of claims 29-35, wherein the host is a eukaryote.

48. The method of claim 47, wherein said eukaryote is from the genii *Aspergillus, Saccharomyces,Pichia, Yarrowia, Issatchenkia, Debaryomyces, Arxula,* or *Kluyveromyces.*

49. The method of claim 47, wherein said eukaryote is *Aspergillus niger.*

50. The method of claim 47, wherein said eukaryote is *Saccharomyces cerevisiae.*

51. The method of claim 47, wherein said eukaryote is *Pichia pastoris*.

52. The method of claim 47, wherein said eukaryote is *Yarrowia lipolytica*.

53. The method of claim 47, wherein said eukaryote is *Issathenkia orientalis*.

54. The method of claim 47, wherein said eukaryote is *Debaryomyces hansenii*.

55. The method of claim 47, wherein said eukaryote is *Arxula adenoinivorans*.

56. The method of claim 47, wherein said eukaryote is *Kluyveromyces lactis*.

57. The method of claim 29, wherein the recombinant host's tolerance to high concentrations of at least one C7 building block is improved through continuous cultivation in a selective environment.

58. The method of any one of claims 29-57, wherein said recombinant host comprises one or more of attenuated enzymes chosen from *polyhydroxyalkanoate synthase*, an *acetyl-CoA thioesterase*, an *acetyl-CoA specific ß-ketothiolase*, a *phosphotransacetylase forming acetate*, an *acetate kinase*, a *lactate dehydrogenase*, a *menaquinol-fumarate oxidoreductase*, a *2-oxoacid decarboxylase* producing isobutanol, an *alcohol dehydrogenase* forming ethanol, a *triose phosphate isomerase*, a *pyruvate decarboxylase*, a *glucose-6-phosphate isomerase*, a *transhydrogenase* dissipating the NADH or NADPH imbalance, an *glutamate dehydrogenase* dissipating the NADH or NADPH imbalance, a NADH/NADPH-utilizing *glutamate dehydrogenase*, a *pimeloyl-CoA dehydrogenase*; an *acyl-CoA dehydrogenase* accepting C7 building blocks and central precursors as substrates; a *glutaryl-CoA dehydrogenase*; or a *pimeloyl-CoA synthetase*.

59. The method of any one of claims 29-58, wherein said recombinant host overexpresses one or more genes encoding an *acetyl-CoA synthetase*, a *6-phosphogluconate dehydrogenase*; a *transketolase*; a *puridine nucleotide transhydrogenase*; a *formate dehydrogenase*; a *glyceraldehyde-3P-dehydrogenase*; a *malic enzyme*; a *glucose-6-phosphate dehydrogenase*; a *fructose 1,6 diphosphatase*; a *L-alanine dehydrogenase*; a *L-glutamate dehydrogenase* specific to the NADH or NADPH used to generate a co-factor imbalance; a

methanol dehydrogenase, a *formaldehyde dehydrogenase*, a *diamine transporter*; a *dicarboxylate transporter*; an *S-adenosylmethionine synthetase* and/or a *multidrug transporter*.

60. A recombinant host for producing pimeloyl-ACP, said recombinant host comprising a deletion of a *bioF gene*.

61. A recombinant host for producing pimeloyl-ACP, wherein said recombinant host expresses a BioF at a level below that of a wild type host.

62. The recombinant host of claim 60, wherein said recombinant host does not express BioF activity.

63. The recombinant host of claim [47], wherein said recombinant host comprises at least one exogenous nucleic acid encoding *bioW* and a CoA-specific *bioF*.

64. A non-naturally occurring organism comprising at least one exogenous nucleic acid encoding at least one polypeptide having the activity of at least one enzyme depicted in any one of FIGs. 1 to 21.

65. A nucleic acid construct or expression vector comprising a polynucleotide encoding a polypeptide having *8-amino-7-oxononanoate synthase* activity, wherein the polynucleotide is operably linked to one or more heterologous control sequences that direct production of the polypeptide and wherein the polypeptide having *8-amino-7-oxononanoate synthase* activity is a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 23.

66. A nucleic acid construct or expression vector comprising a polynucleotide encoding a polypeptide having *thioesterase* activity, wherein the polynucleotide is operably linked to one or more heterologous control sequences that direct production of the polypeptide and wherein the polypeptide having *thioesterase* activity is selected from the group consisting of: (a) a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 1; (b) a polypeptide having at least 70% sequence identity or homology to the

polypeptide of SEQ ID NO: 21; and (c) a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 22.

67. A nucleic acid construct or expression vector comprising a polynucleotide encoding a polypeptide having *carboxylate reductase* activity, wherein the polynucleotide is operably linked to one or more heterologous control sequences that direct production of the polypeptide and wherein the polypeptide having *carboxylate reductase* activity is selected from the group consisting of: (a) a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 2; (b) a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 3; (c) a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 4; (d) a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 5; (e) a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 6; and (f) a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 7.

68. A nucleic acid construct or expression vector comprising a polynucleotide encoding a polypeptide having *ω-transaminase* activity, wherein the polynucleotide is operably linked to one or more heterologous control sequences that direct production of the polypeptide and wherein the polypeptide having *ω-transaminase* activity is selected from the group consisting of: (a) a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 8; (b) a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 9; (c) a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 10 (d) a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 11; (e) a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 12; and (f) a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 13.

69. A nucleic acid construct or expression vector comprising a polynucleotide encoding a polypeptide having *phosphopantetheinyl transferase* activity, wherein the polynucleotide is operably linked to one or more heterologous control sequences that direct production of the polypeptide and wherein the polypeptide having *phosphopantetheinyl transferase* activity is selected from the group consisting of: (a) a polypeptide having at least 70% sequence identity

or homology to the polypeptide of SEQ ID NO: 14 and (b) a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 15.

70. A nucleic acid construct or expression vector comprising a polynucleotide encoding a polypeptide having *pimeloyl-ACP methyl ester esterase* activity, wherein the polynucleotide is operably linked to one or more heterologous control sequences that direct production of the polypeptide and wherein the polypeptide having *pimeloyl-ACP methyl ester esterase* activity is a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 17.

71. A nucleic acid construct or expression vector comprising a polynucleotide encoding a polypeptide having *6-carboxyhexanoate-CoA ligase* activity, wherein the polynucleotide is operably linked to one or more heterologous control sequences that direct production of the polypeptide and wherein the polypeptide having *6-carboxyhexanoate-CoA ligase* activity is a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 18.

72. A nucleic acid construct or expression vector comprising a polynucleotide encoding a polypeptide having *8-amino-7-oxononanoate synthase* activity, wherein the polynucleotide is operably linked to one or more heterologous control sequences that direct production of the polypeptide and wherein the polypeptide having *8-amino-7-oxononanoate synthase* activity is a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 19.

73. A nucleic acid construct or expression vector comprising a polynucleotide encoding a polypeptide having *malonyl-ACP O-methyltransferase* activity, wherein the polynucleotide is operably linked to one or more heterologous control sequences that direct production of the polypeptide and wherein the polypeptide having *malonyl-ACP O-methyltransferase* activity is a polypeptide having at least 70% sequence identity or homology to the polypeptide of SEQ ID NO: 20.

74. A composition comprising the nucleic acid construct or expression vector of any one of claims 65-73.

75. A non-naturally occurring biochemical network comprising a 5-hydroxypentanoyl-CoA, an exogenous nucleic acid encoding a polypeptide having the activity of a β-ketothiolase classified under EC. 2.3.1, and a 3-oxo-7-hydroxyheptanoyl-CoA.

76. A bio-derived, bio-based or fermentation-derived product, wherein said product comprises:

i. a composition comprising at least one bio-derived, bio-based or fermentation-derived compound produced or biosynthesized according to any one of claims 1-75 or or any one of FIGS. 1-21, or any combination thereof;

ii. a bio-derived, bio-based or fermentation-derived polymer comprising the bio-derived, bio-based or fermentation-derived composition or compound of i., or any combination thereof;

iii. a bio-derived, bio-based or fermentation-derived resin comprising the bio-derived, bio-based or fermentation-derived compound or bio-derived, bio-based or fermentation-derived composition of i. or any combination thereof or the bio-derived, bio-based or fermentation-derived polymer of ii. or any combination thereof;

iv. a molded substance obtained by molding the bio-derived, bio-based or fermentation-derived polymer of ii. or the bio-derived, bio-based or fermentation-derived resin of iii., or any combination thereof;

v. a bio-derived, bio-based or fermentation-derived formulation comprising the bio-derived, bio-based or fermentation-derived composition of i., bio-derived, bio-based or fermentation-derived compound of i., bio-derived, bio-based or fermentation-derived polymer of ii., bio-derived, bio-based or fermentation-derived resin of iii., or bio-derived, bio-based or fermentation-derived molded substance of iv, or any combination thereof; or

vi. 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 of i., bio-derived, bio-based or fermentation-derived compound of i., bio-derived, bio-based or fermentation-derived polymer of ii., bio-derived, bio-based or fermentation-derived resin of iii., bio-derived, bio-based or fermentation-derived formulation of v., or bio-derived, bio-based or fermentation-derived molded substance of iv., or any combination thereof.

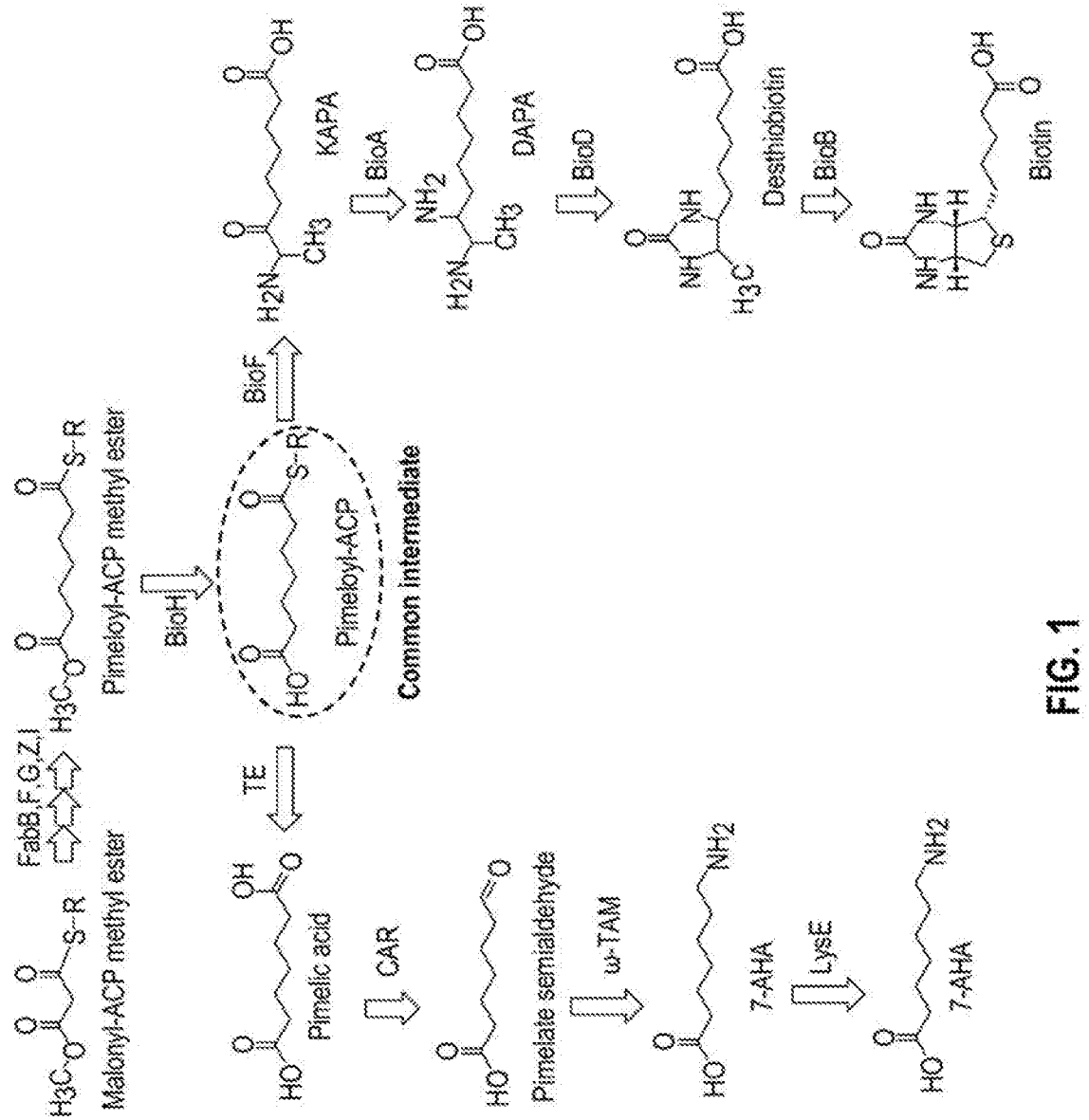
Malonyl-ACP
BioC
Malonyl-ACP methyl ester
FabB,F,G,Z,I
Malonyl-ACP methyl ester
Pimeloyl-ACP methyl ester
BioH
Common intermediate
Pimeloyl-ACP
TE
Pimelic acid
CAR
Pimelate semialdehyde
ω-TAM
7-AHA
LysE
7-AHA
BioF
KAPA
BioA
DAPA
BioD
Desthiobiotin
BioB
Biotin

FIG. 1

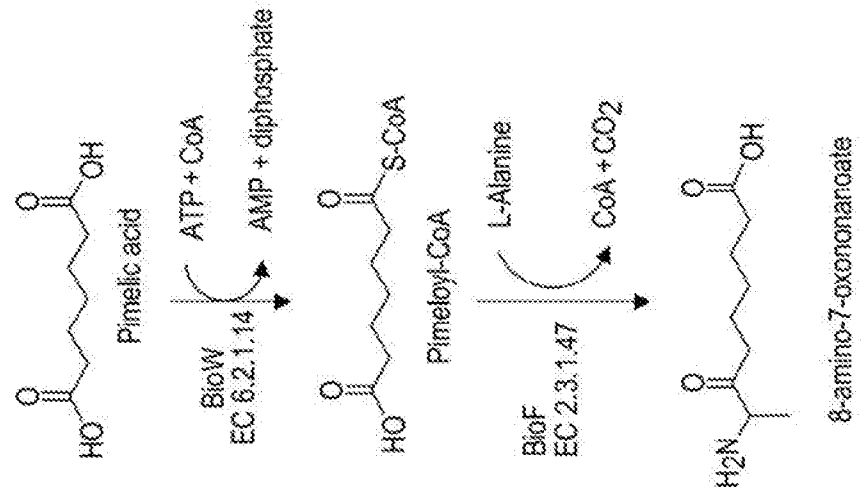
Pimelic acid
ATP + CoA
AMP + diphosphate
BioW
EC 6.2.1.14
Pimeloyl-CoA
L-Alanine
CoA + CO2
BioF
EC 2.3.1.47
8-amino-7-oxononanoate

FIG. 2

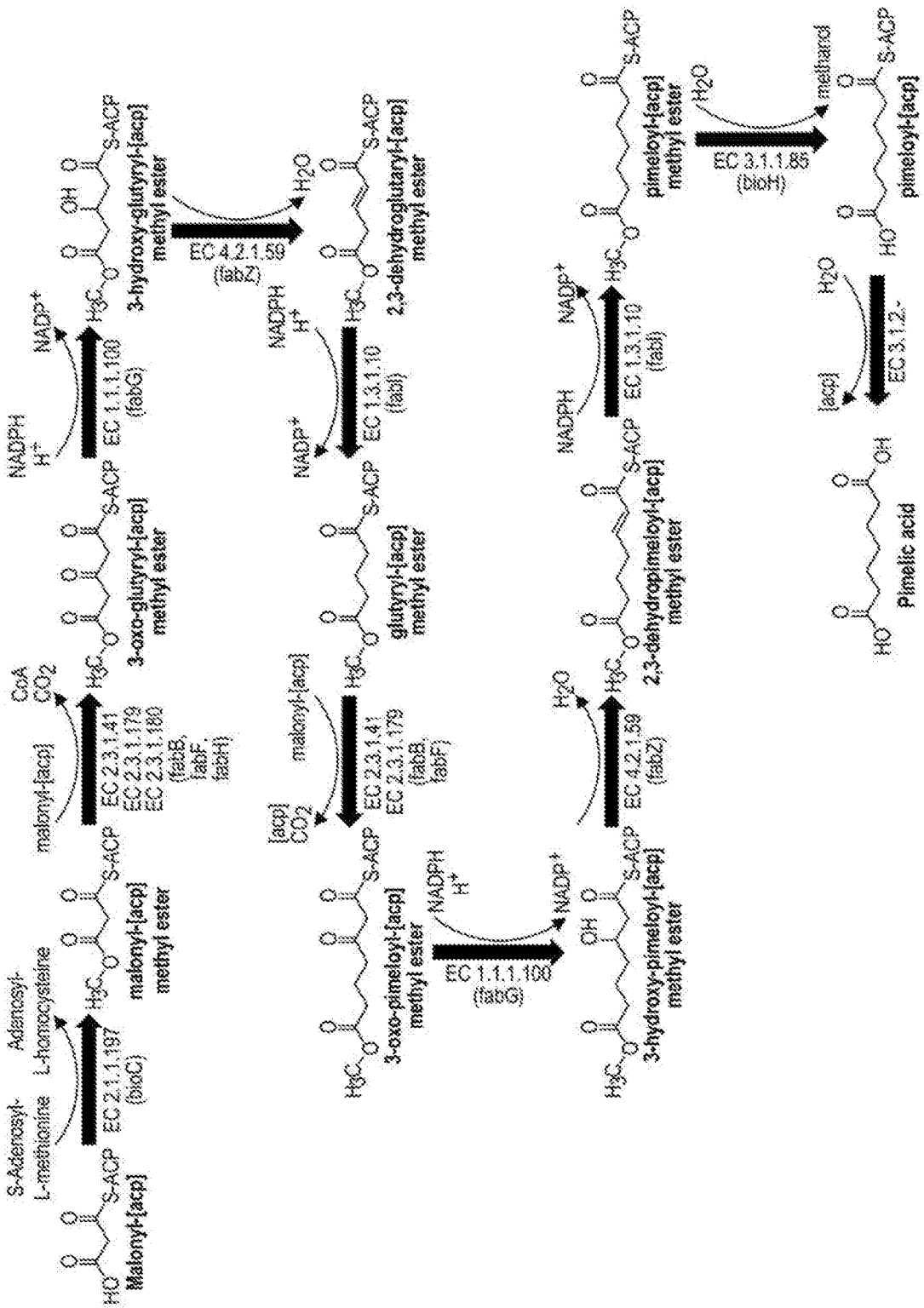
S-Adenosyl-L-methionine
Adenosyl-L-homocysteine
EC 2.1.1.197 (bioC)
Malonyl-[acp]
malonyl-[acp] methyl ester
malonyl-[acp]
CoA
CO2
EC 2.3.1.41
EC 2.3.1.179
EC 2.3.1.180
(fabB, fabF, fabH)
3-oxo-glutyryl-[acp] methyl ester
NADPH H+
NADP+
EC 1.1.1.100 (fabG)
3-hydroxy-glutyryl-[acp] methyl ester
EC 4.2.1.59 (fabZ)
H2O
2,3-dehydroglutaryl-[acp] methyl ester
NADPH H+
NADP+
EC 1.3.1.10 (fabI)
glutyryl-[acp] methyl ester
[acp]
CO2
malonyl-[acp]
EC 2.3.1.41
EC 2.3.1.179
(fabB, fabF)
3-oxo-pimeloyl-[acp] methyl ester
NADPH H+
NADP+
EC 1.1.1.100 (fabG)
3-hydroxy-pimeloyl-[acp] methyl ester
H2O
EC 4.2.1.59 (fabZ)
2,3-dehydropimeloyl-[acp] methyl ester
NADPH
NADP+
EC 1.3.1.10 (fabI)
pimeloyl-[acp] methyl ester
H2O
EC 3.1.1.85 (bioH)
methanol
[acp]
H2O
EC 3.1.2.-
Pimelic acid
pimeloyl-[acp]

FIG. 3

Operon – Pimelic Acid to 7-AHA
Plasmid
7-AHA (mg/L)
araC
PBAD
Bacillus subtilis
10xHis
sfp
Segniliparus rugosus
car
Chromobacterium violaceum uta
Corynebacterium Glutamicum lysE
Term17
PT7
C-Lacta S. rugosus CAR-69 mutant (P368S, G384E, G619L and K692R)
wt Segniliparus rotundus
car
pBBR_BAD_SCTL1-5
0.027
pCDF_T7_S.rug CAR
0.014
pCDF_T7_S.rug CAR-69
0.005
pCDF_T7_S.rot CAR
0.007
pCDF_T7_uta_lysE_sfp
0

FIG. 4

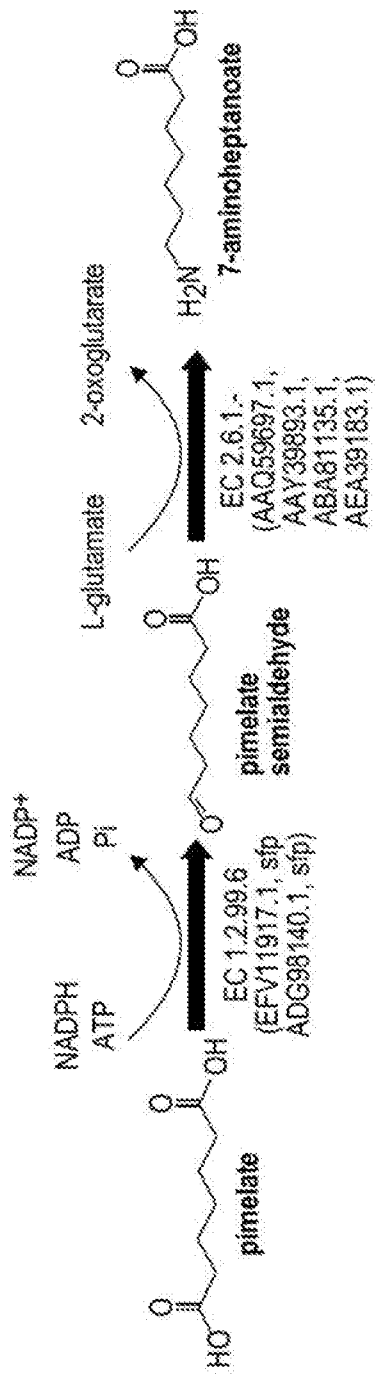
NADPH
ATP
NADP+
ADP
Pi
pimelate
EC 1.2.99.6
(EFV11917.1, sfp
ADG98140.1, sfp)
pimelate
semialdehyde
L-glutamate
2-oxoglutarate
EC 2.6.1.-
(AAQ59697.1,
AAY39893.1,
ABA81135.1,
AEA39183.1)
7-aminoheptanoate

FIG. 5

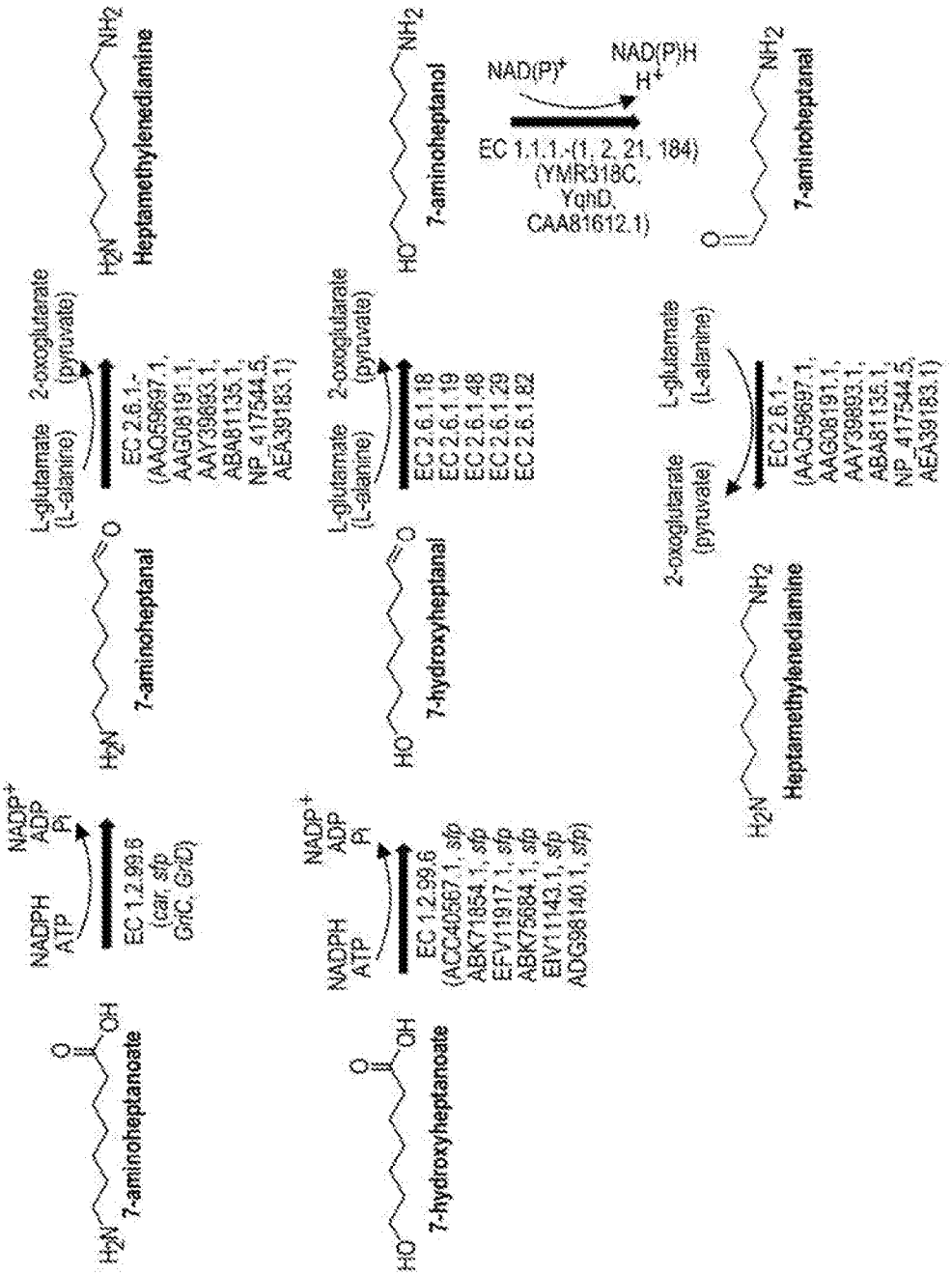
7-aminoheptanoate
NADPH
ATP
NADP+
ADP
Pi
EC 1.2.99.6
(car, sfp
GriC, GriD)
7-aminoheptanal
L-glutamate
(L-alanine)
2-oxoglutarate
(pyruvate)
EC 2.6.1.-
(AAQ59697.1,
AAG08191.1,
AAY39693.1,
ABA81135.1,
NP_417544.5,
AEA39183.1)
Heptamethylenediamine
7-hydroxyheptanoate
EC 1.2.99.6
(ACZ40567.1, sfp
ABK71854.1, sfp
EFV11917.1, sfp
ABK75684.1, sfp
EIV11143.1, sfp
ADG98140.1, sfp)
7-hydroxyheptanal
EC 2.6.1.18
EC 2.6.1.19
EC 2.6.1.48
EC 2.6.1.29
EC 2.6.1.82
7-aminoheptanol
NAD(P)+
NAD(P)H
H+
EC 1.1.1.-(1, 2, 21, 184)
(YMR318C,
YqhD,
CAA81612.1)

FIG. 6A

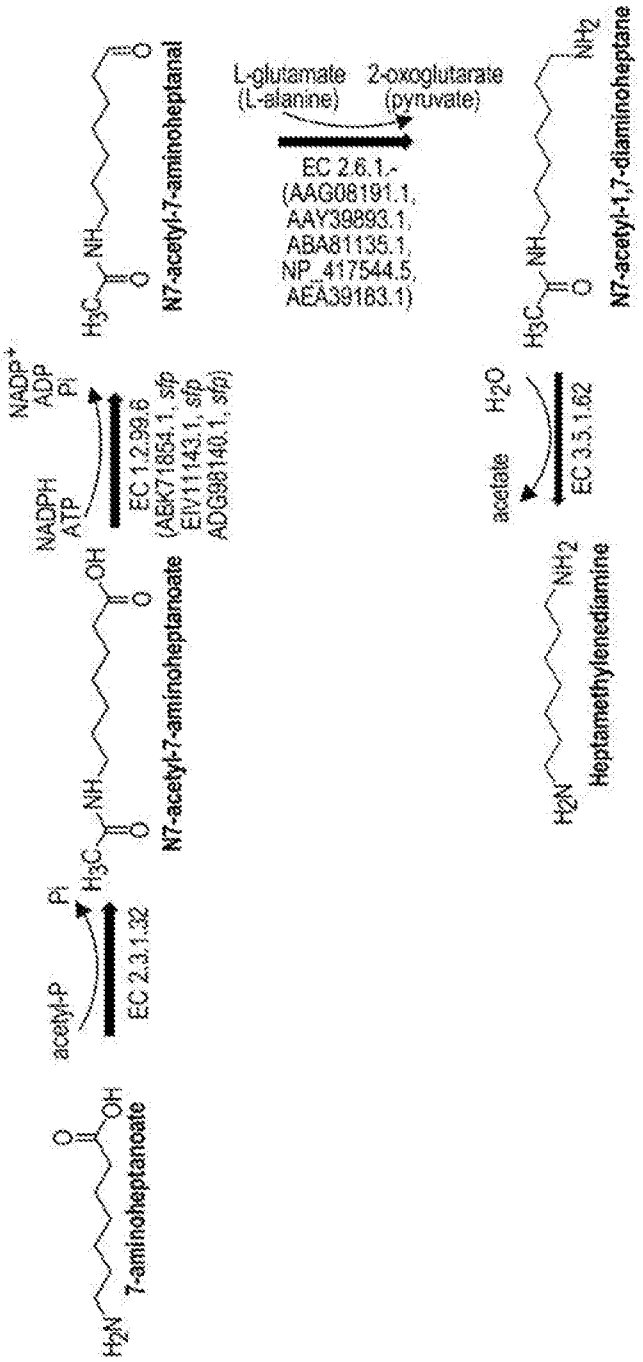
7-aminoheptanoate
acetyl-P
Pi
EC 2.3.1.32
N7-acetyl-7-aminoheptanoate
NADPH
ATP
NADP+
ADP
Pi
EC 1.2.99.6
(ABK71854.1, sfp
EIV11143.1, sfp
ADG98140.1, sfp)
N7-acetyl-7-aminoheptanal
L-glutamate
(L-alanine)
2-oxoglutarate
(pyruvate)
EC 2.6.1.-
(AAG08191.1,
AAY39893.1,
ABA81135.1,
NP_417544.5,
AEA39183.1)
N7-acetyl-1,7-diaminoheptane
H2O
acetate
EC 3.5.1.62
Heptamethylenediamine

FIG. 6B

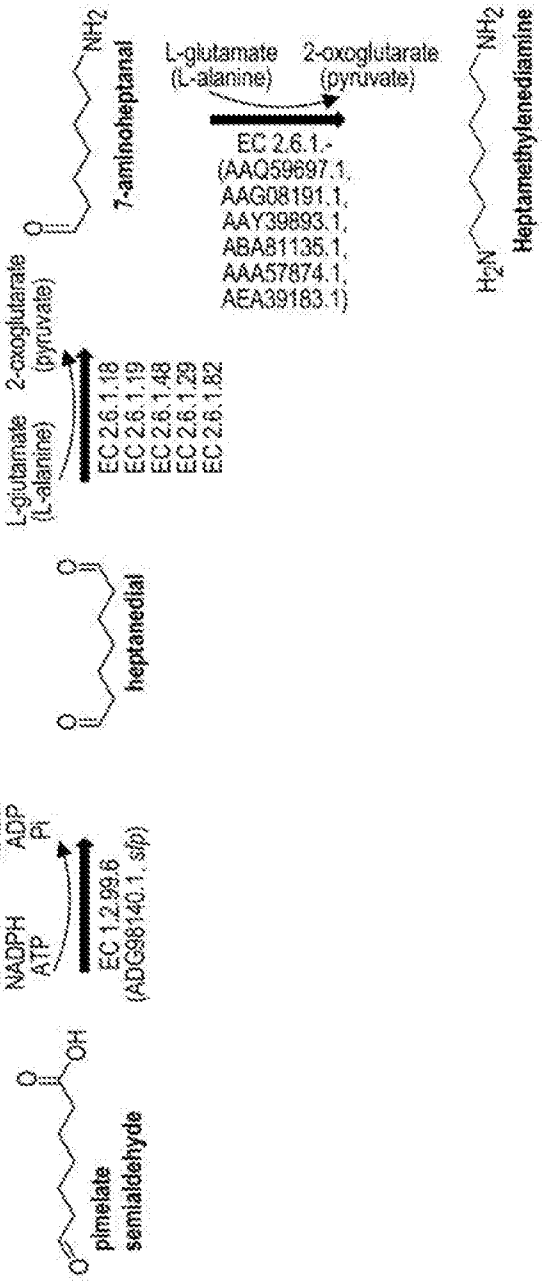
pimelate semialdehyde
NADPH
ATP
NADP+
ADP
Pi
EC 1.2.99.6
(ADG98140.1, slp)
heptanedial
L-glutamate
(L-alanine)
2-oxoglutarate
(pyruvate)
EC 2.6.1.18
EC 2.6.1.19
EC 2.6.1.48
EC 2.6.1.29
EC 2.6.1.82
7-aminoheptanal
L-glutamate
(L-alanine)
2-oxoglutarate
(pyruvate)
EC 2.6.1.-
(AAQ59697.1,
AAG08191.1,
AAY39893.1,
ABA81135.1,
AAA57874.1,
AEA39183.1)
Heptamethylenediamine

FIG. 6C

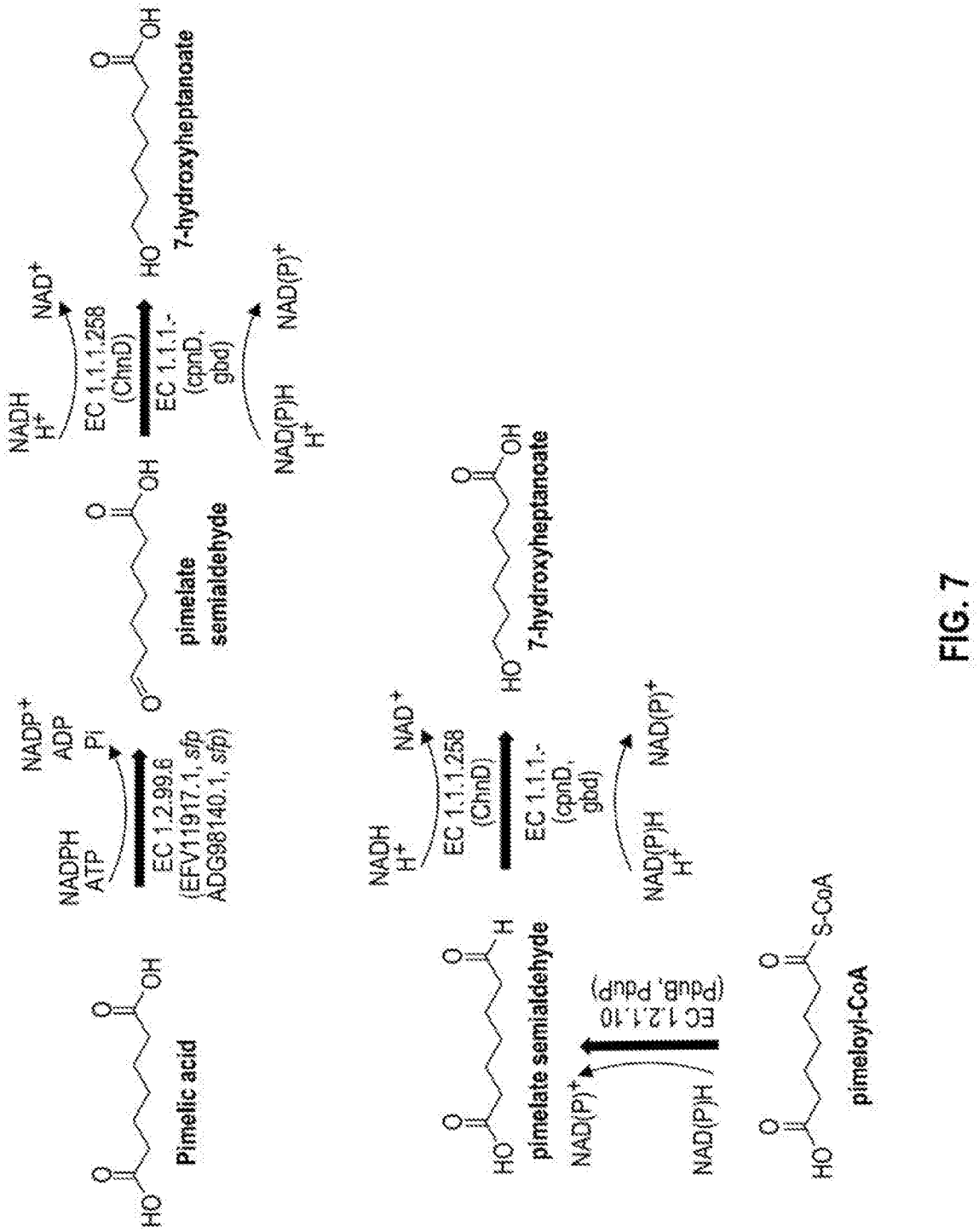
Pimelic acid
NADPH
ATP
NADP+
ADP
Pi
EC 1.2.99.6
(EFV11917.1, sfp
ADG98140.1, sfp)
pimelate
semialdehyde
NADH
H+
NAD+
EC 1.1.1.258
(ChmD)
EC 1.1.1.-
(cpnD,
gbd)
NAD(P)H
H+
NAD(P)+
7-hydroxyheptanoate
pimeloyl-CoA
NAD(P)H
NAD(P)+
EC 1.2.1.10
(PdubB, PdubP)
pimelate semialdehyde
NADH
H+
NAD+
EC 1.1.1.258
(ChmD)
EC 1.1.1.-
(cpnD,
gbd)
NAD(P)H
H+
NAD(P)+
7-hydroxyheptanoate

FIG. 7

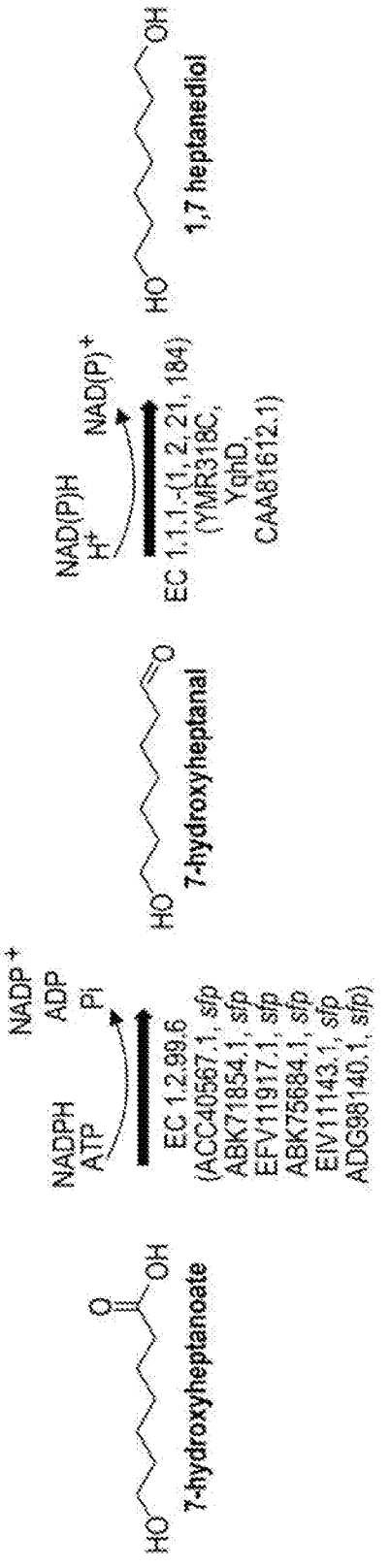
7-hydroxyheptanoate
NADPH
ATP
NADP+
ADP
Pi
EC 1.2.99.6
(ACC40567.1, sfp
ABK71854.1, sfp
EFV11917.1, sfp
ABK75684.1, sfp
EIV11143.1, sfp
ADG98140.1, sfp)
7-hydroxyheptanal
NAD(P)H
H+
NAD(P)+
EC 1.1.1.-(1.2.21.184)
(YMR318C,
YqhD,
CAA81612.1)
1,7 heptanediol

FIG. 8

SEQ ID NO	Organism	GENBANK reference	Amino acid sequence
1	*Escherichia coli*	AAA24665.1	MSQALKNLLTLLNLEKIEEGLFRGQSEDLGLRQVFGGQVVGQALYAAKETVPEERLVHSF HSYFLRPGDSKKPIIYDVETLRDGNSFSARRVAAIQNGKPIFYMTASFQAPEAGFEHQKT MPSAPAPDGLPSETQIAQSLAHLLPPVLKDKFICDRPLEVRPVEFHNPLKGHVAEPHRQV WIRANGSVPDDLRVHQYLLGYASDLNFLPVALQPHGIGFLEPGIQIATIDHSMWFHRPFN LNEWLLYSVESTSASSARGFVRGEFYTQDGVLVASTVQEGVMRNHN
2	*Mycobacterium marinum*	ACC40567.1	MSPITREERLERRIQDLYANDPQFAAAKPATAITAAIERPGLPLPQIIETVMTGYADRPA LAQRSVEFVTDAGTGHTTLRLLPHFETISYGELWDRISALADVLSTEQTVKPGDRVCLLG FNSVDYATIDMTLARLGAVAVPLQTSAAITQLQPIVAETQPTMIAASVDALADATELALS GQTATRVLVFDHHRQVDAHRAAVESARERLAGSAVVETLAEAIARGDVPRGASAGSAPGT DVSDDSLALLIYTSGSTGAPKGAMYPRRNVATFWRKRTWFEGGYEPSITLNFMPMSHVMG RQILYGTLCNGGTAYFVAKSDLSTLFEDLALVRPTELTFVPRVWDMVFDEFQSEVDRRLV DGADRVALEAQVKAEIRNDVLGGRYTSALTCSAPISDEMKAWVEELLDMHLVEGYGSTEA GMILIDGAIRRPAVLDYKLVDVPDLGYFLTDRPHPRGELLVKTDSLFPGYYQRAEVTADV FDADGFYRTGDIMAEVGPEQFVYLDRRNNVLKLSQGEFVTVSKLEAVFGDSPLVRQIYIY GNSARAYLLAVIVPTQEALDAVPVEELKARLGDSLQEVAKAAGLQSYEIPRDFIIETTPW TLENGLLTGIRKLARPQLKKHYGELLEQIYTDLAHGQADELRSLRQSGADAPVLVTVCRA AAALLGGSASDVQPDAHFTDLGGDSLSALSFTNLLHEIFDIEVPVGVIVSPANDLQALAD YVEAARKPGSSRFTFASVHGASNGQVTEVHAGDLSLDKFIDAATLAEAPRLPAANTQVRT VLLTGATGFLGRYLALEWLERMDLVDGKLICLVRAKSDTEARARLDKTFDSGDPELLAHY RALAGDHLEVLAGDKGEADLGLDRQTWQRLADTVDLIVDPAALVNHVLPYSQLFGPNALG TAELLRLALTSKIKPYSYTSTIGVADQIPPSAFTEDADIRVISATRAVDDSYANGYSNSK WAGEVLLREAHDLCGLPVAVFRCDMILADTTWAGQLNVPDMFTRMILSLAATGIAPGSFY ELAADGARQRAHYDGLPVEFIAEAISTLGAQSQDGFHTYHVMNPYDDGIGLDEFVDWLNE SGCPIQRIADYGDWLQRFETALRALPDRQRHSSLLPLLHNYRQPERPVRGSIAPTDRFRA AVQEAKIGPDKDIPHVGAPIIVKYVSDLRLLGLL
3	*Mycobacterium smegmatis*	ABK71854.1	MTSDVHDATDGVTETALDDEQSTRRIAELVATDPEFAAAAPLPAVVDAAHKPGLRLAEIL QTLFTGYGDRPALGYRARELATDEGGRTVTRLLPRFDTLTYAQVWSRVQAVAAALRHNFA QPIYPGDAVATIGFASPDYLTLDLVCAYLGLVSVPLQHNAPVSRLAPILAEVEPRILTVS AEYLDLAVESVRDVNSVSQLVVFDHHPEVDDHRDALARAREQLAGKGIAVTTLDAIADEG AGLPAEPIYTADHDQRLAMILYTSGSTGAPKGAMYTEAMVARLWTMSFITGDPTPVINVN FMPLNHLGGRIPISTAVQNGGTSYFVPESDMSTLFEDLALVRPTELGLVPRVADMLYQHH LATVDRLVTQGADELTAEKQAGAELREQVLGGRVITGFVSTAPLAAEMRAFLDITLGAHI VDGYGLTETGAVTRDGVIVRPPVIDYKLIDVPELGYFSTDKPYPRGELLVRSQTLTPGYY KRPEVTASVFDRDGYYHTGDVMAETAPDHLVYVDRRNNVLKLAQGEFVAVANLEAVFSGA ALVRQIFVYGNSERSFLLAVVVPTPEALEQYDPAALKAALADSLQRTARDAELQSYEVPA DFIVETEPFSAANGLLSGVGKLLRPNLKDRYGQRLEQMYADIAATQANQLRELRRAAATQ PVIDTLTQAAATILGTGSEVASDAHFTDLGGDSLSALTLSNLLSDFFGFEVPVGTIVNPA TNLAQLAQHIEAQRTAGDRRPSFTTVHGADATEIRASELTLDKFIDAETLRAAPGLPKVT TEPRTVLLSGANGWLGRFLTLQWLERLAPVGGTLITIVRGRDDAAARARLTQAYDTDPEL SRRFAELADRHLRVVAGDIGDPNLGLTPEIWHRLAAEVDLVVHPAALVNHVLPYRQLFGP NVVGTAEVIKLALTERIKPVTYLSTVSVAMGIPDFEEDGDIRTVSPVRPLDGGYANGYGN SKWAGEVLLREAHDLCGLPVATFRSDMILAHPRYRGQVNVPDMFTRLLLSLLITGVAPRS FYIGDGERPRAHYPGLTVDFVAEAVTTLGAQQREGYVSYDVMNPHDDGISLDVFVDWLIR AGHPIDRVDDYDDWVRRFETALTALPEKRRAQTVLPLLHAFRAPQAPLRGAPEPTEVFHA AVRTAKVGPGDIPHLDEALIDKYIRDLREFGLI

FIG. 9A

SEQ ID NO	Organism	GENBANK reference	Amino acid sequence
4	*Segniliparus rugosus*	EFV11917.1	MGDGEERAKRFFQRIGELSATDPQFAAAAPDPAVVEAVSDPSLSFTRYLDTLMRGYAERP ALAHRVGAGYETISYGELWARVGAIAAAWQADGLAPGDFVATVGFTSPDYVAVDLAAARS GLVSVPLQAGASLAQLVGILEETEPKVLAASASSLEGAVACALAAPSVQRLVVFDLRGPD ASESAADERRGALADAEEQLARAGRAVVVETLADLAARGEALPEAPLFEPAEGEDPLALL IYTSGSTGAPKGAMYSQRLVSQLWGRTPVVPGMPNISLHYMPLSHSYGRAVLAGALSAGG TAHFTANSDLSTLFEDIALARPTFLALVPRVCEMLFQESQRGQDVAELRERVLGGRLLVA VCGSAPLSPEMRAFMEEVLGFPLLDGYGSTEALGVMRNGIIQRPPVIDYKLVDVPELGYR TTDKPYPRGELCIRSTSLISGYYKRPEITAEVFDAQGYYKTGDVMAEIAPDHLVYVDRSK NVLKLSQGEFVAVAKLEAAYGTSPYVKQIFVYGNSERSFLLAVVVPNAEVLGARDQEEAK PLIAASLQKIAKEAGLQSYEVPRDFLIETEPFTTQNGLLSEVGKLLRPKLKARYGEALEA RYDEIAHGQADELRALRDGAGQRPVVETVVRAAVAISGSEGAEVGPEANFADLGGDSLSA LSLANLLHDVFEVEVPVRIIIGPTASLAGIAKHIEAERAGASAPTAASVHGAGATRIRAS ELTLEKFLPEDLLAAAKGLPAADQVRTVLLTGANGWLGRFLALEQLERLARSGQDGGKLI CLVRGKDAAAARRRIEETLGTDPALAARFAELAEGRLEVVPGDVGEPKFGLDDAAWDRLA EEVDVIVHPAALVNHVLPYHQLFGPNVVGTAEIIRLAITAKRKPVTYLSTVAVAAGVEPS SFEEDGDIRAVVPERPLGDGYANGYGNSKWAGEVLLREAHELVGLPVAVFRSDMILAHTR YTGQLNVPDQFTRLVLSLLATGIAPKSFYQQGAAGERQRAHYDGIPVDFTAEAITTLGAE PSWFDGGAGFRSFDVFNPHHDGVGLDEFVDWLIEAGHPISRIDDHKEWFARFETAVRGLP EAQRQHSLLPLLRAYSFPHPPVDGSVYPTGKFQGAVKAAQVGSDHDVPHLGKALIVKYAD DLKALGLL
5	*Mycobacterium smegmatis*	ABK75684.1	MTIETREDRFNRRIDHLFETDPQFAAARPDEAISAAAADPELRLPAAVKQILAGYADRPA LGKRAVEFVTDEEGRTTAKLLPRFDTITYRQLAGRIQAVTNAWHNHPVNAGDRVAILGFT SVDYTTIDIALLELGAVSVPLQTSAPVAQLQPIVAETEPKVIASSVDFLADAVALVESGP APSRLVVFDYSHEVDDQREAFEAAKGKLAGTGVVVETITDALDRGRSLADAPLYVPDEAD PLTLLIYTSGSTGTPKGAMYPESKTATMWQAGSKARWDETLGVMPSITLNFMPMSHVMGR GILCSTLASGGTAYFAARSDLSTFLEDLALVRPTQLNFVPRIWDMLFQEYQSRLDNRRAE GSEDRAEAAVLEEVRTQLLGGRFVSALTGSAPISAEMKSWVEDLLDMHLLEGYGSTEAGA VFIDGQIQRPPVIDYKLVDVPDLGYFATDRPYPRGELLVKSEQMFPGYYKRPEITAEMFD EDGYYRTGDIVAELGPDHLEYLDRRNNVLKLSQGEFVTVSKLEAVFGDSPLVRQIYVYGN SARSYLLAVVVPTEEALSRWDGDELKSRISDSLQDAARAAGLQSYEIPRDFLVETTPFTL ENGLLTGIRKLARPKLKAHYGERLEQLYTDLAEGQANELRELRRNGADRPVVETVSRAAV ALLGASVTDLRSDAHFTDLGGDSLSALSFSNLLHEIFDVDVPVGVIVSPATDLAGVAAYI EGELRGSKRPTYASVHGRDATEVRARDLALGKFIDAKTLSAAPGLPRSGTEIRTVLLTGA TGFLGRYLALEWLERMDLVDGKVICLVRARSDDEARARLDATFDTGDATLLEHYRALAAD HLEVIAGDKGEADLGLDHDTWQRLADTVDLIVDPAALVNHVLPYSQMFGPNALGTAELIR IALTTTIKPYVYVSTIGVGQGISPEAFVEDADIREISATRRVDDSYANGYGNSKWAGEVL LREAHDWCGLPVSVFRCDMILADTTYSGQLNLPDMFTRLMLSLVATGIAPGSFYELDADG NRQRAHYDGLPVEFIAEAISTIGSQVTDGFETFHVMNPYDDGIGLDEYVDWLIEAGYPVH RVDDYATWLSRFETALRALPERQRQASLLPLLHNYQQPSPPVCGAMAPTDRFRAAVQDAK IGPDKDIPHVTADVIVKYISNLQMLGLL
6	*Mycobacterium massiliense*	EIV11143.1	MTNETNPQQEQLSRRIESLRESDPQFRAAQPDPAVAEQVLRPGLHLSEAIAALMTGYAER PALGERARELVIDQDGRTTLRLLPRFDTTTYGELWSRTTSVAAAWHHDATHPVKAGDLVA TLGFTSIDYTVLDLAIMILGGVAVPLQTSAPASQWTTILAEAEPNTLAVSIELIGAAMES VRATPSIKQVVVFDYTPEVDDQREAFEAASTQLAGTGIALETLDAVIARGAALPAAPLYA PSAGDDPLALLIYTSGSTGAPKGAMHSENIVRRWWIREDVMAGTENLPMIGLNFMPMSHI MGRGTLTSTLSTGGTGYFAASSDMSTLFEDMELIRPTALALVPRVCDMVFQRFQTEVDRR LASGDTASAEAVAAEVKADIRDNLFGGRVSAVMVGSAPLSEELGEFIESCFELNLTDGYG STEAGMVFRDGIVQRPPVIDYKLVDVPELGYFSTDKPHPRGELLLKTDGMFLGYYKRPEV TASVFDADGFYMTGDIVAELAHDNIEIIDRRNNVLKLSQGEFVAVATLEAEYANSPVVHQ IYVYGSSERSYLLAVVVPTPEAVAAAKGDAAALKTTIADSLQDIAKEIQLQSYEVPRDFI IEPQPFTQGNGLLTGIAKLARPNLKAHYGPRLEQMYAEIAEQQAAELRALHGVDPDKPAL ETVLKAAQALLGVSSAELAADAHFTDLGGDSLSALSFSDLLRDIFAVEVPVGVIVSAAND LGGVAKFVDEQRHSGGTRPTAETVHGAGHTEIRAADLTLDKFIDEATLHAAPSLPKAAGI PHTVLLTGSNGYLGHYLALEWLERLDKTDGKLIVIVRGKNAEAAYGRLEEAFDTGDTELL AHFRSLADKHLEVLAGDIGDPNLGLDADTWQRLADTVDVIVHPAALVNHVLPYNQLFGPN VVGTAEIIKLAITTKIKPVTYLSTVAVAAYVDPTTFDEESDIRLISAVRPIDDGYANGYG NAKWAGEVLLREAHDLCGLPVAVFRSDMILAHSRYTGQLNVPDQFTRLILSLIATGIAPG SFYQAQTTGERPLAHYDGLPGDFTAEAITTLGTQVPEGSEGFVTYDCVNPHADGISLDNF VDWLIEAGYPIARIDNYTEWFTRFDTAIRGLSEKQKQHSLLPLLHAFEQPSAAENHGVVP AKRFQHAVQAAGIGPVGQDGTTDIPHLSRRLIVKYAKDLEQLGLL

FIG. 9B

SEQ ID NO	Organism	GENBANK reference	Amino acid sequence
7	*Segniliparus rotundus*	ADG98140.1	MTQSHTQGPQASAAHSRLARRAAELLATDPQAAATLPDPEVVRQATRPGLRLAERVDAIL SGYADRPALGQRSFQTVKDPITGRSSVELLPTFDTITYRELRERATAIASDLAHHPQAPA KPGDFLASIGFISVDYVAIDIAGVFAGLTAVPLQTGATLATLTAITAETAPTLFAASIEH LPTAVDAVLATPSVRRLLVFDYRAGSDEDREAVEAAKRKIADAGSSVLVDVLDEVIARGK SAPKAPLPPATDAGDDSLSLLIYTSGSTGTPKGAMYPERNVAHFWGGVWAAAFDEDAAPP VPAINITFLPLSHVASRLSLMPTLARGGLMHFVAKSDLSTLFEDLKLARPTNLFLVPRVV EMLYQHYQSELDRRGVQDGTREAEAVKDDLRTGLLGGRILTAGFGSAPLSAELAGFIESL LQIHLVDGYGSTEAGPVWRDGYLVKPPVTDYKLIDVPELGYFSTDSPHPRGELAIKTQTI LPGYYKRPETTAEVFDEDGFYLTGDVVAQIGPEQFAYVDRRKNVLKLSQGEFVTLAKLEA AYSSSPLVRQLFVYGSSERSYLLAVIVPTPDALKKFGVGEAAKAALGESLQKIARDEGLQ SYEVPRDFIIETDPFTVENGLLSDARKSLRPKLKEHYGERLEAMYKELADGQANELRDIR RGVQQRPTLETVRRAAAAMLGASAAEIKPDAHFTDLGGDSLSALTFSNFLHDLFEVDVPV GVIVSAANTLGSVAEHIDAQLAGGRARPTFATVHGKGSTTIKASDLTLDKFIDEQTLEAA KHLPKPADPPRTVLLTGANGWLGRFLALEWLERLAPAGGKLITIVRGKDAAQAKARLDAA YESGDPKLAGHYQDLAATTLEVLAGDFSEPRLGLDEATWNRLADEVDFISHPGALVNHVL PYNQLFGPNVAGVAEIIKLAITTRIKPVTYLSTVAVAAGVEPSALDEDGDIRTVSAERSV DEGYANGYGNSKWGGEVLLREAHDRTGLPVRVFRSDMILAHQKYTGQVNATDQFTRLVQS LLATGLAPKSFYELDAQGNRQRAHYDGIPVDFTAESITTLGGDGLEGYRSYNVFNPHRDG VGLDEFVDWLIEAGHPITRIDDYDQWLSRFETSLRGLPESKRQASVLPLLHAFARPGPAV DGSPFRNTVFRTDVQKAKIGAEHDIPHLGKALVLKYADDIKQLGLL
8	*Chromobacterium violaceum*	AAQ59697.1	MQKQRTTSQWRELDAAHHLHPFTDTASLNQAGARVMTRGEGVYLWDSEGNKIIDGMAGLW CVNVGYGRKDFAEAARRQMEELPFYNTFFKTTHPAVVELSSLLAEVTPAGFDRVFYTNSG SESVDTMIRMVRRYWDVQGKPEKKTLIGRWNGYHGSTIGGASLGGMKYMHEQGDLPIPGM AHIEQPWWYKHGKDMTPDEFGVVAARWLEEKILEIGADKVAAFVGEPIQGAGGVIVPPAT YWPEIERICRKYDVLLVADEVICGFGRTGEWFGHQHFGFQPDLFTAAKGLSSGYLPIGAV FVGKRVAEGLIAGGDFNHGFTYSGHPVCAAVAHANVAALRDEGIVQRVKDDIGPYMQKRW RETFSRFEHVDDVRGVGMVQAFTLVKNKAKRELFPDFGEIGTLCRDIFFRNNLIMRACGD HIVSAPPLVMTRAEVDEMLAVAERCLEEFEQTLKARGLA
9	*Pseudomonas aeruginosa*	AAG08191.1	MNARLHATSPLGDADLVRADQAHYMHGYHVFDDHRVNGSLNIAAGDGAYIYDTAGNRYLD AVGGMWCTNIGLGREEMARTVAEQTRLLAYSNPFCDMANPRAIELCRKLAELAPGDLDHV FLTTGGSTAVDTAIRLMHYYQNCRGKRAKKHVITRINAYHGSTFLGMSLGGKSADRPAEF DFLDERIHHLACPYYYRAPEGLGEAEFLDGLVDEFERKILELGADRVGAFISEPVFGSGG VIVPPAGYHRRMWELCQRYDVLYISDEVVTSFGRLGHFFASQAVFGVQPDIILTAKGLTS GYQPLGACIFSRRIWEVIAEPDKGRCFSHGFTYSGHPVACAAALKNIEIIEREGLLAHAD EVGRYFEERLQSLRDLPIVGDVRGMGRFMACVEFVADKASKALFPESLNIGEWVVHLRAQKR GLLVRPIVHLNVMSPPLILTREQVDTVVRVLRESIEETVEDLVRAGHR
10	*Pseudomonas syringae*	AAY39893.1	MSANNPQTLEWQALSSEHHLAPFSDYKQLKEKGPRIITRAEGVYLWDSEGNKILDGMSGL WCVAIGYGREELADAASKQMRELPYYNLFFQTAHPPVLELAKAISDIAPEGMNHVFFTGS GSEGNDTMLRMVRHYWALKGQPNKKTIISRVNGYHGSTVAGASLGGMTYMHEQGDLPIPG VVHIPQPYWFGEGGDMTPDEFGIWAAEQLEKKILELGVENVGAFIAEPIQGAGGVIVPPD SYWPKIKEILSRYDILFAADEVICGFGRTGEWFGSDFYGLRPDMMTIAKGLTSGYVPMGG LIVRDEIVAVLNEGGDFNHGFTYSGHPVAAAVALENIRILREEKIVERVRSETAPYLQKR LRELSDHPLVGEVRGVGLLGAIELVKDKTTRERYTDKGAGMICRTFCFDNGLIMRAVGDT MIIAPPLVISFAQIDELVEKARTCLDLTLAVLQG
11	*Rhodobacter sphaeroides*	ABA81135.1	MTRNDATNAAGAVGAAMRDHILLPAQEMAKLGKSAQPVLTHAEGIYVHTEDGRRLIDGPA GMWCAQVGYGRREIVDAMAHQAMVLPYASPWYMATSPAARLAEKIATLTPGDLNRIFFTT GGSTAVDSALRFSEFYNNVLGRPQKKRIIVRYDGYHGSTALTAACTGRTGNWPNFDIAQD RISFLSSPNPRHAGNRSQEAFLDDLVQEFEDRIESLGPDTIAAFLAEPILASGGVIIPPA GYHARFKAICEKHDILYISDEVVTGFGRCGEWFASEKVFGVVPDIITFAKGVTSGYVPLG GLAISEAVLARISGENAKGSWFTNGYTYSNQPVACAAALANIELMEREGIVDQAREMADY FAAALASLRDLPGVAETRSVGLVGCVQCLLDPTRADGTAEDKAFTLKIDERCFELGLIVR PLGDLCVISPPLIISRAQIDEMVAIMRQAITEVSAAHGLTAKEPAAV
12	*Escherichia coli*	NP_417544.5	MNRLPSSASALACSAHALNLIEKRTLDHEEMKALNREVIEYFKEHVNPGFLEYRKSVTAG GDYGAVEWQAGSLNTLVDTQGQEFIDCLGGFGIFNVGHRNPVVVSAVQNQLAKQPLHSQE LLDPLRAMLAKTLAALTPGKLKYSFFCNSGTESVEAALKLAKAYQSPRGKFTFIATSGAF HGKSLGALSATAKSTFRKPFMPLLPGFRHVPFGNIEAMRTALNECKKTGDDVAAVILEPI QGEGGVILPPPGYLTAVRKLCDEFGALMILDEVQTGMGRTGKMFACEHENVQPDILCLAK ALGGGVMPIGATIATEEVFSVLFDNPFLHTTTFGGNPLACAAALATINVLLEQNLPAQAE QKGDMLLDGFRQLAREYPDLVQEARGKGMLMAIEFVDNEIGYNFASEMFRQRVLVAGTLN NAKTIRIEPPLTLTIEQCELVIKAARKALAAMRVSVEEA

FIG. 9C

SEQ ID NO	Organism	GENBANK reference	Amino acid sequence
13	*Vibrio fluvialis*	AEA39183.1	MNKPQSWEARAETYSLYGFTDMPSLHQRGTVVVTHGEGPYIVDVNGRRYLDANSGLWNMV AGFDHKGLIDAAKAQYERFPGYHAFFGRMSDQTVMLSEKLVEVSPFDSGRVFYTNSGSEA NDTMVKMLWFLHAAEGKPQKRKILTRWNAYHGVTAVSASMTGKPYNSVFGLPLPGFVHLT CPHYWRYGEEGETEEQFVARLARELEETIQREGADTIAGFFAEPVMGAGGVIPPAKGYFQ AILPILRKYDIPVISDEVICGFGRTGNTWGCVTYDFTPDAIISSKNLTAGFFPMGAVILG PELSKRLETAIEAIEEFPHGFTASGHPVGCAIALKAIDVVMNEGLAENVRRLAPRFEERL KHIAERPNIGEYRGIGFMWALEAVKDKASKTPFDGNLSVSERIANTCTDLGLICRPLGQS VVLCPPFILTEAQMDEMFDKLEKALDKVFAEVA
14	*Bacillus subtilis*	WP_003234549.1	MKIYGIYMDRPLSQEENERFMSFISPEKREKCRRFYHKEDAHRTLLGDVLVRSVISRQYQ LDKSDIRFSTQEYGKPCIPDLPDAHFNISHSGRWVICAFDSQPIGIDIEKTKPISLEIAK RFFSKTEYSDLLAKDKDEQTDYFYHLWSMKESFIKQEGKGLSLPLDSFSVRLHQDGQVSI ELPDSHSPCYIKTYEVDPGYKMAVCAAHPDFPEDITMVSYEELL
15	*Nocardia sp. NRRL 5646*	ABI83656.1	MIETILPAGVESAELLEYPEDLKAHPAEEHLIAKSVEKRRRDFIGARHCARLALAELGEP PVAIGKGERGAPIWPRGVVGSLTHCDGYRAAAVAHKMRFRSIGIDAEPHATLPEGVLDSV SLPPEREWLKTTDSALHLDRLLFCAKEATYKAWWPLTARWLGFEEAHITFEIEDGSADSG NGTFHSELLVPGQTNDGGTPLLSFDGRWLIADGFILTAIAYA
16	*Bacillus cereus*	AAS43086.1	MINKTLLQKRFNVAAVSYDQYANVQKKMAHSLLSTLNRRYSTNSSIRILELGCGTGYVTE QLSNLFPKAQITAIDFAESMIAVAKTRQNVNNVTFYCEDIERLRLEETYDVIISNATFQW LNDLKQVITNLFRHLSIEGILLFSTFGQETFQELHASFQRAKEEKNIQNETSIGQRFYSK NQLRHICEIETGDVHVSETCYIERFTEVREFLHSIRKVGATNSNEESYCQSPSLFRAMLR IYERDFTGNEGIMATYHALFVHITKEGKR
17	*Escherichia coli*	AAC76437.1	MNNIWWQTKGQGNVHLVLLHGWGLNAEVWRCIDEELSSHFTLHLVDLPGFGRSRGFGALS LADMAEAVLQQAPDKAIWLGWSLGGLVASQIALTHPERVQALVTVASSPCFSARDEWPGI KPDVLAGFQQQLSDDFQRTVERFLALQTMGTETARQDARALKKTVLALPMPEVDVLNGGL EILKTVDLRQPLQNVSMPFLRLYGYLDGLVPRKVVPMLDKLWPHSESYIFAKAAHAPFIS HPAEFCHLLVALKQRV
18	*Bacillus subtilis*	AAB17457.1	MMQEETFYSVRMRASMNGSHEDGGKHISGGERLIPFHEMKHTVNALLEKGLSHSRGKPDFM QIQFEEVHESIKTIQPLPVHTNEVSCPEEGQKLARLLLEKEGVSRDVIEKAYEQIPEWSDVRGA VLFDIHTGKRMDQTKEKGVRVSRMDWPDANFEKWALHSHVPAHSRIKEALALASKVSRHPA VVAELCWSDDPDYITGYVAGKKMGYQRITAMKEYGTEEGCRVFFIDGSNDVNTYIHDLEKQPI LIEWEEDHDS
19	*Bacillus subtilis*	AAB17459.1	MKIDSWLNERLDRMKEAGVHRNLRSMDGAPVPERNIDGENQTVWSSNNYLGLASDRRLIDA AQTALQQFGTGSSGSRLTTGNSVWHEKLEKKIASFKLTEAALLFSSGYLANVGVLSSLPEKED VILSDQLNHASMIDGCRLSKADTVVYRHIDMNDLENKLNETQRYQRRFIVTDGVFSMDGTIAP LDQIISLAKRYHAFVVVDDAHATGVLGDSGQGTSEYFGVCPDIVIGTLSKAVGAEGGFAAGSA VFIDFLLNHARTFIFQTAIPPASCAAAHEAFNIIEASREKRQLLFSYISMIRTSLKNMGYVVKGDH TPIIPVVIGDAHKTVLFAEKLQGKGIYAPAIRPPTVAPGESRIRITITSDHSMGDIDHLLQTFHSIG KELHII
20	*Escherichia coli*	NP_415298.1	MATVNKQAIAAAFGRAAAHYEQHADLQRQSADALLAMLPQRKYTHVLDAGCGPGWMSRHW RERHAQVTALDLSPPMLVQARQKDAADHYLAGDIESLPLATATFDLAWSNLAVQWCGNLSTA LRELYRVVRPKGVVAFTTLVQGSLPELHQAWQAVDERPHANRFLPPDEIEQSLNGVHYQHHI QPITLWFDDALSAMRSLKGIGATHLHEGRDPRILTRSQLQRLQLAWPQQQGRYPLTYHLFLG VIARE
21	*Lactobacillus brevis*	ABJ63754.1	MAANEFSETHRVVYYEADDTGQLTLAMLINLFVLVSEDQNDALGLSTAFV QSHGVGWVVTQYHLHIDELPRTGAQVTIKTRATAYNRYFAYREYWLLDDA GQVLAYGEGIWVTMSYATRKITTIPAEVMAPYHSEEQTRLPRLPRPDHFD EAVNQTLKPYTVRYFDIDGNGHVNNAHYFDWMLDVLPATFLRAHHPTDVK IRFENEVQYGHQVTSELSQAAALTTQHMIKVGDLTAVKATIQWDNR
22	*Lactobacillus plantarum*	CCC78182.1	MATLGANASLYSEQHRITYYECDRTGRATLTTLIDIAVLASEDQSDALGL TTEMVQSHGVGWVVTQYAIDITRMPRQDEVVTIAVRGSAYNPYFAYREFW IRDADGQQLAYITSIWVMMSQTTRRIVKILPELVAPYQSEVVKRIPRLPR PISFEATDTTITKPYHVRFFDIDPNRHVNNAHYFDWLVDTLPATFLLQHD LVHVDVRYENEVKYGQTVTAHANILPSEVADQVTTSHLIEVDDEKCCEVT IQWRTLPEPIQ
23	*Escherichia coli*	AAA23516.1	MSWQEKINAALDARRAADALRRRYPVAQGAGRWLVADDRQYLNFSSNDYLGLSHHPQIIRA WQQGAEQFGIGSGGSGHVSGYSVVHQALEEELAEWLGYSRALLFISGFAANQAVIAAMMAK EDRIAADRLSHASLLEAASLSPSQLRRFAHNDVTHLARLLASPCPGQQMVVTEGVFSMDGDS APLAEIQQVTQQHNGWLMVDDAHGTGVIGEQGRGSCWLQKVKPELLVVTFGKGFGVSGAA VLCSSTVADYLLQFARHLIYSTSMPPAQAQALRASLAVIRSDEGDARREKLAALITRFRAGVQD LPFTLADSCSAIQPLIVGDNSRALQLAEKLRQQGCWVTAIRPPTVPAGTARLRLTLTAAHEMQ DIDRLLEVLHGNG

FIG. 9D

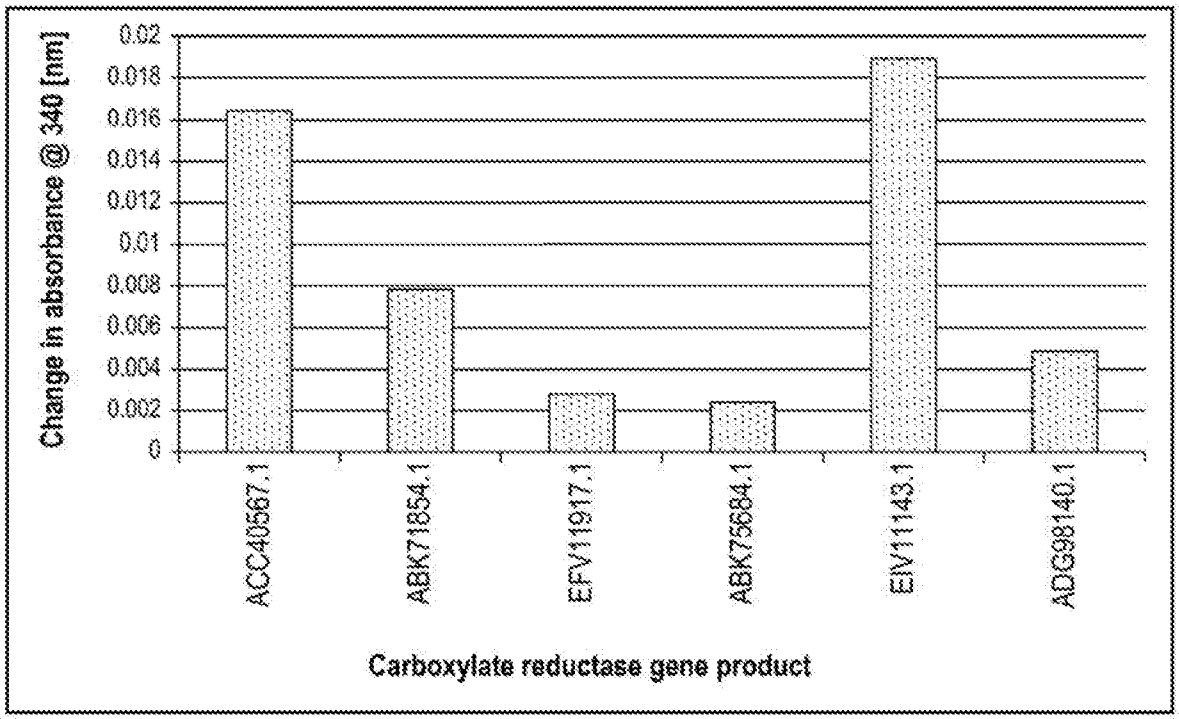
Change in absorbance @ 340 [nm]
0.02
0.018
0.016
0.014
0.012
0.01
0.008
0.006
0.004
0.002
0
ACC40567.1
ABK71854.1
EFV11917.1
ABK75684.1
EIV11143.1
ADG98140.1
Carboxylate reductase gene product

FIG. 10

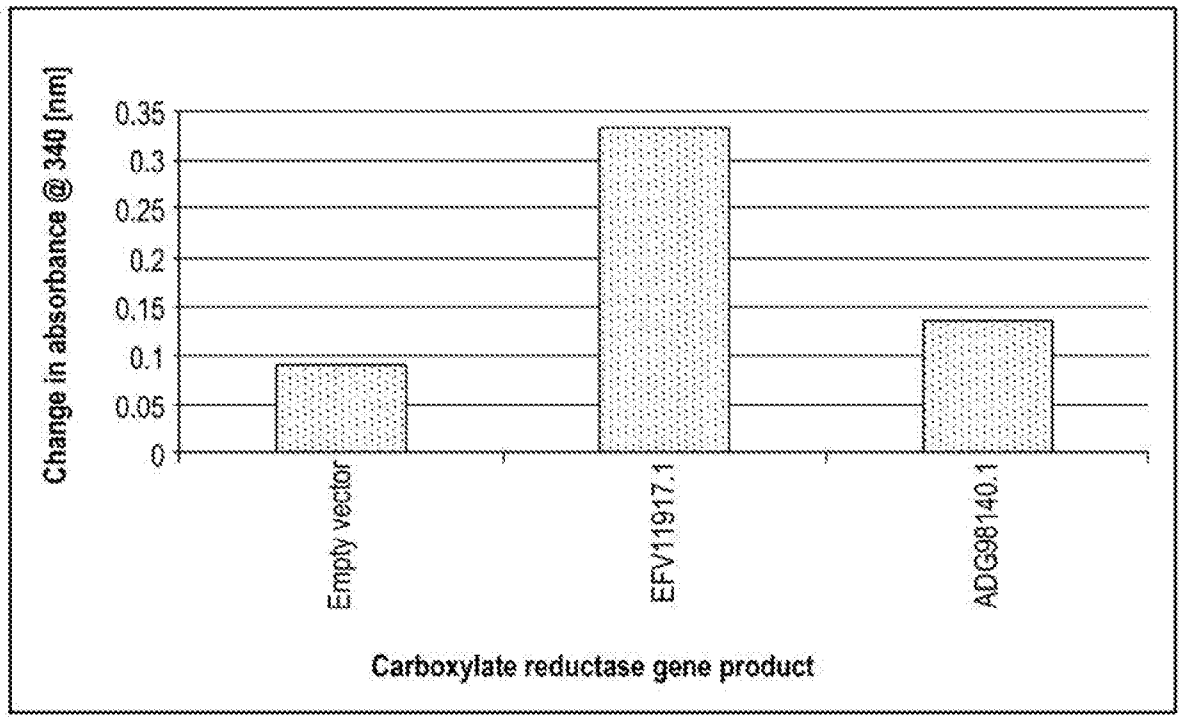
Change in absorbance @ 340 [nm]
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
Empty vector
EFV11917.1
ADG98140.1
Carboxylate reductase gene product

FIG. 11

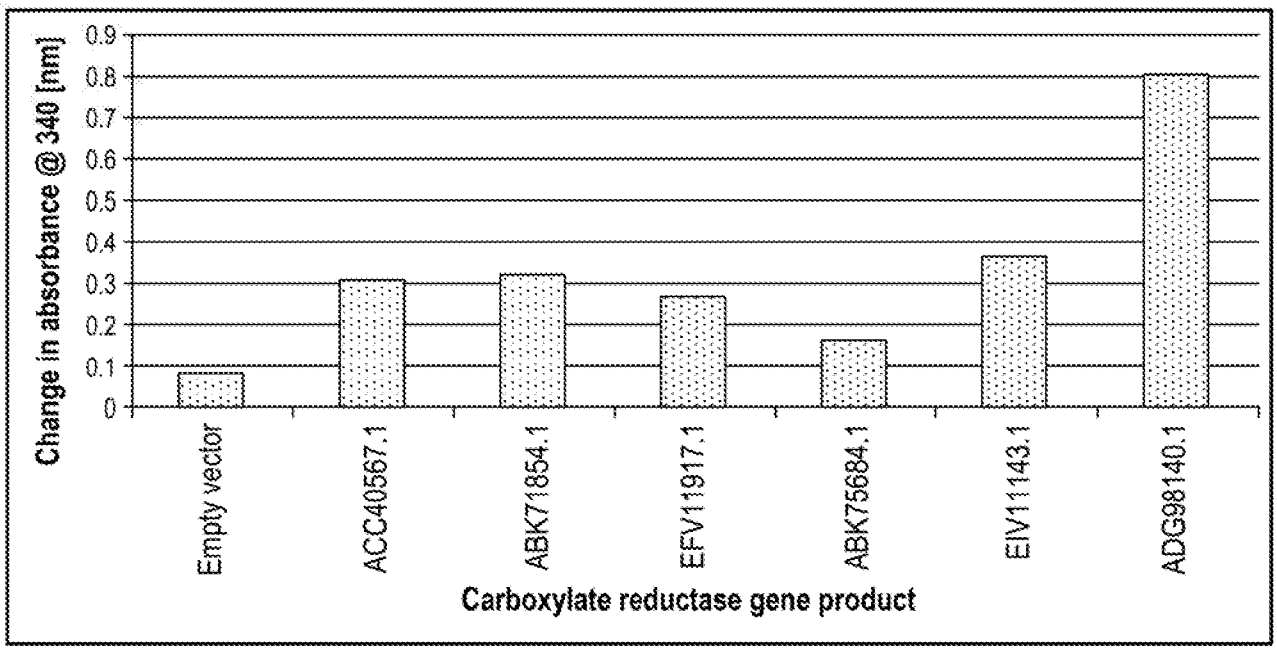
Change in absorbance @ 340 [nm]
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Empty vector
ACC40567.1
ABK71854.1
EFV11917.1
ABK75684.1
EIV11143.1
ADG98140.1
Carboxylate reductase gene product

FIG. 12

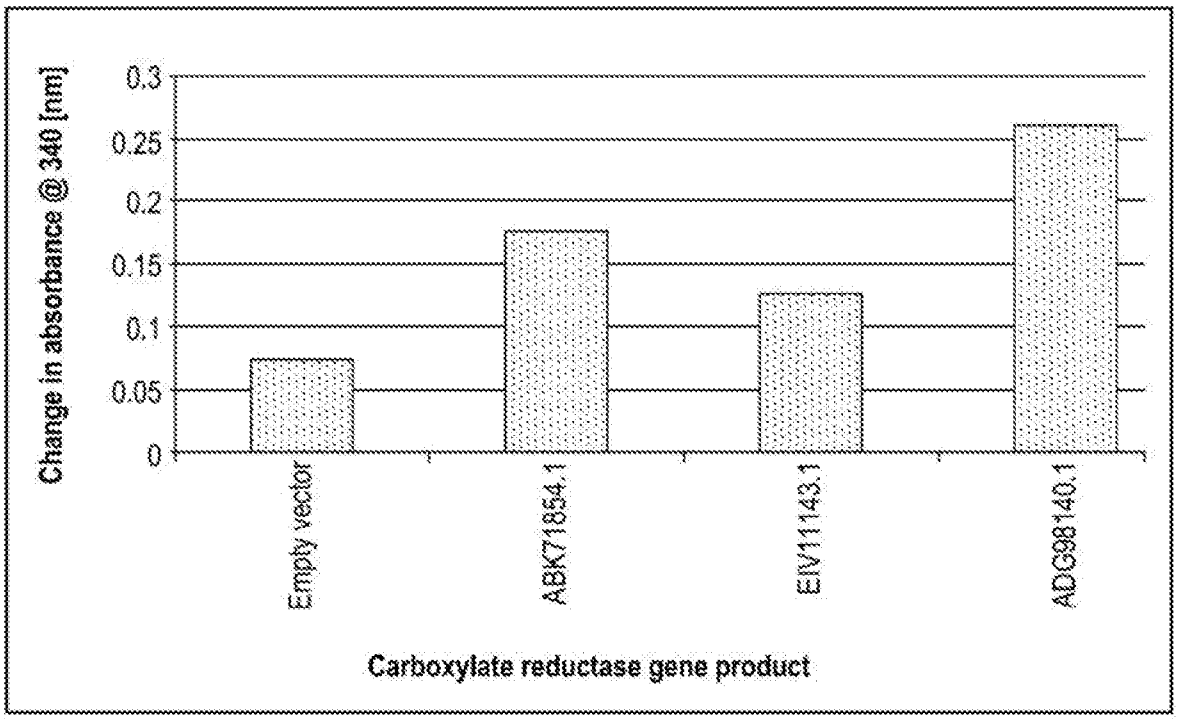
Change in absorbance @ 340 [nm]
0.3
0.25
0.2
0.15
0.1
0.05
0
Empty vector
ABK71854.1
EIV11143.1
ADG98140.1
Carboxylate reductase gene product

FIG. 13

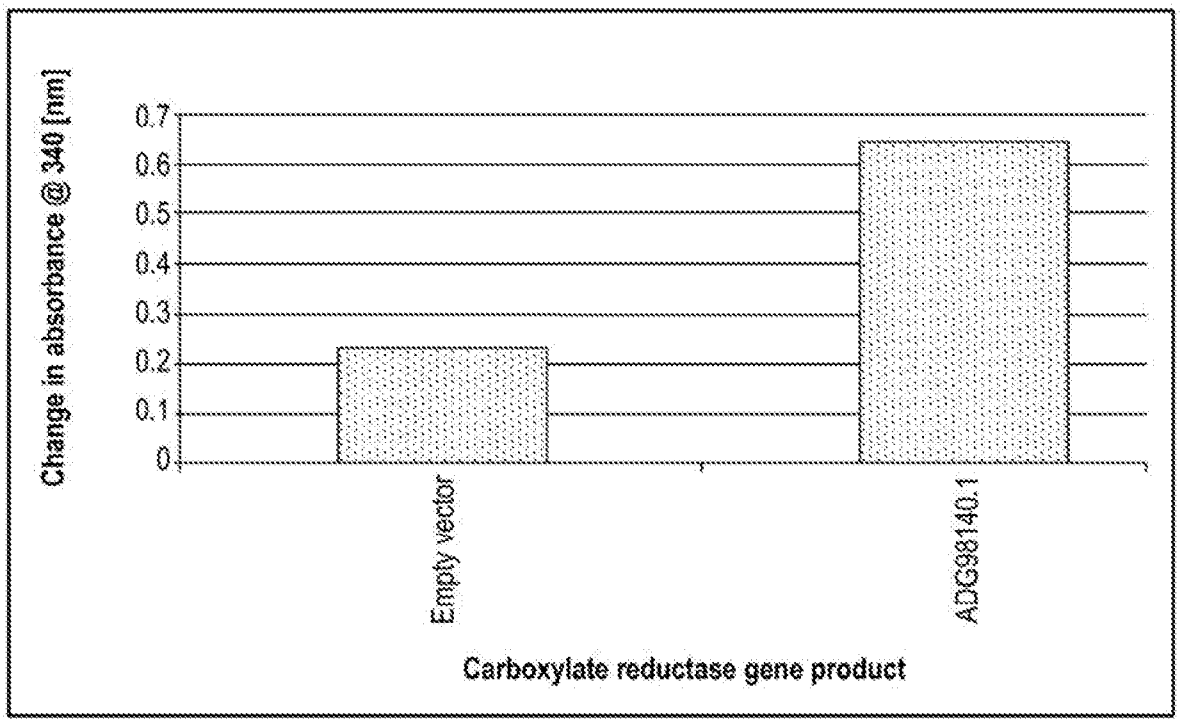
Change in absorbance @ 340 [nm]
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Empty vector
ADG98140.1
Carboxylate reductase gene product

FIG. 14

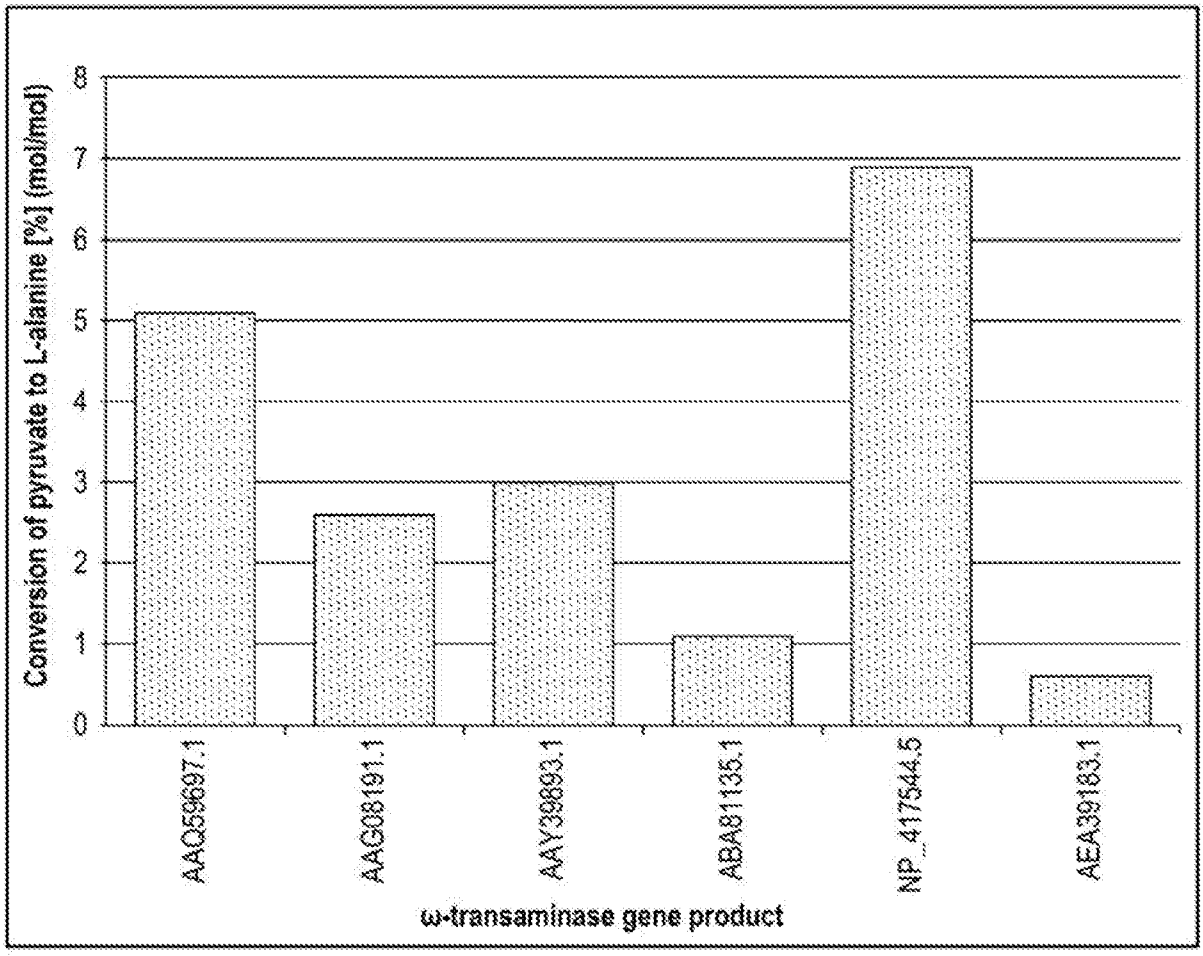
Conversion of pyruvate to L-alanine [%] (mol/mol)
0
1
2
3
4
5
6
7
8
AAQ59697.1
AAG08191.1
AAY39893.1
ABA81135.1
NP_417544.5
AEA39183.1
ω-transaminase gene product

FIG. 15

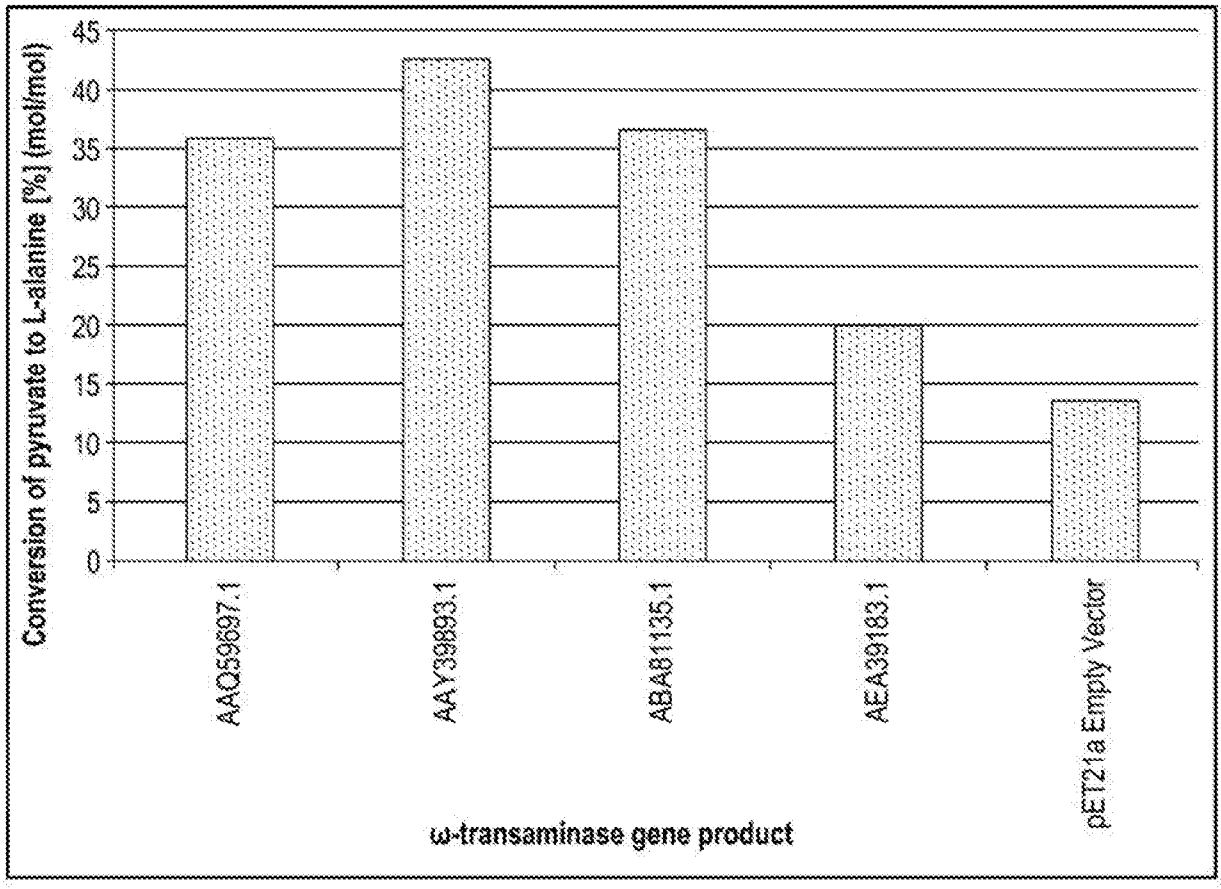
Conversion of pyruvate to L-alanine [%] (mol/mol)
45
40
35
30
25
20
15
10
5
0
AAQ59697.1
AAY39893.1
ABA81135.1
AEA39183.1
pET21a Empty Vector
ω-transaminase gene product

FIG. 16

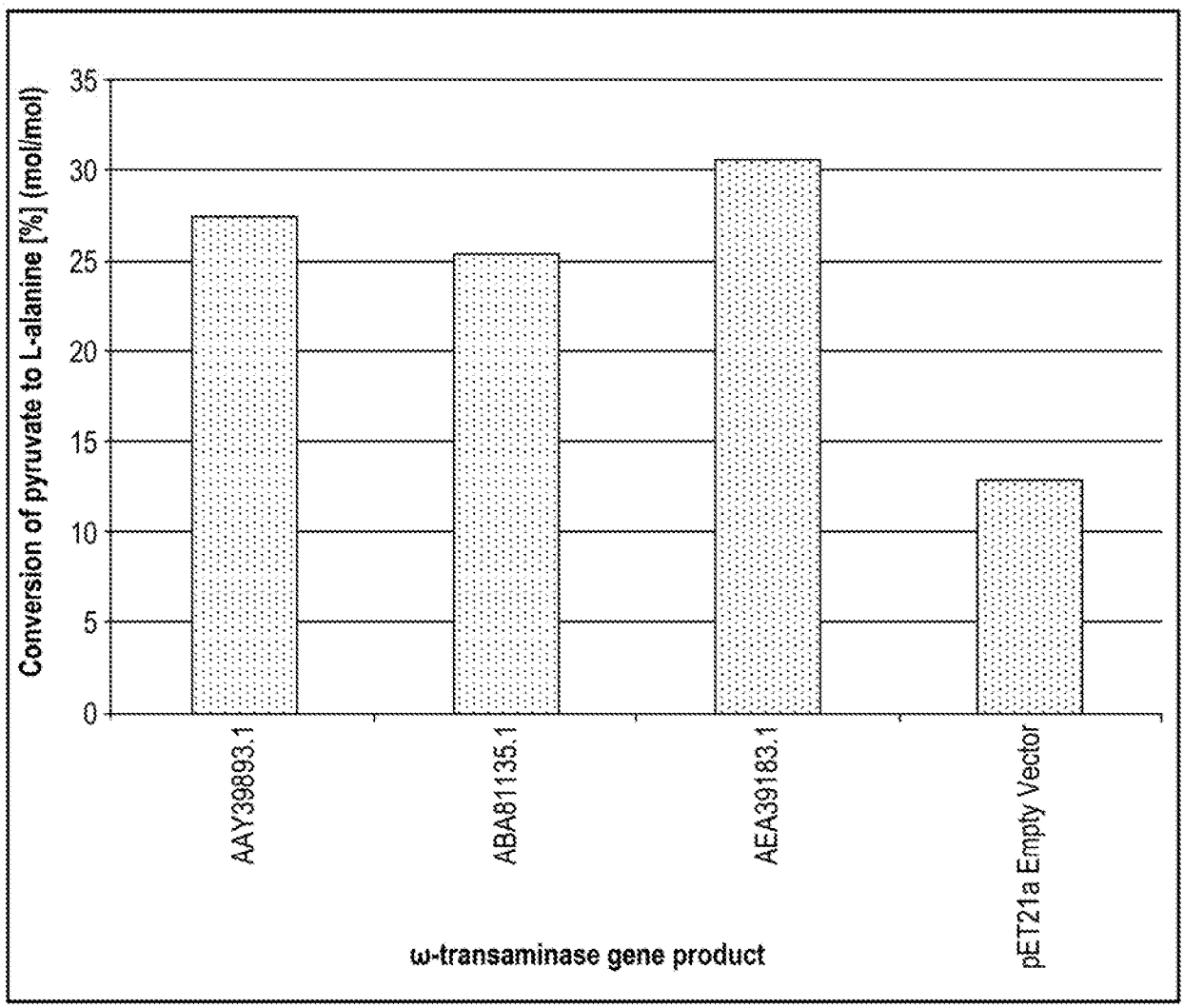
Conversion of pyruvate to L-alanine [%] (mol/mol)
35
30
25
20
15
10
5
0
AAY39893.1
ABA81135.1
AEA39183.1
pET21a Empty Vector
ω-transaminase gene product

FIG. 17

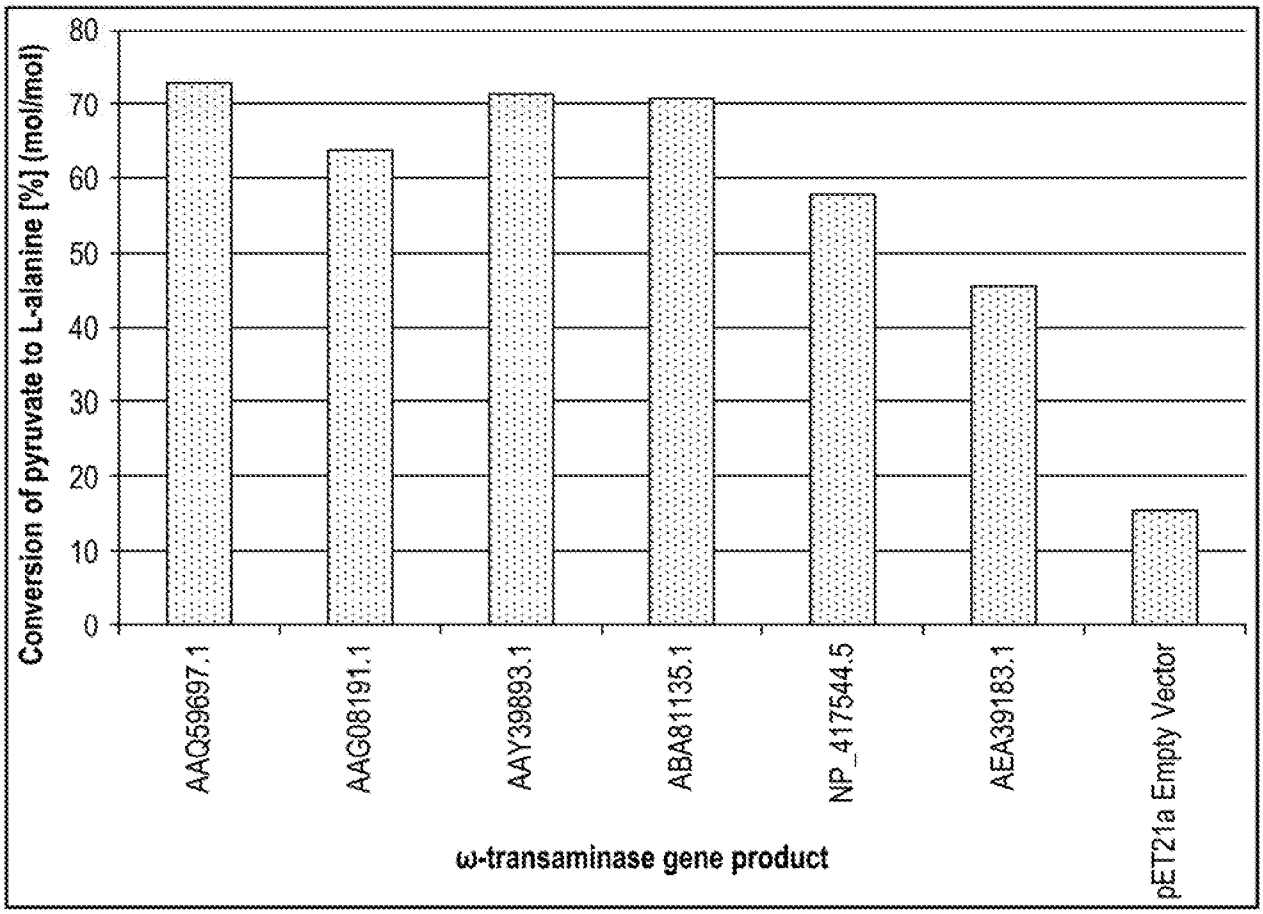
Conversion of pyruvate to L-alanine [%] (mol/mol)
80
70
60
50
40
30
20
10
0
AAQ59697.1
AAG08191.1
AAY39893.1
ABA81135.1
NP_417544.5
AEA39183.1
pET21a Empty Vector
ω-transaminase gene product

FIG. 18

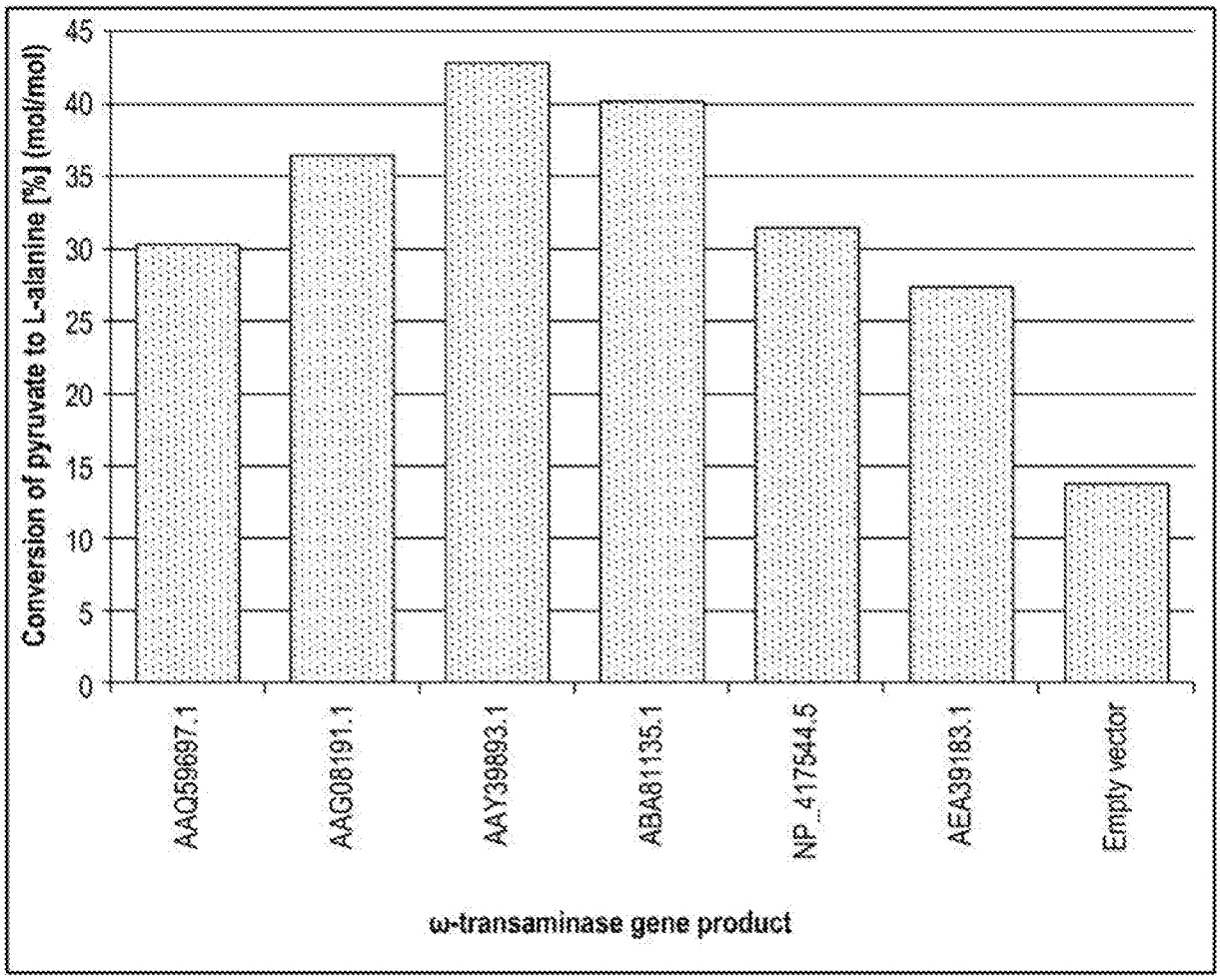
Conversion of pyruvate to L-alanine [%] (mol/mol)
45
40
35
30
25
20
15
10
5
0
AAQ59697.1
AAG08191.1
AAY39893.1
ABA81135.1
NP_417544.5
AEA39183.1
Empty vector
ω-transaminase gene product

FIG. 19

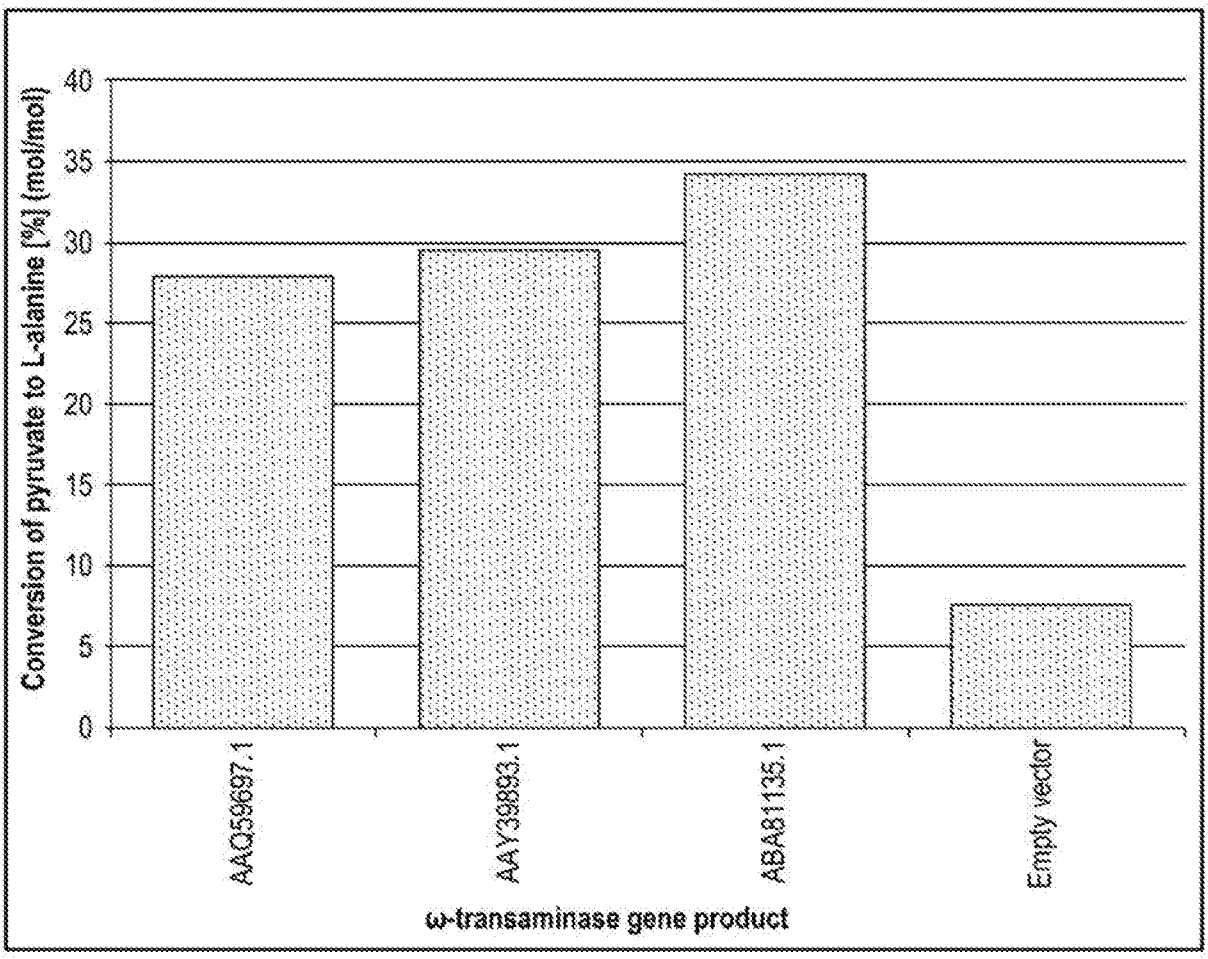
Conversion of pyruvate to L-alanine [%] (mol/mol)
40
35
30
25
20
15
10
5
0
AAQ59697.1
AAY39893.1
ABA81135.1
Empty vector
ω-transaminase gene product

FIG. 20

Sample ID	Analyte	Mwt	Peak Retention Time	Peak Area @ 260nm	Observed Mass (m/z)		Comments
		[g/mol]	[min]	[mAu]	Negative Mode (M-H)	Positive Mode (M+H)	
Reference Standard	pimeloyl-CoA methyl ester	923	6.161	31172.1	922	924	
Biotransformation at 1 [h] time point #1	pimeloyl-CoA methyl ester	923	nd	nd	nd	nd	No substrate detected after 1 [h]
	pimeloyl-CoA	910	5.577	3956.1	908	910	
Biotransformation at 1 [h] time point #2	pimeloyl-CoA methyl ester	923	nd	nd	nd	nd	No substrate detected after 1 [h]
	pimeloyl-CoA	910	5.579	3974.1	908	910	
Biotransformation at 1 [h] time point #3	pimeloyl-CoA methyl ester	923	nd	nd	nd	nd	No substrate detected after 1 [h]
	pimeloyl-CoA	910	5.577	4093.3	908	910	
Substrate only control (no enzyme) at 1 [h] time point	pimeloyl-CoA methyl ester	923	6.333	4577.9	922	924	No product detected after 1 [h]
	pimeloyl-CoA	910	nd	nd	nd	nd	

FIG. 21

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2017/043682

A. CLASSIFICATION OF SUBJECT MATTER
INV. C12N15/52 C12N15/90
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)
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

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO 2014/105788 A2 (INVISTA NORTH AMERICA SARL [US]) 3 July 2014 (2014-07-03) see pages 5-7, 10-17 and figures 1-5, claims 1-2 and 37-42	1-5, 60-62

-/--

[X] Further documents are listed in the continuation of Box C.

[X] 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
20 September 2017	01/12/2017

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	Vix, Olivier

3

Form PCT/ISA/210 (second sheet) (April 2005)

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	V. AGARWAL ET AL: "Structure of the enzyme-acyl carrier protein (ACP) substrate gatekeeper complex required for biotin synthesis", PROCEEDINGS NATIONAL ACADEMY OF SCIENCES PNAS, vol. 109, no. 43, 8 October 2012 (2012-10-08), pages 17406-17411, XP055408191, US ISSN: 0027-8424, DOI: 10.1073/pnas.1207028109 see abstract and fig. 1 and pages 17406-17407	1-5, 60-62
Y	BOWER STANLEY ET AL: "Cloning, sequencing, and characterization of the Bacillus subtilis biotin biosynthetic operon", JOURNAL OF BACTERIOLOGY, AMERICAN SOCIETY FOR MICROBIOLOGY, US, vol. 178, no. 14, 1 July 1996 (1996-07-01) , pages 4122-4130, XP002229556, ISSN: 0021-9193 biotin biosynthetic operon involving the BioF genes	1-5, 60-62
Y	WO 2016/106247 A1 (INVISTA NORTH AMERICA S Á R L [US]; INVISTA TECH SARL [CH]) 30 June 2016 (2016-06-30) see claims and pages 19-25	1-5, 60-62
Y	WO 2015/175698 A1 (INVISTA NORTH AMERICA S Á R L [US]) 19 November 2015 (2015-11-19) see claims 1-12 and pages 3-14	1-5, 60-62
Y	MADELYN M. SHAPIRO ET AL: "Remarkable Diversity in the Enzymes Catalyzing the Last Step in Synthesis of the Pimelate Moiety of Biotin", PLOS ONE, vol. 7, no. 11, 9 November 2012 (2012-11-09), page e49440, XP055113391, DOI: 10.1371/journal.pone.0049440 see abstract and pages e49440 1-2, Fig.1-2 and pages 7-9	1-5, 60-62

-/--

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2017/043682

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 6 083 712 A (BIRCH OLWEN [CH] ET AL) 4 July 2000 (2000-07-04) The query sequence SEQ ID NO:23 has 100.00 % identity (100.00 % similarity) over 384 positions in a common overlap (range (q:s): 1-384:1-384) with subject USPOP:AAE44786 (length: 384) from US6083712-A published on 2000-07-04.	1-5, 60-62
Y	WO 2013/003744 A2 (INVISTA TECHONOLOGIES S A R L [CH]; PEARLMAN PAUL S [US]; CHEN CHANGLI) 3 January 2013 (2013-01-03) see [0073], [00183]-[00188] and Fig. 1-6	1-5, 60-62
A	US 6 277 609 B1 (EDDY CHRISTINA K [US]) 21 August 2001 (2001-08-21) in particular see columns 4, 6-9, 14-15 and claims 1-9, 16, 23-24	1-5, 60-62

3

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2017/043682

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.:

1-5, 60-62

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: 1-5, 60-62

   A method for regulating biosynthesis of at least one C7 building block using a a pathway having a pimeloyl-ACP internediate, said method comprising the step of downregulating the activity of BioF enzyme

   ---

2. claims: 6-59, 63

   A method for regulating biosynthesis of at least one C7 building block using a a pathway having a pimeloyl-ACP internediate, said method comprising the steps of downregulating the activity of BioF enzyme and overexpressing BioW and a CoA-specific BioF.

   ---

3. claims: 65(completely); 64, 74(partially)

   A nucleic acid construct or expression vector comprising a polypeptide having 8-amino-7-oxononanoate synthase activity (having at least 70% sequence identity to SEQ ID N°23).

   ---

4. claims: 66(completely); 64, 74(partially)

   A nucleic acid construct or expression vector comprising a polypeptide having thioesterase activity (having at least 70% sequence identity to SEQ ID N°22).

   ---

5. claims: 67(completely); 64, 74(partially)

   A nucleic acid construct or expression vector comprising a polypeptide having carboxylate reductase activity (having at least 70% sequence identity to SEQ ID N°2-7).

   ---

6. claims: 68(completely); 64, 74(partially)

   A nucleic acid construct or expression vector comprising a polypeptide having ?-transaminase activity (having at least 70% sequence identity to SEQ ID N°8-13).

   ---

7. claims: 69(completely); 64, 74(partially)

   A nucleic acid construct or expression vector comprising a polypeptide having phosphopantetheinyl transferase activity (having at least 70% sequence identity to SEQ ID N°14-15).

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

8. claims: 70(completely); 64, 74(partially)

A nucleic acid construct or expression vector comprising a polypeptide having pimeloyl-ACP methyl ester esterase activity (having at least 70% sequence identity to SEQ ID N°17).

---

9. claims: 71(completely); 64, 74(partially)

A nucleic acid construct or expression vector comprising a polypeptide having 6-carboxyhexanoate-CoA ligase activity (having at least 70% sequence identity to SEQ ID N°18).

---

10. claims: 72(completely); 64, 74(partially)

A nucleic acid construct or expression vector comprising a polypeptide having 8-amino-7-oxononanoate synthase activity (having at least 70% sequence identity to SEQ ID N°19).

---

11. claims: 73(completely); 64, 74(partially)

A nucleic acid construct or expression vector comprising a polypeptide having malonyl-ACP O-methyltransferase activity (having at least 70% sequence identity to SEQ ID N°20).

---

12. claim: 75

A non-naturally occurring biochemical network comprising a 5-hydroxypentanoyl-CoA, an exogenous nucleic acid encoding a polypeptide having the activity of a ?-ketothiolase classified under EC. 2.3.1, and a 3-oxo-7-hydroxyheptanoyl-CoA.

---

13. claim: 76

A bio-derived, bio-based or fermentation-derived product, wherein said product comprises:i. a composition comprising at least one bio-derived, bio-based or fermentation-derived compound produced or biosynthesized according to any one of claims 1-75 or or any one of FIGS. 1-21, or any combination thereof;ii. a bio-derived, bio-based or fermentation-derived polymer comprising the bio-derived, bio-based or fermentation-derived composition or compound of i., or any combination thereof;iii. a bio-derived, bio-based or fermentation-derived resin comprising the bio-derived, bio-based or fermentation-derived compound or bio-derived, bio-based or fermentation-derived composition of i. or any combination thereof or the bio-derived, bio-based or fermentation-derived polymer of ii. or any combination thereof;iv. a molded substance obtained by molding the

bio-derived, bio-based or fermentation-derived polymer of ii. or the bio-derived, bio-based or fermentation-derived resin of iii., or any combination thereof;v. a bio-derived, bio-based or fermentation-derived formulation comprising the bio-derived, bio-based or fermentation-derived composition of i., bio-derived, bio-based or fermentation-derived compound of i., bio-derived, bio-based or fermentation-derived polymer of ii., bio-derived, bio-based or fermentation-derived resin of iii., or bio-derived, bio-based or fermentation-derived molded substance of iv, or any combination thereof; orvi. 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 of i., bio-derived, bio-based or fermentation-derived compound of i., bio-derived, bio-based or fermentation-derived polymer of ii., bio-derived, bio-based or fermentation-derived resin of iii., bio-derived, bio-based or fermentation-derived formulation of v., or bio-derived, bio-based or fermentation-derived molded substance of iv., or any combination thereof.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/US2017/043682

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2014105788 A2	03-07-2014	CN 105189770 A	23-12-2015
		EP 2938731 A2	04-11-2015
		US 2014193862 A1	10-07-2014
		US 2017191096 A1	06-07-2017
		WO 2014105788 A2	03-07-2014
WO 2016106247 A1	30-06-2016	EP 3237634 A1	01-11-2017
		US 2016201097 A1	14-07-2016
		WO 2016106247 A1	30-06-2016
WO 2015175698 A1	19-11-2015	CN 106715701 A	24-05-2017
		EP 3143149 A1	22-03-2017
		US 2017233777 A1	17-08-2017
		WO 2015175698 A1	19-11-2015
US 6083712 A	04-07-2000	AT 163198 T	15-02-1998
		AU 4820293 A	26-04-1994
		CA 2145400 A1	14-04-1994
		CZ 9500809 A3	13-09-1995
		DE 59308149 D1	19-03-1998
		DK 0667909 T3	23-09-1998
		EP 0667909 A1	23-08-1995
		EP 0798384 A1	01-10-1997
		ES 2112995 T3	16-04-1998
		FI 951547 A	31-03-1995
		HU 219827 B	28-08-2001
		JP 3566287 B2	15-09-2004
		JP 3577075 B2	13-10-2004
		JP H08501694 A	27-02-1996
		JP 2004000283 A	08-01-2004
		PL 179166 B1	31-07-2000
		PL 308301 A1	24-07-1995
		SK 42095 A3	08-11-1995
		US 6083712 A	04-07-2000
		WO 9408023 A2	14-04-1994
WO 2013003744 A2	03-01-2013	BR 112013033672 A2	14-03-2017
		CN 104011216 A	27-08-2014
		EP 2726627 A2	07-05-2014
		JP 2014525741 A	02-10-2014
		JP 2017195890 A	02-11-2017
		US 2014242655 A1	28-08-2014
		US 2015337275 A1	26-11-2015
		US 2017298396 A1	19-10-2017
		WO 2013003744 A2	03-01-2013
US 6277609 B1	21-08-2001	NONE	